THE STUDY

What works (and what doesn't) in the stock market

A comprehensive synthesis of what 150 years of market history actually teach us.

Drawn from eighteen research papers; condensed into six principles and twenty algorithms. Updated as new research is completed.

EDITION 2.0 · MAY 2026

Carl C.G Eriksson
FOUNDER · PROREALALGOS

What works (and what doesn't) in the stock market

A comprehensive synthesis of what 150 years of market history actually teach us.

Edition 2.0 · First public edition · May 2026

This document is a living synthesis. It is updated as new research is completed at ProRealAlgos. Major editions are released approximately annually. Minor editions, numbered as point releases, are issued between major editions to incorporate single new papers or significant corrections.

Carl C.G Eriksson, with research contributions from M. Lindberg (2017), P. Andersson (2018), Jon Engström (2024), and J. Wilson (2025).

Stockholm, Sweden.

This is not investment advice.

FOREWORD

Why this study exists

Most material written about systematic trading exists to sell something. Courses, signals, account-management services, software with monthly subscriptions. The financial incentive points one way, and the content drifts toward that direction whether the writer notices or not. After twenty years of teaching, building, and trading systematic strategies, I had accumulated enough irritation about the gap between what was being sold and what the data actually said that I started writing things down.

The first paper was published in February 2017 as a single PDF I sent to about thirty people. The intention was modest, work through the equity risk premium one more time, this time honestly, and put the conclusions somewhere a person could find them. Seventeen papers later, the work had grown into something else. Patterns recurred. A small number of conclusions kept holding up across topics, time periods, and asset classes. A smaller number of methodological convictions emerged from the wreckage of strategies that did not survive walk-forward validation.

This study is the synthesis. It draws on eighteen papers, three of them with co-authors, totalling roughly five hundred pages of detailed empirical work. It compresses that work into a single argument: that the stock market, examined over a hundred and fifty years of out-of-sample data, behaves the way it behaves; that a small number of strategies have survived honest scrutiny while many more have not; that the gap between paper performance and live performance is real, large, and almost always understated by the people selling the paper performance; and that six principles, applied consistently, produce systematic trading results that are not exciting on a monthly basis but that compound credibly over years.

I have tried to keep the writing honest about what we know and what we do not. There are passages where the conclusion is uncomfortable, particularly for someone who makes their living selling algorithmic trading systems. I have not removed those passages.

This is the first public edition. It will not be the last. The research at ProRealAlgos has not stopped at the eighteen papers behind it. New work is in progress; some of it will refine what is here, some of it will revise it, some of it will discontinue lines of inquiry that turned out not to support a publishable conclusion. The structure of this study is built to accommodate that evolution. Future editions will fold new findings into the existing chapters and principles, and a revision history at the back will document what changed and why.

This is a long document. It is meant to be read once and consulted as needed. It is the document that ProRealAlgos is built on.

Carl C.G Eriksson

Stockholm, May 2026

Contents

FRONT MATTER

	Foreword	v
	A note on what this study is and isn't	ix
	How this study is maintained	xi
	Edition history	xiii
	The eighteen papers behind it	xv

PART I · WHAT THE DATA SHOWS

1	The shape of equity returns over fifteen decades	3
2	Bear markets and inflation regimes	9
3	Diversification's actual track record	15
4	What predicts ten-year forward returns	21

PART II · STRATEGIES THAT HAVE SURVIVED

5	Trend-following on equity indices	27
6	Carry across asset classes	33
7	Dollar-cost averaging versus lump sum	39
8	The concentration trap in single stocks	43

PART III · THE SIX PRINCIPLES

I	Survivorship is a lie	49
II	Math doesn't care about feelings	53
III	Position size is the strategy	57
IV	Drawdowns are tuition	61
V	One trader at a time	65
VI	If we don't run it, we don't sell it	69

PART IV · FROM PRINCIPLES TO TWENTY ALGORITHMS

9	The four families and how they compose	75
10	What the algorithms do not do	81

What this study answers, and what it doesn't 85

References · The eighteen papers 89

Forthcoming research and planned editions 93

How to track updates 97

A NOTE BEFORE READING

What this study is, and isn't

It is useful, at the start of a long document, to be honest about what the document does and does not do.

What this study is

A synthesis of eighteen pieces of empirical work conducted between 2017 and 2026. It draws on roughly one hundred and fifty years of stock market data, examined across multiple asset classes, multiple decades, and multiple methodological lenses. It distils six principles that have proven robust across the work, and it points toward the twenty algorithms that ProRealAlgos has built on those principles.

It is opinionated. Where evidence has accumulated in favour of a conclusion, this study states the conclusion without hedging. Where evidence is mixed or weak, it says so. Where I am personally uncertain, I have said that too.

It is written for the reader who wants the conclusion and the work behind the conclusion. There is no shortage of trading literature that delivers a conclusion without the work. There is also no shortage of academic literature that delivers the work without ever quite arriving at a usable conclusion. The intent here is to do both.

What this study is not

It is not investment advice. The work is generic. It does not consider your tax situation, your liquidity needs, your time horizon, or the rest of your portfolio. Use it as a foundation for your own thinking, not as a substitute.

It is not a recipe. The six principles do not, applied mechanically, produce profits. They produce a framework inside which a person might construct strategies that have a chance of working. The construction is still hard.

It is not a complete account of what works. The eighteen papers behind this study examine particular questions in particular ways. Important questions remain unexamined. The conclusions are bounded by the questions we chose to ask.

It is not new. Almost every empirical claim in this study can be traced to earlier work by academics and practitioners I have cited throughout. What is original is the synthesis, the specific weight given to each finding, and the conclusion that emerges when the findings are read together.

A NOTE ON MAINTENANCE

How this study is maintained

This document is a living synthesis. It is not a one-time publication, and the conclusions in it are expected to change as new research is completed.

The intent

The work behind ProRealAlgos has not stopped at the eighteen papers in this edition. New research is continuously in progress. Some of it refines findings that already appear in this study. Some of it points at new principles or reveals limitations in existing ones. Some lines of inquiry will be discontinued because the empirical work does not support a publishable result.

This document is built to accommodate that evolution. Rather than freezing the synthesis at a single point in time and letting subsequent research drift away from it, we maintain the study itself as the canonical reference, updated as the underlying evidence changes.

Edition cadence

Major editions are released approximately annually. A major edition incorporates all the year's completed research, restructures chapters where new findings warrant it, and revises principles where the underlying evidence has shifted. Major editions are numbered as integers (the current edition is 2.0; future major editions will be 3.0, 4.0, and so on) and are accompanied by a substantial revision history.

Minor editions, numbered as point releases (2.1, 2.2, 2.3), are issued between major editions when a single new paper has been completed and merits integration before the next annual cycle. Minor editions are narrower in scope and shorter in revision history but follow the same structural conventions.

How new research is folded in

When a new paper is completed and has passed walk-forward validation, it is integrated into this study in one of three ways. Findings that strengthen or refine an existing chapter are folded into the relevant chapter and noted in the revision history. Findings that contradict a prior conclusion drive a revision to the conclusion, with the old conclusion and the reason for revision documented in the history. Findings that point to a new principle, or that require a structural change to the document's organisation, drive a major edition.

We do not silently revise. Where the conclusion of an earlier edition has been changed, the change is documented. The intent is for readers of a later edition to understand precisely what shifted and why, not to be left guessing whether their memory of an earlier reading is accurate.

Stability of conclusions

The conclusions in this study fall into three categories of expected stability. The empirical findings drawn directly from the long-sample data (Part I and most of Part II) are the most stable. Adding new years of data marginally shifts the point estimates but rarely changes the qualitative direction. The six principles in Part III are the second-most stable; they are methodological convictions that we expect to refine but not to fundamentally revise. The specific implementation choices in Part IV are the least stable; the four families of algorithms will expand and contract as new research informs which return sources warrant production deployment.

Readers can rely on the empirical chapters of the study aging gracefully. The Part IV material is more provisional by construction.

EDITION HISTORY

Six years of revisions

The synthesis behind this study did not arrive whole. It was assembled across six years, with a new internal version issued shortly after each of the eighteen papers was completed. The table below documents the full sequence, from the first internal assembly in March 2020 through edition 2.0 in May 2026, plus the planned cadence going forward. On average across those six years, the document has been updated once every four months.

Versions 0.1, 0.2, and 1.0 through 1.11 were internal. They were not publicly distributed, but they were archived as each new paper was integrated so that the evolution of the synthesis remains traceable. Edition 2.0 is the first edition released publicly. The jump from 0.2 to 1.0 in May 2021 marks the point at which the document stopped being a collection of papers organised together and began functioning as a synthesis built around the principles described in Part III.

Version	Released	Triggered by	Status
0.1	MAR 2020	Initial assembly. Papers 04, 05, 07, 13, 15. Internal only.	SUPERSEDED
0.2	JUL 2020	Paper 06 integrated (DCA versus lump sum).	SUPERSEDED
1.0	MAY 2021	Paper 14 integrated. Principle II first formalised. First version organised around a principle.	SUPERSEDED
1.1	JAN 2022	Paper 08 integrated (single-stock concentration).	SUPERSEDED
1.2	OCT 2022	Paper 02 integrated. Chapter 2 substantially rewritten.	SUPERSEDED
1.3	APR 2023	Paper 10 integrated (60/40 verdict).	SUPERSEDED
1.4	JUN 2023	Paper 03 integrated. Chapter 3 restructured around inflation-regime evidence.	SUPERSEDED
1.5	NOV 2023	Paper 17 integrated. Principle V first formalised.	SUPERSEDED
1.6	FEB 2024	Paper 16 integrated. Principle IV first formalised.	SUPERSEDED
1.7	SEP 2024	Paper 09 integrated. Chapter 6 (cross-asset carry) introduced.	SUPERSEDED
1.8	AUG 2025	Paper 11 integrated (gold as hedge). Chapter 3 extended.	SUPERSEDED
1.9	OCT 2025	Paper 18 integrated. Principle VI formalised. Major rewrite of Part IV.	SUPERSEDED
1.10	JAN 2026	Paper 12 integrated (CAPE predictors). Chapter 4 added.	SUPERSEDED
1.11	APR 2026	Pre-publication review draft. Content frozen pending external review.	SUPERSEDED
2.0	MAY 2026	**First public edition.** Paper 01 integrated. Full polish, structure final, references and forthcoming research sections added.	CURRENT
2.1	Q3 2026	Planned. Papers 19 and 20 (multi-asset trend; volatility-targeted sizing).	PLANNED
2.2	Q1 2027	Planned. Papers 21 and 22, plus implementation-shortfall refresh.	PLANNED
3.0	MAY 2027	Annual major refresh. Cumulative research May 2026 – April 2027.	PLANNED

Conclusion-level revisions between public editions

This section documents revisions where a conclusion stated in an earlier public edition has been changed in a later one. Edition 2.0 is the first public edition; no public revisions have occurred yet. Future editions will populate this section as revisions accumulate. Internal-version revisions are summarised in the table above rather than enumerated here.

> REVISIONS IN THIS EDITION
>
> None. This is the first public edition.

Readers wanting to be notified when a new edition is released can register at `prorealalgos.com/study`. Notification is opt-in and limited to study updates; we do not use the list for marketing.

BEHIND THIS STUDY

The eighteen papers

The work below is the empirical scaffolding underneath this study. Each paper addresses a single, narrow question. Full references appear at the back of this document.

01	MAY 2026	The "best days" myth *Why missing the worst days matters more than missing the best*
02	OCT 2022	A century and a half of bear markets *Depth, duration, and recovery across fifteen U.S. equity drawdowns*
03	JUN 2023	Equities, bonds, and the inflation regime *What 152 years of return data reveal about the asset-class consequences of price-level instability*
04	MAR 2019	The 12-month moving average as a regime filter *Returns, volatility, and drawdowns conditional on a single trend signal*
05	FEB 2017	Where the equity risk premium comes from *A Gordon-style decomposition of U.S. equity returns*
06	JUL 2020	Lump sum, dollar cost, and the cost of waiting *1,097 ten-year simulations on what drip-feeding actually buys you*
07	NOV 2018	A century and a half of trend-following *What a 12-month signal applied to U.S. equities does for return, volatility, and drawdown*
08	JAN 2022	The concentration problem in single-stock returns *Why most stocks underperform Treasury bills*
09	SEP 2024	Carry, everywhere *What thirty years of cross-asset evidence reveals about a single risk premium*
10	APR 2023	The 60/40 verdict *Why a portfolio that worked for forty years failed in 2022, and what to do about it*
11	AUG 2025	Gold as hedge: fact and folklore *What 50 years of post-Bretton-Woods data reveal about the metal's actual role in a portfolio*
12	JAN 2026	What predicts the next decade *CAPE, real rates, and other long-run forecasters tested across 145 years*
13	AUG 2018	The graveyard of strategies *A survivorship audit of trading systems published between 1995 and 2010*
14	MAY 2021	Override and underperform

The empirical cost of human discretion in rule-based trading systems

15	MAR 2020	Position sizing dominates strategy selection *10,000 simulated equity curves on what determines whether a strategy enriches or ruins its trader*
16	FEB 2024	The drawdown-abandonment curve *When traders quit, and what it costs them*
17	NOV 2023	Knowledge transfer in expert systems *What novice users need to know to capture the documented returns of a designed system*
18	OCT 2025	Implementation shortfall in retail systematic strategies *Signal versus execution: paper-traded and live-traded performance compared*

Part I

What the data shows

Before we discuss what to do about the market, it is worth being clear about what the market actually looks like across a hundred and fifty years.

1 The shape of equity returns over fifteen decades

2 Bear markets and inflation regimes

3 Diversification's actual track record

4 What predicts ten-year forward returns

CHAPTER ONE

The shape of equity returns over fifteen decades

The U.S. equity market produces a long-run real return of around six and a half percent a year, but that average is the least informative thing about it. The interesting facts are about the distribution.

DRAWN FROM Paper 01 (May 2026) · Paper 05 (Feb 2017) · Paper 02 (Oct 2022)

A reader new to systematic investing typically encounters a single number first, the annualised real return of the U.S. equity market, somewhere between six and seven percent depending on the start date and the dataset. The number is correct. It is also, taken on its own, the least useful piece of information one can have about the market, because almost nothing about how an investor actually experiences equity returns is captured by an average.

The first useful question is: where does that six and a half percent come from? The Gordon-style decomposition in our 2017 paper splits the long-run U.S. equity return into three additive components, dividend yield (the cash flow you receive while holding), real earnings growth (the rate at which the cash flows expand), and the change in valuation multiple (how much more or less the market is willing to pay per dollar of earnings at the end of the period than at the start). Across the full 1871–2016 sample, the average dividend yield contributes 4.3 percent, real earnings growth contributes 1.9 percent, and PE10 revaluation contributes 0.5 percent, summing to approximately 6.7 percent.

The relative weight of these components has shifted dramatically across the sample. Before 1980, equity returns were predominantly cash flows. Dividend yields averaged above 4 percent and earnings grew with the broad economy. After 1980, the contributions inverted. Dividend yields fell to under 2 percent. Multiple expansion, which had contributed essentially nothing across the prior century, became the dominant driver of total returns during the 1980s and 1990s. The post-2000 period is more ambiguous. PE10 has hovered between 24 and 38, well above the historical median of 16. Whether the rest of the twenty-first century delivers anything resembling the post-1980 return profile depends on how much further multiples can expand from already-stretched levels.

The distribution of years

The second useful question is: what is the variance around the average? Across 152 calendar years of U.S. equity returns (1872–2023), the median year produces a real total return of approximately 9 percent. But the standard deviation is roughly 18 percent. The full empirical distribution stretches from about minus 38 percent (1931) to about plus 53 percent (1933, an unusual recovery year). Of those 152 years, 42 produced negative real returns, twelve produced losses exceeding 15 percent, and four produced losses exceeding 25 percent.

This matters because investor experience is not captured by the average. It is captured by the path. A strategy with a real expected return of 6.5 percent and a 20 percent annual standard deviation produces, across a ten-year holding period, a wide range of realised outcomes. The 5th-percentile ten-year cumulative return is approximately -5 percent. The 95th-percentile is approximately 320 percent. The investor who holds equities through one specific decade gets one specific outcome from this distribution, and the variance around the expected value is wider than most investors emotionally prepare for.

The asymmetry of best days and worst days

A claim the financial industry routinely makes is that missing a small number of "best days" destroys long-term returns. The most common version uses 25 years of S&P 500 monthly total-return data: if you had held the index from January 2000 through December 2024, your annualised return is approximately 7.1 percent, but if you missed the ten single best trading days, your return drops to approximately 3.2 percent. The implication offered alongside this statistic is that investors should never time the market because the cost of missing the upside is too high.

The 2026 paper revisits this claim using 155 years of S&P 500 monthly total-return data (1871–2026). The headline statistic replicates. Across the long sample, the ten best months explain roughly 64 percent of all positive months' contribution, and missing them reduces a buy-and-hold portfolio's terminal value by 81 percent. But the paper continues, and the continuation is what the industry's version of the chart usually omits.

BEST 10 MONTHS MISSED	WORST 10 MONTHS MISSED	BOTH MISSED
81%	**7.1×**	**98%**
terminal value reduction	terminal value multiplier	of buy-and-hold preserved

If, instead of missing the best ten months, an investor missed the worst ten months, their terminal value would increase by a factor of approximately seven. The asymmetry is real: the worst days punish portfolios more than the best days reward them, because the same percentage move down requires a larger percentage move up to recover. Excluding both the best ten and the worst ten months produces a portfolio with 98 percent of the buy-and-hold's terminal value. The "best days" claim is empirically correct but it is the wrong moral. The moral is that the cluster of large positive months and the cluster of large negative months substantially offset, and any framework that systematically reduces participation during high-volatility regimes captures more of the upside-without-downside than it costs in missed best-day exposure.

This is the empirical foundation underneath the trend-following work in Part II. The market's worst months are not randomly distributed across time. They cluster during regimes that a coarse trend signal can identify, after the fact, with reasonable consistency.

What investors actually need to know

From the 152 years of return data, four facts seem to us to deserve more emphasis than they typically receive.

The first. The equity premium is not a continuous monthly subsidy. It is a payout for tolerating the empirical distribution of monthly outcomes. Across the long sample, that distribution includes routine months of plus or minus 5 percent, occasional months of plus or minus 15 percent, and rare months that exceed 25 percent in either direction. The investor who emotionally requires a smooth equity curve will exit at the wrong time and will earn substantially less than the long-run premium. This is documented in detail in paper 16.

The second. The average is a poor description of what investors experience. The mean of 6.5 percent real is a statistical artefact of averaging across many decades. Individual decades have produced realised real returns ranging from approximately minus 1 percent annualised (1970s) to approximately plus 14 percent annualised (1990s). The investor's actual experience is whichever decade they happened to live through. Forecasting which decade you are in matters more than knowing the long-run mean.

The third. Drawdowns are deeper and longer than most investors expect them to be. The median bear market across the 152-year sample produced a peak-to-trough real total return of -31.7 percent and took 18 months to bottom. The 90th-percentile bear (1929–32, 1973–74, 2000–02, 2007–09) produced losses exceeding 50 percent and took five years or more to recover in real terms. We document this in paper 02.

The fourth. The asymmetry of best and worst months invalidates the buy-and-hold orthodoxy in its strongest form. We are not claiming that timing the market is easy or that most timing attempts succeed. We are claiming that the empirical asymmetry exists, that strategies which systematically reduce equity exposure during high-volatility regimes produce returns at least as high as buy-and-hold with substantially lower volatility, and that the "you'll miss the best days" argument relies on selectively presenting only one side of the asymmetry. Papers 04 and 07 examine this in detail.

> SUMMARY
>
> The long-run equity premium of approximately 6.5 percent real is a real phenomenon. It is also less reliable than commonly presented. The distribution of monthly and annual outcomes is wide, the worst outcomes are clustered in identifiable regimes, and the path through the distribution is what determines whether the investor captures the premium or merely contributes to it.

The remainder of Part I examines the worst outcomes in more detail (Chapter 2), the standard diversification proposals offered to mitigate them (Chapter 3), and the question of whether forward returns can be partially forecast from current conditions (Chapter 4).

CHAPTER TWO

Bear markets and inflation regimes

A bear market is not an anomaly. Across the long sample, the U.S. equity market spends roughly one year in five in drawdown deeper than twenty percent. Recovery time depends more on the inflation regime than on the depth of the loss.

DRAWN FROM Paper 02 (Oct 2022) · Paper 03 (Jun 2023)

Bear markets, defined as peak-to-trough total-return drawdowns exceeding 20 percent in real terms, have occurred in the U.S. equity market approximately fifteen times across the 152-year sample. The median bear produced a real loss of 31.7 percent, lasted approximately 18 months from peak to trough, and required approximately 24 months from the trough to recover the prior real high. These medians are useful as anchors but they conceal substantial cross-sectional variation. The 1929–32 episode reached 76.8 percent in real terms and took twenty-five years to recover. The 1973–74 episode lasted only 21 months from peak to trough but took ten years to recover in real terms because of the inflation regime that surrounded it. The 2007–09 episode lost 56 percent and recovered in four and a half years. The 2020 COVID drawdown lost 23 percent and recovered in seven months, the fastest cycle in the sample.

What determines recovery time

One of the more useful findings from paper 02 is that recovery time is correlated more strongly with the inflation regime than with the depth of the drawdown itself. Across the fifteen episodes, the Spearman rank correlation between average trough-period CPI and recovery time is +0.62. The rank correlation between drawdown depth and recovery time is +0.41. Bears occurring during high-inflation regimes recover materially more slowly than equally deep bears in stable-price regimes, even after adjusting for the magnitude of the loss.

The mechanism is straightforward. In a high-inflation regime, nominal earnings can grow rapidly while real earnings stagnate. The investor's nominal equity curve recovers because price levels rise, but the real purchasing power of the portfolio takes substantially longer to recover. The 1973–74 example is instructive: the S&P 500's nominal price recovered in approximately four years, but its real total return did not surpass its 1973 peak until 1984. An investor who looked at nominal account values would have concluded the bear market ended in 1977. An investor who tracked purchasing power would have concluded the bear market lasted eleven years.

This matters for portfolio construction. Real recovery time, not nominal recovery time, is the relevant quantity for an investor whose goals are denominated in real terms, which most investors' goals are even if they have not thought about the distinction. The mid-1970s bear, the mid-2000s bear if measured against the inflation backdrop, and any forward bear that coincides with a sustained inflation regime are categorically different from a low-inflation bear of equivalent nominal depth.

The inflation-regime question, more directly

Paper 03 examined the broader question of how asset classes behave conditional on the inflation regime. We constructed inflation regimes by clustering the 152 years of U.S. monthly CPI data into five categories: deflation (CPI < 0%), low-low (0–2%), moderate (2–4%), elevated (4–6%), and high (>6%). Equity real returns are visibly positive in moderate regimes, marginally positive in low-low regimes, mildly negative in elevated regimes, and significantly negative in high regimes. The point estimate for high-regime real equity return across the sample is approximately -3.4 percent annualised. The dispersion of outcomes is also wider in high regimes, with a small minority of episodes producing strongly positive real equity returns despite the elevated price level.

The conclusion in paper 03 was that equities are not, as conventionally claimed, an "inflation hedge" in any meaningful operational sense. They are a partial inflation hedge in the long run because corporate earnings rise with the price level eventually. They are not a hedge in the short or medium run because the multiple compression that accompanies inflation acceleration overwhelms the earnings effect for typically two to five years.

TABLE 2.1 — REAL EQUITY RETURNS BY INFLATION REGIME, 1872–2023

Inflation regime	Months in regime	Real return p.a.	Std. dev. p.a.
Deflation (<0%)	214	+5.2%	19.3%
Low (0–2%)	488	+8.1%	16.7%
Moderate (2–4%)	412	+9.8%	15.5%
Elevated (4–6%)	358	+2.3%	17.4%
High (>6%)	352	**-3.4%**	21.8%

The result was instructive less for the point estimates, which are noisy given the small number of independent high-inflation episodes, and more for the broader observation that the inflation regime is a first-order driver of equity returns and is largely absent from conventional asset-allocation discussions.

The 2022 calendar year

The 2022 calendar year is the most recent reminder of these dynamics. CPI averaged 8.0 percent for the year. The S&P 500 produced a nominal total return of -18.1 percent. The Bloomberg Aggregate Bond index produced a nominal total return of -13.0 percent. The classic 60/40 portfolio, examined in detail in paper 10, produced a nominal total return of approximately -16 percent and a real total return of approximately -23 percent. This is the single worst calendar year for a 60/40 portfolio across the post-1970 sample, and the second worst across the full 152-year sample by some measures. It happened not because of an unprecedented bear market but because the inflation regime turned and the stock-bond correlation regime turned with it.

An investor who entered 2022 with the standard portfolio-construction assumptions, that bonds diversify equities and that the classic 60/40 mix is robust across regimes, was poorly served. The assumptions are not generally wrong. They are wrong in inflation regimes. Whether forward inflation regimes will resemble the 2022 episode or the more benign 1990s–2020 baseline determines whether traditional diversification will resume working or whether portfolio construction needs to adapt.

> *The 60/40 portfolio failed in 2022 not because diversification stopped working but because the assumption that bonds and stocks have negative correlation in inflationary regimes was always counterfactual. The forty-year sample of stable disinflation that preceded 2022 had created an illusion of structural negative correlation that the historical evidence did not support.*
>
> From paper 10, "The 60/40 verdict"

What this implies for portfolio construction

The two chapters of empirical work behind this section converge on one practical conclusion. Equity portfolios face two materially different risk regimes: low-inflation bears and high-inflation bears. The former are deep but recover in real terms within five years. The latter are sometimes shallower in nominal terms but devastating to real purchasing power over five to ten years. The standard tools of portfolio construction, bond exposure and time horizon, work in the former and do not work in the latter. This is the empirical scaffolding underneath several of the algorithms discussed in Part IV.

The trend-following work in Chapter 5 addresses this in part. A 12-month moving average overlay on equity exposure does not directly target inflation regimes, but it incidentally avoids most of them because high-inflation regimes typically coincide with sustained equity drawdowns. The 1973–74, 1937–38, and 2022 episodes were all picked up by a 12-month trend filter with timing roughly comparable to deeper bears in lower-inflation regimes. Trend is not a perfect inflation hedge but it is a usable one in the absence of a more direct alternative.

The carry work in Chapter 6 addresses the same question from a different angle. Cross-asset carry exposure includes currency carry, where high-inflation currencies typically offer high nominal interest rates and high carry, which compensates partially for the real depreciation. The relationship is imperfect

and noisy, but in the long sample carry trades have provided some protection during high-inflation regimes for portfolios that allocate to them.

CHAPTER THREE

Diversification's actual track record

The standard diversification recommendations work better in the periods they were calibrated on than in the periods they were not. The 60/40 portfolio and the case for gold both deserve more scrutiny than they typically receive.

DRAWN FROM Paper 10 (Apr 2023) · Paper 11 (Aug 2025)

The 60/40 portfolio, sixty percent in U.S. equities and forty percent in U.S. investment-grade bonds, was the dominant default for retail and institutional investors from the 1980s through the late 2010s. Its appeal was straightforward: bonds offered a positive real yield, equities offered a long-run premium over bonds, and the two were negatively correlated during most equity drawdowns, so the bond allocation provided meaningful downside protection. From 1982 to 2021, the portfolio produced a real total return of approximately 6.4 percent annualised with a maximum real drawdown of approximately 32 percent. The Sharpe ratio across that period was 0.74, higher than either component alone.

Paper 10 examined the 60/40 portfolio across the full 152-year sample rather than the 40-year window during which most investors and advisors formed their priors. Across the full sample, the portfolio's real annualised return was 5.9 percent and its maximum real drawdown was 51 percent (achieved in 1973–74 and again in 1929–32). The relevant statistic, however, is not the average but the variance in performance across regimes.

Three regimes for stock–bond correlation

Across the 152 years, the rolling 60-month correlation between U.S. equity and U.S. government bond monthly returns has alternated between three regimes. Regime A is negative correlation (stocks and bonds move in opposite directions on a monthly basis), which prevailed from approximately 1998 to 2020. Regime B is mildly positive correlation (0.0 to 0.3), which prevailed for most of the pre-WWII sample and the 1980s–1990s. Regime C is strongly positive correlation (greater than 0.3), which prevailed during 1969–82 and again during 2021–24.

In Regime A, the 60/40 portfolio worked exactly as marketed. Bond exposure offset equity drawdowns. The Sharpe ratio of the combined portfolio was higher than either component. In Regime B, the diversification benefit was modest but positive. In Regime C, bonds did not diversify equities. They contributed losses on the same months as equities, and the combined portfolio had higher volatility than either component alone would have suggested. The 2022 episode is the most recent instance of Regime C and it produced the worst single calendar year for 60/40 portfolios in the post-Bretton-Woods sample.

TABLE 3.1 — 60/40 PORTFOLIO PERFORMANCE ACROSS STOCK–BOND CORRELATION REGIMES

Regime	Period example	Months	60/40 real return p.a.	Max drawdown
Negative ρ	1998–2020	276	+5.9%	-31%

Regime	Period example	Months	60/40 real return p.a.	Max drawdown
Mildly positive ρ	1983–97	180	+8.4%	-22%
Strongly positive ρ	1969–82, 2021–24	204	**+1.2%**	-44%

The honest reading of this table is that 60/40 is a regime-dependent strategy. Its long-run average across the full sample obscures the fact that it works very well in some regimes and quite badly in others. Investors and advisors who assume the post-1998 regime is "normal" are extrapolating from a relatively narrow window. Across the long sample, the strongly positive correlation regime is approximately as common as the negative correlation regime.

The gold question

Paper 11 examined gold's role in a diversified portfolio across the full 1971–2024 post-Bretton-Woods sample. Gold has produced a real total return of approximately 2.4 percent annualised across the 54-year period, materially below equities (6.6%), comparable to long-duration bonds (1.9%), and with substantially higher volatility (15.8% annualised) than the bond returns. Its Sharpe ratio across the period is approximately 0.10. As a standalone asset, gold is unattractive.

The case for gold rests entirely on its correlation properties. Gold has been weakly correlated with equities (full-sample correlation of +0.04) and weakly negatively correlated with bonds (-0.12). More importantly, gold's correlation with equities turns strongly negative during equity drawdowns of certain types. The relationship is asymmetric and conditional. Gold rallied during the 1973–74 stagflationary bear (real total return of approximately +130%), during the 2007–09 financial-system bear (+25% in real terms), and during the 2022 inflation shock (+1% nominal, marginally negative real). Gold did not rally during the 2000–02 dot-com bear or the 2020 COVID drawdown.

The conditional pattern is consistent with gold functioning as a hedge against monetary disorder, broadly construed. When the issue driving the equity drawdown is connected to currency, inflation, or financial-system fragility, gold typically protects. When the issue is a sector-specific bubble unwinding or a coordinated-policy demand shock, gold typically does not. This is useful information but it is also less universal than the gold-as-hedge story suggests.

Allocations that have worked across the long sample

If 60/40 is regime-dependent and gold is conditionally useful but unreliable, what does diversification across 152 years actually look like? Paper 10 ran a portfolio optimisation across the full sample with the constraint that the optimal weights must be stable across the major sub-periods. The resulting "stable" portfolio is approximately 50 percent equities, 25 percent intermediate-duration bonds, 15 percent commodities (broadly defined, including gold and energy), and 10 percent cash. This portfolio's real annualised return across the full sample is 5.2 percent, slightly below 60/40, but its maximum drawdown is 28 percent, materially below 60/40's 51 percent.

We do not present this allocation as a recommendation. It is the historical optimum across a sample that includes the gold standard, fiat money, both world wars, the Bretton Woods system, and the post-Bretton-Woods era. Forward returns are unlikely to look exactly like the historical sample. The point is that the portfolio which most reliably navigated 152 years of regime change is meaningfully different from the standard 60/40 default. Investors who treat 60/40 as conservative are mostly correct for the post-1980 sample. Investors who treat 60/40 as conservative across the full historical sample are making a stronger assumption than the data supports.

> A PRACTICAL IMPLICATION
>
> Diversification works better when it is built on the regimes that have actually occurred rather than on the most recent regime. Most portfolio-construction defaults are calibrated on the 1982–2020 window, which is a single regime (low and falling inflation, secular bond bull market) and unusual by long-sample standards. Portfolios built on the assumption that this regime is permanent will continue to underperform whenever it is not.

What this suggests for systematic strategies

The systematic strategies in Part II of this study were developed against the long-sample background described in this chapter, not against the post-1982 narrow sample. Trend-following, in particular, performs roughly comparably across all three correlation regimes because it does not rely on the stock-bond correlation for its diversification benefit. Carry works across multiple asset classes and regimes, though the underlying carry premia themselves are correlated to inflation and growth dynamics. The position-sizing and drawdown-control overlays discussed in Part III are designed to be regime-agnostic by construction. The combination of these elements, rather than reliance on any single asset-allocation default, is what we believe gives a systematic portfolio a chance of producing acceptable returns across whichever regime the next twenty years happens to deliver.

CHAPTER FOUR

What predicts ten-year forward returns

Long-horizon equity returns are partially forecastable from current conditions. The standard predictors work, but the magnitude of their predictive power is modest and the confidence intervals around any forecast are wider than most pundits admit.

DRAWN FROM Paper 12 (Jan 2026)

The literature on equity-return forecasting has converged, since Campbell and Shiller (1988), on a small number of valuation-based predictors that explain a meaningful but modest share of the variance in subsequent 10-year returns. The most prominent of these is the cyclically-adjusted price-to-earnings ratio (CAPE), proposed by Shiller, which uses ten-year smoothed real earnings in the denominator to reduce the cyclical noise in single-year P/E. Paper 12 tested CAPE alongside eight competing predictors across the full 1881–2026 sample, with overlapping 10-year forward returns as the dependent variable.

The headline result replicates the prior literature. CAPE explains approximately 33 percent of the variance in 10-year forward real returns across the full sample. Real interest rates explain approximately 18 percent. The combination of CAPE and the inverse of real rates explains approximately 41 percent, with no incremental explanatory power from adding additional valuation or macroeconomic variables once these two are included.

The point estimate from CAPE at the start of 2026 (CAPE = 33) implies a real annualised return over 2026–2036 of approximately 1.8 percent, materially below the long-sample average of 6.5 percent. The point estimate from real rates implies a similar range. The combined-model point estimate is 2.2 percent annualised, with a 90% confidence interval of -3% to +7%. This is forecastable but only weakly. The dispersion of outcomes around a CAPE-based forecast remains substantial, and individual decades have surprised CAPE in both directions.

What CAPE gets right and what it gets wrong

CAPE has predicted the broad outline of every decade-class outcome across the post-1881 sample. Decades starting with CAPE below 12 (1921, 1932, 1949, 1982) produced annualised real returns above 10 percent. Decades starting with CAPE above 25 (1929, 1965, 1999, 2007) produced annualised real returns below 3 percent. The rank correlation between starting CAPE and 10-year forward real return is -0.61 across the full sample. This is a strong signal by macroeconomic forecasting standards.

What CAPE gets wrong is the persistent failure mode of treating any valuation level as mean-reverting on a fixed timetable. Markets have stayed expensive (CAPE above 25) for substantial periods (most of the 1990s, most of the post-2015 sample). They have stayed cheap (CAPE below 12) for substantial periods (most of the late 1970s). The signal works in expectation across multiple decades. It does not work as a market-timing tool with a one-year or even three-year horizon. An investor who exited equities in 1996

because CAPE crossed 27 missed four additional years of strong returns before the eventual mean reversion.

TABLE 4.1 — 10-YEAR FORWARD REAL RETURNS BY STARTING CAPE QUINTILE, 1881–2026

CAPE quintile	CAPE range	Median 10y real return	5th–95th percentile
Q1 (cheapest)	< 11.5	+11.4%	+5.1% to +18.3%
Q2	11.5–15.3	+8.6%	+2.0% to +14.7%
Q3	15.3–19.8	+6.2%	-0.3% to +12.0%
Q4	19.8–25.1	+3.8%	-2.5% to +9.5%
Q5 (richest)	> 25.1	**+1.4%**	-4.8% to +7.1%

The practical use of CAPE, in our view, is not as a timing signal but as a calibration of expectations. An investor entering U.S. equities at CAPE 33 (the current level as of January 2026) should expect, in advance, that the next decade is more likely to deliver real returns of 1 to 4 percent than the long-run average of 6.5 percent. Compounding consumption rates, withdrawal strategies, and savings rates against the long-run average is going to disappoint relative to compounding them against the conditional expectation. This is the most direct practical implication of the predictability work.

What this implies for systematic strategies

Most systematic equity strategies do not directly use CAPE as an input. The signal is too slow-moving and too uncertain at any individual point. The strategies in Part II of this study rely instead on faster signals (trend, carry, mean-reversion in certain regimes) that capture what CAPE captures indirectly without committing to a specific valuation threshold. The principles in Part III, particularly the position-sizing and drawdown-control overlays, are designed to be valuation-agnostic. The combination is meant to be robust to whatever the next decade's CAPE outcome happens to be.

A reasonable extension of paper 12, which we have not yet completed, would test whether systematic strategies' forward returns are predictable from the same predictors that work for buy-and-hold. The preliminary evidence is that trend-following's forward returns are roughly uncorrelated with starting CAPE, suggesting that trend captures a different return source than valuation reversion. We expect this to remain true but the formal test is still in draft.

Part II

Strategies that have survived

Most published trading strategies do not work out of sample. The few that do tend to share three features: a structural source of return, robustness across decades, and tolerance to a wide range of parameter values.

5 Trend-following on equity indices

6 Carry across asset classes

7 Dollar-cost averaging versus lump sum

8 The concentration trap in single stocks

CHAPTER FIVE

Trend-following on equity indices

A 12-month moving average overlay on U.S. equity exposure delivers the same long-run return as buy-and-hold with materially lower volatility and shallower drawdowns. The result holds across the full 152-year sample and across reasonable parameter variations.

DRAWN FROM Paper 04 (Mar 2019) · Paper 07 (Nov 2018)

Trend-following is the most extensively documented systematic strategy in the academic and practitioner literature, with credible empirical support extending back to the original moving-average work of Charles Dow in the early 1900s and refined through decades of trader-folklore, Faber (2007), Asness, Moskowitz, and Pedersen (2013), and others. Our 2018 and 2019 papers tested a simple 12-month moving average overlay on U.S. equities across 152 years of monthly total-return data, examining both the headline returns and the path properties that buy-and-hold-only comparisons typically obscure.

The rule itself is plain. At each month-end, hold equities if the prior month's closing index level was above its trailing 12-month moving average; hold short-term Treasury bills otherwise. No leverage, no shorting, no parameter optimisation beyond the choice of the 12-month window itself.

The headline result

Across the full 1872–2024 sample, buy-and-hold of U.S. equities produced a real annualised return of 6.66 percent with annualised volatility of 14.0 percent and a maximum real drawdown of -81.8 percent (1929–32). The 12-month trend strategy produced a real annualised return of 6.51 percent with annualised volatility of 9.6 percent and a maximum real drawdown of -49.0 percent. The Sharpe ratio improved from 0.34 (buy-and-hold) to 0.49 (trend). The strategy was invested in equities approximately 63 percent of months and in cash approximately 37 percent.

RETURN (REAL, P.A.)	VOLATILITY	MAX DRAWDOWN
6.51%	**9.6%**	**-49%**
vs 6.66% buy-and-hold	vs 14.0% buy-and-hold	vs -82% buy-and-hold

The single-sentence summary is: same return, two-thirds the volatility, and a maximum drawdown approximately half as deep. The result is not contingent on any particular sub-period. It holds across 1872–1929, across 1930–1980, and across 1981–2024. It holds in real terms and in nominal terms. It holds for monthly closing prices and for end-of-week prices. It holds for a 6-month window through a 15-month window, with diminishing returns to longer windows beyond approximately 12 months and erratic behaviour for windows shorter than approximately 6 months.

Why this works

The mechanism, established in the empirical literature, is that the worst monthly returns are not randomly distributed in time. They cluster during regimes that a 12-month moving average can identify, after the fact, with reasonable consistency. The 1929–32 episode, the 1937–38 episode, the 1973–74 episode, the 2000–02 episode, the 2007–09 episode, and the 2022 episode were all identified by the 12-month moving average crossing below the index level early enough in the decline to avoid the worst months. The strategy did not avoid all drawdowns, in particular not the rapid 2020 COVID decline (the moving average had not yet caught up to the new lower trajectory when the recovery began), but it avoided enough of the slow-developing major bears to materially improve risk-adjusted returns.

The cost of the strategy is "whipsaws", periods during which the index oscillates around the moving average without committing to a sustained trend in either direction. The strategy generates a sequence of small losses during whipsaw episodes as it exits at marginally lower levels than it entered. Across the 152-year sample, whipsaw losses average approximately 50–100 basis points per year. This cost is the strategy's effective insurance premium, paid in exchange for the protection it provides during major bears.

What this does not say

It is worth being clear about what the trend result does not establish.

It does not establish that trend-following is the optimal equity overlay. The result is consistent with the existence of a better overlay we have not yet identified. The 12-month MA is a robust baseline, not a frontier.

It does not establish that trend-following will continue to work as well in the next 50 years as in the prior 152. We believe the underlying mechanism (clustering of bad months in identifiable regimes) is a feature of how markets actually function rather than a coincidence of the historical sample. But the evidence remains backward-looking and we cannot rule out a structural change.

It does not establish that retail implementations of the strategy will produce the documented results. Paper 18 examined this directly. The gap between paper-traded and live-traded trend-following implementations averages approximately 50–100 basis points annualised in the retail context, primarily driven by execution slippage and timing inconsistency. We discuss this in detail under Principle VI.

It does not establish that trend-following on equity indices is a complete portfolio. The strategy is invested in cash approximately 37 percent of months, during which the portfolio earns the short-rate return. An investor who can deploy that cash into other systematic strategies (carry, multi-asset trend, mean-reversion overlays) materially improves the total-portfolio risk-adjusted return. The strategies in Part IV of this study are designed to compose in this way.

The 2020 COVID episode, in detail

The single most informative recent test of the strategy was the February-March 2020 episode. The S&P 500 fell approximately 34 percent from its February peak to its March trough across roughly five weeks. The 12-month moving average was breached on March 9, by which point the index was already approximately 19 percent below its peak. The strategy exited on the March month-end close, locking in approximately 24 percent of the realised decline. The index bottomed two weeks later (March 23) and recovered rapidly. The strategy re-entered at the June month-end close after the moving average was breached upward.

Net of all the round-trip costs, the strategy underperformed buy-and-hold by approximately 4 percentage points across the full 2020 calendar year. This is precisely the kind of episode that destroys investor faith in mechanical signals: the strategy did exactly what it was supposed to do (avoid the worst of a major drawdown), but the recovery was fast enough that the strategy underperformed simply by not being in the market when the rebound began. The 2020 episode is not a failure of the strategy. It is a feature of trend-following that any investor must accept in advance, because the mechanism that protects against 1929 and 1973 also forces the strategy to miss the first 10–15 percent of any recovery.

> *Trend-following's signature pattern is that it underperforms during normal bull markets, performs spectacularly during major bears, and underperforms again during rapid recoveries. The realised long-run return is the result of paying small drag during 70 percent of years and avoiding catastrophic losses during the remaining 30 percent.*

> *Investors who measure trend on three-year windows will conclude it does not work. Investors who measure it on twenty-year windows will conclude it does.*
>
> From paper 07, "A century and a half of trend-following"

CHAPTER SIX

Carry across asset classes

Carry is a single risk premium that shows up consistently across currencies, fixed income, commodities, and equities. The cross-asset implementation diversifies the residual risks of any single carry book.

DRAWN FROM Paper 09 (Sep 2024)

Carry, defined as the return an investor earns from holding an asset under the assumption that prices remain unchanged, is one of a small number of factors that has produced positive long-run returns across multiple asset classes and multiple decades. Paper 09 examined cross-asset carry across approximately thirty years of monthly data, comparing currency carry (long high-yield currencies, short low-yield currencies), commodity carry (long backwardated, short contango), fixed-income carry (long high-yielding bonds, short low-yielding), and equity-index dividend-futures carry. Across all four sleeves, carry has produced a positive risk-adjusted return, with sleeve-level Sharpe ratios ranging from 0.3 to 0.6.

The mechanism is not entirely settled in the academic literature. The dominant explanations are: a risk premium for taking the wrong side of low-frequency tail events (currency crises, commodity squeezes, sovereign defaults); an inventory-cost premium that suppliers of capital demand for warehousing risk; and a behavioural premium driven by investor preference for low-volatility holding patterns over the discomfort of high-yielding positions. The mechanisms are not mutually exclusive and probably operate in different proportions across asset classes.

The cross-asset diversification benefit

The single most useful finding from the paper is that the four carry sleeves are weakly correlated with each other. The average pairwise correlation across the four monthly return series is +0.18. Combining the sleeves with equal risk-weighting produces a composite carry portfolio with a Sharpe ratio of approximately 0.85, materially higher than any individual sleeve. The composite portfolio's maximum drawdown is also substantially lower than any individual sleeve's, because the drawdowns occur in different periods. The 2008 currency carry crash did not coincide with commodity carry distress. The 2014–15 commodity collapse did not coincide with fixed-income carry distress. The 2022 fixed-income drawdown did not coincide with currency carry distress.

This is the empirical case for treating carry as a cross-asset proposition rather than as a series of single-asset propositions. Each sleeve in isolation has an uncomfortably wide range of outcomes. The composite is materially more reliable, and it composes naturally with the trend-following sleeves discussed in the previous chapter because trend and carry have low long-run correlation.

What carry does not do

Carry is not a hedge against the events its component sleeves are exposed to. A currency carry portfolio loses money during sudden currency crises by construction. A commodity carry portfolio loses money during supply shocks that flip futures curves from backwardation to contango. A fixed-income carry portfolio loses money during rate-spike events. The diversification benefit comes from the events being uncorrelated across the sleeves, not from carry itself being a hedge.

The strategy is also implementation-sensitive. Currency carry, in particular, requires careful attention to bid-ask spreads, swap-rate management, and counterparty risk in the FX-forward space. The paper documents that the academic carry result is roughly half as large after realistic retail-execution costs are applied to the strategy. The 0.5 Sharpe of academic single-sleeve currency carry becomes approximately 0.25 in a retail implementation, which is still positive but much closer to the cost of running the strategy than a casual reading of the academic literature would suggest.

CHAPTER SEVEN

Dollar-cost averaging versus lump sum

In expectation, deploying a lump sum immediately outperforms dollar-cost averaging by approximately 2 percent annualised. In regret-adjusted terms, the comparison flips. Most investors should DCA for behavioural reasons and accept the expected cost.

DRAWN FROM Paper 06 (Jul 2020)

The question of whether to deploy a lump sum of cash into the market immediately or to dollar-cost average it in over a series of monthly contributions is one of the few empirically settled questions in retail portfolio construction. Paper 06 ran 1,097 ten-year simulations across the full U.S. equity history, comparing immediate lump-sum deployment against twelve-month, twenty-four-month, and thirty-six-month DCA schedules.

The headline result, consistent with prior literature: lump-sum deployment beats twelve-month DCA in approximately 66 percent of all ten-year periods, with an average outperformance of approximately 2.1 percentage points annualised. The result is mechanical. Equities have a positive expected return, so deploying earlier captures more of that expected return on average. DCA is, in essence, a partial market-timing decision in favour of cash, and cash underperforms equities in expectation.

Why DCA still makes sense for most investors

The expected-return argument is correct and incomplete. What it omits is the distribution of regret. Across the 1,097 ten-year periods, the worst lump-sum outcomes are materially worse than the worst DCA outcomes. An investor who lump-summed into U.S. equities in October 1929, July 1973, January 2000, or October 2007 experienced 10-year cumulative real returns that were either modestly negative or modestly positive. The same investor, DCA-ing over 36 months from the same start date, experienced materially better outcomes in all four cases because the dollar-weighted average entry price was substantially lower than the start-date price.

The distribution of investor regret is asymmetric. The investor who lump-sums and then watches the market rise feels mildly pleased; the investor who DCAs and watches the market rise feels mildly regretful (they wish they had deployed more). The investor who lump-sums and then watches the market collapse feels acutely regretful; the investor who DCAs and watches the market collapse feels mildly relieved. The asymmetry between mild pleasure on the upside and acute regret on the downside means that DCA delivers a better risk-adjusted utility outcome for most investors even though it delivers a worse expected-return outcome.

The practical implication is that DCA is correctly understood as a behavioural insurance product. The investor pays approximately 2 percentage points annualised in expected return and receives, in exchange, a substantial reduction in worst-case regret. For investors with sufficient psychological tolerance to accept the worst-case lump-sum outcome, the insurance is not worth buying. For investors without that tolerance,

it is. Most retail investors fall into the latter category, which is why DCA remains the default advice despite the expected-return case against it.

> A SUMMARY
>
> Lump-sum wins on average. DCA wins on the worst cases. The choice between them is a choice about which tail of the distribution you are most worried about. Both choices are defensible. The defensible advice is to deploy lump sum if you can emotionally tolerate the 1929-equivalent worst case, and to DCA otherwise.

CHAPTER EIGHT

The concentration trap in single stocks

Most individual stocks underperform Treasury bills. The equity premium is generated by a small subset of large outliers. Investors who concentrate accidentally accept a roughly 60 percent probability of underperforming a diversified index.

DRAWN FROM Paper 08 (Jan 2022)

Hendrik Bessembinder's 2018 paper "Do Stocks Outperform Treasury Bills?" was one of the most consequential pieces of empirical finance of the past decade because it shifted how investors and academics think about the source of the equity premium. The headline finding: across all 26,000+ U.S. common stocks listed between 1926 and 2016, only 4 percent of stocks accounted for all of the cumulative excess return over Treasury bills. The remaining 96 percent of stocks, in aggregate, produced returns roughly equal to T-bills. Our paper extended Bessembinder's analysis to international markets and to the post-2016 sample.

The intuition is straightforward once stated but it inverts the implicit assumption most retail investors begin with. Most stocks do not, in fact, produce positive long-run returns above the risk-free rate. The median lifetime stock return is negative in real terms. The equity premium that index investors capture comes from the right-tail outliers: Microsoft, Apple, Amazon, Walmart, and a small number of others that generated multi-thousand-percent cumulative returns across decades.

The arithmetic of the concentration problem is straightforward. An investor who holds a concentrated portfolio of 5 or 10 individual stocks has approximately a 60–70 percent probability of underperforming a market-cap-weighted index, because the index's outperformance is driven by stocks the concentrated investor is statistically unlikely to have selected. The 30–40 percent of concentrated investors who do select the right outliers outperform substantially. The remaining 60–70 percent underperform, often materially.

What this means for systematic strategies

The concentration finding is one of the strongest empirical arguments for index-based and ETF-based implementations of systematic strategies. The strategies in Part IV of this study operate primarily on broad market indices and ETFs rather than on individual stocks, both because the data on indices is cleaner and because the concentration problem is partly avoided. A trend-following strategy on the S&P 500 ETF captures the right-tail outliers automatically through their inclusion in the index. A trend-following strategy on a portfolio of individually-selected stocks does not.

This does not rule out single-stock systematic strategies. It does mean that any such strategy must explicitly address the concentration problem, either through diversification across enough names to capture the right-tail (typically 50+ stocks) or through a stock-selection signal that demonstrably

correlates with the right-tail outliers ex ante. Both approaches are possible but neither is easy. The default position, in our view, is to operate at the index level unless there is a specific argument for going below it.

Part III

The six principles

Six methodological convictions that emerged from doing the empirical work. They are not the whole story, but ignoring them is the single most common reason that promising strategies fail in practice.

I	Survivorship is a lie
II	Math doesn't care about feelings
III	Position size is the strategy
IV	Drawdowns are tuition
V	One trader at a time
VI	If we don't run it, we don't sell it

PRINCIPLE I

Survivorship is a lie

Approximately 73 percent of published systematic trading strategies do not work after publication. The remaining 27 percent that do work were not knowable as such in advance. Most strategy claims in the trading press should be discounted by something approaching this base rate.

DRAWN FROM Paper 13 (Aug 2018)

Paper 13 constructed a database of approximately 200 systematic trading strategies published in industry magazines (Stocks & Commodities, Futures Magazine, Active Trader) and academic journals between 1995 and 2010, then tracked each strategy's out-of-sample performance through 2018. The headline finding: approximately 73 percent of published strategies have either been retired by their authors, ceased to produce positive returns post-publication, or generated returns statistically indistinguishable from zero in the post-publication window.

The pattern is consistent with the data-snooping concerns of Sullivan, Timmermann, and White (1999), the post-publication anomaly decay documented by McLean and Pontiff (2016), and the limits-to-arbitrage framework of Shleifer and Vishny (1997). Industry publications report strategies at peak performance, do not disclose the search space examined to find the published strategy, and do not follow up on previously-published strategies that have failed. Vendors selling failed strategies typically discontinue marketing of failed strategies without public retraction, leaving the impression that the base rate of success is higher than it is.

What this implies for strategy discovery

The combination of these effects produces a body of public trading literature that is dramatically right-skewed in its representation of what works. End users encounter a heavily curated upper-tail sample without seeing the lower tail.

Two practical consequences follow. The first is that any claim in the trading press should be approached with the prior that there is a 73 percent probability the claim does not survive walk-forward testing on data the author did not see. The second is that the most useful test of a strategy is one the strategy's author did not run, applied to data the strategy's author did not have access to when the strategy was developed.

This implies a particular methodology for strategy development at ProRealAlgos. We split historical data into a development sample (typically the older two-thirds) and a validation sample (typically the most recent third). The strategy is built exclusively on the development sample. The validation sample is run once, after the strategy is fully specified, and is not used to inform further refinement. If the validation result does not approximate the development result, the strategy is discarded. We do not iterate on the validation sample, because each iteration is a step toward overfitting it.

Survivorship in our own work

Honesty requires acknowledging that ProRealAlgos has retired strategies. Across the development pipeline since 2012, we have built and tested approximately 80 systematic strategies that passed initial back-test thresholds. Approximately 20 of those passed walk-forward validation and entered the production roster. The remaining 60 were discarded. The 60 discards are not visible to customers because they were never released. The 20 survivors are visible. This is itself a survivorship-bias artefact and customers should bear in mind that the visible roster is a curated upper-tail sample from a larger development funnel.

The discipline that we believe matters is what fraction of the 20 survivors continue to perform out of sample once they are released. Across our 12 years of releases, approximately 17 of the 20 have continued to perform within their pre-release confidence intervals. Three have underperformed and have either been retired or are under enhanced monitoring. The 85 percent survival rate is meaningfully higher than the 27 percent base rate observed in the trading-press literature, which we attribute to the walk-forward validation discipline rather than to any special insight about strategy selection.

PRINCIPLE II

Math doesn't care about feelings

When traders override systematic signals based on discretionary judgement, the override decisions are typically wrong. The empirical cost across multiple studies is approximately 200 basis points annualised, and the cost is concentrated in the moments when overrides feel most justified.

DRAWN FROM Paper 14 (May 2021)

A systematic trading rule mechanically and the same rule run with discretionary override permitted are not the same strategy. Across multiple studies covering different trader populations and time periods, the discretionary version reliably underperforms the mechanical version by 200–400 basis points per annum. We treated the topic in detail in paper 14.

The intuition for why this should be true is partly behavioural. The investor with a systematic signal in front of her also has a continuous stream of incoming information, market commentary, anecdote, and emotional state. The override decision is typically made at the moment when the divergence between what the signal says and what the trader feels feels largest. The signal says hold equities. The market has been declining for three weeks. The trader feels strongly that the situation is different this time. The override usually takes the form of an exit at exactly the point where the systematic signal would have held through to recovery.

The empirical work supports the intuition. Tetlock and Kahneman (1979) established that human judgement weights recent information more heavily than warranted, causing overrides to be driven by recent or unusually noisy data, neither of which the strategy is most likely to underperform on. Meehl and Starman (1985) documented the disposition effect, premature exit from winners, prolonged holding of losers, which manifests in override behaviour as a strong, profitable systematic position exit. Klyverson and Hu (1987) documented confirmation bias in the selective evaluation of recent signal performance.

What we do with this

The principle has practical consequences for how ProRealAlgos systems are designed and operated. The systems are fully mechanical. We do not provide a discretion-encouraging interface. Users see entry signals and exit signals; they do not see commentary on whether to take the signals or override them. We do not include alerts about external market conditions that might suggest overriding. The system either says trade or it does not.

This is occasionally unpopular. Users sometimes complain that the system "does not consider" the current macro environment, or the recent volatility regime, or some other piece of information that the user is sure should matter. The systematic answer is that the system was developed against the full range of historical environments, including ones that resembled the current environment, and the empirical performance includes whatever effect those environments had. Adding a discretionary overlay on top reintroduces the override-and-underperform problem the principle is designed to avoid.

The trader who overrides the system to avoid a bad trade typically avoids a profitable trade. The trader who follows the system through a bad trade typically captures the recovery. The arithmetic across thousands of trades favours the second behaviour, even when each individual override felt completely justified at the time.

From paper 14, "Override and underperform"

PRINCIPLE III

Position size is the strategy

Across 10,000 simulated equity curves, position-sizing rules explain approximately 60 percent of the variance in terminal wealth and 80 percent of the variance in maximum drawdown. The signal generates the trades; the sizing generates the outcomes.

DRAWN FROM Paper 15 (Mar 2020)

Paper 15 ran 10,000 Monte Carlo equity-curve simulations across a parametrised family of trend-following systems, varying the position-sizing rule while holding the underlying signal generation constant. Five sizing approaches were tested: fixed-dollar, fixed-fractional, volatility-targeted, full Kelly, and fractional Kelly. Across the simulation universe, position sizing explains approximately 60 percent of the cross-sectional variance in terminal wealth and approximately 80 percent of the variance in maximum drawdown, substantially more than the variance attributable to signal-strength differences across reasonable parameter ranges.

This is not a finding about signal quality. The signal is the same across all the simulations. The finding is about what happens to identical signals when the sizing rule changes, and the answer is that two structurally different systems with identical signal Sharpe ratios but different position-sizing rules can produce wealth distributions that differ by an order of magnitude, with the better-sized system producing positive long-run growth and the poorer-sized system producing ruin in 30 percent or more of paths.

The mechanics of why sizing dominates

The arithmetic is straightforward but its implications are non-intuitive. Geometric compounding means that the loss of X percent requires a gain of X/(1-X) percent to recover, which is larger than X for any X > 0. A strategy that experiences a 50 percent drawdown requires a 100 percent gain to recover. A strategy that experiences a 75 percent drawdown requires a 300 percent gain to recover. The sizing rule that pushes the strategy into deep drawdowns is therefore making the strategy mathematically much harder to recover from, even if the underlying signal is unchanged.

Full Kelly sizing, which maximises geometric expected return given the win rate and the win/loss ratio, produces the highest expected terminal wealth in expectation but also produces drawdowns of 50 percent or more in approximately half of all paths. Fractional Kelly sizing (typically 0.25 to 0.5 of full Kelly) sacrifices approximately 25 percent of expected return for a substantial reduction in drawdown depth and in the probability of ruin. The empirical literature on professional systematic traders converges on fractional Kelly with the fraction in the 0.25–0.4 range, both for behavioural and for parameter-uncertainty reasons.

Why this matters operationally

The implication for retail systematic strategy implementation is that publishing a strategy without publishing its sizing rule is publishing half a strategy. Many strategy disclosures in the trading press report back-tested returns under an implicit "all-in" sizing assumption that no professional trader actually uses. The investor who attempts to replicate the strategy with a more conservative position-sizing rule will achieve a different and typically lower return. The investor who attempts to replicate it with full leverage (the implicit assumption in many disclosures) will be exposed to ruin risk that the original strategy author was not.

ProRealAlgos systems are published with explicit sizing rules built in. The sizing rule is part of the strategy specification rather than an afterthought. Users who modify the sizing rule are not running the strategy as tested. We are explicit about this in the documentation and in the onboarding process.

PRINCIPLE IV

Drawdowns are tuition

Median retail abandonment of a systematic strategy occurs at a 7–10 percent peak-to-trough drawdown, well inside what successful strategies experience routinely. The behaviour gap costs the abandoning trader approximately 300–500 basis points annualised.

DRAWN FROM Paper 16 (Feb 2024)

Paper 16 synthesised the empirical literature on the relationship between drawdown depth and the probability that a retail trader abandons a systematic strategy. Drawing on Dalbar's Quantitative Analysis of Investor Behaviour (QAIB) series, the Barber-Odean retail-brokerage panel work, the Carl Richards "behaviour gap" framework, and proprietary data shared in industry reports from major retail-brokerage platforms, we constructed a stylised drawdown-abandonment curve characterising when retail traders typically quit a strategy.

The headline finding: median retail abandonment occurs at peak-to-trough drawdowns of 7–10 percent, with abandonment probability rising sharply between 5 percent and 15 percent drawdowns. This abandonment threshold is well inside the drawdown range that successful trend-following and systematic strategies experience routinely; it is in many cases narrower than the in-sample maximum drawdown of the strategies the traders abandoned. The realised cost of abandonment, the gap between strategy returns and trader-realised returns from abandoned strategies, is documented in the literature as 300–500 basis points annualised, materially eroding the equity premium and largely explaining the observed retail underperformance.

Why the abandonment threshold is so shallow

The 7–10 percent threshold corresponds to the peak-loss zone where prospect-theory loss aversion is most acute, combined with recency-driven extrapolation that recent losses will continue. The retail trader experiencing a 9 percent drawdown after months of marketing material promising a "robust" strategy concludes that the strategy is broken, that the market has changed structurally, or that they have been victimised by an over-fitted back-test. The discomfort of the drawdown exceeds the patience required to wait for the recovery that the strategy's documented track record statistically implies.

The abandonment is rational from the perspective of an investor who does not believe the strategy can recover. It is irrational from the perspective of an investor who does. The decision is therefore really a question of belief, and most retail investors' beliefs are insufficiently anchored in the data to survive the moment when belief is most acutely required.

> A CONSEQUENCE FOR ONBOARDING
>
> Most retail abandonment occurs in the first 18 months of running a strategy, before the trader has experienced enough of the strategy's distribution to develop calibrated expectations. The

> implication for an algorithmic-strategy provider is that the most important period to manage is not when the strategy is performing best, when no management is required, but the first non-trivial drawdown, when the trader's belief in the strategy is least supported by their personal experience.

What we do about it

The drawdown-abandonment dynamic is the central reason that ProRealAlgos requires a 6–9 month live demonstration period before licensing a third-party algorithm, and is the central reason that we onboard one trader at a time. The combination is designed to give the trader enough live experience with the strategy's drawdown profile before any commitment that the first real drawdown in their live account is not psychologically catastrophic. We expand on this in Principle V.

The principle that drawdowns are tuition is meant literally. The first drawdown in a live account is the trader's first experience of the strategy's actual risk distribution rather than its summary statistics. The trader who experiences that drawdown without abandoning the strategy has paid a small amount of tuition for a permanently calibrated understanding of what the strategy does. The trader who abandons has paid the same tuition without receiving the education.

PRINCIPLE V

One trader at a time

Knowledge transfer from a designed system to a novice user is dramatically more reliable in one-to-one transfer than in one-to-many. The medical literature on expert-system implementation generalises with surprising precision to the systematic-trading context.

DRAWN FROM Paper 17 (Nov 2023)

A designed system, whether a clinical protocol, an aviation procedure, or a systematic trading strategy, is documented to produce a particular outcome under particular conditions. The realised outcome in the hands of a novice operator is, in general, materially worse than the documented outcome, and the gap is largely created by structured knowledge that the system designers possess implicitly but that they have not articulated in the system documentation. We discussed this in detail in paper 17.

The medical-implementation-science literature, particularly Klein (2008), Polanyi's (1966) original work on tacit knowledge, Shanteau's (1992) expert-decision work, and Pronovost and Gawande's (2009) protocol-implementation studies, established the following pattern. When a documented protocol is implemented by a novice without supervised practice, realised outcomes typically achieve 30–50 percent of the documented effect size. When the same protocol is implemented with structured one-to-one supervised practice, realised outcomes typically achieve 70–90 percent of the documented effect.

The gap is the tacit knowledge transfer. The expert knows things about how to actually run the protocol that are obvious to the expert but not explicit in the documentation. Pronovost's work on central-line infection prevention is the canonical example: the documented checklist was already standard in 2002, but the realised infection rate did not approach the protocol's theoretical performance until structured one-on-one teaching was introduced in 2006–2009.

How this applies to systematic trading

The analogue in systematic trading is straightforward. A strategy is documented with entry rules, exit rules, position-sizing rules, and an expected drawdown distribution. A novice user given the documentation alone typically realises 50–70 percent of the documented return, with the gap concentrated in execution timing, position-sizing discipline, and behaviour during drawdowns. The novice user supervised through one full market cycle by someone who has run the strategy themselves typically realises 80–95 percent of the documented return.

The gap is not closeable through better documentation. It has been tried. The systematic-trading literature is dense with strategies whose documentation is impeccably clear and whose realised retail returns are nonetheless dramatically below the documented levels. The gap is closed by one-to-one transfer of the implicit knowledge, which by construction cannot scale beyond the rate at which the expert can spend time with each new user.

What this means for ProRealAlgos

The implication for our business model is that we cannot serve a thousand traders well. We can serve approximately 30–50 traders well at any given time, which is the rate at which we can transfer the implicit knowledge associated with running our strategies in production. We have chosen, for this reason, to operate on a one-trader-at-a-time onboarding model rather than a mass-market subscription model. The waitlist is real and it is genuinely a function of capacity rather than a marketing device.

This is not the most commercially attractive model. A subscription product at scale would generate more revenue with less labour. We have run the experiment in both directions across the company's history and the realised outcomes for users in the mass-market version were substantially worse than the realised outcomes for users in the one-at-a-time version, consistent with the literature. We made the choice to optimise for user outcomes rather than for company revenue. We do not claim this is the only defensible choice. We claim it is the choice consistent with the principle.

PRINCIPLE VI

If we don't run it, we don't sell it

Paper-traded back-tests systematically over-state what live traders achieve. The realised implementation shortfall is approximately 180 basis points annualised, and the gap is largely independent of the user's competence or effort. We require every algorithm we sell to be one we are running ourselves.

DRAWN FROM Paper 18 (Oct 2025)

Paper 18 compared paper-traded and live-traded performance for a sample of publicly-available retail systematic trading strategies over the 2018–2024 period. The headline finding: median annualised implementation shortfall, defined as the gap between paper-traded performance and the same strategy run with live execution, is approximately 180 basis points across the sample. The shortfall components are roughly: slippage 60–90bp, missed signals 30–50bp, broker-specific fill differences 20–40bp, and parameter drift between paper and live implementations 20–30bp.

The result is consistent with Perold's (1988) original implementation-shortfall framework and with the Frazzini, Israel, and Moskowitz (2018) institutional-execution-cost study, scaled to the retail context. The implication is that backtests and paper-traded results systematically over-state what end users will achieve, unless the strategy is being run, in parallel, on real capital by the same operator who is selling it.

The "in parallel" qualifier matters. A strategy provider who runs their own version of the strategy in production, but on a different broker, with different position sizes, with different rebalance timing, is not running the same strategy that their customers are running. The published track record will be the strategy provider's track record, not the customer's, and the gap can be material. The honest comparison is between the customer's live result and a strategy-provider implementation that uses the same broker, same trade timing, same position-sizing rules, and same data feed as the customer.

What we do

ProRealAlgos requires that every strategy in the production roster has been run for at least 6 months on real capital in the operator's own account before it is offered to any customer. We run the same strategy, on the same broker, at the same time, on the same data feed. The customer is not getting a different implementation than the operator is.

This has two practical implications. The first is that we cannot offer strategies we have not yet personally validated in live conditions. The waitlist for new strategy releases reflects the time required to accumulate the live track record, not the time required to develop the strategy itself. The strategy is typically built in 4–6 weeks. The live validation period is 6–9 months. The bottleneck is the validation, not the development.

The second is that we cannot scale a strategy beyond the operator's own capital. If we have, say, $500,000 of capital running a particular strategy ourselves, we cannot sell that strategy to customers whose aggregate deployment would push the total participating capital materially beyond what the operator's

own implementation can absorb. This is a capacity constraint that we believe most strategy providers do not honour, and it shows up in the implementation-shortfall data as a slow degradation of strategy performance as the customer base grows.

> *A strategy that is sold without being personally run by the seller is not a strategy. It is a back-test with an invoice attached.*
>
> A direct quote from paper 18

The cost of this discipline

The discipline is expensive. It limits the strategy roster to roughly twenty algorithms, it limits the customer count to roughly fifty active accounts at any given time, and it requires that the operator maintain a meaningful personal capital allocation to each strategy. The alternative model, selling subscriptions to strategies the seller does not personally run, is materially more profitable per unit of operator time. We have chosen the discipline anyway because we believe it is the only honest way to claim that customer outcomes are connected to the documented strategy performance. The principle is the constraint that ProRealAlgos is built around.

Part IV

From principles to twenty algorithms

Twenty algorithms organised into four families. Each one is built on the principles above. Each one is run by us, on our own capital, before it is offered to anyone else.

9 The four families and how they compose

10 What the algorithms do not do

CHAPTER NINE

The four families and how they compose

The twenty algorithms cluster into four families: long-bias trend, cross-asset carry, regime-aware mean reversion, and portfolio overlays. Each family addresses a particular return source. The portfolio benefit is largely about the residual correlations between the families being low.

The twenty algorithms in the ProRealAlgos roster are not twenty unrelated strategies. They are twenty specific implementations of four broad return sources, applied across instruments and time horizons that we have personally validated and that we continue to operate. The four families and their approximate composition are described below. We are deliberately vague about the specific algorithm names and parameter values in this section. The reason is not commercial. The reason is that disclosing the specifics here would create a survivorship-bias problem: readers would over-weight the strategies whose names they could read and under-weight the underlying principles. We discuss the families and their roles in a portfolio rather than the individual strategies.

Family one: long-bias trend (six algorithms)

Six algorithms in the roster are variants of the trend-following work described in Chapter 5. Two operate on broad U.S. equity exposure (S&P 500 and Nasdaq 100). One operates on broad European equity exposure. One operates on Swedish equity exposure (the OMXS30, our home market). Two operate on multi-asset baskets including equities, fixed income, and commodities, providing some natural diversification within the family itself.

The six algorithms differ in their signal calibration (the moving-average windows range from 8 to 12 months), in their position-sizing rules (volatility-targeted versus fixed-fractional), and in their entry-exit conventions (month-end versus weekly, with hysteresis around the signal threshold to reduce whipsaw). The pairwise correlations between the six are in the 0.5–0.8 range, meaningfully correlated but not redundant. A trader running the full family of six experiences a smoother aggregate equity curve than any individual algorithm produces.

Family two: cross-asset carry (five algorithms)

Five algorithms implement the carry work described in Chapter 6, with implementations spanning currency carry, commodity carry, fixed-income carry, equity-dividend carry, and a composite cross-asset carry portfolio. The five sleeves are roughly the same as the academic-literature sleeves but with retail-execution considerations built in: bid-ask spread caps, position-size limits per sleeve, and a re-balancing protocol that minimises round-trip costs.

The composite cross-asset carry algorithm is the most relied-on of the five for new users. The single-sleeve carry algorithms are made available primarily for users who want explicit exposure to particular carry premia (typically traders who have a thesis about a specific asset class).

Family three: regime-aware mean reversion (four algorithms)

Four algorithms exploit short-horizon mean reversion in liquid index futures and ETFs, conditional on a regime filter that suppresses the strategy during trending environments where mean reversion does not work. The regime filter is essentially the same 12-month moving average from family one, applied as a meta-signal: mean reversion is allowed during sideways and trending-up regimes, suppressed during trending-down regimes.

The four algorithms in this family include a 4-day RSI strategy on the OMXS30 (one of the earliest production algorithms, originally published 2014 and continuously operated since), an overbought-oversold strategy on S&P 500 ETFs, a volatility-conditioned mean reversion strategy on Nasdaq futures, and an end-of-week gap-reversion strategy. Pairwise correlations within the family are low (0.2–0.4), reflecting different time horizons and trigger conditions.

Family four: portfolio overlays (five algorithms)

The remaining five algorithms are overlays rather than primary strategies. They operate on top of the family-one, family-two, and family-three algorithms, modifying position sizes based on portfolio-level state. The five include a volatility-targeted sizing overlay (scales positions inversely to recent realised volatility), a drawdown-control overlay (reduces aggregate position size during drawdowns exceeding configurable thresholds), a correlation overlay (reduces position sizes when intra-portfolio correlation spikes), an inflation-regime overlay (modifies the trend-following window during high-inflation regimes), and a regime-bridge overlay (smooths between full-exposure and zero-exposure regimes rather than binary switching).

The overlays are operationally meaningful but conceptually subordinate to the primary strategies. A trader running the full twenty-algorithm portfolio receives the overlay benefits automatically. A trader running only one or two algorithms can use a subset of the overlays optionally.

How the four families compose

The composition is the point. Trend-following, carry, and mean-reversion are individually useful but their cross-family correlations are low: family one and family two have a long-run correlation of approximately +0.15, family one and family three approximately +0.05, family two and family three approximately +0.10. A portfolio that allocates risk evenly across the three primary families achieves a Sharpe ratio of approximately 0.9 in our internal long-sample testing, compared to family-level Sharpe ratios of approximately 0.5 each. The composition is the source of the realised risk-adjusted return; the individual families are merely the building blocks.

This is what we mean when we say that the algorithms are not a list of strategies but a portfolio. A trader who buys access to one of the algorithms in isolation is buying access to one building block. A trader who deploys the composite portfolio of all twenty is buying access to the construction. The construction is what produces the realised risk-adjusted return we document in the track record. We make both available, but we are explicit with new users that the composite portfolio is what the documentation describes and the individual algorithms in isolation will produce more volatile and lower-Sharpe outcomes than the composite.

> A COMPOSITION EXAMPLE
>
> A reasonable starter portfolio for a new trader, after the validation period, is one trend-following algorithm from family one, one composite carry algorithm from family two, and two of the portfolio overlays from family four (the volatility-targeted sizing and the drawdown-control overlay). This combination captures most of the family-level diversification benefit while remaining operationally manageable. We typically expand to additional family-three algorithms and additional family-one regional algorithms after the trader has completed at least 12 months of live operation and demonstrated the discipline to follow the existing algorithms without override.

What the portfolio does not look like

The portfolio we have just described is not the portfolio that an outside observer might expect from a "twenty algorithms" claim. There is no algorithm that picks individual stocks based on news sentiment. There is no algorithm that trades the FTX-style or the Robinhood-style retail-meme trade. There is no algorithm that uses leverage above 2x. There is no algorithm that promises uncorrelated returns to traditional asset classes in the sense that "alternative" funds typically promise. We have deliberately avoided these categories because the underlying research does not support them.

The portfolio we describe is conservative in the sense that the underlying return sources are well-documented in the academic and practitioner literature, the implementations have been live-traded for years, the position sizes are deliberately moderate, and the expected risk-adjusted return is high enough to be attractive but low enough to be believable. We think this is the right level of ambition for a systematic

strategy portfolio. The trader who wants 30 percent annualised returns is not our target customer and we are honest about that.

CHAPTER TEN

What the algorithms do not do

A complete picture requires being explicit about the situations in which our strategies do not work, the risks they cannot eliminate, and the assumptions they rest on.

No systematic-trading portfolio is robust to every imaginable environment. The ProRealAlgos portfolio is robust to most of the environments that have occurred across the historical sample we have tested it on, but there are categorical situations where the portfolio either does nothing, or where it can lose money on a sustained basis. We list them explicitly below because the alternative is the marketing-driven implication that no such situations exist, and that implication is dishonest.

Situations where the portfolio does nothing useful

The portfolio is invested in cash approximately 30–40 percent of the time across the full long-sample test. During these periods it earns roughly the short-term Treasury return. A trader expecting the portfolio to produce equity-like returns continuously is going to be disappointed during the long-tail periods when the portfolio is largely in cash. This is a feature, not a flaw, because the cash periods coincide with the regimes the trend-following strategies are designed to avoid. But the trader needs to understand that "the algorithm is in cash this quarter" is the normal mode during certain market regimes, not a malfunction.

Situations where the portfolio can lose

The portfolio can experience drawdowns exceeding 20 percent. Our internal simulation across 152 years suggests the median maximum drawdown is approximately 14 percent and the 95th-percentile maximum drawdown is approximately 26 percent. A trader who cannot psychologically tolerate a 25 percent drawdown should size the portfolio smaller than otherwise reasonable. This is the central application of Principle III (position size is the strategy).

The portfolio can underperform a simple buy-and-hold benchmark for sustained periods. Across the long-sample test, there are 5-year windows in which the portfolio's cumulative return is 10–15 percentage points below buy-and-hold's. The longest such window is 7 years. A trader who measures performance only against buy-and-hold will be tempted to abandon the portfolio during these windows. The portfolio recovers the relative underperformance during the next major bear market, which is when it was always going to outperform buy-and-hold.

The portfolio can lose money during rapid recovery rallies that follow short, sharp drawdowns. The 2020 COVID episode is the canonical example. The trend-following sleeves exit equities after the moving average is breached, by which point the index has typically already fallen 15–20 percent. The exit locks in the loss. The recovery is then missed for the first 10–15 percent of the rebound. This produces a calendar year of significant underperformance against buy-and-hold even when the strategy worked exactly as designed.

Categorical risks not addressed

The portfolio does not protect against broker default, counterparty default, regulatory action against the trader's account, identity theft of the trader's brokerage credentials, or any of the operational risks of running a retail brokerage account. The strategies are signal generators. The trader is responsible for the operational integrity of their execution venue.

The portfolio does not generate alpha relative to the academic literature's published factor models in any rigorous sense. The strategies use academic-literature factors (trend, carry, mean reversion) implemented with retail-execution considerations. The realised Sharpe ratios are roughly consistent with what the academic literature documents for those factors after costs. We do not claim to have discovered uncorrelated returns to standard factor models. We claim to have made the factors accessible to retail-capacity traders with execution costs and operational practicalities that the academic literature usually abstracts away.

The portfolio does not solve for the trader's tax situation, withdrawal needs, currency translation if the trader's home currency is not USD or SEK, or any of the personal financial planning questions that determine whether the strategies are actually useful to the trader. The strategies are an input to the trader's overall financial plan, not a substitute for one.

IN CLOSING

What this study answers, and what it doesn't

A long document deserves a short summary. The reader who has made it this far is entitled to a clear statement of what the study concludes and what it leaves open. We have tried throughout to avoid the marketing trap of claiming more than the data supports. The closing summary is meant to be honest about the boundary.

What the study concludes

One. The U.S. equity market produces a long-run real return of approximately 6.5 percent annualised, but the distribution around that average is wider than typically presented. Individual decades have produced returns from approximately -1 percent to approximately +14 percent annualised. The investor's realised outcome is the outcome of whichever decade they invest through, not the long-run average.

Two. Bear markets are deeper, longer, and more dependent on the inflation regime than typically presented. The recovery time from a deep bear market in a high-inflation regime can exceed ten years in real terms. The 2022 episode is a recent reminder that this regime can return.

Three. Trend-following on equity indices, applied as a simple 12-month moving-average overlay, produces approximately the same long-run return as buy-and-hold with two-thirds the volatility and half the maximum drawdown. This result is robust across the full 152-year sample and across reasonable parameter variations.

Four. Carry, applied across asset classes, produces a positive risk-adjusted return uncorrelated with trend. The cross-asset composite is more useful than any single-asset implementation.

Five. Position sizing dominates strategy selection. Across thousands of simulated equity curves, sizing rules explain more variance in terminal wealth than signal-strength differences.

Six. Drawdowns drive most retail abandonment. The drawdown that triggers abandonment is typically narrower than the drawdown the strategy was designed to tolerate. The cost of abandonment is approximately 300–500 basis points annualised.

Seven. Knowledge transfer from a designed system to a novice user is dramatically better in one-to-one supervised practice than in documentation alone. The gap is the implicit knowledge that documentation cannot capture.

Eight. The implementation-shortfall gap between paper-traded and live-traded strategies is approximately 180 basis points annualised on average. Closing this gap requires the strategy operator to personally run the same strategy in parallel, on the same broker, with the same data feed.

What the study does not conclude

The study does not conclude that systematic trading is the right approach for every investor. Most investors are better served by a low-cost index fund and a strict savings discipline than by any active

systematic approach, because most investors do not have the time, attention, or temperament to follow a systematic strategy through its full distribution.

The study does not conclude that the strategies that have worked across 152 years will continue to work across the next 50. Markets evolve, factor premia decay, and the relationship between historical and forward returns is partial at best. We believe the underlying mechanisms (the clustering of bad months, the cross-asset carry risk premium, the position-sizing arithmetic, the implementation shortfall) are structural features of how markets actually function. But "we believe" is not "we know".

The study does not conclude that ProRealAlgos is the only or best implementation of the principles described. Other systematic-trading providers operate by largely similar principles, and any reader who concludes after reading this study that they want to pursue systematic trading should evaluate the available providers on the merits rather than because of this document. We have tried to write a study that would be useful to a reader regardless of which provider they ultimately choose, including no provider at all.

IN CLOSING, CONTINUED

Who this is for, and who it isn't

A study that does not specify its audience risks being useful to no one. We have written this for a particular kind of reader and it will not be useful to others. Honesty about the audience is owed.

This study is useful for

The reader who wants to understand systematic trading from first principles before deciding whether to do any of it. The empirical foundations are described in enough detail to support an informed decision in either direction.

The reader who is already running systematic strategies and wants a second opinion on the foundations they have built on. Many of the principles described here are non-controversial in the professional systematic-trading community but the synthesis in one document is unusual. The principles operate as a checklist.

The reader evaluating ProRealAlgos as a potential provider and wanting to understand what the company actually believes about how markets work. The study is the most complete statement of our view that exists. The marketing material is downstream of the study, not the other way around.

The reader who wants to write critical commentary on the systematic-trading industry. We have tried to present the evidence in a form that allows informed criticism. The references at the back are real, the data sources are documented in the underlying papers, and the empirical claims are individually testable.

This study is not useful for

The reader who wants a quick answer to whether systematic trading "works". The honest answer is that some forms of it work, modestly, with disciplines that most retail traders do not have the patience to apply. This is not the answer most quick-answer-seekers want.

The reader who is looking for the optimal trading strategy. There is no such thing. The strategies described here are robust, evidence-based, and personally validated, but they are not optimal in any rigorous sense. A reader looking for optimality is looking for a thing that does not exist.

The reader who needs the trading approach to be intellectually exciting on a weekly basis. Systematic trading, done well, is dull on a weekly basis. The interesting work is done in advance, in the strategy design and validation. The trading itself is mostly a matter of executing the same rules week after week and not interfering. Readers who want excitement should find a different hobby.

The reader who needs to make trading profits in the next 90 days. The strategies described here operate on horizons measured in years. They reliably underperform on horizons measured in weeks or months. A reader with a 90-day need should not be in any active trading approach, systematic or otherwise; they should be in cash.

> *The most honest statement we can make about ProRealAlgos is that we have built a portfolio of strategies that we personally run, that we*

believe will produce a positive risk-adjusted return over the next decade with reasonable probability, and that we have spent twenty years learning what does not work in order to arrive at the small set of things that we believe do. We make no guarantees. We do not need to.

Carl C.G Eriksson, Stockholm 2026

REFERENCES

The eighteen papers

Full citations for the research papers that this study is built on. Each paper is available separately on request from ProRealAlgos.

---

[01]	2026.05	Eriksson, C.C.G. "The 'best days' myth: why missing the worst days matters more than missing the best, and why investors realistically can't do either." ProRealAlgos Research Note 01, May 2026.
[02]	2022.10	Eriksson, C.C.G. "A century and a half of bear markets: depth, duration, and recovery across fifteen U.S. equity drawdowns, 1871–2022." ProRealAlgos Research Note 02, October 2022.
[03]	2023.06	Eriksson, C.C.G. "Equities, bonds, and the inflation regime: what 152 years of return data reveal about the asset-class consequences of price-level instability." ProRealAlgos Research Note 03, June 2023.
[04]	2019.03	Eriksson, C.C.G. "The 12-month moving average as a regime filter: returns, volatility, and drawdowns conditional on a single trend signal, 1871–2018." ProRealAlgos Research Note 04, March 2019.
[05]	2017.02	Eriksson, C.C.G. & Lindberg, M. "Where the equity risk premium comes from: a Gordon-style decomposition of U.S. equity returns into yield, growth, and revaluation, 1871–2016." ProRealAlgos Research Note 05, February 2017.
[06]	2020.07	Eriksson, C.C.G. "Lump sum, dollar cost, and the cost of waiting: 1,097 ten-year simulations on what drip-feeding actually buys you." ProRealAlgos Research Note 06, July 2020.
[07]	2018.11	Eriksson, C.C.G. "A century and a half of trend-following: what a 12-month signal applied to U.S. equities does for return, volatility, and drawdown over 147 years." ProRealAlgos Research Note 07, November 2018.
[08]	2022.01	Eriksson, C.C.G. "The concentration problem in single-stock returns: why most stocks underperform Treasury bills, and what this implies for active management." ProRealAlgos Research Note 08, January 2022.
[09]	2024.09	Eriksson, C.C.G. & Engström, J. "Carry, everywhere: what thirty years of cross-asset evidence reveals about a single risk premium." ProRealAlgos Research Note 09, September 2024.
[10]	2023.04	Eriksson, C.C.G. "The 60/40 verdict: why a portfolio that worked for forty years failed in 2022, and what to do about it." ProRealAlgos Research Note 10, April 2023.
[11]	2025.08	Eriksson, C.C.G. "Gold as hedge: fact and folklore. What 50 years of post-Bretton-Woods data reveal about the metal's actual role in a portfolio." ProRealAlgos

Research Note 11, August 2025.

[12] 2026.01 Eriksson, C.C.G. "What predicts the next decade: CAPE, real rates, and other long-run forecasters tested across 145 years." ProRealAlgos Research Note 12, January 2026.

[13] 2018.08 Eriksson, C.C.G. & Andersson, P. "The graveyard of strategies: a survivorship audit of trading systems published in industry and academic literature, 1995–2010." ProRealAlgos Research Note 13, August 2018.

[14] 2021.05 Eriksson, C.C.G. "Override and underperform: the empirical cost of human discretion in rule-based trading systems." ProRealAlgos Research Note 14, May 2021.

[15] 2020.03 Eriksson, C.C.G. "Position sizing dominates strategy selection: 10,000 simulated equity curves on what determines whether a strategy enriches or ruins its trader." ProRealAlgos Research Note 15, March 2020.

[16] 2024.02 Eriksson, C.C.G. "The drawdown-abandonment curve: when traders quit, and what it costs them. A synthesis of behaviour-gap evidence in retail systematic trading." ProRealAlgos Research Note 16, February 2024.

[17] 2023.11 Eriksson, C.C.G. "Knowledge transfer in expert systems: what novice users need to know to capture the documented returns of a designed system." ProRealAlgos Research Note 17, November 2023.

[18] 2025.10 Eriksson, C.C.G. & Wilson, J. "Implementation shortfall in retail systematic strategies: signal versus execution, paper-traded and live-traded performance compared across publicly-available retail systems, 2018–2024." ProRealAlgos Research Note 18, October 2025.

FORTHCOMING

What's next in the pipeline

The following papers are in active development or planned for the next twelve to eighteen months. Their conclusions are not stated in this edition because the underlying work is not yet complete. They will be folded into editions 1.1, 1.2, and 2.0 as they reach publication.

A living document should be honest about what it does not yet contain. The papers below represent the next eight pieces of research in the ProRealAlgos pipeline. Some are in active drafting. Some are scoped but unstarted. Their inclusion here is a commitment to transparency about the direction of the research programme, not a guarantee that each paper will be completed exactly as described.

Forthcoming research is by definition speculative. Some of these papers will be published approximately as outlined. Others will be substantially revised in scope, methodology, or conclusion once the empirical work is done. Some will be discontinued because the data does not support a publishable result. We will not publish a paper merely because we said we would, and the absence of a forthcoming paper from a future edition is not, by itself, a sign that anything went wrong with it.

In active development

19	Q3 2026	Multi-asset trend-following DRAFT *Extending paper 07 from U.S. equities to a 12-asset cross-sectional universe. Co-authored with M. Lindberg.*
20	Q4 2026	Volatility-targeted position sizing DRAFT *A 25,000-path study comparing five sizing rules across multiple Sharpe-ratio regimes. Extends paper 15.*
22	Q1 2027	Retail behaviour during the 2022 bear DRAFT *Brokerage-panel data on retail abandonment rates during the 2022 stock–bond drawdown. Co-authored with P. Andersson. Extends paper 16.*
24	Q3 2026	Implementation shortfall, refreshed 2018–2026 DRAFT *A two-year update of paper 18, with the larger sample now permitting cohort-level comparisons across broker types.*

Scoped but unstarted

21	2026	Currency carry in high-inflation regimes PLANNED *A regime-conditional study of currency carry, focusing on the 1970s and 2021–2024 episodes. Co-author TBD.*

23	Q1 2027	Knowledge-transfer cohort study PLANNED *Empirical follow-up to paper 17 using ProRealAlgos onboarding cohorts as the data sample. Will test the one-trader-at-a-time hypothesis against an alternative cohort-based delivery model.*
25	2027	Index rebalance flows as a return source PLANNED *A new direction. Examines whether the predictable order flows around major index rebalances constitute a tradeable signal at retail capacity.*
26	2027	Sector rotation in trend-following overlays PLANNED *Tests whether layering a sector-rotation signal on top of the broad trend-following framework improves risk-adjusted returns enough to justify the additional implementation complexity.*

> **A NOTE ON CO-AUTHORS AND EXTERNAL RESEARCH**
>
> Several forthcoming papers are co-authored with researchers outside ProRealAlgos. We deliberately seek external collaborators on topics where the work benefits from a second methodological perspective, and we credit them prominently. Reader feedback on forthcoming topics is welcome at `research@prorealalgos.com`. We do not commit to incorporating any specific feedback, but we read all of it.

FOR THE READER

How to track updates

The mechanics of staying current with this living document.

Edition notifications

When a new edition (major or minor) is released, registered readers receive a single email noting the edition number, the date, and a brief summary of what changed. The email contains a link to the new edition and to a summary of the revision history since the prior edition. We do not use the list for marketing, for promotion of other ProRealAlgos products, or for anything other than study updates.

To register, visit `prorealalgos.com/study` and provide an email address. To unsubscribe at any time, follow the link at the bottom of any notification email or write to `study@prorealalgos.com`.

Past editions

All public editions of this study, from 2.0 onward, are archived and available on request. A reader wanting to verify what a previous edition concluded on a particular topic can request the relevant edition by writing to `study@prorealalgos.com` with the edition number. Internal drafts (versions 0.x and 1.x) are not made publicly available.

Citing this study

Readers who wish to cite material from this study in their own work should include the edition number alongside the citation, because conclusions may have changed between editions. The recommended citation format for this edition is:

```
Eriksson, C.C.G. (2026). What works (and what doesn't) in the stock market: a
comprehensive synthesis of what 150 years of market history actually teach
us. Edition 2.0. ProRealAlgos, Stockholm.
```

Reporting errors

We have tried to be careful with the numbers, citations, and historical claims throughout this document. Errors nonetheless will occur. A reader who finds a factual error, a misattributed citation, or a logical inconsistency is invited to write to `errata@prorealalgos.com`. Corrections that survive review will be incorporated into the next edition with attribution to the reader who identified them, unless the reader prefers to remain anonymous.

What this is not

The notification list is not a paid subscription. There is no fee to receive updates. The study itself is freely distributed in PDF form, and we do not paywall future editions. We make the work available because

credibility in this field requires showing the work, and the work loses its credibility if it is hidden behind a paywall. Commercial revenue at ProRealAlgos comes from the algorithm-licensing business described in Part IV, not from this document.

> *The study is the artefact of the work, not the product. The product is the systematic-trading portfolio we operate ourselves and license to a small number of traders. The study exists to show that the portfolio is built on something real.*
>
> Carl C.G Eriksson, Stockholm 2026

End of edition 2.0.

What works (and what doesn't) in the stock market.

A comprehensive synthesis of what 150 years of market history actually teach us.

Drawn from 18 research papers; condensed into six principles and twenty algorithms.

Carl C.G Eriksson

ProRealAlgos · Stockholm · May 2026

Edition 2.0 · First public edition

Next planned edition: 2.1 · Q3 2026

prorealalgos.com/study