Slow to Catch Up: Return Predictability on the Paris Bourse, 1845–2025*

Jérémy Fouliard†

This version: November 19, 2026

Abstract

I study the real-time predictability of stock returns over 180 years with a new security-level dataset for the Paris Bourse (DFIH, 1845–1976, matched with EUROFIDAI, 1977–2025). Forecasts combine models as they become available through Bernstein Online Aggregation, a rule with a bounded regret that assumes nothing about the return-generating process. The market index is barely predictable: macro-financial predictors explain 0.5% of its half-monthly variance, mostly in crises, and the larger predictability of lagged index returns is a stale-price artefact. Individual stocks are more predictable, by about 0.6%, mostly because they react to market-wide news with a delay that is strongest in small stocks, not explained by stale prices, and unchanged over 155 years. Momentum earns 7% a year net of costs in almost every decade since 1900, while the historical volatility and beta premia vanish after 1977. Markets became better at pricing risk, not faster at absorbing information.

Keywords: return predictability, online learning, expert aggregation, momentum, lead-lag effects, financial history, Paris Bourse.

JEL: G12, G14, G17, N23, N24, C53.

*I am very grateful to Pierre Gaillard, Pierre-Cyrille Hautcoeur, Eric Monnet, Hélène Rey and Angelo Riva as well as to seminar participants at Paris School of Economics. I am also grateful to the DFIH team and especially to David Smadja, for their help with the data. I also thank Eurofidai (European Financial Institute, UAR CNRS/ ESSEC 3390) for access of the data. I am highly indebted to London Business School and especially to Chris Le Grice and Helene Rey for granting me access to high-performance computing resources. The views expressed herein are those of the author and do not necessarily reflect the views of the institution.

†jeremy.fouliard@insee.fr

1 Introduction

Whether stock returns can be predicted has fascinated academics, practitioners and policymakers for more than a century. Louis Bachelier, a student of Henri Poincaré, formalised the idea that speculative prices should be unpredictable in 1900, using data from the Paris Bourse. For investors, predictability is the promise of profit; for policymakers, it is a warning signal, because the classic narrative of financial crises rests on the joint dynamics of credit and asset prices (Kindleberger, 1978; Minsky, 1986; Jordà et al., 2015; Greenwood et al., 2022; Hennig et al., 2023). For academics, it is the empirical content of the Efficient Market Hypothesis, the "organizing principle for empirical financial economics" (Cochrane, 2014): if prices fully reflect available information (Fama, 1970), predictable variation in returns can only reflect predictable variation in required compensation for risk.

The empirical literature on return predictability is immense, and it has a striking generational structure. A first generation related market returns to valuation ratios, interest rates and macroeconomic variables in predictive regressions; a second responded with economic restrictions, forecast combinations, diffusion indices and regime-switching models; the current generation uses machine learning and text.[1] The pattern itself is informative: each generation's dominant model tended to fade once it was known, because competition among asset managers ensures that a predictable pattern, once identified and exploited, tends to disappear (McLean and Pontiff, 2016; Haddad et al., 2025). Forecasting models must therefore be renewed continually, and predictability, where it exists, is concentrated in "pockets" (Timmermann, 2018; Farmer et al., 2023). Most of this evidence, however, is in-sample: re-estimated recursively, as a real-time investor would have, almost none of these predictors reliably beats the historical mean out of sample (Welch and Goyal, 2008; Goyal et al., 2024)[2]

Three hypotheses summarise what this literature leads one to expect. First, competition keeps the signal-to-noise ratio of any predictive model very low, so that adding predictors is unlikely to raise out-of-sample

[1] For the first generation, see, among many others, Rozeff (1984), Campbell and Shiller (1988a) and Fama and French (1988) on the dividend–price ratio; Kothari and Shanken (1997) and Pontiff and Schall (1998) on the book-to-market ratio; Breen et al. (1989), Campbell (1987) and Fama and French (1989) on interest rates and spreads; Nelson (1976) and Fama and Schwert (1977) on inflation; and Lamont (1998), Baker and Wurgler (2000), Boudoukh et al. (2007), Lettau and Ludvigson (2001), Guo (2006), Santos and Veronesi (2006), Rangvid (2006), Cooper and Priestly (2009), Driesprong et al. (2008) and Hai Hong and Stein (2007) on payout ratios, issuance, the consumption–wealth ratio, volatility, labour income, output, oil prices and lagged industry returns. For the second, see Campbell and Thompson (2008) and Ferreira and Santa-Clara (2011) on economic restrictions; Elliott and Timmermann (2004), Rapach et al. (2010) and Lin et al. (2018) on combinations; Ludvigson and Ng (2007), Neely et al. (2014) and Kelly and Pruitt (2015) on diffusion indices; and Guidolin and Timmermann (2007), Henkel et al. (2011) and Dangl and Halling (2012) on regime switching. For the third, see Gu et al. (2020), Kelly et al. (2024), Sirignano and Cont (2019), Qiu et al. (2020) and Steyn et al. (2020).

[2] Welch and Goyal (2008) find that most univariate predictors would have done worse than the prevailing historical mean, and that even their in-sample significance is fragile. Revisiting the original 17 predictors and 29 proposed after 2008, with data to 2021, Goyal et al. (2024) find that more than a third are no longer significant even in sample, half of the rest perform poorly out of sample, and only a handful predict both in and out of sample.

accuracy much, even though small gains can be economically meaningful.[3] Second, whatever predictability exists should be concentrated in particular segments of the market, in particular periods, or both: in securities that attract little attention, and in periods when risk is mispriced. Third, markets should have become more efficient over time, because electronic trading, lower transaction costs, financial integration and the growth of the analyst industry all reduce the opportunities for predictable patterns (Hendershott et al., 2011; Foucault et al., 2023).

Testing these hypotheses requires data that span several generations of models, several market structures and, ideally, the full cross-section of listed securities rather than a single index. Almost all the evidence comes from the United States after 1926, and almost all of it is about the aggregate market or about portfolios formed within the CRSP universe. This paper uses a new dataset that meets these requirements. The Data for Financial History (DFIH) project at the Paris School of Economics has digitised the official quotation lists of the Paris Bourse from 1845 onwards (Lagneau-Ymonet and Riva, 2010; Ducros et al., 2018): prices and sometimes dividends for every listed security at a half-monthly frequency, together with information on the issuing firms. I match this archive, for the first time, with the EUROFIDAI database of daily prices from 1977 to 2025, and obtain a continuous security-level panel of about 15,000 securities and 3 million half-monthly returns over 180 years. Before 1914 Paris was, with London, one of the two main international financial centres (Hautcœur and Riva, 2012); in 1913 its stock market capitalisation was 78% of French GDP, twice the ratio of the United States (Rajan and Zingales, 2003). Its history spans the gold standard, two world wars, financial repression, nationalisations, deregulation in the 1980s, the move to electronic trading in 1986–89, and the European integration of exchanges. It is also a market for which the microstructure (the monopoly of the *agents de change*, the parallel *Coulisse*, the forward market, the tax on stock-exchange transactions) is well documented (Hautcœur et al., 2010; Hautcœur and Riva, 2012), which makes it possible to ask *why* predictability changes, not only whether it does.

To test these assumptions, I adopt a framework originating from the online learning literature, a body of statistical theory which originates in the work on repeated games of Robbins (1951), Hannan (1957) and Blackwell (1956) and which studies problems in which information is revealed sequentially. It is therefore particularly well suited to address overfitting concerns. Online learning has been a main topic of research in the theory of machine learning and has become the infrastructure of large-scale forecasting systems in various fields[4]. While applications in economics [5] remain limited (Amat et al., 2018; Fouliard et al., 2021),

[3] Campbell and Thompson (2008) show that a monthly out-of-sample R^2 of 0.5% can raise the average excess return of a mean-variance investor's portfolio substantially. There may also be a publication bias working against predictability: a researcher who finds a profitable model has an incentive to trade on it rather than publish it.

[4] see Fouliard et al. (2021) for a more complete literature review

[5] The online learning literature shares several distinctive features with many fields in economics, especially the real time combination literature. Since the seminal work of Bates and Granger (1969), forecast combinations are viewed as a simple and

the online learning framework has been extensively used in finance. A well-established literature studies sequential investment problems, starting with the seminal contribution of Cover (1966). This line of research re-emerged in the 1990s with Shafer and Vovk (2005) and has since generated a substantial body of work (Jun et al., 2017; Jang et al., 2023). In particular I adopt the framework of sequential predictions with expert advice as the forecasting environment of this paper, a framework which was developed by Vovk (1990), Littlestone and Warmuth (1994) and Cesa-Bianchi and Lugosi (2006). At each date, a forecaster combines the forecasts of several models ("experts") [7]with weights that depend only on their past performance; after the return is revealed, the weights are updated.

I use the Bernstein Online Aggregation (BOA) rule of Wintenberger (2017), whose adaptive learning rates deliver a second-order bound on the regret, the difference between the cumulative loss of the aggregate and that of the best expert in hindsight. In the online learning framework, the regret measures how much the forecaster loses by not knowing in advance which model would turn out to be best: a bound that grows more slowly than the number of periods guarantees that, on average, the aggregate does as well as that model, whatever the sequence of returns. Three features make this framework well suited to the question. It makes no assumption about the return-generating process, which may be stationary, non-stationary or even adversarial, as one would expect when prices are set by agents who themselves try to forecast them. It aggregates models *as they become available*: predictive regressions are used from the start, forecast combinations and diffusion indices only once they were proposed, so that the forecaster never uses a model that did not yet exist, a form of look-ahead bias that is rarely addressed. And it is designed against overfitting, since weights are updated by a rule rather than estimated, and the only estimated objects are the experts themselves.

Two methodological contributions are needed to make this framework work on a long security-level panel. The first is a *pooled* aggregation rule. Security by security, a few hundred half-monthly observations are not enough for any rule to learn which of forty experts is best when the signal is as weak as it is in returns: the weights never leave the historical mean. I therefore learn a *single* weight vector for all securities, updated at each date with the cross-sectional average of the linearised loss; forecasts remain security-specific. The second is a set of real-time, cross-sectionally pooled experts: predictive regressions whose slopes are estimated on all securities (within-security, with security-specific intercepts), a shrinkage expert equal to

effective way to perform better than individual models (Timmermann (2006), Elliott and Timmermann (2008), Diebold and Shin (2019)). [6]

[7]I am aware that this terminology could be misleading. In the real-time forecast combinations literature (Capistrán and Timmermann (2009), experts often refer to individual survey forecasters. The term "expert" we use throughout the paper stands instead for any forecasting model, variable, or individuals survey forecasts. It comes from the online learning literature and reminds us that the problem is not a standard estimation problem but that it can be viewed as a "meta-statistical" approach. It is also worth pointing out that the only potential source of overfitting in our analysis could come from the estimation of individual forecasting models but not from the aggregation rule (designed to avoid overfitting, since overfitted experts will quickly get low weights in the aggregation rule)). Our contribution is not in the estimation of individual experts but in their optimal aggregation .

the mean return of all securities up to t, and a characteristics expert built from real-time Fama–MacBeth regressions on momentum, reversal, volatility, beta, illiquidity and dividend yield, in the spirit of Gu et al. (2020).

The paper also confronts two measurement problems that are specific to long historical data and that generate predictability no investor could have earned. The first is stale pricing. Historical equal-weighted indices average prices that are not all updated every period, so index returns are positively autocorrelated ($\rho_1 = 0.32$ at the half-monthly frequency) in a way that cannot be traded (Scholes and Williams, 1977; Lo and MacKinlay, 1990a). With the lagged index among the predictors, the market-level out-of-sample R^2 is 9.6%; with macro-financial predictors only, it is 0.5%. I exclude predictors built from stock prices from the market-level analysis and report the artefact separately. The second is the mechanical link between a noisy price and the next return. A dividend–price ratio measured at t contains P_t; if P_t is recorded with error, the ratio is too high exactly when the next return is too high, which produces predictability that is pure measurement error (Blume and Stambaugh, 1983; Asparouhova et al., 2013). I lag the ratio by one period, which breaks the link.

The results are organised in three levels, from the market to the cross-section.

At the market level, with macro-financial predictors only (interest rates, bond yields, money and credit aggregates, exchange rates and, in the modern period, the VIX and oil prices), the Bernstein Online Aggregation rule, which combines the forecasts of all models with weights updated each period according to their past accuracy, attains an out-of-sample R^2 of 0.48% per half-month against the historical mean over 1870–2025 and 1.31% against zero; the Clark–West statistic is 3.1, the Diebold–Mariano statistic 1.2. The gain is 0.30% before 1977 and 0.76% after, it is worth 1.9% per year in certainty-equivalent return to a mean-variance investor, and it is entirely concentrated in half-months of high recent volatility, a state known in real time (R^2_{OOS} of 1.3% against 0.0%), and in crises (1.0% against 0.1%). All aggregation rules give the same order of magnitude. This is the upper bound that Campbell and Thompson (2008) regard as economically meaningful, and it is one twentieth of what the stale-price artefact would have suggested.

At the security level, the pooled BOA forecasts beat the security's historical mean by a pooled R^2_{OOS} of 0.69% ($t = 6.4$), the zero forecast by 0.63% ($t = 3.0$) and the mean of all securities, a benchmark that already removes the noise in each security's own mean, by 0.62% ($t = 3.2$). A timing portfolio that holds each security only when its forecast is positive has a Sharpe ratio of 0.61 against 0.11 for equal-weighted buy and hold. The aggregation rule reveals where this predictability comes from: it gives 60% of its weight before 1977 and 80% after to a single expert, the regression of a security's return on the lagged return of the market, 14% and 7% to the characteristics expert, 10% and 7% to the zero forecast, and almost nothing

to the forty other macro-financial predictors or to the combination and diffusion-index experts. Individual securities react to market-wide information with a delay, the lead-lag effect of Lo and MacKinlay (1990b), Brennan et al. (1993) and Hou (2007): the predictability is strongest for small and long-listed securities, it rises with market volatility (the annual R^2_{OOS} reaches 4.3% in 1987, 5.6% in 2008 and about 3% in 1927 and 1930) and shows no structural break, no business-cycle pattern and no response to any institutional reform in 155 years. Measured against the all-securities mean, the predictability of the aggregate is 0.29% before 1977 and 0.75% after, so that electronic trading did not remove it, and the lead-lag expert accounts for about four fifths of it; without that expert, the other experts deliver 0.11%. It is not stale pricing. Restricting the evaluation to securities quoted on consecutive dates, which removes the observations in which a stale quote catches up with the market, raises the R^2_{OOS} to 0.71%; and with the daily EUROFIDAI trading data for 1977–2025, which say whether a security actually traded on the day its price is sampled, the R^2_{OOS} is 0.90% ($t = 3.4$) among securities traded on both sampling dates, where no price can be stale, against 0.13% ($t = 0.6$) among securities whose sampled price is a stale quote. The delayed reaction is a property of prices formed by trades, and it is concentrated in small, actively traded securities: 1.3% in the smallest quintile of market capitalisation and nothing in the largest.

At the cross-sectional level, the characteristics expert is the only one whose forecasts can be turned into a long–short portfolio, and momentum (Jegadeesh and Titman, 1993) is its main ingredient. A decile portfolio formed on past 12-month returns (skipping the last month) earns 10.8% per year gross ($t = 5.5$) over 1900–2025 and 7.1% net of a 0.5% cost per transaction ($t = 3.6$), with a positive return in twelve of thirteen decades, the 1970s being the exception. The premium is 11.6% in the historical period and 10.1% in the modern one, it survives size-neutral sorts (8.2%), penny-stock filters, costs up to 1% per transaction and, in the modern period, costs measured by each security's own bid–ask spread; it is, however, a premium of small and mid-capitalisation securities, since a value-weighted version earns 1.7% per year over 1977–2025 ($t = 0.4$), a size gradient steeper than the one Hong et al. (2000) document for the United States; and it crashes, as in the United States (Daniel and Moskowitz, 2016), in the rebounds that follow bear markets: the fifteen worst months are all market rebounds, and the interaction of a bear-market indicator with the market return has a t-statistic of -3.3. The French and US momentum factors have a correlation of 0.32 (0.49 after 1977) and the same mean. In contrast, the premia on high volatility and high beta are 16.6% and 14.3% per year before 1977 and zero afterwards, the long-term reversal premium is negative throughout (De Bondt and Thaler, 1985), and an illiquidity premium exists only in the modern period. The time variation of the momentum premium is explained by lagged market volatility and, weakly, by exchange controls; neither the transaction tax nor the closures of the forward market nor the coexistence of the official and parallel markets move it.

This paper asked whether stock returns could have been predicted in real time, by a forecaster using only the models and the information available at each date. Over 155 years of the Paris Bourse, they are predictable, but not everywhere and not always profitably. The market as a whole is barely predictable, and only in periods of monetary turmoil.[8] Individual stocks are more predictable, mostly because small stocks react to market-wide news with a delay that has not shrunk in 155 years, not even with electronic trading. Across stocks, past winners have kept outperforming past losers in almost every decade since 1900, whereas the premium once earned by risky stocks disappeared when diversification became cheap.[9] Markets have become better at pricing risk, but not faster at absorbing information.

These results support two conclusions. First, the prices of individual stocks are not a random walk, and this is not the mechanical predictability created by stale prices that explained much of the earlier lead-lag evidence: it is as strong among stocks that traded on both dates as in the full sample. It is not, however, evidence of inefficiency, since it cannot be traded profitably after costs.[10] Second, for individual stocks, the lagged market return beats the predictors that have dominated the literature for four decades: it receives 60 to 80% of the weight of the aggregation rule, and all other experts together, from the dividend–price ratio to the characteristics model, deliver less than a fifth of the predictability.[11]

The paper is related to four strands of work. First, the long-run predictability literature: Golez and Koudijs (2018) use four centuries of dividend yields in several countries, Goetzmann and Huang (2018) and Geczy and Samonov (2016) study momentum in Imperial Russia and in the United States back to 1800, Chabot et al. (2014) in Victorian Britain, and Baltussen et al. (2021) global factor premia since 1800; all use index or portfolio data. The security-level French panel used here is, to my knowledge, the longest continuous one that runs from the nineteenth century to the present day in a single market: it covers every security of the official quotation list at a half-monthly frequency, whereas most historical security-level datasets cover a subset of securities, at a monthly frequency, and end before the modern period. Second, the French financial-history literature that built and documented the DFIH data and the long-run Paris index (Le Bris and Hautcœur, 2010; Hautcœur and Riva, 2012; Hautcœur et al., 2010). Third, the real-time forecasting literature on the aggregate market (Welch and Goyal, 2008; Campbell and Thompson, 2008; Rapach et al., 2010; Rapach and

[8] Macro-financial variables explain half a percent of the variance of the half-monthly index return, concentrated in a few episodes of large monetary shocks.

[9] Can investors make money from this? Hardly from the delay, which is concentrated in small stocks at a horizon where trading costs absorb the gains; somewhat from the market forecasts, worth about 2% a year to a mean-variance investor; and clearly from momentum, which earns about 7% a year after costs, but only in small and mid-sized stocks and at the price of occasional severe crashes.

[10] Any test of efficiency is a joint test with a model of expected returns (Fama, 1970). In the sense of Jensen (1978), a market is efficient with respect to an information set if no profit can be made from it net of trading costs; the half-monthly long–short portfolios of Section 6 turn a gross return of 10% a year into a net loss.

[11] Against the all-securities mean, the aggregate attains 0.62% with the lead-lag expert and 0.11% without it (Appendix J); no macro-financial predictor receives more than 0.6% of the weight.

Zhou, 2013) and on the cross-section with machine learning (Gu et al., 2020; Kelly et al., 2024); relative to it, I add a sequential-prediction rule with regret guarantees, the real-time introduction of models, and the pooled aggregation and pooled experts. Fourth, the literatures on lead-lag effects and on the decay of anomalies (Lo and MacKinlay, 1990b; Chordia and Swaminathan, 2000; Hou, 2007; McLean and Pontiff, 2016; Chordia et al., 2014; Linnainmaa and Roberts, 2018), to which I add 180 years of French evidence that the delayed reaction of individual securities to the market and the momentum premium did not decay at all.

Section 2 describes the data, the cleaning rules and the two measurement problems. Section 3 presents the sequential-prediction framework, the experts, the pooled aggregation rule and the evaluation design. Section 4 reports the market-level results, Section 5 the security-level results and their evolution over 155 years, Section 6 the cross-sectional results, and Section 7 the French factor dataset, momentum crashes, the international comparison and the mechanisms. Section 8 concludes.

2 Data

2.1 Sources

The dataset combines two sources that had never been matched. From 1845 to 1976 it draws on the DFIH database, a long-run stock exchange dataset for the French markets built from two archival corpora: the official quotation lists of the Paris Bourse (over 600,000 pages) and the yearbooks published annually by the *Compagnie des agents de change* or by private publishers (about 300,000 pages). From the quotation lists, information on issuers, securities, prices and dividends was digitised, preserving dates, names and currencies; the yearbooks add information on the companies (sector, head office, factories). Figure A.1 in Appendix A shows an excerpt. The collection and cleaning of the DFIH database required extensive work (Ducros et al., 2018), and it has not been used for return forecasting before. DFIH records prices twice a month, on the 1st and the 15th. From January 1977, prices are taken from the EUROFIDAI–ESSEC daily stock database, which provides daily prices, adjustment coefficients for corporate events and dividends for all securities listed in Paris.

The two sources were matched using the SICOVAM code, the numerical identifier assigned to each security listed in Paris by the *Société interprofessionnelle pour la compensation des valeurs mobilières*, created in 1949 to allow the transfer of securities between accounts. Securities issued before 1949 have no SICOVAM code, and a security may leave the market for many reasons (bankruptcy, delisting, suspension), so the matching is necessarily incomplete: 29.2% of the DFIH securities quoted after 1 December 1976 are matched to a EUROFIDAI security, and 46% of the EUROFIDAI securities quoted before 15 January 1977

have a DFIH counterpart. Unmatched securities are treated as distinct securities. This conservative choice can only lower measured predictability, since it shortens the learning history available for some securities, whereas false matches would distort it. Section A.3 in the Appendix describes the sample of securities in more detail

2.2 A common half-monthly calendar

All the analysis is conducted on a common half-monthly calendar over 1845–2025: the 1st and the 15th of each month, the native frequency of DFIH. The convention matters for the stale-price problem discussed below: a price several days old on a calendar date behaves like a stale quote. For 1977–2025, each trading day is assigned to the next calendar date (day 1 to the 1st, days 2 to 15 to the 15th, days 16 to the end of the month to the 1st of the next month); the price of a security on a calendar date is its last adjusted closing price of the half-month, and its return is the sum of its daily log returns over the half-month, that is, the log change between the last prices of two consecutive half-months[12]. Staying at the DFIH frequency throughout, rather than using the daily data after 1977, keeps the learning problem comparable across the two sources (the same number of observations per year, the same horizon, the same scale of returns) and avoids a change of frequency that would coincide exactly with the change of source. Half-monthly returns are long enough for bid–ask bounce to matter less than at the daily frequency, and short enough to preserve 24 observations per year, which is what makes 180 years of learning possible.

2.3 Returns and corporate events

EUROFIDAI provides a price series adjusted for all corporate events (splits, rights issues, spin-offs, distributions and dividends). DFIH provides adjustment coefficients for the corporate events detected in the archives, but the list is partial. I therefore rescale a DFIH price series whenever a corporate event is suspected: if the absolute relative price variation between two consecutive dates exceeds 80%, all prices before the event are multiplied by $s = 1 + v$, where v is the relative variation. This affects slightly more than 100 securities, listed in Appendix A. Returns are continuously compounded,

$$r_{i,t} = \log\left(P_{i,t}/P_{i,t-1}\right), \tag{1}$$

[12] A security with no daily record in a half-month has no price on that calendar date, and its next return spans several periods; the last record of a half-month is therefore never more than a half-month old. EUROFIDAI also records a price on days without trading, however, so that the last price of a half-month need not be a transaction price: this is the case for 8 to 11% of securities in 1977–1985, below 1% in 1995 and 20 to 25% since 2005 (Section 2.6), and Appendix J shows that the predictability of individual securities does not come from these observations. Covariates follow the same calendar: the daily market return is summed over the half-month, and the other covariates take their last value known at the date, so that no information posterior to the date is used.

where $P_{i,t}$ is the adjusted closing price (the opening price when the closing price is missing) and $t-1$ denotes the previous date at which security i was quoted. When a security is not quoted on consecutive calendar dates, its return spans several periods and is known only when the next quote appears; Section 3 explains how such returns are handled so that no information is used before it is available.

2.4 Cleaning rules

Historical price data contain transcription errors that would dominate any squared-error evaluation if left in. I apply four rules, fixed before any forecasting was done, once, to the half-monthly series, so that every result in the paper rests on the same data. (1) *Round trips*: a log return larger than 1 in absolute value (a price multiplied or divided by 2.7) that is immediately reversed by the next return is a quotation error; the intermediate price is removed together with the two returns around it. (2) *Hampel filter*: a price more than three times above or below the median of its 13 neighbouring quotes (±6 dates) is removed, with the two surrounding returns; genuine and persistent changes of level are not affected because the local median follows them. (3) *Cap*: an increase of more than 300% in a half-month is a data error and is removed; large declines (bankruptcies) are kept. (4) *Mixed series*: a few identifiers alternate between two price levels for years, which happens when a share and a certificate of the same firm were recorded under one identifier; a series is removed entirely when more than 10% of consecutive return pairs exceed 30% in absolute value with opposite signs and its first-order autocorrelation is below −0.3, or when more than 20% of pairs alternate. Table A.1 in Appendix A reports the effect: 3,723 returns are removed out of 2.96 million (0.13%), namely 572 round trips, 1,680 prices caught by the Hampel filter and 407 capped increases; 784 securities are touched by at least one rule, and six series (875 observations) are removed as mixed. The two sources are affected in the same proportion: the EUROFIDAI period holds 78% of the observations and 73% of the removed returns. The equal-weighted market index before 1977 is recomputed on the cleaned series.

2.5 Market index and predictors

Two equal-weighted market indices serve as the market return $R_{m,t}$: for 1845–1976, the cross-sectional average of the half-monthly log returns of all securities quoted at t and $t-1$,

$$\bar{r}_t = \frac{1}{S_t} \sum_{i=1}^{S_t} r_{i,t}, \tag{2}$$

where S_t is the number of securities with a return at t; for 1977–2025, the EUROFIDAI equal-weighted index of common stocks sampled on the same calendar.[13]

The macro-financial predictors are chosen as variables that a forecaster living at the time could have observed, at least indirectly, and would plausibly have used. Levy-Garboua and Monnet (2016) provide monthly series from 1800 for the Banque de France's discount rate, the money-market rate, the regulated savings rate (*Livret A*) and the long-term government bond yield. Bazot et al. (2022) provide the first standardised monthly dataset of French macro-financial series from 1891: the Banque de France's metallic reserves, international portfolio, short-term lending, monetary-policy operations, deposits, bond yields, the exchange rate on London and a stock price index. For the modern period I add OECD short and long rates, the BIS real effective exchange rate, the OECD share price index, Brent oil, the VIX and the Nasdaq France technology index from FRED. Table A.2 in Appendix A lists the variables, their frequency, their first and last observation and their source. Predictors enter in log differences. Most are monthly or lower frequency; they are brought to the half-monthly calendar by a bridge-equation nowcast that uses the stock index, observed at the half-monthly frequency, and only information available at the forecast date (Appendix A.2). Each security's own dividend–price ratio is also a predictor, lagged by one period for the reason explained next, and so is the lagged market return $R_{m,t}$ itself.

2.6 Trading data for the modern period

For 1977–2025, the daily EUROFIDAI files contain more than prices: the number of shares traded and their value, the number of shares outstanding and the market capitalisation, the daily high and low, and, for a small subset of security-days, the closing bid and ask. These variables are not available before 1977 and are therefore not used by the forecasting experts, whose inputs must exist over the whole sample; they are used to test, on the modern period, what the prices alone cannot establish. Five measures are built from them at each half-monthly sampling date. (i) A *trading indicator*: whether the security traded on the day on which its sampled price was set, so that the price is a transaction price rather than a carried-forward quote. (ii) *Turnover*, the number of shares traded over the half-month divided by shares outstanding. (iii) The *Amihud* illiquidity ratio, the average absolute daily return per million euros traded. (iv) The *quoted spread*, $(a - b)/m$ with $m = (a + b)/2$, when bid and ask are both recorded, and the *high–low spread* estimator of Corwin and Schultz (2012) on consecutive trading days otherwise. (v) The *market capitalisation*, used in Section 5 to form real-time size quintiles and in Section 6 to value-weight portfolios. Table A.3 in Appendix A reports

[13] Because each security contributes equally at each date, the historical index approximates an equal-weighted index rebalanced every period whose composition varies with entries and exits. A value-weighted index requires market capitalisation, which DFIH does not provide directly; Section 7 uses the number of listed securities of a firm as a size proxy before 1977.

the coverage of each variable and Figure A.2 the evolution of liquidity. Market capitalisation is available for 95 to 99% of the half-monthly observations in every year; volumes for 70 to 100% until 2000 and for 50 to 65% since 2005, when the many small listings of the Euronext Growth and Access compartments enter the database; the high–low spread for 45 to 99%; quoted bid and ask prices for essentially none of the observations before 2005 and for at most 8% of them afterwards, so that spread-based results rest on the high–low estimator, which is floored at zero and has a median of zero before 1999.[14] The share of securities whose sampled price is not a transaction price is 8 to 11% in 1977–1985, below 1% in 1995, and 20 to 25% since 2005: stale quotes are not a problem of the archival period only.

2.7 Two measurement problems in long-run data

Two features of historical price data generate predictability that is spurious, in the sense that no investor could have earned it, and that would contaminate the results if ignored.

Stale prices and the index. Not every security is traded on every quotation date, and the quotation lists then carry the last price. An equal-weighted average of such prices adjusts to common news with a lag, so that its return is positively autocorrelated even when every underlying security follows a random walk (Scholes and Williams, 1977; Lo and MacKinlay, 1990a). Table F.6 reports the autocorrelation of the half-monthly index return: $\rho_1 = 0.32$ over 1870–2025, 0.30 before 1977 and 0.34 after, with $\rho_2 = 0.19$ and $\rho_3 = 0.04$, implying an in-sample R^2 of $\rho_1^2 \approx 10\%$ from the lagged index alone. A forecaster who includes the lagged index return, or any predictor that is itself built from stock prices (the stock index of Bazot et al. (2022), the OECD share price index), obtains an out-of-sample R^2 of 9.6% at the market level, which an investor trading the index could not have realised because the stale constituents could not be bought at the stale prices. The main market-level specification therefore uses macro-financial predictors only; the "all predictors" specification is reported for comparison and measures the size of the artefact. At the security level, the lagged market return is kept as a predictor, because the delayed reaction of an individual security to market-wide information is a genuine and much-studied feature of returns (Lo and MacKinlay, 1990b; Hou, 2007); Section 5 shows that it is in fact the dominant source of security-level predictability, and discusses which part of it is stale pricing.

Noisy prices and price-scaled predictors. A predictor that divides by the current price, like the dividend–price ratio d_t/P_t, inherits the measurement error of P_t: when the recorded price is too low (a stale quote, a

[14] The Corwin–Schultz estimator sets the spread to zero whenever the two-day range is not larger than the two one-day ranges, which happens for most quiet days of a low-volatility security; it understates the spreads of the early period. The costs computed from it in Section 6 are therefore a lower bound, and the uniform 0.5% cost per transaction used throughout the paper is the conservative figure.

bid, a transcription error), the ratio is too high *and* the next return is too high because the error reverses. The resulting predictive slope is positive, significant and entirely spurious; it is the time-series analogue of the bias that Blume and Stambaugh (1983) and Asparouhova et al. (2013) document in portfolio returns. In this dataset it produced a security-level out-of-sample R^2 several times larger than anything else before it was identified. Lagging the ratio by one period, d_{t-1}/P_{t-1}, removes the common price and the bias, at the cost of a slightly older signal. All results use the lagged ratio.

3 Sequential prediction with expert advice

3.1 The forecasting game

Let $r_{i,1}, \dots, r_{i,T}$ denote the half-monthly log returns of security $i = 1, \dots, S$. No assumption is made about the process that generates them. At each date t the forecaster observes the forecasts $f_{i,j,t}$ of N models, the *experts* $j = 1, \dots, N$, each a function of information available at t, and forms the aggregate forecast of $r_{i,t+1}$,

$$\hat{r}_{i,t+1} = \sum_{j=1}^{N} w_{j,t}\, f_{i,j,t}, \qquad w_t \in \mathcal{P} = \left\{ w \in \mathbb{R}_+^N : \textstyle\sum_j w_j = 1 \right\}. \tag{3}$$

The return is then revealed, the forecaster incurs the quadratic loss $\ell(\hat{r}, r) = (\hat{r} - r)^2$, and the weights are updated. Performance is measured by the *regret* with respect to the best expert, or the best fixed combination, in hindsight:

$$\mathrm{Regret}_T = \sum_{t=1}^{T} \ell(\hat{r}_{t+1}, r_{t+1}) - \min_{q \in \mathcal{P}} \sum_{t=1}^{T} \ell\Big(\textstyle\sum_j q_j f_{j,t},\; r_{t+1} \Big). \tag{4}$$

Because the sequence of returns is not assumed to be a realisation of a stationary process, the performance of a forecasting strategy cannot be judged by the distance between forecasts and outcomes alone; the regret compares the forecaster to the best strategy that could have been chosen with the benefit of hindsight, for the realised sequence, whatever it is. A rule whose regret grows sublinearly in T does, asymptotically, as well as the best expert on *every* sequence.

3.2 Bernstein Online Aggregation

Weights are chosen by the Bernstein Online Aggregation rule (BOA; Wintenberger, 2017). With the gradient trick of Cesa-Bianchi and Lugosi (2006), the quadratic loss is linearised around the aggregate forecast, so

that the quantity that matters for expert j at date t is its instantaneous regret

$$\rho_{j,t} = \nabla \ell(\hat{r}_{t+1}, r_{t+1})\,(\hat{r}_{t+1} - f_{j,t}) = 2\,(\hat{r}_{t+1} - r_{t+1})(\hat{r}_{t+1} - f_{j,t}), \tag{5}$$

which is positive when expert j was closer to the outcome than the aggregate. Let $V_{j,t} = \sum_{s \le t} \rho_{j,s}^2$ be the quadratic variation of the regret and $E_{j,t} = 2^k$, with k the smallest integer such that $\max_{s \le t} |\rho_{j,s}| \le 2^k$, its range rounded up to a power of two. BOA with adaptive learning rates sets

$$w_{j,t} \;\propto\; \eta_{j,t}\, \exp\!\big(\eta_{j,t} \widetilde{R}_{j,t}\big)\, \pi_j, \qquad \eta_{j,t} = \min\left\{ \frac{1}{E_{j,t}},\, \sqrt{\frac{\log(1/\pi_j)}{V_{j,t}}} \right\}, \tag{6}$$

with uniform prior weights $\pi_j = 1/N$ and the regularised cumulative regret

$$\widetilde{R}_{j,t} = \sum_{s \le t} \tfrac{1}{2}\Big(\rho_{j,s} - \eta_{j,s}\, \rho_{j,s}^2 + E_{j,s}\, \mathbf{1}\{\eta_{j,s} \rho_{j,s} > 1/2\} \Big). \tag{7}$$

Two features of the learning rate $\eta_{j,t}$ matter. The first term, $1/E_{j,t}$, is a safeguard: when losses are large, the learning rate is small and the weights do not overreact to an extreme observation. The second term adapts to the data: when all experts perform poorly, $V_{j,t}$ grows and the rule becomes more cautious; when losses are consistently small, the second term keeps the rule learning instead of letting the learning rate collapse. The $-\eta\rho^2$ term in the regularised regret is what makes the bound "second order": an expert whose regret is volatile is penalised relative to one whose regret is steady. Wintenberger (2017) shows that the cumulative regret of BOA with respect to any expert j is bounded, for every sequence, by a quantity of order $\sqrt{V_{j,T} \log N} + E_{j,T} \log N$, up to a log log term, so that

$$\limsup_{T \to \infty} \frac{1}{T} \operatorname{Regret}_T \le 0. \tag{8}$$

The aggregate therefore converges, in average loss, to the best expert on the realised sequence; with the gradient trick the guarantee extends to the best convex combination of experts. Algorithm 2 [15] in Appendix B gives the full recursion[16].

[15] The implementation uses the `opera` package (Gaillard and Goude, 2016).

[16] Section 4 also reports, for the market index, four alternative rules with the same experts: BOA with the Campbell and Thompson (2008) restriction that negative forecasts are set to zero, the polynomially weighted average with multiple learning rates of Gaillard et al. (2014) (MLpol), the fixed-share rule of Herbster and Warmuth (1998), and the equal-weighted average of the experts.

3.3 Experts

To avoid the look-ahead bias of using a model before it existed, each expert is introduced at the date at which a forecaster could have constructed it. All experts are estimated recursively, at every date, on the information available at that date. Forecasts of the regression-based experts are bounded to the interval $[\bar{r}_{i,t} - 3s_{i,t},\ \bar{r}_{i,t} + 3s_{i,t}]$, where $\bar{r}_{i,t}$ and $s_{i,t}$ are the mean and standard deviation of the security's past returns, a loose version of the Campbell and Thompson (2008) restriction that prevents a single unstable regression from dominating the loss.

Reference experts. Three experts require no estimation. The *historical mean* $\bar{r}_{i,t}$ of the security's own past returns, the forecast implied by a random walk with drift for prices and the benchmark of Welch and Goyal (2008) and Campbell and Thompson (2008). The *zero* forecast, the benchmark of Gu et al. (2020), which is harder to beat than the historical mean for an individual security because the latter is a very noisy estimate of a small quantity. And the *all-securities mean*, the average of all returns of all securities dated t or earlier, which pools the information of the whole cross-section about the common component of returns and is, in effect, a shrinkage target; it is available from 1845 and it is also the most demanding benchmark used in the evaluation.

Predictive regressions (first generation). For each predictor x available at t, an analyst could have run the predictive regression of Campbell and Shiller (1988b),

$$r_{i,s} = \alpha_i + \beta\, x_{i,s-1} + \varepsilon_{i,s}, \qquad s \le t, \tag{9}$$

and used $\hat{r}_{i,t+1} = \hat{\alpha}_{i,t} + \hat{\beta}_t\, x_{i,t}$. The slope is estimated by pooling all securities with security fixed effects,

$$\hat{\beta}_t = \frac{\sum_i \sum_{s \le t} (x_{i,s-1} - \bar{x}_{i,t})(r_{i,s} - \bar{r}_{i,t})}{\sum_i \sum_{s \le t} (x_{i,s-1} - \bar{x}_{i,t})^2}, \qquad \hat{\alpha}_{i,t} = \bar{r}_{i,t} - \hat{\beta}_t\, \bar{x}_{i,t}, \tag{10}$$

where the means are taken over the pairs of security i dated t or earlier. For a macro-financial predictor, common to all securities, the pooled slope is the within-security slope of returns on the predictor; for the security's own dividend–price ratio it is a panel slope with fixed effects. Pooling trades the flexibility of security-specific slopes for the precision that a long panel allows; security-by-security estimation is the natural robustness check (Appendix J). Regressions are computed by exact recursive least squares and re-estimated at every date. The lagged market return is one of the predictors, so that one of the bivariate experts is the lead-lag regression $r_{i,t+1} = \alpha_i + \beta R_{m,t} + \varepsilon$.

Forecast combination and diffusion index (second generation). From 2008, the forecaster also has access to the combination of Rapach et al. (2010), the equal-weighted average of the bivariate forecasts, each estimated on the sample starting in 2008 and restricted to predictors without missing values on that sample,

$$\hat{r}^{C}_{i,t+1} = \bar{r}_i + \frac{1}{K}\sum_{k=1}^{K} \hat{\beta}_k \, (x_{k,t} - \bar{x}_k), \tag{11}$$

and to the diffusion index of Ludvigson and Ng (2007): the predictors are standardised, the first $q = 2$ principal components $F_{1,t}, F_{2,t}$ are extracted, and $r_{i,s}$ is regressed on F_{s-1}; the components and the regression are re-estimated every 24 periods and held fixed in between. The starting date of 2008 is the date at which these methods became established in the return-forecasting literature; they are used only from that date, with a sample that starts at that date, exactly as a forecaster adopting them in 2008 would have done.[17]

Characteristics expert. The last expert is a real-time cross-sectional model of the kind that defines the current generation (Gu et al., 2020), built in Section 6. At each date, eight characteristics of each security known at t (12-month momentum skipping the last month, short-term reversal, long-term reversal, 12-month volatility, 24-month beta on the equal-weighted index, the share of zero returns as an illiquidity measure (Lesmond et al., 1999), the lagged dividend–price ratio and the number of listed securities of the firm as a size proxy) are ranked cross-sectionally into $[-1, 1]$. A Fama–MacBeth regression (Fama and MacBeth, 1973) of the next relative return, $r_{i,t+1} - \bar{r}_{t+1}$, on the ranks is run at each date, and the forecast at T uses the average of the slopes of all dates before T, after at least five years of cross-sections. The expert's forecast is the all-securities mean plus the predicted relative return. Like the others, it uses only information dated t or earlier.

The set of experts deliberately stops short of the latest generation of forecasting models, machine-learning methods and sentiment analysis (Gu et al., 2020; Sirignano and Cont, 2019; Steyn et al., 2020), for three reasons. First, the real-time principle of this paper introduces each model only from the date at which a forecaster could have used it; for these models that date is around 2018–2020, which leaves too few half-months to estimate them recursively and to evaluate them with any power. Second, sentiment measures would require an intensive text-mining effort on historical news and documents, which is beyond the scope of this paper. Third, the information that machine-learning models extract from the cross-section is largely captured by the characteristics expert: in Gu et al. (2020), the predictors that dominate across all methods are price trends, liquidity and volatility, which are precisely the characteristics used here, combined through

[17] Gradient-boosted trees (Chen and Guestrin, 2016) were also implemented, to be introduced from 2020. They are excluded from the results reported here, because the sample on which they could be evaluated is too short to say anything; the code allows them to be switched on.

real-time Fama–MacBeth slopes that impose the shrinkage these methods achieve by regularisation. The framework is open, however, and any such model can be added as an expert once enough data exist to learn its weight.

3.4 Pooled aggregation

Applying BOA security by security does not work, and it is worth saying why. For a typical security there are a few hundred observations, the signal is weak, and the experts' losses differ by a fraction of a percent of the loss itself. The learning rate in (6) is then tiny, the weights stay at their prior and the aggregate reproduces the historical mean: in a first version of this study the security-level aggregate had an R^2_{OOS} of essentially zero against the historical mean, while the constant forecast of zero beat the historical mean by 0.36%. The information needed to rank the experts exists, but it is spread across securities.

The solution is to learn a *single* weight vector for all securities. At each date t, let I_t be the set of securities whose return $r_{i,t+1}$ is realised and known by the next calendar date. The instantaneous regret of expert j is the cross-sectional average of its security-level regrets,

$$\rho_{j,t} = \frac{1}{|I_t|} \sum_{i \in I_t} 2\,(\hat{r}_{i,t+1} - r_{i,t+1})(\hat{r}_{i,t+1} - f_{i,j,t}), \tag{12}$$

and the BOA recursion (6) is applied to these averages. Forecasts remain security-specific, $\hat{r}_{i,t+1} = \sum_j w_{j,t} f_{i,j,t}$, but the weights are updated with hundreds or thousands of observations per date instead of one. This is the aggregation analogue of pooling slopes across securities: it assumes that the *ranking* of experts is common across securities at a given date, not that their forecasts are. When an expert is unavailable for a security (the combination before 2008, a characteristic that cannot be computed), its forecast is replaced by the security's historical mean, so that it is neutral for that security. The weights used at t depend only on returns realised at t or earlier; returns that span several calendar periods, because a security was not quoted on the next date, are left out of the weight update. The regret bound of BOA applies to the cross-sectional average loss, with respect to the best common combination of experts (Appendix B.1).

3.5 Evaluation

Out-of-sample R^2. The main metric is the pooled out-of-sample R^2 of Campbell and Thompson (2008) and Gu et al. (2020),

$$R^2_{\mathrm{OOS}} = 1 - \frac{\sum_i \sum_t (r_{i,t+1} - \hat{r}_{i,t+1})^2}{\sum_i \sum_t (r_{i,t+1} - b_{i,t})^2}, \tag{13}$$

computed over all forecasts of a group of securities and dates, against three benchmarks $b_{i,t}$: the security's historical mean (the standard benchmark), zero, and the all-securities mean. The third benchmark is the one I emphasise. The historical mean of an individual security is so noisy that beating it mostly rewards shrinkage, not information; the all-securities mean removes that noise, and the R^2_{OOS} against it measures the contribution of the *information* in the experts, over and above what a forecaster who simply knew the market's average return would have achieved. I also report the R^2_{OOS} with securities weighted by the inverse of their return variance, so that volatile securities do not dominate, and after trimming the 1% most volatile securities.

Statistical tests. For the pooled R^2_{OOS}, I follow Gu et al. (2020): at each month, the cross-sectional average of the loss differential $(r - b)^2 - (r - \hat{r})^2$ is computed, and the mean of this monthly series is tested with a Newey–West t-statistic with 12 lags (Newey and West, 1987); the test is one-sided (H_1: the aggregate is more accurate). For individual securities, I report Diebold–Mariano tests (Diebold and Mariano, 1995) on the loss differential against the historical mean and the Clark–West test (Clark and West, 2007), which corrects the bias of the squared-error comparison when the larger model nests the benchmark, as it does here since the historical mean is one of the experts; the share of securities significant at 5% is reported raw and after the false-discovery-rate correction of Benjamini and Hochberg (1995), since several thousand tests are run. Appendix C gives the details.

Economic value. Statistical accuracy is an incomplete measure of economic value (Cenesizoglu and Timmermann, 2012; Kelly et al., 2024). For individual securities, I form a *timing* portfolio: each security is held during a half-month if its aggregate forecast is positive and held in cash otherwise, and the positions are averaged across securities with equal weights, month by month. It is compared with the same rule applied to the historical-mean forecast and with equal-weighted buy and hold, in terms of annualised return, volatility, Sharpe ratio, maximum drawdown and the utility gain of a mean-variance investor with risk aversion $\gamma = 3$. Security-level Sharpe ratio differences are tested with a Wilcoxon test. For the market index, I follow Campbell and Thompson (2008): the investor allocates $w_t = \hat{r}_{t+1}/(\gamma \hat{\sigma}^2_t)$ to the index, with $\gamma = 3$, $\hat{\sigma}^2_t$ the variance of the past five years and w_t bounded in $[0, 1.5]$, and I report the gain in certainty-equivalent return relative to the historical-mean allocation.

Time variation. To study how predictability evolves, I compute for each year the pooled R^2_{OOS} of all forecasts made that year, with a bootstrap confidence interval obtained by resampling securities, the share of securities for which the aggregate beats the historical mean in that year and over the trailing ten years, and the sign accuracy. The annual series is decomposed into a structural component (a Hodrick–Prescott trend

with $\lambda = 100$) and a cyclical one, tested for structural breaks with the procedure of Bai and Perron (2003) (BIC-selected number of breaks, 15% trimming), compared across recession and expansion years and around institutional events (five years before versus five years after, with permutation p-values), and regressed on market conditions with Newey–West standard errors.

4 Is the Paris Bourse predictable? Evidence from the index

I start at the level at which the literature has asked the question: can the return on the market be predicted out of sample with macro-financial variables? The target is the equal-weighted index return $R_{m,t+1}$ at the half-monthly frequency. Experts are the historical mean, zero, the recursive bivariate regressions on each macro-financial predictor, and the combination and diffusion index from 2008; the first forecast is made after ten years of learning, the first year of forecasts is a burn-in, and the evaluation runs from 1870 to 2025, 3,693 half-months. All t-statistics use Newey–West standard errors with 24 lags (one year).

4.1 The stale-price artefact

Table F.6 in Appendix F.1 documents the problem described in Section 2.7. The first-order autocorrelation of the half-monthly index return is 0.32 over the full sample, 0.30 in the DFIH period and 0.34 in the EUROFIDAI period; the second-order autocorrelation is 0.19 and the third 0.04. A first-order autoregression alone would have an in-sample R^2 of 9 to 12%. The autocorrelation is not smaller in the modern period, which indicates that it comes from the composition of the equal-weighted index (many small, infrequently traded securities) rather than from the technology of quotation. In the "all predictors" specification, which includes the lagged index return and the price-based indices, BOA attains an R^2_{OOS} of 9.6% against the historical mean over 1870–2025 (Diebold–Mariano 5.8), 8.3% before 1977 and 11.6% after; MLpol and fixed share give about 10%, and the allocation of Section 3.5 would have earned a certainty-equivalent gain of 11.7% per year (Table F.7 in Appendix F.1). Almost all of it is attributable to the AR(1) expert, and it disappears when the price-based predictors are removed: the equal-weighted average of experts, in which the AR(1) expert has a weight of one in thirty, attains only 2.0%. I treat this as an artefact and report the macro-financial specification as the main result.

4.2 Out-of-sample accuracy

Table 1 reports the R^2_{OOS} of BOA and of the alternative aggregation rules against the historical mean and against zero, with Diebold–Mariano and Clark–West statistics, for the full sample and by source. Over

1870–2025 BOA attains an R^2_{OOS} of 0.48% against the historical mean and 1.31% against zero. The Clark–West statistic, the appropriate one for a model that nests the benchmark, is 3.1; the Diebold–Mariano statistic, which ignores the estimation noise of the larger model, is 1.2. The gain is 0.30% before 1977 (Clark–West 1.8) and 0.76% after (Clark–West 2.6; 2.6% against zero).

Three observations follow. First, the magnitude is the one that the US literature regards as the upper bound of what macro-financial predictors can deliver: Campbell and Thompson (2008) argue that a monthly R^2_{OOS} of 0.5% is economically meaningful, and Rapach et al. (2010) obtain 0.3 to 1% with combinations. At the half-monthly frequency, 0.5% is on the high side of that range. Second, the gain does not come from learning the weights, but from combining the experts. No macro-financial predictor beats the historical mean by more than 0.35% on its own (the VIX[18], over the years in which it exists), and only 12 of them beat it at all; the equal-weighted average of all experts, a uniform rule with fixed weights and no learning, attains 0.44%, and BOA is not more accurate than this uniform rule (Diebold–Mariano $t = 0.14$, Table F.8 in Appendix F.2). BOA with the non-negativity restriction gives 0.46% and MLpol 0.55%; the fixed-share rule gives 1.49%, but only in the historical period (2.53% before 1977, −0.19% after). With macro-financial predictors, the predictability of the index is the diversification gain of combining many weak forecasts (Rapach et al., 2010), which a uniform rule captures as well as a learning one. The contrast with the "all predictors" specification shows when the rule matters: there, one expert, the regression on the lagged index, attains 9.95% on its own, BOA identifies it and attains 9.6%, and the uniform average, which gives it a weight of one in thirty, attains 2.0%; BOA is more accurate than the uniform rule with a Diebold–Mariano statistic of 5.4. Learning the weights pays when one expert stands out, and not otherwise. Third, the R^2_{OOS} against zero is larger than against the historical mean, because the historical mean of the index is itself a noisy forecast early in the sample; the zero forecast is the harder benchmark for individual securities, the easier one for the index.

[18]The VIX is available only from 1990. Before that date, its forecast is replaced by the historical mean, which leaves the squared error unchanged, so that its 0.35% over 1870–2025 comes entirely from the years after 1990; measured over those years alone, its out-of-sample R^2 is higher.

Table 1: Market index: out-of-sample R^2 by aggregation rule and period

Rule	Sample	Obs.	R^2_{OOS} vs. HM	R^2_{OOS} vs. 0	DM t	DM p	CW t	CW p
BOA	All	3693	0.48	1.31	1.25	0.11	3.09	0.00
BOA	DFIH (1870–1976)	2550	0.30	0.49	0.81	0.21	1.80	0.04
BOA	EUROFIDAI (1977–2025)	1143	0.76	2.61	0.96	0.17	2.63	0.00
BOA (CT)	All	3693	0.46	1.29	2.59	0.00	3.43	0.00
BOA (CT)	DFIH (1870–1976)	2550	0.22	0.41	1.36	0.09	1.82	0.03
BOA (CT)	EUROFIDAI (1977–2025)	1143	0.85	2.70	2.27	0.01	3.04	0.00
MLpol	All	3693	0.55	1.38	2.21	0.01	3.45	0.00
MLpol	DFIH (1870–1976)	2550	0.27	0.46	1.13	0.13	1.76	0.04
MLpol	EUROFIDAI (1977–2025)	1143	1.00	2.84	1.95	0.03	3.19	0.00
Fixed Share	All	3693	1.49	2.31	1.01	0.16	4.79	0.00
Fixed Share	DFIH (1870–1976)	2550	2.53	2.72	2.91	0.00	3.90	0.00
Fixed Share	EUROFIDAI (1977–2025)	1143	-0.19	1.67	-0.05	0.52	3.14	0.00
Equal weights	All	3693	0.44	1.27	2.44	0.01	3.11	0.00
Equal weights	DFIH (1870–1976)	2550	0.37	0.56	1.61	0.05	1.90	0.03
Equal weights	EUROFIDAI (1977–2025)	1143	0.54	2.39	1.92	0.03	2.68	0.00

Half-monthly forecasts of the equal-weighted index return, 1870–2025, macro-financial predictors only. R^2_{OOS} in percent against the historical mean (HM) and against zero. DM: Diebold–Mariano t-statistic of the loss differential against the historical mean; CW: Clark–West statistic for nested models; both Newey–West with 24 lags, one-sided p-values. Rules: BOA; BOA (CT), BOA with forecasts truncated at zero; MLpol, polynomially weighted average with multiple learning rates; Fixed Share; equal-weighted average of the experts. Source: index_forecast.R, table TI1.

Figure 1 shows the cumulative difference in squared errors between the historical mean and BOA, the diagnostic of Welch and Goyal (2008): it rises when BOA is more accurate. The gains accrue in four episodes, and only four. The first is the war of 1870–71 and the Commune. The second runs from the end of the First World War to 1929, the years of post-war inflation, the depreciation of the franc and its stabilisation in 1926–28; the cumulative gain jumps in 1919–20 and keeps rising until 1929, and the rolling ten-year R^2_{OOS} of Figure 2 is about 2%, with a confidence band above zero, for the windows ending in the 1920s. The third is the Liberation of 1944–45. The fourth runs from 1993 to 2008, the years of disinflation, convergence to the euro, the dot-com cycle and the global financial crisis, with rolling R^2_{OOS} of 2 to 4.5% for the windows ending in 2003–2013. In between, under the gold standard before 1914, in the 1930s and during the post-war decades of administered rates and exchange controls, the curve is flat or declining: for long periods the macro-financial experts do worse than the historical mean, and the windows ending in 1893–1902 are significantly negative. The isolated spikes of 1987 and 2008 are single half-months that reverse immediately.

The four episodes have a common feature: they are periods of large monetary shocks and changes of monetary regime, in which inflation, interest rates and the exchange rate move far from their usual range. The experts that carry the weight in each episode are precisely the monetary ones: the money-market rate in the 1870s, the long-term and money-market rates in the 1920s, the exchange rate on London around the Second World War (Figure 3). When discount rates and the price of risk move most, the variables that measure monetary conditions say something about the next return; when the monetary regime is stable, they add only

noise. This is consistent with the concentration of predictability in high-volatility states in Table F.10, and with the evidence that return predictability is concentrated in recessions and periods of stress (Rapach et al., 2010; Henkel et al., 2011; Dangl and Halling, 2012); its natural interpretation is a time-varying risk premium that becomes visible only when shocks are large, and it is therefore not by itself evidence against market efficiency.

The profile is the one that Welch and Goyal (2008) emphasise for the United States, where much of the apparent predictability comes from the oil shock of 1973–75: out-of-sample performance is concentrated in a few episodes and unstable in between, so that an investor in 1919 or in 1985 would have observed decades in which the model did worse than the historical mean, and would have had no reason to trust it at the moment it was about to work. The French evidence differs in one respect: the cumulative gain remains positive over 155 years and does not vanish after the 1970s, because it comes from the combination of many weak experts rather than from a single predictor. The predictability of the index is episodic, in line with the pockets of Farmer et al. (2023).

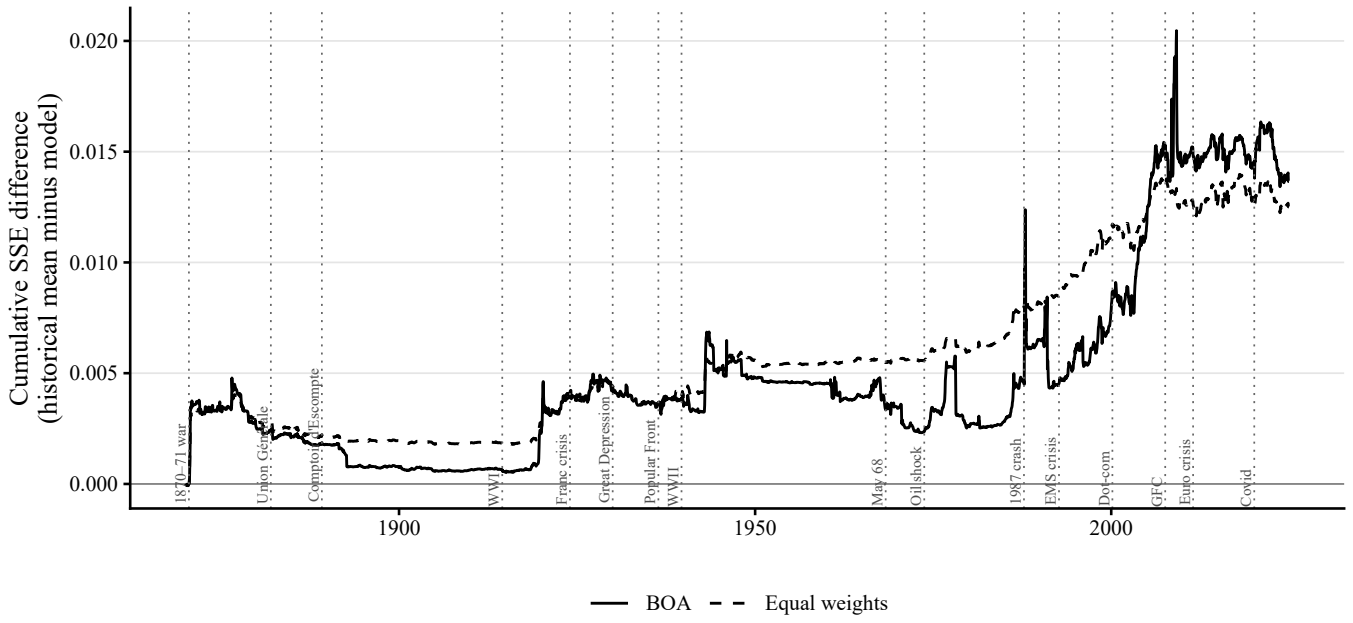


Figure 1: Market index: cumulative squared-error difference, historical mean minus model. Solid line: BOA; dashed: equal-weighted average of experts. An upward slope means the model is more accurate than the historical mean. Dotted vertical lines mark the main French crises.

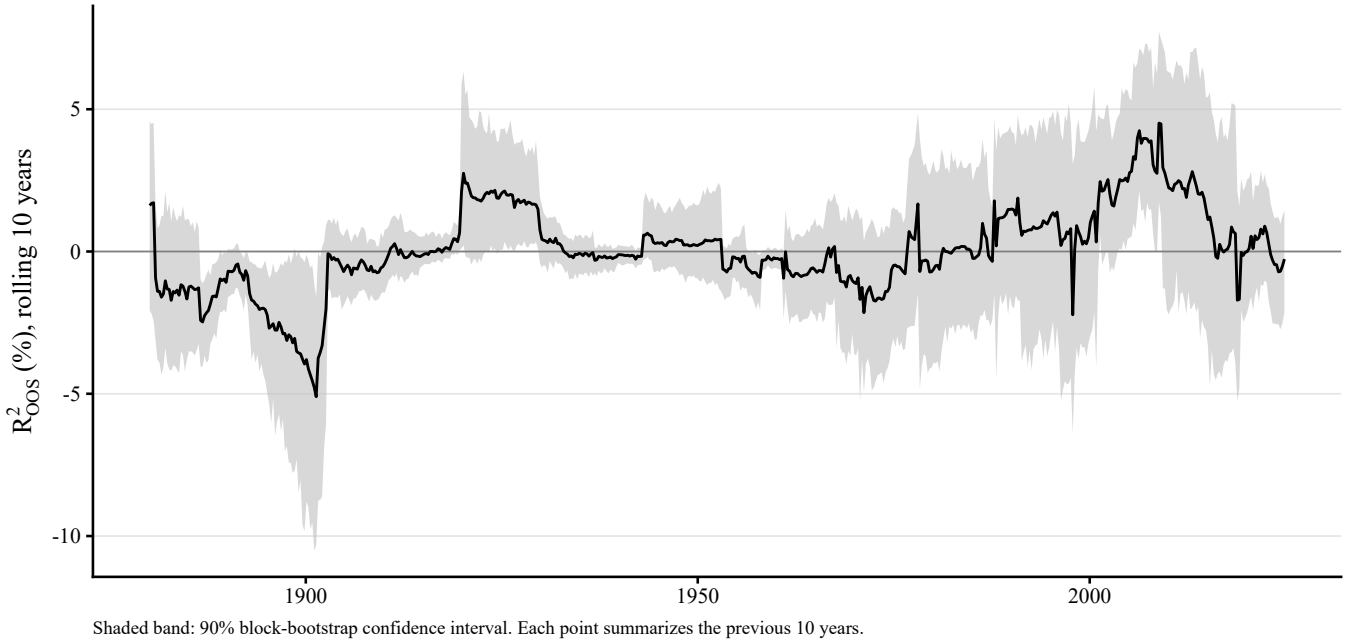


Figure 2: Market index: out-of-sample R^2 of BOA against the historical mean over rolling ten-year windows (%), with a 90% circular block-bootstrap band (six-month blocks).

4.3 Economic value and market states

Table F.9 in Appendix F.3 translates accuracy into the certainty-equivalent gain of the mean-variance investor. Over 1870–2025, the BOA allocation raises the certainty-equivalent return by 1.9% per year relative to the historical-mean allocation, and the Sharpe ratio from 0.51 to 0.66; the gain is 0.7% per year before 1977 and 4.6% after, where the Sharpe ratio rises from 0.90 to 1.12. The average equity weight is 0.72. The other rules give gains between 1.3 and 1.6% per year, again with the exception of fixed share.

Table F.10 in Appendix F.3 splits the evaluation by market state. The first split is known in real time: a half-month is in the "high volatility" state when the variance of the past six months exceeds the 80th percentile of its own history up to that date. The R^2_{OOS} of BOA is 1.31% in the 697 high-volatility half-months (Clark–West 2.6) and zero in the 2,996 others, and the pattern holds for MLpol (1.22% against 0.16%), for the equal-weighted average of experts (0.82% against 0.22%) and, more weakly, for BOA with the non-negativity restriction (0.60% against 0.38%). Since the variance of returns is itself larger in volatile periods, a higher R^2_{OOS} in those periods means that the predictable component of returns rises more than proportionally with their total variance. In calm periods, expected returns barely move and the historical mean is the best available forecast; when volatility is high, the risk premium moves, and the variables that measure monetary conditions become informative. This is what models with time-varying risk aversion predict (Campbell and Cochrane, 1999), and it matches the concentration of predictability in recessions and periods of stress in the United States (Henkel et al., 2011; Dangl and Halling, 2012); it is consistent with market efficiency, although the

data cannot distinguish a rational risk premium from mispricing in periods of stress. Because the state is known in real time, an investor could have known when the forecasts were worth following.

The second split, by crisis episodes dated ex post (the 1870–71 war, the Union Générale crash of 1882, the Comptoir d'Escompte in 1889, the two world wars, the franc crisis of 1924–26, the Great Depression, the Popular Front, May 1968, the oil shock, the 1987 crash, the EMS crisis, the dot-com bust, the global financial crisis, the euro crisis and Covid), is descriptive and less robust. BOA attains 1.01% in crises and 0.06% outside, but BOA with the non-negativity restriction, MLpol and the equal-weighted average all do better outside crises. Since the only difference between BOA and its restricted version is that negative forecasts are set to zero, the gain of BOA in crises comes from forecasts of falling prices: correct on average, but noisy enough that rules which smooth them do not benefit. Figure 3 shows the five experts with the largest average weight over the whole sample: the money-market rate and the long-term rate in the nineteenth century, the exchange rate on London around the Second World War; in the modern period none of them receives more than 10%, the weight being spread over the many experts available after 1960. and the interest-rate experts, none above 10%.

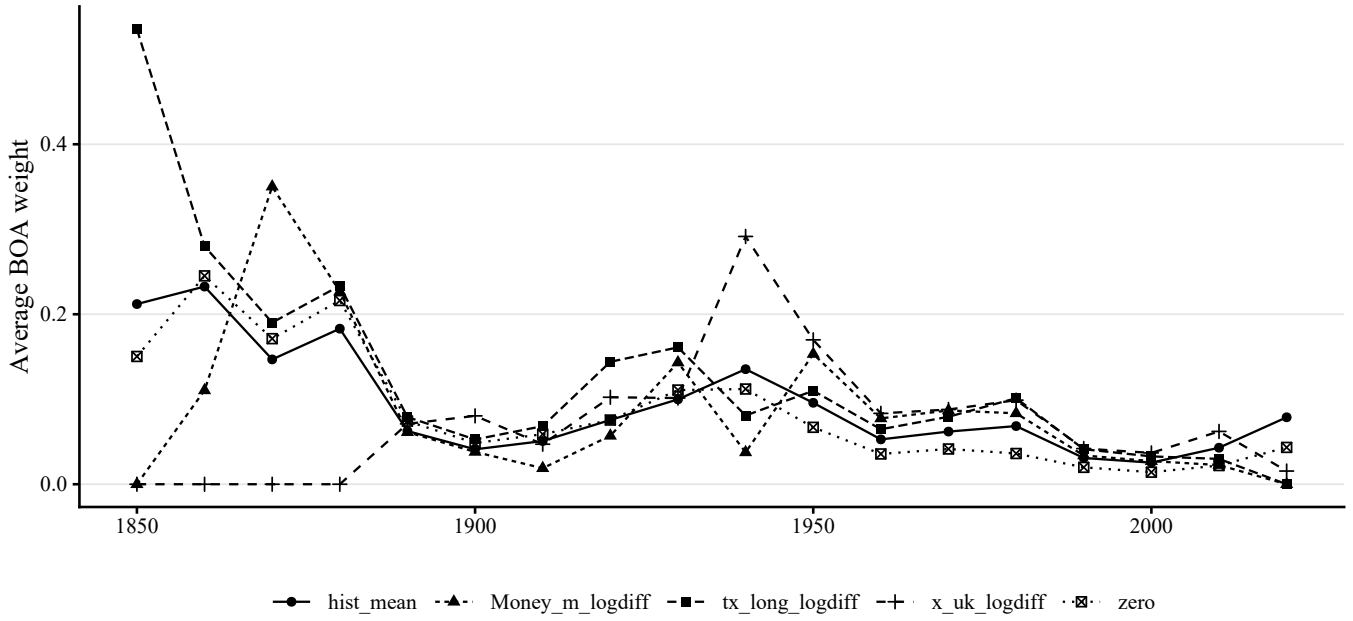


Figure 3: Market index: average BOA weight by decade of the five experts with the largest average weight (the first decade is the learning period and is not evaluated).

5 Predictability of individual securities

The market index is the wrong object on which to look for predictability in a historical market: Section 4 shows that its apparent predictability is mostly a stale-price artefact, and that what remains, with macro-financial predictors only, is half a percent per half-month. This section turns to the 7,382 individual securities that

make up the market and asks the same question of each of them, with the pooled aggregation of Section 3.4: a single weight vector learned across all securities, applied to security-specific expert forecasts. The forecasts are evaluated against three benchmarks that answer different questions. The security's own historical mean is the benchmark of the literature, but it is noisy for a security with a few hundred observations, so that beating it partly rewards shrinkage; zero and the mean return of all securities up to t remove that advantage and measure the information that the experts add. The section first measures accuracy and its statistical significance, then asks which experts carry it, what it is worth to an investor, in which securities it is found, and how it has evolved over 155 years. The answer to the second question organises the rest: most of the predictability of individual securities is their delayed reaction to the return of the market as a whole.

5.1 Can the aggregate beat the historical mean?

Table 2 is the central table of this section. Over 1870–2024, with 7,382 securities and 1.55 million forecasts, the pooled BOA forecasts have an R^2_{OOS} of 0.69% against the security's historical mean ($t = 6.4$), 0.63% against zero ($t = 3.0$) and 0.62% against the all-securities mean ($t = 3.2$). Weighting securities by the inverse of their variance gives an R^2_{OOS} against the historical mean of 0.55%, and trimming the most volatile percent 0.77%: the gain is not driven by a few volatile securities.

The split by period is instructive. Against the historical mean, the gain is the same in both periods, 0.70% and 0.69%. Against the two benchmarks that do not reward shrinkage, it is not: 0.25% against zero and 0.29% against the all-securities mean in 1870–1976, neither significant ($t = 1.0$ and 1.1); 0.77% and 0.75% in 1977–2024, both significant ($t = 3.1$ and 3.2). In the historical period, most of what the aggregate gains over a security's own mean is the shrinkage of that mean towards zero or towards the market; in the modern period, the experts contain information beyond it. The magnitudes are comparable to those of Gu et al. (2020), who report monthly R^2_{OOS} against zero of 0.3 to 0.4% for US stocks with machine-learning models on more than ninety characteristics; here they are obtained with far fewer inputs, at a higher frequency, over a much longer sample, and with a rule rather than a fitted model.

Two readings of these magnitudes are useful. The first is in units of correlation: an R^2_{OOS} of 0.6% means that the forecasts have a correlation of about 0.08 with the next return, that is, that 99.4% of the variance of a security's half-monthly return is unforecastable noise. This is why the question cannot be settled security by security, and why the pooled aggregation and the pooled experts of Section 3 are needed: the signal is only visible when it is averaged over hundreds of securities and thousands of dates. The second reading is in units of risk-adjusted return. By the bound of Campbell and Thompson (2008), a forecast with out-of-sample R^2 equal to R^2 raises the maximal squared Sharpe ratio from S_0^2 to $(S_0^2 + R^2)/(1 - R^2)$ per period; with the

Sharpe ratio of 0.11 of equal-weighted buy and hold, an R^2_{OOS} of 0.62% per half-month implies an annualised Sharpe ratio of about 0.4 for an investor who uses the forecasts optimally and trades without cost. A small R^2 at a short horizon is therefore a large economic quantity, and Section 5.3 shows that a crude timing rule already captures most of it. The split between the two periods adds a third point: in the historical period, nearly all of what the experts know about a security is what the cross-section knows about the common component of returns, which the all-securities mean already contains; the information specific to the security, the lead-lag and the characteristics, is measurable only in the modern period, where the number of securities per date is three to six times larger and the pooled quantities are estimated with correspondingly less noise.

Table 2: Pooled out-of-sample R^2 of the aggregate forecast, by period

Sample	Securities	Obs.	R^2_{OOS} vs. HM	t_{NW}	p	R^2_{OOS} vs. 0	t_{NW}	p	R^2_{OOS} vs. CS mean	t_{NW}	p	R^2_{OOS} (1/var. weighted)	R^2_{OOS} (trimmed)
DFIH (1870–1976)	1545	443353	0.70	5.71	0	0.25	1.01	0.16	0.29	1.15	0.13	0.74	0.71
EUROFIDAI (1977–2025)	6034	1104415	0.69	4.03	0	0.77	3.11	0.00	0.75	3.24	0.00	0.48	0.79
All	7382	1547768	0.69	6.40	0	0.63	3.03	0.00	0.62	3.19	0.00	0.55	0.77

Half-monthly forecasts of individual security returns, evaluated from 1870 after a 12-month burn-in per security. R^2_{OOS} in percent, pooled over all forecasts of the group, against three benchmarks: the security's historical mean (HM), zero, and the mean of all securities up to t (CS mean). t_{NW}: Newey–West t-statistic (12 lags) of the mean of the monthly cross-sectional average loss differential, one-sided p-value. The last two columns weight securities by the inverse of their return variance (excluding the 1% least volatile) and exclude the 1% most volatile securities. Source: paper_analysis.R, table T2.

The distribution across securities is wide. Table G.11 and Figure G.9 in Appendix G.1 show that the median security has an R^2_{OOS} of 0.60% against its historical mean, with an interquartile range from −0.99 to 1.78%, and that 62% of securities have a positive one (81% in the historical period, 59% in the modern one). Table G.12 reports the share of securities for which the gain is statistically significant among the 6,513 securities with at least 30 forecasts: 21.5% by the Diebold–Mariano test and 34.7% by the Clark–West test at the 5% level, 3.8% and 17.4% after the Benjamini–Hochberg correction. The reverse test, that the historical mean is significantly *more* accurate, rejects for 6.6% of securities raw, close to the size of the test, and for a single security after the correction. Table G.13 in Appendix G.1 gives the same shares by period: 35.8% raw and 12.5% corrected in 1870–1976, 17.1% and 2.1% in 1977–2024. Predictability is significant for a minority of securities, well above what chance would produce, and the aggregate is essentially never significantly worse than the benchmark, which is the practical content of the regret bound. The asymmetry is the economically important feature of the distribution. A fitted model with forty predictors, estimated security by security, would overfit for some securities and be significantly worse than the historical mean for more than the size of the test; a rule whose weights are learned across securities and updated by past performance behaves, for each security, like insurance: it rarely hurts and sometimes helps. This is also what makes the figures that follow interpretable: a share of 62% of securities with a positive R^2_{OOS} is not the outcome of a selection among specifications, since there is only one.

5.2 Which experts predict?

Table 3 reports the average BOA weight of each family of experts by period, Table G.14 in Appendix G.2 the five bivariate predictors with the largest weight, and Figure 4 the weights over time. The result is stark. The family of bivariate regressions receives 65% of the weight in 1870–1976 and 80% in 1977–2024, and within that family a single expert, the regression of the security's return on the lagged return of the equal-weighted market, receives 60% and 80%; the next bivariate predictor (the Banque de France's international portfolio, the money-market rate, the metallic reserves, the discount rate) receives less than 0.6%. The characteristics expert receives 14% and 7%, the zero forecast 10% and 7%, the all-securities mean 10% and 6%. The historical mean, the forecast combination and the diffusion index receive 0.2% or less. Table E.5 in Appendix E tells the same story security by security: the ex-post best expert is the lagged market return for 31% of securities, zero for 14%, the characteristics expert for 9%, and no macro-financial predictor for more than 7%.

Table 3: Average BOA weight by expert family and period

Expert family	DFIH (1870–1976)	EUROFIDAI (1977–2025)
All-securities mean	0.100	0.059
Bivariate regressions	0.652	0.803
Characteristics	0.142	0.071
Diffusion index	0.002	0.000
Forecast combination	0.002	0.000
Historical mean	0.002	0.000
Zero forecast	0.100	0.067

Weights of the pooled aggregation, averaged over the forecasts of each period (an expert that is unavailable for a security has weight zero for that security). Each column sums to one. Source: paper_analysis.R, table T7b.

Three things follow. First, the macro-financial information that the first generation of predictive regressions was built on, and that the second generation combined, is of almost no use for predicting the return of an individual security at a half-month horizon, in either period; the index-level R^2_{OOS} of 0.5% in Section 4 is the whole of it. Second, the second-generation methods, introduced in 2008, never earn a weight: by the time they became available, there was nothing in the macro-financial predictors for them to combine. Third, the predictability of individual securities over 155 years has two sources, in this order: the delayed reaction of a security's price to the return of the market as a whole, and the security's own characteristics relative to those of the others. The contrast with Section 4 is instructive, because the index is the average of these securities and the same macro-financial experts deliver 0.5% for it. A security's return is the common component plus an idiosyncratic component several times more volatile; a predictor that explains half a percent of the variance of the common component explains a small fraction of a percent of the variance of the security's return, and the aggregation rule, which ranks experts by their cross-sectional average loss, finds

that contribution dominated by that of the lead-lag expert, whose signal is about the common component itself. Macro-financial predictability is a property of the market, and it does not survive the passage to individual securities; the lead-lag effect is a property of individual securities, and it disappears in the index, where every security's delay is averaged with the others' into the stale-price autocorrelation of Section 2.7.

The first source is the lead-lag effect documented by Lo and MacKinlay (1990b) for the United States, and attributed to the slow diffusion of common information by Brennan et al. (1993), Chordia and Swaminathan (2000) and Hou (2007): individual securities react with a delay to the return of the market as a whole. Figure 4 shows the weight of the lead-lag expert rising in the 1920s and 1930s, after 1985, and in the volatile years 1987, 2008 and 2020. Part of this effect could be the stale-price mechanism of Section 2.7 seen from the side of the security: a quote that has not been updated will "catch up" with the market, which makes the next recorded return predictable but not tradable at the recorded price. With prices alone the two mechanisms cannot be fully separated, but two facts show that the effect is neither an artefact of the historical quotation technology nor mainly a matter of quotes that are not updated, and the trading data of the modern period then separate them. Measured against the all-securities mean, it is more than twice as strong in the modern period, with daily data sampled on the same calendar, as in the historical one. And when the evaluation is restricted to forecasts whose return is realised on the next calendar date, that is, to securities quoted on two consecutive dates, which removes the multi-period returns in which a stale quote catches up with the market, the R^2_{OOS} against the all-securities mean *rises*: from 0.62 to 0.71% ($t = 3.4$), from 0.29 to 0.47% ($t = 1.7$) in the historical period and from 0.75 to 0.77% ($t = 3.1$) in the modern one, on 91% of the forecasts (Appendix J, Table J.35). The observations most exposed to stale pricing are the ones in which the forecasts do worst, not best: the delayed reaction is a property of traded prices.

The trading data of Section 2.6 settle the question for 1977–2025, because they say whether the sampled price of a security is a transaction price or a carried-forward quote. The stale-price mechanism requires the price at t to be stale; conditioning on a trade on the sampling day removes it. Table J.38 and Figure J.12 in Appendix J report the result. Among the forecasts whose price at t was set by a trade (670,000 of the 813,000 modern-period forecasts for which volumes are observed), the R^2_{OOS} against the all-securities mean is 0.89% ($t = 3.6$), and the lead-lag expert alone attains 0.89%; among those traded on both sampling dates, where neither the price at t nor the price at $t + 1$ can be stale, 0.90% ($t = 3.4$) and 0.90%. Among the 143,000 forecasts whose sampled price is a stale quote, the R^2_{OOS} is 0.13% ($t = 0.6$) and the lead-lag expert alone −0.11%: a stale quote does not catch up with the market in a way that the lagged market return predicts, presumably because a security that did not trade on one sampling day often does not trade on the next either, so that its recorded return is zero or is realised much later. The same holds by turnover: the aggregate attains

0.06% (t = 1.1) in the lowest tercile, where the lead-lag expert is at −0.14%, and 1.16% (t = 3.5) in the highest, where it reaches 1.24%. The delayed reaction is a property of prices formed by trades, and it is stronger among them than in the full sample.

The second source is the security's own characteristics relative to those of the other securities. The characteristics expert takes the weight from the lead-lag expert in 1890–1920, in 1965–1985 and since 2015, the calmest periods of the sample (Figure 4). Its contribution can be measured by removing the lead-lag expert from the aggregation (Appendix J, Table J.37): the characteristics expert then receives 70% of the weight, and the remaining experts deliver 0.11% against the all-securities mean, significant only in the modern period (0.15%, t = 3.2). The lead-lag effect thus accounts for about four fifths of the information content of the security-level forecasts, and the characteristics for the rest. At a half-month horizon, the characteristics are a small source of predictability; their value is realised through long–short portfolios held for several months, which Section 6 studies.

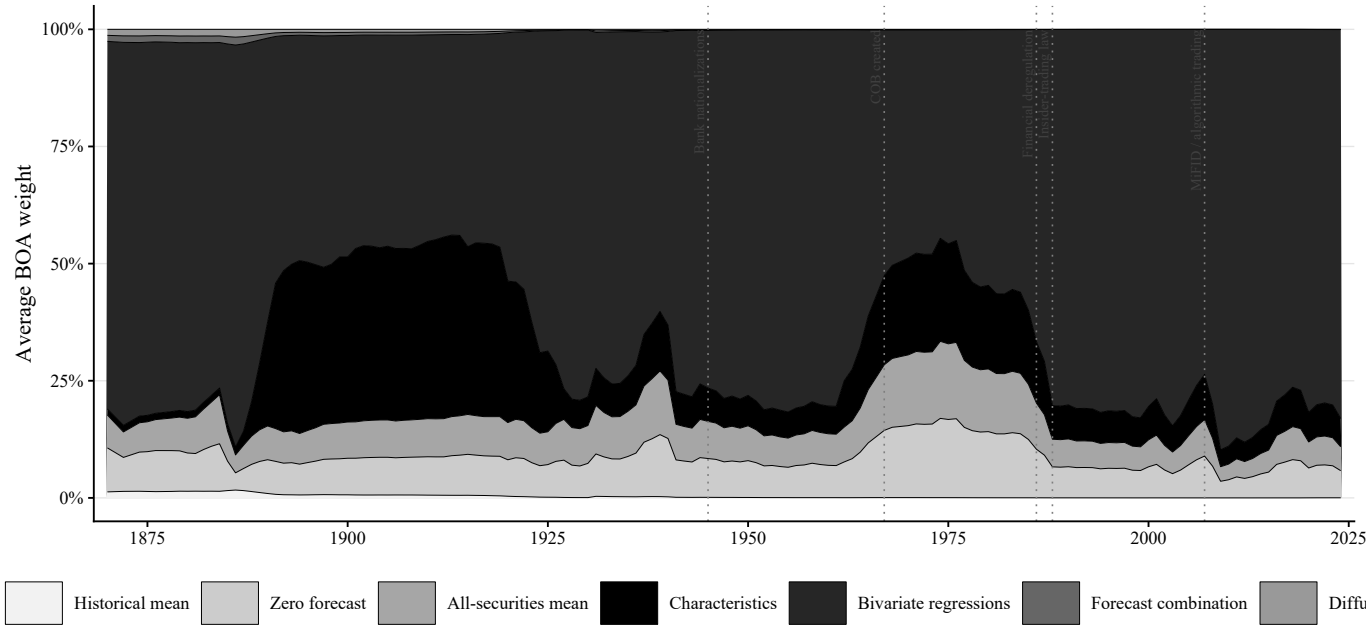


Figure 4: Average BOA weight by expert family and year, 1870–2024. The bivariate-regression family is almost entirely the lagged market return. Dotted lines mark institutional events.

5.3 Economic value

Figure 5 plots the cumulative log return of the equal-weighted timing portfolio, of the same rule applied to the historical-mean forecast, and of equal-weighted buy and hold; Table G.15 reports their performance by period. Over 1870–2024, equal-weighted buy and hold of all listed securities earns 1.6% per year in log terms with a volatility of 14.2%, a Sharpe ratio of 0.11 and a maximum drawdown of 80%; the low return reflects

the many small securities that are delisted at a loss, since the portfolio holds every security listed in Paris.[19] The timing portfolio on the BOA forecast earns 5.2% per year with a volatility of 8.6%, a Sharpe ratio of 0.61, a maximum drawdown of 35% and a utility gain of 5.6% per year over buy and hold; the timing portfolio on the historical mean earns 0.4% with a Sharpe ratio of 0.05. The gain is present in both periods: a Sharpe ratio of 0.49 against 0.19 in 1870–1976, 0.87 against −0.01 in 1977–2024. Table G.16 in Appendix G.3 tests it security by security: the median difference between the Sharpe ratio of a security's BOA timing portfolio and that of its historical-mean timing portfolio is 0.18, positive for 74% of securities (Wilcoxon $p < 10^{-100}$), and 0.11 against buy and hold, positive for 64%; Figure E.5 in Appendix E shows the distribution.

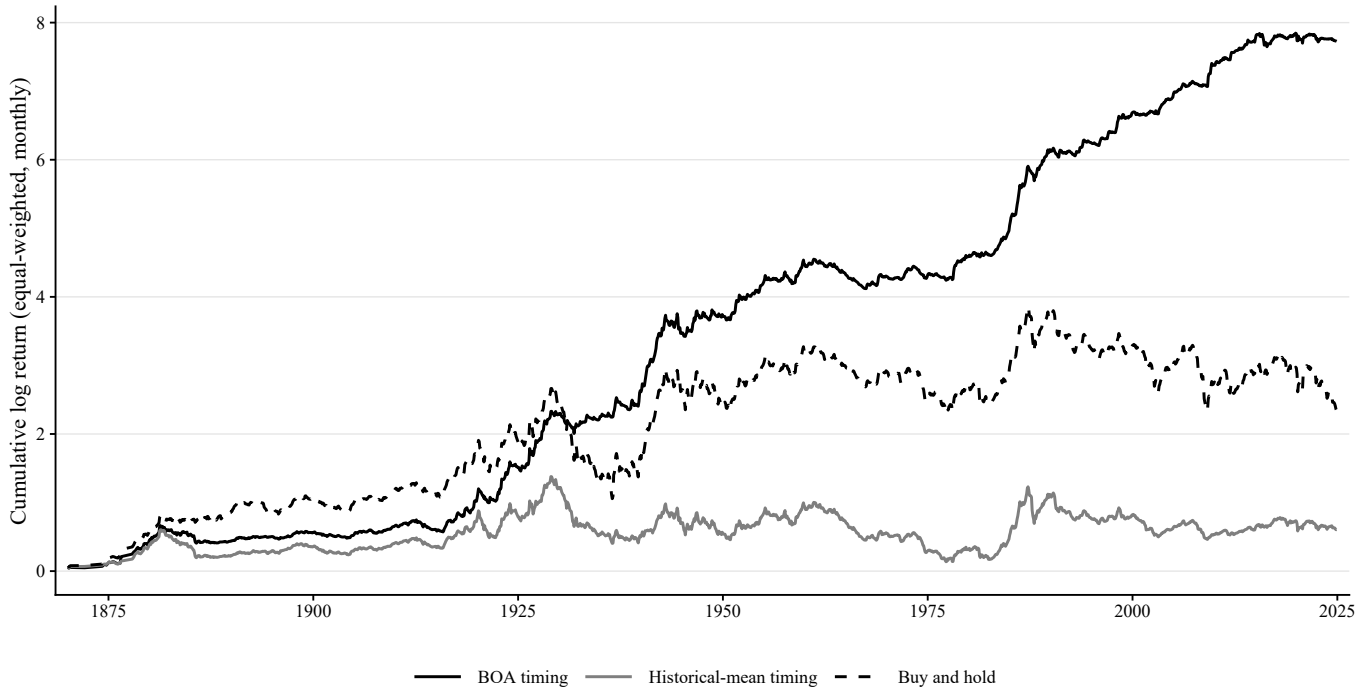


Figure 5: Cumulative log return of the equal-weighted timing portfolios (long a security when its forecast is positive, in cash otherwise) based on the BOA forecast and on the historical mean, and of equal-weighted buy and hold, monthly, 1870–2024.

These numbers must be read with Section 5.2 in mind. The timing rule is a blunt instrument, long or in cash with no shorting, no leverage and no transaction costs, so that this exercise measures the directional content of the forecasts rather than an implementable strategy. More importantly, since the forecast is dominated by the lagged market return, the timing portfolio is in effect long the cross-section after the market has risen and in cash after it has fallen; its jumps in Figure 5, in 1941–44 and in 1986–88, are the half-months in which the market moved most. To the extent that the delayed reaction of individual securities is stale pricing, the recorded return of such a portfolio overstates what could have been earned at tradable prices. On the consecutive-date sample of Appendix J, which removes the observations most exposed to this problem,

[19] Returns are averages of log returns, which penalise the dispersion of security returns; the equal-weighted market of Section 7, computed from simple returns of the securities used in the factor portfolios, earns 7.8% per year.

the timing portfolio keeps a Sharpe ratio of 0.56 against 0.11 for buy and hold (0.42 against 0.17 in the historical period, 0.87 against 0.00 in the modern one), so the directional content of the forecasts is not a recording artefact; it remains a directional content, not a trading strategy. The equal-weighted index of Section 4 is subject to the same caveat, which is why the market-level results exclude the lagged index.

Whether this directional content could have been turned into profit is a different question, and the answer is largely no. A rule that trades a security every time the sign of its forecast changes turns over its positions at a rate comparable to the half-monthly long–short portfolios of Section 6, which costs of 0.5% per transaction turn from a gross gain of 10% per year into a net loss; and Section 5.4 shows that the predictability is concentrated in the smallest securities, which are the most expensive to trade in size. The predictability of individual securities is therefore of the kind that an efficient market with transaction costs permits (Jensen, 1978; Grossman and Stiglitz, 1980): unambiguous statistically, informative about how prices are formed, and not a free lunch. This is also why it can persist for 155 years without being arbitraged away, which Section 5.5 documents.

5.4 Where is predictability? Size and listing duration

The second hypothesis of the introduction locates predictability in the securities that attract the least attention, and the lead-lag literature in the small securities that react last to common information. Table 4 reports the median security-level R^2_{OOS}, the share of positive values and the share of significant Diebold–Mariano tests by quintile of a size proxy (the number of listed securities of the issuing firm times their average price) and by tercile of listing duration. The median R^2_{OOS} is 0.77 to 0.89% in the three smallest size quintiles, 0.49% in the fourth and −0.32% in the largest; the share of securities with a positive R^2_{OOS} falls from 67–71% to 45%, and the share with a significant test from 5.0% to 2.4%. A Kendall test of the trend across quintiles rejects at any level ($\tau = -0.14$, $p < 0.0001$).

The gradient comes entirely from the lead-lag expert. When that expert is removed from the aggregation (Appendix J), the median R^2_{OOS} is between 0.34 and 0.42% in every size quintile, with no trend. The hypothesis is therefore confirmed in a specific form: small securities are more predictable because they absorb market-wide information last, not because neglected securities are mispriced in general; the predictability that comes from security characteristics is the same for small and large securities. The negative median R^2_{OOS} of the largest securities is the cost of a common weight vector: the weights are learned on the average security, which is small, and impose on the largest securities a delayed reaction that they do not have. This is the counterpart of the guarantee of the pooled rule, which is relative to the best *common* combination of experts (Appendix B.1).

Listing duration does not test the attention hypothesis. If it did, recently listed securities, which are less followed, should be the more predictable; the opposite holds: securities in the shortest tercile have a median R^2_{OOS} of −0.28% and no significant test, those in the longest 0.96% and 16% of significant tests. This gradient, too, disappears without the lead-lag expert (0.42, 0.43 and 0.36% across terciles), and it is largely mechanical: the security-specific intercept of the lead-lag regression is estimated more precisely when the history is longer, and the pooled experts have more time to learn. The size proxy is coarse, and Table 5 repeats the test for 1977–2025 with the market capitalisation itself.

Table 4: Predictability across securities: size and listing duration

(a) By size quintile (1 = smallest)

Sample	N	Median R^2_{OOS} (%)	Share > 0	Share signif. (BH)
1	1477	0.768	0.686	0.050
2	1477	0.890	0.708	0.054
3	1476	0.831	0.671	0.028
4	1476	0.488	0.604	0.039
5	1476	-0.321	0.455	0.024

(b) By listing duration tercile (1 = shortest)

Sample	N	Median R^2_{OOS} (%)	Share > 0	Share signif. (BH)
1	2461	-0.283	0.457	0.000
2	2461	0.553	0.615	0.004
3	2460	0.959	0.802	0.157

Median security-level R^2_{OOS} against the historical mean (%), share of securities with a positive R^2_{OOS}, and share with a significant Diebold–Mariano test after the Benjamini–Hochberg correction. Size proxy: number of listed securities of the issuing firm times their average price. Source: paper_analysis.R, tables T12a and T12b.

Table 5 sorts the modern-period forecasts into quintiles of market capitalisation at the forecast date, formed in real time, and reports the pooled R^2_{OOS} of the aggregate against the three benchmarks, together with the R^2_{OOS} of the lead-lag expert and of the characteristics expert taken alone against the all-securities mean. The gradient is the same as with the proxy, and sharper. Against the all-securities mean, the aggregate attains 1.32% ($t = 4.0$) in the smallest quintile, 0.92% and 0.85% in the second and third, 0.15% ($t = 1.8$) in the fourth and −0.57% ($t = 0.0$) in the largest. The lead-lag expert alone goes from 1.43% in the smallest quintile to −0.93% in the largest, which is the cost of the common weight vector described above, and the characteristics expert from 0.15% to −0.01%: both sources of security-level predictability are concentrated in small securities, and for the fifth of the market that carries most of its capitalisation nothing predicts the next half-month's return. In the fourth quintile the aggregate is positive while the lead-lag expert is not, which is the shrinkage towards zero and towards the market at work. This is the size pattern of the lead-lag literature (Lo and MacKinlay, 1990b; Hou, 2007), now measured security by security with real capitalisations, and it

bears on the economic value of Section 5: the timing portfolio earns its return on the securities that are the hardest to trade in size.

The gradient also discriminates between the two explanations that the lead-lag literature offers. Lo and MacKinlay (1990a) and Boudoukh et al. (1994) attribute a large part of the cross-autocorrelation of US portfolio returns to non-synchronous trading, which is the stale-price mechanism of Section 2.7; Brennan et al. (1993), Hong and Stein (1999) and Hou (2007) attribute it to the slow diffusion of common information towards the securities that receive the least attention. The two explanations make opposite predictions about where the effect should be strongest. Non-synchronous trading predicts that it is strongest where prices are stale; Table J.38 in Appendix J shows that it is absent there and strongest where prices are fresh, among small securities that trade on both sampling dates, and that it increases with turnover rather than decreasing with it. Slow diffusion predicts that it is strongest where attention is scarce, that is, in small firms, which Table 5 confirms with real capitalisations. The size gradient of predictability in Paris is the attention gradient of Hong et al. (2000), not the trading-frequency gradient of Lo and MacKinlay (1990a).

Table 5: Predictability by market-capitalisation quintile, 1977–2024

			R^2_{OOS} (%) vs.				Expert alone, vs. CS mean	
	Obs.	Securities	HM	Zero	CS mean	t_{NW}	Lead-lag	Characteristics
All (1977-2024)	1,104,440	6,059	0.69	0.77	0.75	3.24	0.70	0.07
Size quintile (market cap., real time)								
Q1 (smallest)	217,402	2,304	0.49	1.37	1.32	3.96	1.43	0.15
Q2	217,851	2,634	1.22	0.95	0.92	3.84	0.88	0.07
Q3	217,878	2,423	1.18	0.86	0.85	2.79	0.78	0.01
Q4	217,867	2,096	0.70	0.14	0.15	1.79	-0.04	-0.02
Q5 (largest)	218,320	1,419	-0.04	-0.62	-0.57	0.04	-0.93	-0.01

Forecasts of 1977–2024 sorted at each date into quintiles of market capitalisation known at the forecast date (EUROFIDAI). Pooled R^2_{OOS} (%) of the aggregate against the historical mean, zero and the all-securities mean; t_{NW}: Newey–West t-statistic (12 lags) on the monthly cross-sectional loss differential against the all-securities mean; last two columns: R^2_{OOS} against the all-securities mean of the lead-lag expert (regression on the lagged market return) and of the characteristics expert, each taken alone. Source: eurofidai_extra.R, table TE1.

5.5 One hundred and fifty-five years of predictability

Figure 6 is the main figure of the paper. Panel A plots the annual pooled R^2_{OOS} of the aggregate against the historical mean with its bootstrap band, Panel B the share of securities for which the aggregate beats the historical mean in the year and over the trailing ten years, and Panel D the annual R^2_{OOS} against the all-securities mean. Against the historical mean, the annual R^2_{OOS} is positive in 131 of the 152 evaluated years, with a confidence interval above zero in 87 of the 108 years after 1917; its level is around 1% from 1917 to 1950 and around 0.7% from 1950 to 1985, with spikes at 3.1% in 1927, 3.2% in 1930, 2.8% in 1941,

4.3% in 1987, 2.4% in 1998, 5.6% in 2008, 3.3% in 2009 and 3.2% in 2012. The share of securities beaten in the year hovers between 55 and 75%, and the trailing ten-year share between 70 and 90% from 1920 to 2013. Panel D removes the shrinkage: against the all-securities mean, the R^2_{OOS} is zero or negative in almost every year before 1917, around zero in 1945–1985, and positive in pockets that are exactly the volatile years listed above, with the largest at 6.6% in 2008 and 3.5% in 1987. The last decade is different: since 2014, when the number of securities more than doubles, the R^2_{OOS} is negative in seven years out of eleven against the historical mean and in five against the all-securities mean, and the trailing ten-year share falls below 50% for the first time since the 1880s.Figure G.10 separates the structural (HP trend) and cyclical components: the trend is flat, and the variation is cyclical.

Three formal exercises confirm what the figure shows: predictability has no trend and no institutional history, only a volatility cycle. Table G.17 in Appendix G.4 finds no structural break in the mean of the annual R^2_{OOS}, of the annual share or of the trailing share by the BIC criterion of Bai and Perron (2003). Table G.18 in Appendix G.4 finds no difference between recession and expansion years: the mean R^2_{OOS} is 1.03% in the 33 recession years and 0.71% in the 119 others ($p = 0.26$). Table G.19 in Appendix G.4 finds no institutional event, among nine from the Union Générale crash of 1882 to the global financial crisis, after which the five-year mean of the R^2_{OOS} or of the share differs from the five-year mean before by more than chance; the largest difference, +1.2 points after MiFID and the advent of algorithmic trading in 2007, has a permutation p-value of 0.08, and it coincides with 2008. Table G.20 in Appendix G.4 regresses the annual measures on a recession indicator, market volatility and the log number of securities: each percentage point of annualised market volatility raises the annual R^2_{OOS} by 0.05 point ($p = 0.004$) and the share of securities beaten by 0.6 point ($p = 0.001$); recessions have no effect, and the breadth of the market has none on the R^2_{OOS} and a small negative one on the share of securities beaten (−2.1 points per log unit, $p = 0.04$). The volatility of the market is the only state variable that forecasts how predictable individual securities will be, which is what a delayed reaction to common information implies: the more the market moves, the more there is to catch up with.

The stability of the effect over 155 years is at odds with the view that a predictable pattern, once known and exploitable, disappears (McLean and Pontiff, 2016; Chordia et al., 2014). The lead-lag effect has been documented since Lo and MacKinlay (1990b) and has not weakened since; it did not weaken either with electronic trading in 1986–89, which removed the mechanical delays of open outcry, nor with the growth of institutional ownership and of the analyst industry after 1980. This is what one expects if the effect is not an arbitrage opportunity but the price of the attention and the capital that trading small securities requires: Section 5.3 showed that it cannot be collected at plausible costs, and Section 6 shows that the related

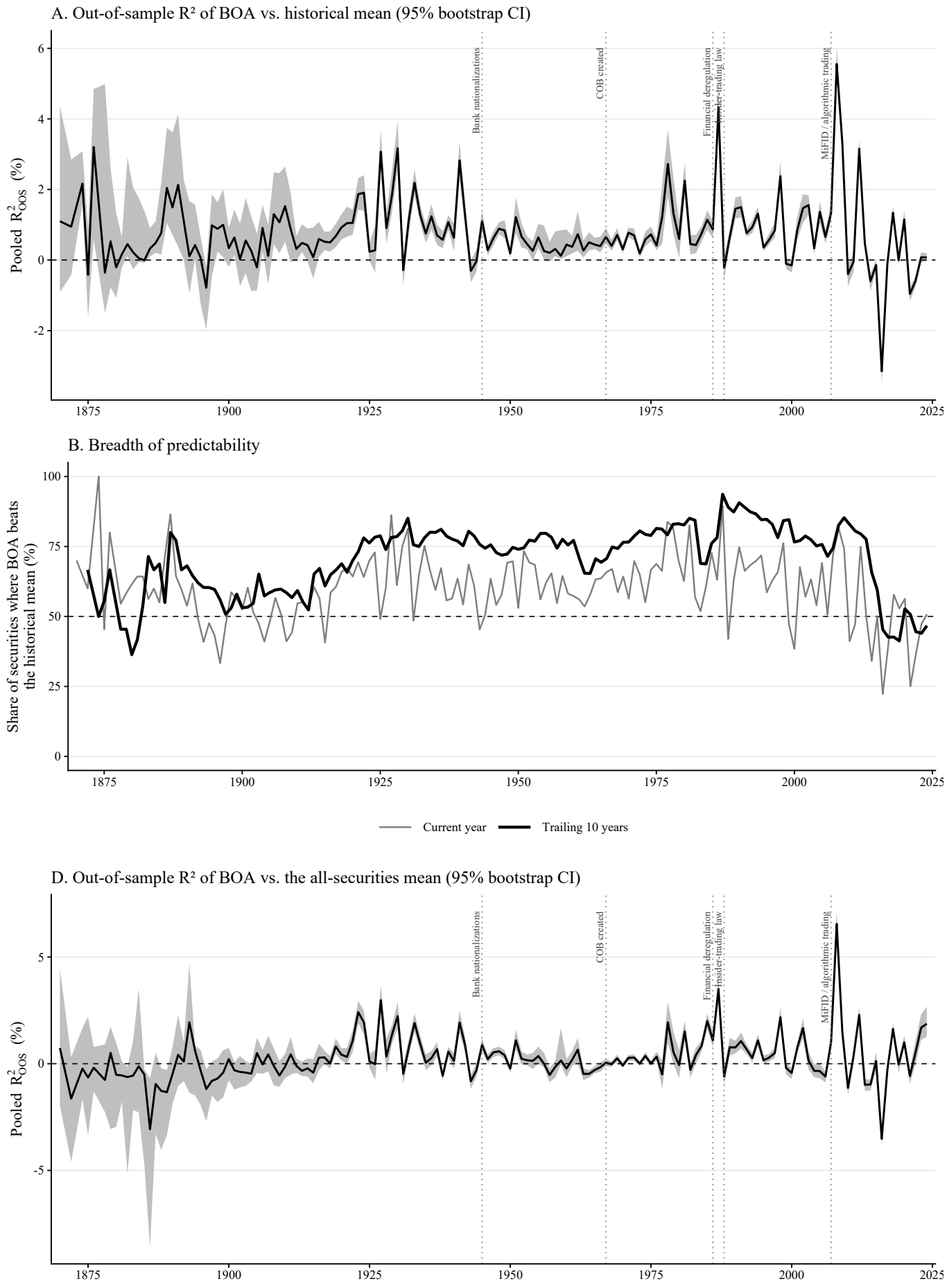


Figure 6: Predictability of individual securities, 1870–2024. A: annual pooled R^2_{OOS} of the aggregate against the historical mean (%), with a 95% bootstrap confidence interval obtained by resampling securities. B: share of securities for which the aggregate has a lower squared error than the historical mean, in the year and over the trailing ten years. D: annual pooled R^2_{OOS} against the all-securities mean. Years with fewer than 10 securities are omitted. Dotted lines mark institutional events.

momentum premium disappears in the largest capitalisations, which are the only ones in which arbitrage capital operates at scale. Patterns that competition can remove are removed; this one is not of that kind. The deterioration since 2014 is the only candidate exception, and it admits two readings. One is decay, with the growth of passive and algorithmic trading. The other is composition: the number of securities doubles in those years with the entry of the small listings of the Euronext Growth and Access compartments, whose sampled prices are stale quotes in 20 to 25% of cases (Section 2.6), and Appendix J shows that the aggregate does worst precisely on such quotes. The second reading is the more likely, but a decade is too short to decide, and it is the one question about the lead-lag effect that the next decade of data will answer.

6 The cross-section: characteristics and momentum

The characteristics expert is the second source of security-level predictability, and the only one whose forecasts can be traded as a long–short portfolio that is immune to the market-wide component. This section opens it up: which characteristics carry the information, how much of the relative return they explain out of sample, and what the resulting portfolios earn.

6.1 Real-time Fama–MacBeth regressions

Table H.21 in Appendix H.1 reports the time-series averages of the Fama–MacBeth slopes of the next half-monthly relative return on the cross-sectional ranks of the characteristics, in percent per unit of rank, with Newey–West t-statistics (24 lags), over 3,221 cross-sections. Momentum has by far the largest and most stable slope: 0.28% per half-month between the top and the median security ($t = 9.3$), 0.20% ($t = 5.3$) in 1870–1976 and 0.42% ($t = 9.7$) in 1977–2024. The share of zero returns, the illiquidity measure, comes second: 0.11% ($t = 4.0$), driven by the modern period (0.20%, $t = 3.9$, against 0.06%, $t = 1.9$). Volatility and beta have negative slopes in the modern period (−0.19%, $t = -3.1$, and −0.09%, $t = -2.2$) and slopes indistinguishable from zero in the historical one. Short-term reversal, long-term reversal, the dividend yield and the number of listed securities have no significant slope in either period. Table H.22 in Appendix H.1 reports the out-of-sample R^2 of the real-time relative-return forecasts against zero: 0.08% over 2.7 million forecasts, 0.007% in the historical period and 0.09% in the modern one, with correlations between forecast and outcome of 0.01 and 0.04. These are small numbers, as they are in Gu et al. (2020); what matters is whether they can be traded. The slopes say how much. A slope of 0.28% per half-month per unit of rank is 0.56% between the two extreme securities, about 13% per year, which is the order of magnitude of the decile portfolios below, whose positions are less extreme; and a momentum slope that doubles between the

two periods while the equal-weighted decile return is of the same order in both periods (11.6 and 10.1%, Section 6.3) says that the relation between past and future returns has become more linear across the ranks, not stronger at the tails.

6.2 Long–short portfolios

Accuracy becomes economic value through decile portfolios. At each date, securities in the bottom decile of price (penny stocks, whose returns are dominated by bid–ask bounce) are excluded; the remaining securities are sorted on the real-time forecast; the portfolio is long the top decile and short the bottom decile, equal-weighted, and held for one half-month (rebalanced every period) or for six months (rebalanced every twelve periods). Transaction costs of 0.5% per transaction are charged on the turnover of each leg. A date is used only if at least half of the securities in each leg have an observed return, to avoid returns computed on a handful of survivors. Table H.23 in Appendix H.2 reports the results and Figure 7 the cumulative return with six-month holding. With six-month holding, the portfolio earns 10.5% per year gross ($t = 5.9$, Sharpe ratio 0.66) and 6.8% net ($t = 3.8$), with a turnover of 7.8% of each leg per half-month; 10.0% gross and 5.1% net in 1870–1976, 11.0% and 8.2% in 1977–2024. With half-monthly rebalancing, the gross return is the same, 10.4%, but the turnover of 43% per half-month turns it into a loss of 10% net: at this horizon the forecasts are right, but they are not worth trading on.

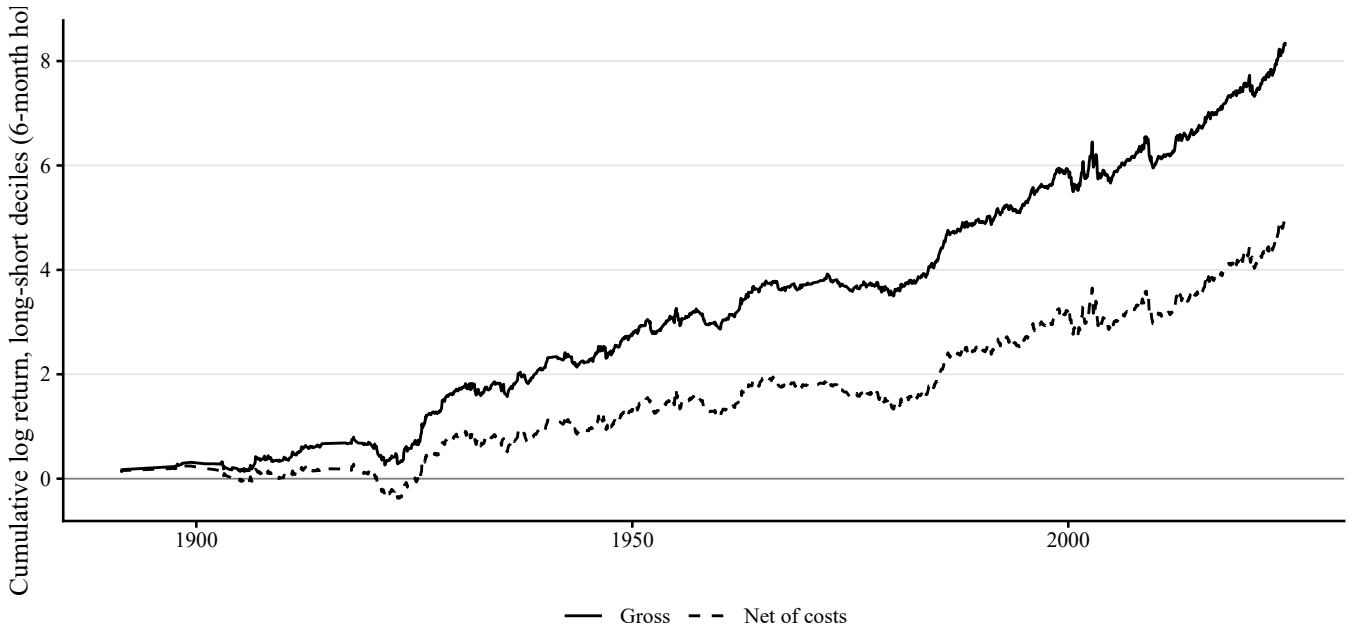


Figure 7: Cumulative log return of the long–short decile portfolio on the characteristics forecast, six-month holding, gross and net of costs.

6.3 One characteristic at a time

Table H.24 in Appendix H.3 sorts on each characteristic separately, with six-month holding, which requires no model and starts as soon as a cross-section has fifty securities, around 1900 for momentum. The results draw a clear map of the French cross-section over 125 years.

Momentum earns 10.8% per year gross ($t = 5.5$, Sharpe ratio 0.63) and 7.1% net ($t = 3.6$) over 1900–2024; the premium is 11.6% gross and 6.9% net in the historical period, 10.1% and 7.3% in the modern one. Figure 8 shows the cumulative return and Figure 9 the annualised return by decade with its confidence interval: it is positive in twelve of thirteen decades, between 5 and 25% per year, and negative only in the 1970s. It is an equal-weighted premium: Section 6.4 shows that it survives within every size tercile but that a value-weighted version earns 1.7% per year over 1977–2025, so that the largest capitalisations carry no momentum. This extends the US evidence of Jegadeesh and Titman (1993), Geczy and Samonov (2016) and Goetzmann and Huang (2018), the Victorian evidence of Chabot et al. (2014) and the global evidence of Baltussen et al. (2021) to a continuous security-level record on the European continent. It is also a direct test of the decay hypothesis of McLean and Pontiff (2016): in the three decades after the publication of Jegadeesh and Titman (1993), the premium is 12% per year in the 1990s, 2% in the 2000s, 15% in the 2010s and 19% in the 2020s, not lower than in the century before. France is a clean laboratory for this test, because the premium was discovered on US data and the trading of French securities on the basis of that discovery could only have started once European evidence existed (Rouwenhorst, 1998); if publication removed it, the French premium should have fallen after 1998, and it did not.

Long-term reversal is negative throughout: securities with the highest returns over months 13 to 36 subsequently underperform by 5.8% per year gross ($t = -3.4$), 7.6% in the historical period and 4.3% in the modern one, as in De Bondt and Thaler (1985). *Short-term* past returns (the return at $t - 1$, skipping t) continue rather than reverse, by 6.3% gross, but the signal is gone net of costs (1.8%, $t = 1.1$). *Volatility* and *beta* are the most striking: high-volatility securities outperform low-volatility ones by 16.6% per year ($t = 4.4$) before 1977 and by 0.3% after; high-beta securities outperform low-beta ones by 14.3% ($t = 3.7$) before 1977 and underperform by 2.3% after. In the historical Paris market, volatility and beta were rewarded; in the modern one, neither the low-volatility pattern of Ang et al. (2006) nor the betting-against-beta pattern of Frazzini and Pedersen (2014) appears, and neither does their opposite. *Illiquidity*, measured by the share of zero returns, earns 4.2% per year gross ($t = 2.9$) and 3.2% net in the modern period, and nothing before 1977 (1.5%, $t = 1.1$), when zero returns are mostly stale quotations rather than the absence of trading. The *dividend yield* and the *number of listed securities* of a firm earn nothing in either period.

The reversal of the volatility and beta premia is the largest change in the French cross-section in 155 years, and it has a natural reading. In a market in which portfolios were concentrated, collective investment did not exist before the 1960s and short sales were confined to the securities admitted to the forward market, idiosyncratic volatility could not be diversified away and was priced, as in Merton (1987); the size-neutral sorts of Section 6.4 show that most of the premium was a reward to small, single-security firms, which is exactly where under-diversification binds. Once diversification became cheap, the premium vanished. It did not reverse into the low-volatility and betting-against-beta anomalies of the United States, which suggests that those owe something to institutional features, benchmarking and leverage constraints (Frazzini and Pedersen, 2014), that the French market did not share in the same form. Momentum and long-term reversal, by contrast, are unchanged across the two periods. The premia that compensated a risk that investors could not diversify disappeared with the technology of diversification; the premia that reflect how prices incorporate information did not. This is the form in which the third hypothesis of the introduction, that markets became more efficient, is confirmed, and the form in which it is not.



Figure 8: Momentum in Paris, 1900–2024: cumulative log return of the winners-minus-losers decile portfolio, six-month holding, gross and net of costs.

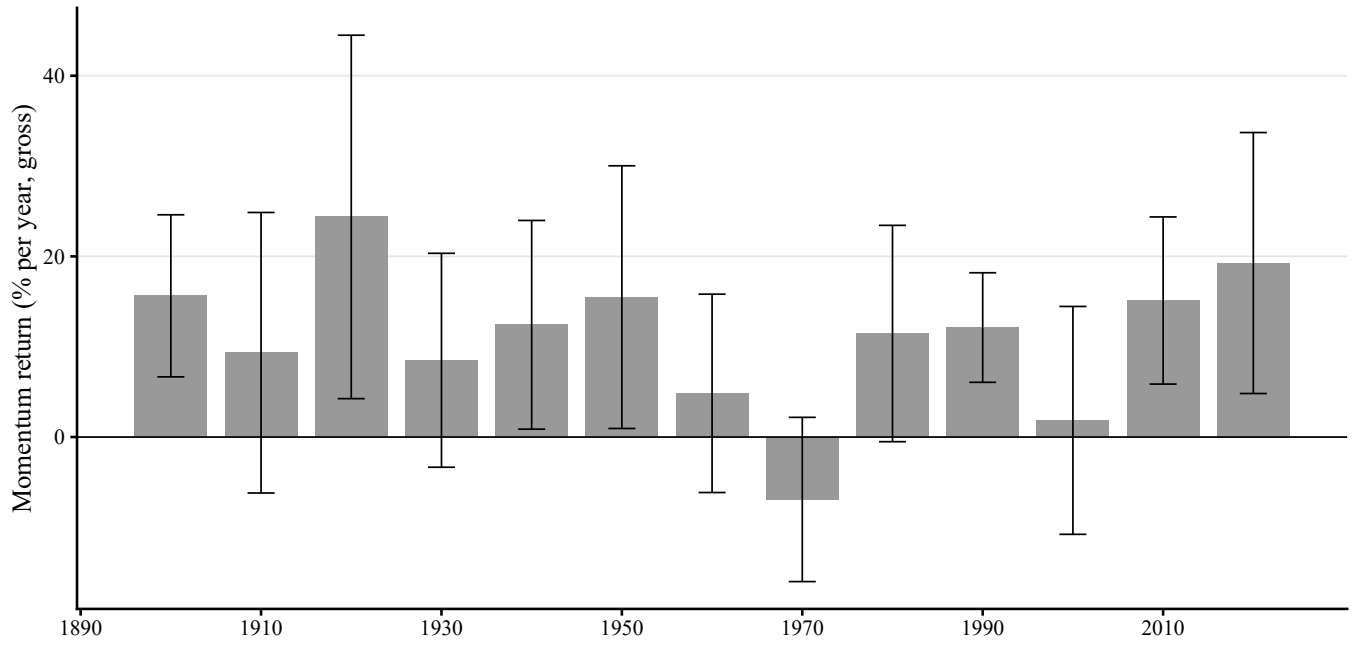


Figure 9: Annualised gross momentum return by decade (%), with 95% Newey–West confidence intervals.

6.4 Robustness

Three concerns are specific to historical long–short portfolios: costs, which were higher in the past; thin trading, which can leave a leg with few observed returns; and size, since the extreme deciles are populated by small and illiquid securities. Table H.25 in Appendix H.4 reports the net momentum and combined-model returns for costs of 0.25, 0.5, 1 and 2% per transaction: with six-month holding the momentum premium is 8.9% net at 0.25%, 7.1% at 0.5%, 3.5% at 1% ($t = 1.7$) and −3.8% at 2%; the break-even cost is about 1.3% per transaction, 1.1% in the historical period and 1.6% in the modern one. Table H.26 in Appendix H.4 lowers the coverage threshold of each leg from one half to one third of its securities, with no material effect (10.2% against 10.8% gross for momentum). Table H.27 in Appendix H.4 repeats the sorts within size groups, so that the long and short legs have the same size composition: market-capitalisation terciles when capitalisation is available, and otherwise the number of listed securities of the firm (one, two, three or more), which separates the large multi-security companies of the historical period from the rest. Momentum survives in every size group, at 8.2% gross ($t = 5.5$) and 5.1% net, with the same magnitude in both periods. The historical volatility and beta premia are considerably reduced by the size control, from 16.6% to 5.7% and from 14.3% to 5.5% ($t = 2.2$ for both): a large part of the reward to volatility and beta in the historical market was a reward to small, single-security firms. Long-term reversal and the modern illiquidity premium are unaffected.

The market capitalisations and spreads of Section 2.6 allow the size and cost concerns to be addressed directly for 1977–2025. Table H.28 in Appendix H.4 reports, for momentum and for the other characteristics,

the gross return, the return net of the uniform 0.5% cost per transaction, and the return net of each security's own half-spread charged on the shares it trades at each rebalancing. Four results stand out. First, costs measured by the securities' own spreads are lower than the uniform cost: the equal-weighted momentum portfolio with six-month holding earns 11.0% gross and 10.4% net of its own spreads, against 8.2% net of the uniform cost, with an average annual cost of 0.6%; with one-month holding, whose turnover is six times larger, it earns 12.9% gross, 11.4% net of own spreads and 5.4% net of the uniform cost. Since the high–low estimator understates spreads (Section 2.6), the own-spread figures are an upper bound on net returns and the uniform-cost figures remain the conservative ones; the premium is well inside both. Figure H.11 in Appendix H.4 plots the three cumulative returns. Second, momentum is a premium of small and mid-capitalisation securities. The value-weighted decile portfolio earns 1.7% per year ($t = 0.4$) and nothing net of costs; the size-neutral portfolio, long–short quintiles within capitalisation terciles, earns 7.4% gross ($t = 3.9$) and 6.9% net of own spreads ($t = 3.6$); the equal-weighted portfolio excluding the smallest capitalisation quintile earns 6.0% ($t = 2.1$). This is the pattern of Hong et al. (2000) for the United States, in which momentum profits decline sharply with size, in a steeper form: in Paris the largest capitalisations carry no momentum at all. Third, the premium to size itself is 9.1% per year ($t = 3.1$) over 1977–2025 with monthly rebalancing, and 8.7% net of own spreads; the modern illiquidity premium is 9.5% ($t = 3.8$) when illiquidity is measured by the share of zero returns and 5.5% ($t = 2.6$) when measured by the Amihud ratio, but the latter has an annual cost of 2.0% at its own spreads and is a loss net of the uniform cost. Fourth, the volatility and beta premia, with real capitalisations and in the modern period, are 3.3% ($t = 1.0$) and -4.5% ($t = -1.3$) gross; the size-neutral volatility premium is 2.0% ($t = 0.8$). The disappearance of the two historical premia is confirmed with the proper size control.

The two main findings of the paper are thus located in the same securities. The delayed reaction to market-wide returns of Section 5 and the momentum premium are both concentrated in small, actively traded securities, both are absent from the largest capitalisations, and neither is stale pricing. This is what the gradual-information-diffusion model of Hong and Stein (1999) predicts: when information travels slowly across investors, prices underreact to it, which produces a lead-lag effect at the horizon of weeks and momentum at the horizon of months, and both are strongest where information travels slowest, in the small firms that Hong et al. (2000) identify. The French evidence adds two things to that model: the two effects are measured on the same securities in the same market, and neither decayed in 155 years, which is what a mechanism rooted in the cost of attention, rather than in a temporary mispricing, implies.

7 A long-run view: factors, crashes, other markets and mechanisms

Sections 5 and 6 establish where predictability lies; this section uses the length of the sample to ask questions that shorter samples cannot answer. Some of them concern rare events: momentum crashes follow bear markets, and a sample that contains the Great Depression, two world wars, the 2001–03 and 2008–09 bear markets and the Covid rebound has enough of them to test whether the crash profile documented for the United States is a general property of the strategy. Others concern comparisons across markets and institutions: whether the French premia are those of the United States and Europe, and whether the many changes in the organisation of the Paris Bourse over 155 years, from the transaction tax and the closures of the forward market to exchange controls and electronic trading, changed the amount of predictability. The section first presents the French factor series built from the security-level data, then studies momentum crashes, compares the French momentum factor with its US and European counterparts, and finally relates annual predictability and the momentum premium to the state and the structure of the market.

7.1 A French factor dataset, 1870–2025

As a by-product, the paper provides monthly and half-monthly return series for a set of French equity factors: the equal-weighted market, momentum, long-term reversal, volatility, beta and illiquidity, each the return of a decile long–short portfolio rebalanced monthly, with the bottom price decile excluded. Table I.29 reports their annualised mean, volatility, Sharpe ratio and t-statistic by period. The equal-weighted market earns 7.8% per year with a volatility of 15.7% and a Sharpe ratio of 0.50, the same in both periods. Momentum earns 10.7% per year ($t = 6.3$, Sharpe ratio 0.59), 8.8% before 1977 and 13.3% after; volatility 10.9%, all of it before 1977 (16.7%, $t = 5.2$, against 3.6%, $t = 1.1$); beta 4.5%, with 11.6% before 1977 ($t = 4.0$) and −4.1% after; illiquidity 3.3%, all of it after 1977 (5.9%, $t = 3.6$); long-term reversal −4.2% ($t = -2.5$). The series are available with the paper and are, to my knowledge, the longest security-based factor series for a continental European market. For 1977–2025, the dataset also contains the value-weighted market and the value-weighted, size-neutral and ex-micro-cap momentum factors, the size factor and the Amihud illiquidity factor of Table H.28. Table I.30 in Appendix I.1 compares the two market returns over 1977–2025: the equal-weighted market earns 7.8% per year with a volatility of 14.8% (Sharpe ratio 0.53), the value-weighted market 6.4% with a volatility of 21.5% (Sharpe ratio 0.30). The lower volatility of the equal-weighted index is in part the stale-price smoothing of Section 2.7, since a fifth of its constituents carry a stale quote on a typical date after 2005; the higher mean is the size premium.

7.2 Momentum crashes

Daniel and Moskowitz (2016) show that US momentum crashes in the months that follow bear markets, when the market rebounds and the short leg, made of high-beta losers, rallies. Table I.31 in Appendix I.1 estimates their regression on French data: the monthly momentum return on a bear-market indicator (the cumulative market return over the past 24 months is negative, known at $t-1$), its interaction with the contemporaneous market return, and the market return. The interaction coefficient is −0.56 ($t=-3.3$): in a bear market, a 10% monthly rise of the market costs the momentum portfolio 5.6 percentage points, while outside bear markets the momentum return is unrelated to the market (coefficient 0.01, $t=0.2$) and the bear-market indicator itself has no effect on its level. Table I.32 lists the fifteen worst momentum months and Figure 10 plots the monthly series. The worst months are May 2003 (−29%), November 2020 (−29%), April 2009 (−24%), October and November 2002, February 1988, October and November 2001, December 1931, January 1931, February 1932 and March 1935; in twelve of the fifteen, the bear-market indicator is on and the market rises, by 6 to 25% in the month. The French market has the same optionality in momentum as the United States, in a sample that includes the Great Depression, the 2001–03 and 2008–09 bear markets and the Covid rebound. The optionality is what the mechanism predicts: in a bear market the short leg is made of the securities that fell most, which are the ones with the highest betas, so that the strategy is short a call option on the market, and it pays for that option when the market rebounds. Finding the same pattern in an independent market, over a sample that includes the 1930s, addresses the concern that a result documented on one market over a few episodes is specific to those episodes (Harvey et al., 2016); it also says that the crash risk is part of what the momentum premium compensates, and not an accident of US data.

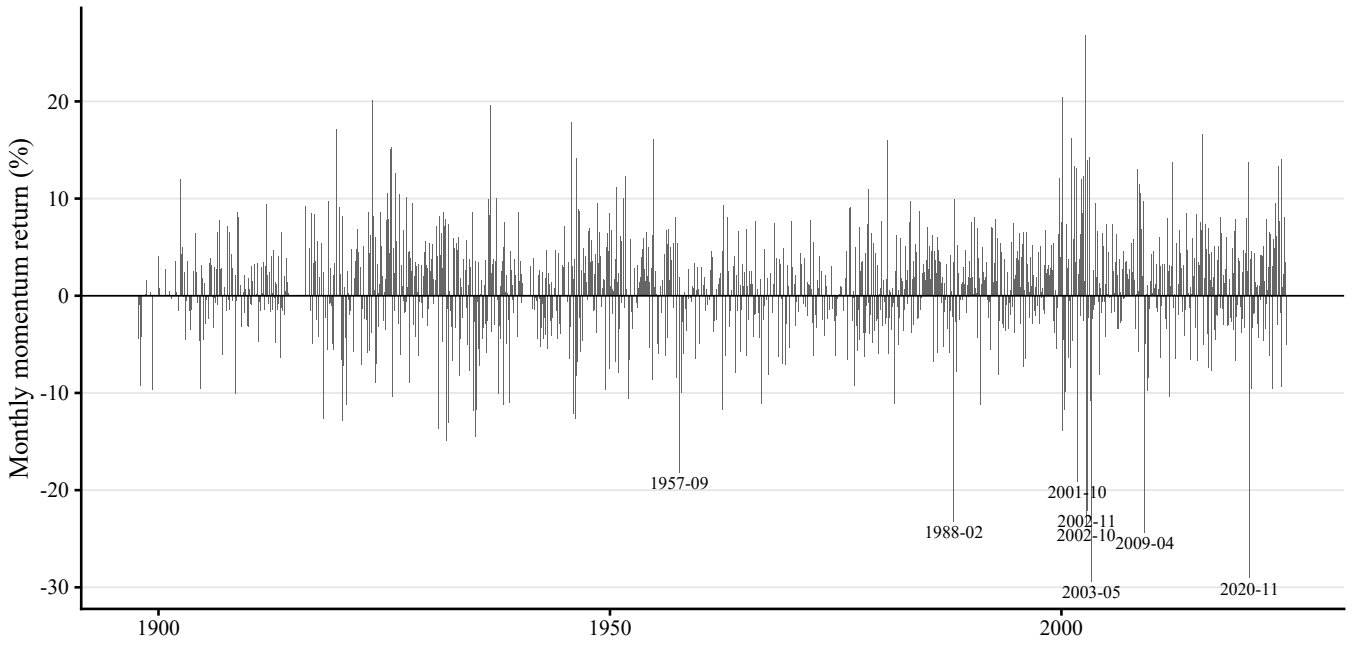


Figure 10: Monthly momentum returns in Paris, 1900–2024 (%); the eight worst months are labelled.

7.3 France, the United States and Europe

Table I.33 in Appendix I.2 compares the French momentum factor with the US factor of Kenneth French's data library from 1927 and with the European factor from 1990: correlation, annualised means and the t-statistic of the difference on common months. Over 1,067 common months since 1927, the French and US factors have a correlation of 0.32, 0.14 before 1977 and 0.49 after; France earns 10.3% per year and the United States 7.3%, a difference that is not significant ($t = 1.5$). Over 410 common months since 1990, the correlation with the European factor is 0.70, and France earns 14.7% against 10.2% ($t = 2.0$). The comparison is not like for like, however: the French factor is an equal-weighted decile portfolio, whereas the US and European factors of the data library are value-weighted within size halves. With the capitalisations of Section 2.6, the French size-neutral momentum portfolio (Table H.28), the closest counterpart of that construction, earns 7.4% per year over 1977–2025 against 6.9% for the United States over the same period (Table I.33); the value-weighted French portfolio earns 1.7%. On a comparable basis, France does not earn more than the United States in the modern period; it earns the same, and the excess in Table I.33 is the small-capitalisation component of the French factor. Figure 11 plots rolling ten-year means. Two implications follow. The premium is not a US peculiarity nor an artefact of CRSP; and since the French and US factors share only a third of their variance and crash in different months, an investor diversified across the two would have borne markedly less crash risk than the US evidence alone suggests. The rise of the correlation from 0.14 before 1977 to 0.49 after, and 0.70 with Europe, is the comovement of momentum returns across integrated markets that Asness et al. (2013) document; the low historical correlation is the more informative number. Momentum earned the same premium in two markets that were, before 1977, largely segmented and whose factors shared 2% of their variance. A premium that exists independently in segmented markets is hard to attribute to a single global risk factor, and easy to attribute to a mechanism that operates market by market, which is again the slow diffusion of information among investors.[20]

7.4 Mechanisms: predictability and market structure

Why does predictability vary? Table I.34 in Appendix I.3 regresses the annual momentum premium and the annual R^2_{OOS} of the aggregate against the all-securities mean on characteristics of the market in that year, namely the log number of listed securities, illiquidity (the average share of zero returns), the cross-sectional dispersion of returns, lagged market volatility and lagged market return, and a post-1977 indicator, and on five variables of market structure assembled from the institutional history of the Paris Bourse: an indicator

[20] [The French rolling mean is only drawn from 1937 in Figure 11 because the rolling window returns NA after any missing month; recompute with a minimum number of observations so that the French series starts in 1910.]

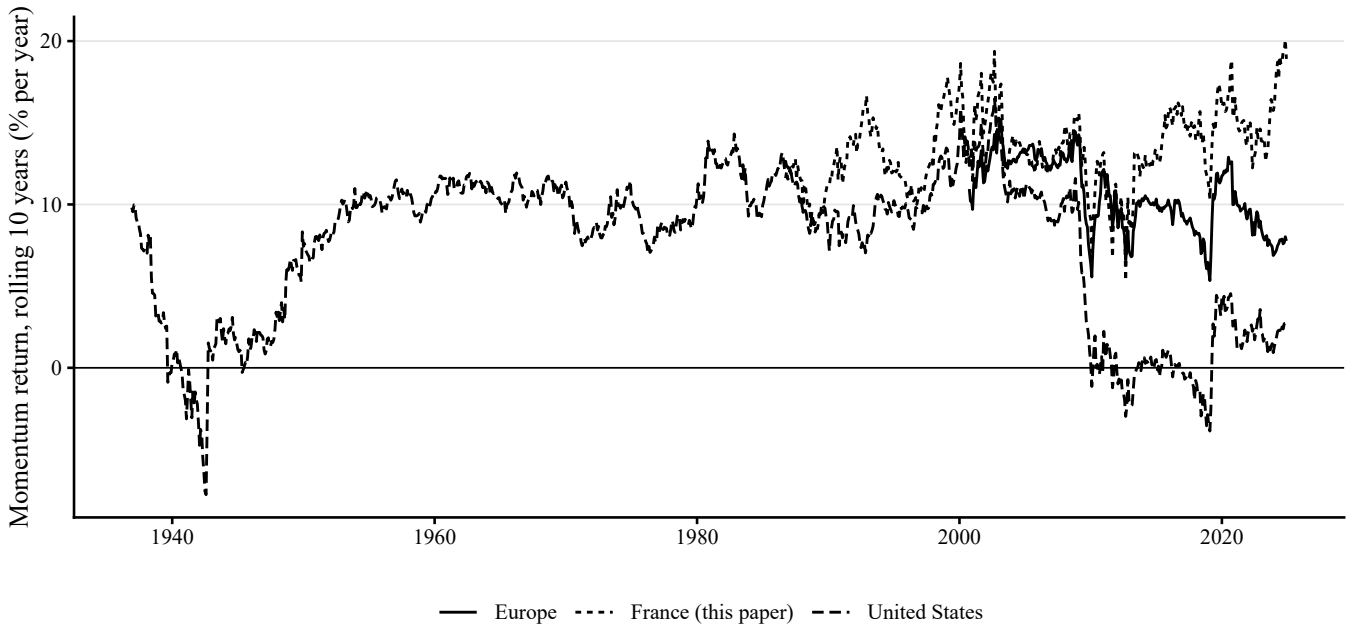


Figure 11: Momentum returns over rolling ten-year windows (% per year): France, the United States and Europe.

for a tax on securities transactions (the *impôt sur les opérations de bourse* created in 1893 and repealed in 2008, and the financial-transaction tax from 2012), the fraction of the year during which the forward market was open (it was closed from July 1914 to 1920 and from May 1940 to 1949), an indicator for the coexistence of the official market of the *agents de change* and the *Coulisse* (until their merger in 1961), an indicator for exchange controls (1940–1989) and an indicator for electronic trading (from 1986). Appendix D documents the sources and coding.

For the momentum premium, one variable matters: lagged market volatility, with a coefficient of −0.47 ($t = -2.5$), that is, a year in which annualised volatility is 10 points higher is followed by a momentum return 4.7 points lower, the annual counterpart of the crashes of Table I.31. Exchange controls are associated with a premium 7 points lower ($t = -1.8$), weakly. The transaction tax, the closures of the forward market, the dual market, electronic trading, the number of securities, illiquidity and dispersion have no effect. For the R^2_{OOS} against the all-securities mean, nothing is significant at 5%; illiquidity has a negative coefficient ($t = -1.9$), consistent with the stale-price reading of zero returns in the historical period. The absence of any effect of the institutional variables is itself informative: in 155 years, no reform of the Paris Bourse changed the amount of predictability in individual securities or the momentum premium in a way that the data can detect. The identification is coarse, however, since the step indicators are collinear with the post-1977 indicator and with each other; a sharper test would use security-level variation in the ability to sell short, which the forward market provided for some securities and not others, and is left for future work.

8 Conclusion

This paper asks an old question with a new dataset and a forecasting framework built for it. The dataset follows securities listed in Paris, twice a month, from 1845 to 2025, across more market structures and more generations of forecasting models than any other security-level record. The framework, sequential prediction with expert advice, aggregates forecasting models as they become available, assumes nothing about the process that generates returns, and comes with a guarantee that the aggregate does, in the long run, as well as the best model in hindsight. Two methodological additions make it work on a long panel: learning a common weight vector across securities, and building experts that pool the cross-section in real time. Two measurement problems of historical data, stale prices in indices and noisy prices in price-scaled predictors, are identified and removed; both generate predictability that is large and spurious, and both are likely to matter for other long-run studies.

The answer that emerges is layered. At the level of the market, macro-financial predictors deliver an out-of-sample R^2 of half a percent per half-month, the order of magnitude the US literature regards as an upper bound, worth about 2% per year to a mean-variance investor, and entirely concentrated in volatile and crisis periods; the second-generation methods of combination and diffusion add nothing to it, and aggregation rules does not perform better than the simple average of experts. At the level of individual securities, a single set of weights learned across the market beats the historical mean, zero and the all-securities mean by 0.6 to 0.7% per half-month; the aggregation rule shows that this predictability is, first, the delayed reaction of individual securities to market-wide returns, strongest for small and long-listed securities and in volatile years, and second, their characteristics. The delayed reaction has no trend, no break and no institutional history over 155 years, it is not weaker in the era of electronic trading, and it is not stale pricing: with daily trading data, it is as strong among securities traded on both sampling dates as in the full sample, absent from stale quotes, and concentrated in small, actively traded securities. At the level of the cross-section, momentum has earned about 10% per year gross in every decade but one since 1900, net of plausible costs and of the securities' own spreads, within every size group but not among the largest capitalisations, with the same crash profile as in the United States and no decay after its publication; the volatility and beta premia of the nineteenth and early twentieth century have disappeared, while an illiquidity premium has appeared.

The three hypotheses of the introduction fare differently. The first, that signal-to-noise ratios are low and more predictors or more sophisticated combinations do not help much, is confirmed at every level: forty macro-financial predictors, two generations of combination methods and four aggregation rules add nothing to one lagged market return, one momentum sort and an interest rate or two. The second, that predictability

lives in particular segments and periods, is confirmed: in small securities, in the cross-section rather than in the aggregate, and in the volatile years in which prices have the most common information to absorb. The third, that markets have become more efficient, is confirmed for the premia that reflected a compensation for risk that could not be diversified away in a market of concentrated portfolios, volatility and beta, and refuted for momentum and for the speed at which individual securities absorb common information, which did not improve with the technology of trading. The common thread is that the two forms of predictability that survive every test, the delayed reaction and momentum, live in the same securities, small and actively traded, and have the signature of slow information diffusion rather than of stale prices or of a premium for risk.

References

Amat, C., Michalski, T., and Stoltz, G. (2018). Fundamentals and exchange rate forecastability with simple machine learning methods. *Journal of International Money and Finance*, 88:1–24.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299.

Asness, C. S., Moskowitz, T. J., and Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance*, 68(3):929–985.

Asparouhova, E., Bessembinder, H., and Kalcheva, I. (2013). Noisy prices and inference regarding returns. *Journal of Finance*, 68(2):665–714.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

Baker, M. P. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *Journal of Finance*, 55(5):2219–2257.

Baltussen, G., Swinkels, L., and van Vliet, P. (2021). Global factor premiums. *Journal of Financial Economics*, 142(3):1128–1156.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Bazot, G., Monnet, E., and Morys, M. (2022). Taming the global financial cycle: central banks as shock absorbers in the first era of globalization. *The Journal of Economic History*, 82(3):801–839.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300.

Blackwell, D. (1956). An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8.

Blume, M. E. and Stambaugh, R. F. (1983). Biases in computed returns: An application to the size effect. *Journal of Financial Economics*, 12(3):387–404.

Boudoukh, J., Michaely, R., Richardson, M., and Roberts, M. R. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *The Journal of Finance*, 62(2):877–915.

Boudoukh, J., Richardson, M. P., and Whitelaw, R. F. (1994). A tale of three schools: Insights on autocorrelations of short-horizon stock returns. *Review of Financial Studies*, 7(3):539–573.

Breen, W., Glosten, L. R., and Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of finance*, 44(5):1177–1189.

Brennan, M. J., Jegadeesh, N., and Swaminathan, B. (1993). Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies*, 6(4):799–824.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of financial economics*, 18(2):373–399.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251.

Campbell, J. Y. and Shiller, R. J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors. *The review of financial studies*, 1(3):195–228.

Campbell, J. Y. and Shiller, R. J. (1988b). Stock prices, earnings, and expected dividends. *the Journal of Finance*, 43(3):661–676.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Capistrán, C. and Timmermann, A. (2009). Forecast combination with entry and exit of experts. *Journal of business & economic statistics*, 27(4):428–440.

Cenesizoglu, T. and Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance*, 36(11):2974–2987.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, learning, and games*. Cambridge university press.

Chabot, B., Ghysels, E., and Jagannathan, R. (2014). Momentum trading, return chasing, and predictable crashes. Working Paper 20660, National Bureau of Economic Research.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Chordia, T., Subrahmanyam, A., and Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58(1):41–58.

Chordia, T. and Swaminathan, B. (2000). Trading volume and cross-autocorrelations in stock returns. *Journal of Finance*, 55(2):913–935.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. (2014). Eugene f. fama, efficient markets, and the nobel prize. *Chicago Booth Review*.

Cooper, I. and Priestly, M. (2009). Output gaps and time-varying predictability. *Journal of Money, Credit and Banking*, 41(7):1415–1434.

Corwin, S. A. and Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance*, 67(2):719–760.

Cover, T. M. (1966). *Behavior of sequential predictors of binary sequences*. Number 7002. Stanford University, Stanford Electronics Laboratories, Systems Theory . . . .

Dangl, T. and Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1):157–181.

Daniel, K. and Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2):221–247.

De Bondt, W. F. M. and Thaler, R. (1985). Does the stock market overreact? *Journal of Finance*, 40(3):793–805.

Diebold, F. X. and Mariano, R. S. (1995). Com paring predictive accu racy. *Journal of Business and Economic Statistics*, 13(3):253–263.

Diebold, F. X. and Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting*, 6(1):21–40.

Diebold, F. X. and Shin, M. (2019). Machine learning for regularized survey forecast combination: Partially-egalitarian lasso and its derivatives. *International Journal of Forecasting*, 35(4):1679–1691.

Driesprong, G., Jacob, B., and Oberlechner, B. (2008). Striking oil: Another puzzle? *Journal of Financial Economics*, 89(2):307–327.

Ducros, J., Grandi, E., Hékimian, R., Prunaux, E., Riva, A., and Ungaro, S. (2018). Collecting and storing historical financial data: the dfih project. Technical report, HAL.

Elliott, G. and Timmermann, A. (2004). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics*, 122(1):47–79.

Elliott, G. and Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature*, 46(1):3–56.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, 22(1):3–25.

Fama, E. F. and French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of financial economics*, 25(1):23–49.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Fama, E. F. and Schwert, G. W. (1977). Asset returns and inflation. *Journal of financial economics*, 5(2):115–146.

Farmer, L. E., Schmidt, L., and Timmermann, A. (2023). Pockets of predictability. *The Journal of Finance*, 78(3):1279–1341.

Ferreira, M. A. and Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3):514–537.

Foucault, T., Pagano, M., and Röell, A. (2023). *Market liquidity: theory, evidence, and policy*. Oxford University Press.

Fouliard, J., Howell, M., Rey, H., and Stavrakeva, V. (2021). Answering the queen: Machine learning and financial crises. Technical report, National Bureau of Economic Research.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

Gaillard, P. and Goude, Y. (2016). opera: Online prediction by expert aggregation. R package, https://CRAN.R-project.org/package=opera.

Gaillard, P., Stoltz, G., and van Erven, T. (2014). A second-order bound with excess losses. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, volume 35 of *JMLR Workshop and Conference Proceedings*, pages 176–196.

Geczy, C. C. and Samonov, M. (2016). Two centuries of price-return momentum. *Financial Analysts Journal*, 72(5):32–56.

Goetzmann, W. N. and Huang, S. (2018). Momentum in Imperial Russia. *Journal of Financial Economics*, 130(3):579–591.

Golez, B. and Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics*, 127(2):248–263.

Goyal, A., Welch, I., and Zafirov, A. (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 37(11):3490–3557.

Greenwood, R., Hanson, S. G., Shleifer, A., and Sørensen, J. A. (2022). Predictable financial crises. *The Journal of Finance*, 77(2):863–921.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5):2223–2273.

Guidolin, M. and Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11):3503–3544.

Guo, H. (2006). Common variation in national stock index returns and volatility. *Journal of Business & Economic Statistics*, 24(3):356–371.

Haddad, V., Huebner, P., and Loualiche, E. (2025). How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. *American Economic Review*, 115(3):975–1018.

Hai Hong, J. D. L. and Stein, J. D. (2007). A unified framework for exploring predictability in returns, order imbalance, volatility, and liquidity. *Journal of Financial Economics*, 84(2):437–468.

Hannan, J. (1957). Approximation to bayes risk in repeated plays. dresher m, tucker a, wolfe p, eds. contributions to the theory of games, vol. iii.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Hautcœur, P.-C., Rezaee, A., and Riva, A. (2010). How to regulate a financial market? The impact of the 1893 regulatory reform on the Paris Bourse. Working Paper 2010-01, Paris School of Economics.

Hautcœur, P.-C. and Riva, A. (2012). The Paris financial market in the nineteenth century: Complementarities and competition in microstructures. *Economic History Review*, 65(4):1326–1353.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of finance*, 66(1):1–33.

Henkel, S. J., Martin, J. S., and Nardari, F. (2011). Time-varying short-horizon predictability. *Journal of financial economics*, 99(3):560–580.

Hennig, T., Iossifov, M. P. K., Iossifov, P., and Varghese, M. R. (2023). *Predicting financial crises: The role of asset prices*. International Monetary Fund.

Herbster, M. and Warmuth, M. K. (1998). Tracking the best expert. *Machine Learning*, 32(2):151–178.

Hong, H., Lim, T., and Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, 55(1):265–295.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54(6):2143–2184.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies*, 20(4):1113–1138.

Jang, K., Jun, K.-S., Kuzborskij, I., and Orabona, F. (2023). Tighter pac-bayes bounds through coin-betting. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 2240–2264. PMLR.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2–3):95–101.

Jordà, O., Schularick, M., and Taylor, A. M. (2015). Leveraged bubbles. *Journal of Monetary Economics*, 76:S1–S20.

Jun, K.-S., Orabona, F., Wright, S., and Willett, R. (2017). Online learning for changing environments using coin betting.

Kelly, B., Malamud, S., and Zhou, K. (2024). The virtue of complexity in return prediction. *The Journal of Finance*, 79(1):459–503.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

Kindleberger, C. (1978). Manias, panic and crashes. *New York Basic Books*.

Kothari, S. P. and Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial economics*, 44(2):169–203.

Lagneau-Ymonet, P. and Riva, A. (2010). Nouvelles des archives. aux sources de l'histoire boursière: les archives de la compagnie des agents de change. *Entreprises et histoire*, 58(1):179–182.

Lamont, O. A. (1998). Earnings and expected returns. *Journal of Finance*, 53(6):1563–1584.

Le Bris, D. and Hautcœur, P.-C. (2010). A challenge to triumphant optimists? A blue chips index for the Paris stock exchange, 1854–2007. *Financial History Review*, 17(2):141–183.

Lesmond, D. A., Ogden, J. P., and Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of Financial Studies*, 12(5):1113–1141.

Lettau, M. and Ludvigson, S. C. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, 56(3):815–849.

Levy-Garboua, V. and Monnet, É. (2016). Les taux d'intérêt en france: une perspective historique. *Revue d'économie financière*, 121(1):35–58.

Lin, H., Wu, C., and Zhou, G. (2018). Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. *Management Science*, 64(9):4218–4238.

Linnainmaa, J. T. and Roberts, M. R. (2018). The history of the cross-section of stock returns. *Review of Financial Studies*, 31(7):2606–2649.

Littlestone, N. and Warmuth, M. K. (1994). The weighted majority algorithm. *Information and computation*, 108(2):212–261.

Lo, A. W. and MacKinlay, A. C. (1990a). An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45(1–2):181–211.

Lo, A. W. and MacKinlay, A. C. (1990b). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2):175–205.

Ludvigson, S. C. and Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of financial economics*, 83(1):171–222.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *Journal of Finance*, 71(1):5–32.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42(3):483–510.

Minsky, H. (1986). Global consequences of financial deregulation. *Washington University Working Paper*, (96).

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management science*, 60(7):1772–1791.

Nelson, C. R. (1976). Inflation and rates of return on common stocks. *The journal of Finance*, 31(2):471–483.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Pesaran, H. and Timmermann, A. (2005). Real-time econometrics. *Econometric Theory*, 21(1):212–231.

Pontiff, J. and Schall, L. D. (1998). Book-to-market ratios as predictors of market returns. *Journal of financial economics*, 49(2):141–160.

Qiu, J., Wang, B., and Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PloS one*, 15(1):e0227222.

Rajan, R. G. and Zingales, L. (2003). The great reversals: The politics of financial development in the twentieth century. *Journal of Financial Economics*, 69(1):5–50.

Rangvid, B. (2006). Macro predictors and the stock market: Evidence from denmark. *Scandinavian Journal of Economics*, 108(2):269–287.

Rapach, D. and Zhou, G. (2013). Forecasting stock returns. In Elliott, G., Granger, C., and Timmermann, A., editors, *Handbook of Economic Forecasting, Volume 2*, pages 328–383. Elsevier.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862.

Robbins, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, pages 131–149. University of California Press.

Rouwenhorst, K. G. (1998). International momentum strategies. *Journal of Finance*, 53(1):267–284.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio management*, pages 68–75.

Santos, P. and Veronesi, P. (2006). Labor income and predictable stock returns. *Review of Financial Studies*, 19(1).

Scholes, M. and Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5(3):309–327.

Shafer, G. and Vovk, V. (2005). *Probability and finance: it's only a game!*, volume 491. John Wiley & Sons.

Sirignano, J. and Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*, 19(9):1359–1390.

Steyn, D. H., Greyling, T., Rossouw, S., Mwamba, J. M., et al. (2020). Sentiment, emotions and stock market predictability in developed and emerging markets. Technical report, Global Labor Organization (GLO).

Timmermann, A. (2006). Chapter 4: Forecast combinations. volume 1 of *Handbook of Economic Forecasting*, pages 135–196. Elsevier.

Timmermann, A. (2018). Forecasting methods in finance. *Annual Review of Financial Economics*, 10(1):449–479.

Vovk, V. G. (1990). Aggregating strategies. In *Proceedings of the third annual workshop on Computational learning theory*, pages 371–386.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Wintenberger, O. (2017). Optimal learning with bernstein online aggregation. *Machine Learning*, 106:119–141.

A Data appendix

COMPAGNIE DES AGENTS DE CHANGE

COURS AUTHENTIQUE

SEUL OFFICIEL

PARAISSANT TOUS LES JOURS DE BOURSE

BOURSE DE PARIS

Lundi 16 Mars 1885

Prix d'Abonnement. — Paris : un an, 32 fr.; six mois, 16 fr.; trois mois, 8 fr. — Départements : un an, 42 fr.; six mois, 21 fr.; trois mois, 11 fr. — Etranger: frais de poste en sus

Reports Comptant liq.	Reports Liq. pr.	Reports Liq. à l'autre	Taux d'émiss.	Désignation des valeurs	Jouissance	Au comptant	A terme	1er cours	Plus haut	Pl. bas	Dernier cours	Derniers cours cotés précédemment Compt.	Derniers cours cotés précédemment Terme	Intérêts et dividendes.
				FONDS D'ÉTAT FRANÇAIS										
14			Divers	3 %	1er avril 1885	81f35 40 45	en liq…	81 47½			81 60	82 20	82 20	3f
							fin cour.							
							Pr. fin c.		d 2f					
							Pr. fin c.		d 1f	81 65	81 80 d 50			
							Pr. fin c.	81 87½	81 90 d 25					
							Pr. fin p.		d 1f		d 50			
							Pr. fin p.		d 2f					
							Pr. fin p.		d 25		d 3f			
			Divers	3 % amortissable (remboursable à 500 fr., par coupure de 15 fr. de rente, de 1878 à 1953)	16 janv. 85.	83f85 90	en liq…	83 75	83 90	83 75	83 85	83 85	83 75	3f
							fin cour.		d 1f		84 .. d 50			
							Pr. fin c.		d 25					
							Pr. fin p.		d 1f		d 50			
							Pr. fin p.		d 25					
			…	4 %	22 mars 1885		en liq…					102 ..	98 ..	4f
							fin cour.							
			…	4 ½ %	22 mars 1885	105f 104f90 75 105f	en liq…					105 ..	107 37½	4f50
							fin cour.							
19			…	4 ½ % 1883	16 févr. 1885	110f10 15 05 10 05 110f	en liq…					110 ..	110 12½	4f50
							fin cour.	110 22½	110 25	110 20	110 20			
							Pr. fin c.		d 2f					
							Pr. fin c.		d 1f		110 32½ d 50			
							Pr. fin c.		110 45 d 25					
							Pr. fin p.		d 1f					
							Pr. fin p.		d 2f		110 85 d 50			
							Pr. fin p.		d 25		d 2f			
			…	Promesses d'inscriptions 4 ½ % 1883	16 août 1883	0.24						0.24		
			…	Bons du Trésor								3¾ % de 10 à 11 mois.		
			…	Bons du Trésor, 4 %, éch. le 1er sept. 1885. (coup. de 500 f.)	mars 1885	Coup. de 5000 f.						509 ..		4 %
				d° d° d° (coup. de 1000 f.)		Coup. de 10000 f.						1017 50		4 %
				d° d° éch. 1er mars 1886 (coup. de 500 f.)	mars 1885	Coup. de 5000 f.						502 50		4 %
				d° d° d° (coup. de 1000 f.)		Coup. de 10000 f.						1005 25		4 %
				d° d° éch. 1er sept. 1886 (coup. de 500 f.)	mars 1885	Coup. de 5000 f.						502 50		4 %
				d° d° d° (coup. de 1000 f.)		Coup. de 10000 f.						1005 25		4 %
				d° d° éch. 1er mars 1887 (coup. de 1000 f.)	mars 1885	Coup. de 5000 f.						1024 25		4 %
				d° d° d° (coup. de 10000 f.)								10262 50		
				d° d° éch. 1er sept. 1887 (coup. de 1000 f.)	mars 1885	Coup. de 5000 f.	fin cour.					1027 ..	1020 ..	
				d° d° d° (coup. de 10000 f.)								10275 ..		
				d° d° éch. 1er sept. 1888 (coup. de 500 f.)	mars 1885	Coup. de 5000 f. 5035						507 50		
				d° d° d° (coup. de 10000 f.)								10250 ..		
			440	Obligations du Trésor, int. 20 f., r. à 500 fr.	20 janvier 85	507 509 507	en liq…					509 ..	435 ..	20f
			470	d° (annuités finiss. en 1889,) t. p.	16 déc. 1884	509 510 511	fin cour.					510 ..	497 50	20f
			500	d° (annuités finiss. en 1907,) t. p.	15 janv. 1885	523 50 522 523	fin cour.					522 50	480 ..	25f
				Bons de liquid. 5 % (ém. 1874-1875), r. 500 f.										
				EMPRUNTS (Seine et Villes)										
			205	Dép. de la Seine, obl. 1857, 4 %, r. à 225 f. t. p.	janvier 1885.	275 25 278						276 25		9f
			450	Obl. 1855-1860, 3 %, remb. à 500 f. t. p.	mars 1885 ..	510 509 ..	fin cour.					510 ..	380 ..	15f
			345	d° 1865, 4 %, remb. à 500 fr., t. p.	févr. 1885	518 520 522	fin cour.					520 ..	447 50	20f
				d° 1869, 3 %, remb. à 400 fr., t. p.	31 janv. 1885	407 406	fin cour.					405 50	400 ..	12f
			277	d° 1871, 3 %, remb. à 400 fr., t. p.	janvier 1885.	399 398 Samedi, 399.	en liq…					398 75	398 75	12f
			69f75	d° quarts 3 %, remb. à 100 fr.	janvier 1885.	122 50 121 50 121 25	fin cour.					121 50		3f
			…	d° séries sorties (unités)								800 ..		
			…	d° séries sorties (série entière)								10000 ..		
			…	d° 1875, 4 %, remb. à 500 fr., t. p.	15 oct. 1884.	519 50	en liq…					519 ..	526 25	20f
			465	d° 1876, 4 %, remb. à 500 fr., t. p.	15 oct. 1884.	518 50 519 25 519 50	fin cour.							
			500	Bons de liquid. 5 %, r. 500 f., t. p. (*garantis et payables par l'Etat, exempts d'impôt*)	20 oct. 1884.	523 50 530 532	en liq…					519 ..	477 50	20f
							Pr. fin c.							
							fin cour.					532 ..	496 25	25f
			352 50	Ville de Marseille 1877, 3 %, remb. à 400 f., tout payé	31 janvier 85	379 379 50	en liq…					379 ..	383 75	12f
							fin cour.							
							Pr. fin c.		d 5f		d 10			
				VALEURS FRANÇAISES										
			500	La Foncière (Cie d'assurances mobilières et immobilières), act. 500 f., 125 f. p. (nominatives)	23 mai 1877	430	en liq…					430 ..	425 ..	
							Pr. au 15							
							Pr. au 30		d 20		d 10			
			500	La Grande Compagnie d'assurances, act. 500 fr., 125 fr. p. (nominatives)	16 juin 1882		Pr. au 15		d 5f		d 10			
							en liq…					375 ..	720 ..	ex. 81. 5f20
							au 30…							
			500	La Métropole (Cie d'assurances mobilières et immobilières), act. 500 fr., 250 fr. p. (nominatives)			Pr. au 30		d 5f		d 10			
							Pr. au 15		d 5f		d 10			
							en liq…					280 ..	625 ..	
							au 30…							
							Pr. au 15		d 10		d 10			
			Divers	Banque de France (nominatives)	janvier 1885	5200 5225 5240 5245	en liq…					5220 ..	5220 ..	ex. 84, 220f61.
							fin cour.	5240 ..						
							Pr. fin c.		d 50		d 20			
							Pr. fin p.		d 50		d100			
		1 25	Divers	Banque d'Escompte de Paris, a. 500 f., 250 f. p. (ex-coupon 11).	janvier 1885.	580 577 50 578 25	en liq…	578 75	578 75	578 75	577 50	575 ..	577 50	ex. 83, 12f50..
		75					au 31…	575 ..	581 25	575 ..	577 50			à-c. 84, 12f50.
							Pr. au 31		d 5f		d 10			
							Pr. au 15		d 5f		d 10			
		50	500	Banque de Paris et des Pays-Bas, act. 500 f. tout payé (ex-coup. 25)	janvier 1885	800 802 50 800	en liq…	797 50	800 ..	797 50	797 50			
		1 75					au 31…	800 ..	801 25	797 50	798 75	798 75	795 ..	ex. 83. 50f..
							Pr. au 31		d 20		d 10			à-c. 84. 20f..
							Pr. au 15		d 20		d 10			
		15					en liq…							
		Pair	…	Banque Transatlantique, act. 500 f., 250 f. p. (coupon 3 attaché).	4 avril 1884…	433 75 430	au 31…					440 ..	440 ..	ex. 83. 14f..
							Pr. au 15							
							Pr. au 31		d 5f		d 10			
			Divers	Cie Algérienne, act. 500 r., t. p. ex-c. 13)	janvier 1885		en liq…							
							au 30…					425 ..	460 ..	ex. 83. 25f..
							Pr. au 15		d 5f		d 10			à-c. 84. 10f..
							en liq…							
			Pair	Cie Foncière de France, a. 500 fr., 250 f. p. (ex-coupon 4).	janvier 1885.	382 50 380	au 31…					382 50	390 ..	ex. 83. 18f..
							Pr. au 15		d 5f		d 10			à-c. 84. 5f..
							Pr. au 31		d 5f		d 10			
							en liq…							
		1 25	Divers	Comptoir d'Escompte, act. 500 fr., t. p. (ex-coupon 59)	février 1885.	990	au 30…					990 ..	930 ..	ex. 84. 48f..
		75					Pr. au 15		d 5f		d 10			
							en liq…							
		Pair	500	Crédit Algérien, act. de 500 fr., 250 fr. p. (ex-coupon 4).	janvier 1885		au 15…					440 ..	460 ..	ex. 83. 15f..
		75					Pr. au 15		d 5f		d 10			à-c. 84. 7f50
							Pr. au 31				d 20			
			Divers	Crédit de France (anc. Soc. Gén. Française de Crédit), act. 500 fr., tout payé	21 nov. 1881		en liq…							
							au 30…					1 25	50 ..	ex. 80. 50f..
							Pr. au 30		d 20		d 10			à-c. 81. 15f..
							Pr. au 15		d 20		d 10			
							en liq…							
			500	Crédit de Paris, act. 500 fr., 300 fr. payés.	27 mai 1881.		au 30…					201 25	245 ..	
							Pr. au 30		d 10		d 10			
							Pr. au 15				d 5f			
			…	Crédit Foncier et Agricole d'Algérie, act. 500 fr., 125 fr. p. (nominatives)	janvier 1885.	487 50 490	en liq…							
							au 15…					487 50	495 ..	ex. 83. 7f50
							Pr. au 15		d 5f		d 10			à-c. 84. 3f10
							Pr. au 31		d 10		d 10			
			500	Crédit Foncier Colonial, act. 500 f., 300 f. p. (nominatives)	juillet 1884.		au 15…					295 ..	380 ..	ex. 83. 10f309

Pour tout ce qui concerne le *Cours authentique*, abonnements, changements d'adresse, etc., s'adresser à la Chambre syndicale de la Compagnie des Agents de change de Paris, rue Menars, 6.

Figure A.1: Excerpt from the official quotation list of the Paris Bourse, 16 March 1885 (DFIH).

Table A.1: Effect of the cleaning rules

		Returns removed by rule					Mixed series	
Sample	Observations	Round trips	Hampel	Cap	Total removed	Securities touched	Series	Obs.
All (1845–2025)	2,961,725	572	1,680	407	3,723	784	6	875
DFIH (1845–1976)	643,132	355	268	35	1,007	149	2	700
EUROFIDAI (1977–2025)	2,318,593	217	1,412	372	2,716	653	4	175

Rules 1 to 4 of Section 2, applied once to the half-monthly series, by source. Observations: half-monthly returns before cleaning. Round trips: returns larger than 1 in absolute value and immediately reversed; Hampel: prices more than three times above or below the median of their 13 neighbouring quotes; Cap: increases of more than 300% in a half-month; each removal also drops the surrounding returns, so that the total exceeds the sum of the three columns. Securities touched: securities with at least one removed return. Mixed series: identifiers alternating between two price levels, removed entirely, with the number of observations they contained. Source: clean_timeseries.R, cleaning_report.csv.

Table A.2: Predictors: description, frequency, coverage and source

Variable	Description	Freq.	First	Last	Source
metal_total	Metallic reserves of the Banque de France	M	1891	2022	Bazot et al. (2022)
international_portfolio	International portfolio of the Banque de France	M	1891	2022	Bazot et al. (2022)
short-term_loans	Short-term lending of the Banque de France	M	1891	2022	Bazot et al. (2022)
monetary_policy_operations	Monetary-policy operations	M	1891	2022	Bazot et al. (2022)
dep_tot_nongvt	Non-government deposits	M	1891	2022	Bazot et al. (2022)
total_deposit	Total deposits	M	1891	2022	Bazot et al. (2022)
bond_yield	Government bond yield	M	1891	2022	Bazot et al. (2022)
x_uk	Exchange rate on London	M	1891	2022	Bazot et al. (2022)
rn, rb	[describe]	M	1891	2022	Bazot et al. (2022)
total_stock_index, non_bank_stock_index	Stock price indices (excluded from the index analysis)	M	1891	2022	Bazot et al. (2022)
Money_m	Money-market rate	M	1863	2022	Levy-Garboua and Monnet (2016)
tx_long	Long-term interest rate	M	1800	2022	Levy-Garboua and Monnet (2016)
livreta	Regulated savings rate (Livret A)	M	1802	2022	Levy-Garboua and Monnet (2016)
esc_bdf	Discount rate of the Banque de France	M	[1800?]	2022	Levy-Garboua and Monnet (2016)
IR3TIB01FRM156N	3-month interbank rate	M	1970	2024	OECD (FRED)
IRLTLT01FRM156N	10-year government bond yield	M	1960	2024	OECD (FRED)
RBFRBIS	Real broad effective exchange rate	M	1994	2024	BIS (FRED)
SPASTT01FRM661N	Share price index (excluded from the index analysis)	M	1955	2024	OECD (FRED)
DCOILBRENTEU	Brent crude oil price	D	1987	2024	EIA (FRED)
IHLIDXFRTPBAFI	Banking and finance job postings, France	D	2020	2024	Indeed (FRED)
NASDAQNQFR10T	Nasdaq France Technology index (excluded from the index analysis)	D	2001	2024	Nasdaq (FRED)
VIXCLS	VIX	D	1990	2024	CBOE (FRED)
dp_lag	Security's dividend–price ratio, lagged one period	HM	1862	2025	DFIH, EUROFIDAI
R_m	Equal-weighted market return (excluded from the index analysis)	HM	1845	2025	DFIH, EUROFIDAI

M: monthly; D: daily; HM: half-monthly. Monthly and daily series are brought to the half-monthly calendar by the nowcasting procedure of Appendix A.2 and enter the regressions in log differences. [Complete the descriptions of rn and rb and the first observation of esc_bdf.]

A.1 Corporate-event adjustments in DFIH

See Internet Appendix

A.2 Nowcasting of low-frequency predictors

Let v_m be a predictor observed once per month (or at a lower frequency) and z_t the equal-weighted stock index, observed on the half-monthly calendar. For each month m, a bridge equation

$$v_m = a + b_0\, z_m + \sum_{\ell=1}^{L} b_\ell\, z_{m-\ell} + u_m \tag{14}$$



Figure A.2: Liquidity of the Paris market, 1977–2025, annual medians across securities: Corwin–Schultz high–low spread (%), share of securities not traded on the sampling day (%), and turnover (% of shares outstanding per half-month). Source: eurofidai_extra.R, figure FE1.

Table A.3: EUROFIDAI trading data, 1977–2025: coverage and liquidity

			Coverage (% of obs.)				Median spread (%)		Not traded	Median
Year	Securities	Obs.	Mkt cap.	Volume	Quoted	High–low	Quoted	High–low	at t (%)	turnover (%)
1977	337	7,521	99	74	0	72	–	0.00	11	0.15
1980	392	8,976	98	72	0	70	–	0.00	8	0.20
1985	430	10,125	99	73	0	69	–	0.00	8	0.41
1990	1,319	27,575	77	95	0	93	–	0.11	7	0.27
1995	1,264	23,688	98	99	0	98	–	0.00	0	0.14
2000	1,723	33,175	97	86	0	79	–	0.07	12	0.22
2005	1,797	37,414	98	67	8	57	0.00	0.18	24	0.19
2010	3,332	65,881	95	61	6	55	0.00	0.14	22	0.05
2015	3,891	78,943	98	64	0	56	0.00	0.27	21	0.05
2020	5,403	108,214	99	62	0	54	0.00	0.26	22	0.02
2025	5,138	5,138	99	50	0	44	0.00	0.17	20	0.01

Selected years. Securities and half-monthly observations in the year; share of observations with a market capitalisation, a volume, a quoted bid–ask spread and a high–low (Corwin–Schultz) spread; median quoted and high–low spreads in percent of the price; share of securities not traded on the day their price is sampled; median turnover in percent of shares outstanding per half-month. The high–low estimator is floored at zero. Source: eurofidai_extra.R, table TE0. [The median quoted spread is zero in every year in which bid and ask are recorded, which suggests that the two fields are often equal (or both equal to the close): check their definition in the EUROFIDAI documentation before using Table J.39.]

is estimated on the months for which both v and the monthly aggregate of z are observed, using only months before the forecast date. Within the current month, the nowcast replaces the full-month aggregate z_m by the partial aggregate computed with the half-months observed so far. Published values are assigned to the first calendar date at which they were available, and nowcasts are produced only for months in which v_m is eventually observed. The procedure yields a half-monthly series that updates as the index moves and that never uses information posterior to the forecast date.

A.3 Sample description

The learning period starts with the first quotations in 1845, but forecasts are evaluated only from 1870, when the market is large enough for the pooled quantities to be well estimated and when the macro-financial predictors cover at least the interest-rate block. For each security, the first forecast is made after 100 half-monthly observations (about four years) and the first 12 months of forecasts are treated as a burn-in of the aggregation rule and excluded from all evaluations. Table A.4 describes the evaluation sample: 7,382 securities and 1.55 million evaluated forecasts, 1,545 securities and 443,000 forecasts in 1870–1976, 6,034 securities and 1.10 million forecasts in 1977–2024; 197 securities have evaluated forecasts in both sources; the median security has 137 evaluated half-months. Figure A.3 shows the number of securities per year, from 10 in 1870 to 50 in 1890, 500 in the late 1930s, 700 in 1970, 1,000 in 2005 and 2,000 to 2,500 since 2014.

The most visible feature of the sample is the break at the junction of the two sources: the number of securities with evaluated forecasts falls from 590 in 1976 to 37 in 1977 and only returns to its previous level

Table A.4: Evaluation sample

Sample	Securities	OOS obs.	First year	Last year	Median obs./security
DFIH (1870–1976)	1545	443353	1870	1976	187
EUROFIDAI (1977–2025)	6034	1104415	1977	2024	109
Total	7382	1547768	1870	2024	137
Matched across sources	197	NA	NA	NA	NA

Out-of-sample observations are half-monthly forecasts of security returns evaluated from 1870, after at least 100 observations of learning per security and a 12-month burn-in. "Matched across sources" is the number of securities with evaluated forecasts in both DFIH and EUROFIDAI. Source: paper_analysis.R, table T1.

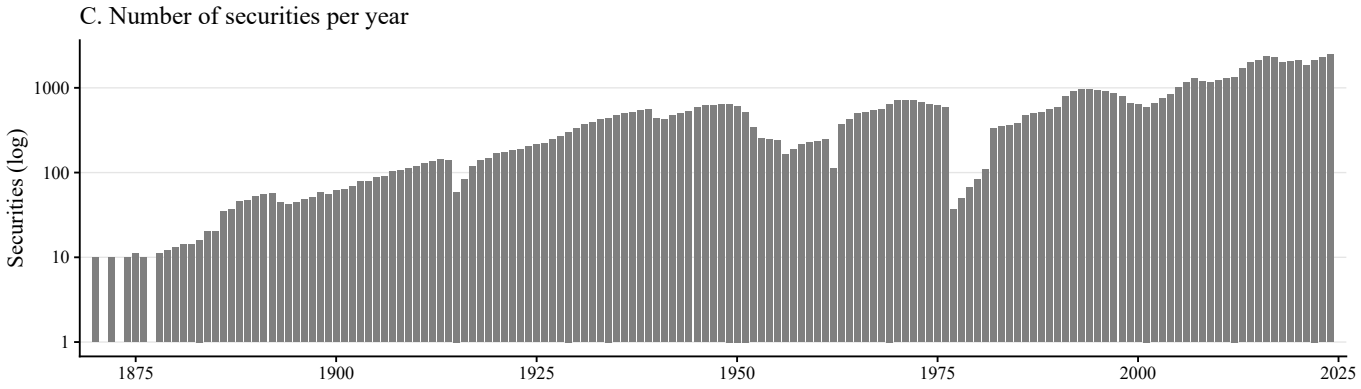


Figure A.3: Number of securities with evaluated forecasts, by year (log scale).

in 1982 (Figure A.3). DFIH is exhaustive, recording every security on the official quotation lists, whereas EUROFIDAI does not cover the whole of the Paris market in its early years; more importantly, the securities that could not be matched across the two sources enter the sample in 1977 without any history, and are evaluated only after the 100 half-months of learning and the 12-month burn-in required of every security, that is, from about 1982.

B Algorithms

Algorithm 1 The sequential prediction game

Parameters: experts $j = 1, \dots, N$, prior weights $\pi \in \mathcal{P}$.
for $t = t_0, \dots, T$ **do**
    Each expert outputs $f_{i,j,t}$ from information dated t or earlier.
    The forecaster outputs $\hat{r}_{i,t+1} = \sum_j w_{j,t} f_{i,j,t}$.
    The return $r_{i,t+1}$ is revealed; the forecaster incurs $\ell(\hat{r}_{i,t+1}, r_{i,t+1})$.
    The weights are updated to w_{t+1}.
end for

Algorithm 2 Bernstein Online Aggregation with pooled weights

Initialise $\widetilde{R}_j = V_j = B_j = 0$, $\eta_j = 1$, $\pi_j = w_j = 1/N$ for $j = 1, \dots, N$.
for each date t **do**
  For each security i quoted at t: $\hat{r}_{i,t+1} = \sum_j w_j f_{i,j,t}$ (an unavailable expert is replaced by the security's historical mean).
  Let I_t be the securities whose return $r_{i,t+1}$ is known by the next calendar date.
  For each expert j: $\rho_j = \dfrac{1}{|I_t|} \displaystyle\sum_{i \in I_t} 2\,(\hat{r}_{i,t+1} - r_{i,t+1})(\hat{r}_{i,t+1} - f_{i,j,t})$.
  $B_j \leftarrow \max(B_j, |\rho_j|)$, $\quad V_j \leftarrow V_j + \rho_j^2$, $\quad E_j = 2^{\lceil \log_2 B_j \rceil}$.
  $\eta_j = \min\{1/E_j,\ \sqrt{\log(1/\pi_j)/V_j}\}$ (unchanged if $V_j = 0$).
  $\widetilde{R}_j \leftarrow \widetilde{R}_j + \frac{1}{2}\big(\rho_j - \eta_j \rho_j^2 + E_j \mathbf{1}\{\eta_j \rho_j > 1/2\}\big)$.
  $w_j \propto \eta_j\, \pi_j \exp(\eta_j \widetilde{R}_j)$, normalised to sum to one.
end for

For the market index (Section 4) there is a single series, $I_t = \{m\}$, and the algorithm reduces to the BOA of Wintenberger (2017) with the gradient trick, as implemented in `opera` (Gaillard and Goude, 2016); with several securities, the recursion is the same as in `opera`, applied to the cross-sectional average of the instantaneous regrets.

B.1 The regret bound of the pooled aggregation

Let I_t be the set of securities whose return $r_{i,t+1}$ is known by the next calendar date, $f_{i,t} = (f_{i,1,t}, \dots, f_{i,N,t})'$ the vector of expert forecasts for security i, and define the *cross-sectional average loss* of a weight vector $q \in \mathcal{P}$ at date t,

$$L_t(q) = \frac{1}{|I_t|} \sum_{i \in I_t} \big(q' f_{i,t} - r_{i,t+1}\big)^2. \tag{15}$$

Assumption. Returns and expert forecasts are bounded: $|r_{i,t+1}| \le M$ and $|f_{i,j,t}| \le M$ for all i, j, t.[21]

Proposition. Let w_t be the weights produced by Algorithm 2, with the adaptive learning rates of Wintenberger (2017). For every $q \in \mathcal{P}$,

$$\sum_{t=1}^{T} L_t(w_t) - \sum_{t=1}^{T} L_t(q) \;\le\; \max_{1 \le j \le N} \sum_{t=1}^{T} \rho_{j,t} \;\le\; C \max_{1 \le j \le N} \Big(\sqrt{V_{j,T}\, \log N} + B_T \log N \Big) + C\, B_T \log\log(1 + V_{j,T}/B_T^2), \tag{16}$$

where $\rho_{j,t}$ is the averaged instantaneous regret of Section 3.4, $V_{j,T} = \sum_{t \le T} \rho_{j,t}^2$, $B_T = \max_{j,t} |\rho_{j,t}| \le 8M^2$, and C is a universal constant. In particular, since $V_{j,T} \le 64 M^4 T$, the average regret satisfies

[21] In the data, expert forecasts are clipped to the security's historical mean plus or minus three standard deviations, and returns are bounded after the cleaning rules of Section 2; M is finite in any finite sample. No other assumption is made on the sequence of returns, which may be non-stationary or chosen by an adversary.

$\limsup_{T\to\infty} \frac{1}{T}\big(\sum_t L_t(w_t) - \min_{q\in\mathcal{P}} \sum_t L_t(q)\big) \le 0.$

Proof. *Step 1 (convexity).* L_t is a convex function of q, as an average of squares of affine functions. Its gradient at w_t is

$$\nabla L_t(w_t) = \frac{1}{|I_t|} \sum_{i\in I_t} 2\,(\hat{r}_{i,t+1} - r_{i,t+1})\, f_{i,t}, \qquad \hat{r}_{i,t+1} = w_t' f_{i,t}. \tag{17}$$

By convexity, for every $q \in \mathcal{P}$,

$$L_t(w_t) - L_t(q) \le \nabla L_t(w_t)'(w_t - q). \tag{18}$$

Step 2 (linearisation). Since $\sum_j q_j = 1$, $\nabla L_t(w_t)'(w_t - q) = \sum_j q_j\, \nabla L_t(w_t)'(w_t - e_j)$, where e_j is the j-th unit vector, and

$$\nabla L_t(w_t)'(w_t - e_j) = \frac{1}{|I_t|} \sum_{i\in I_t} 2\,(\hat{r}_{i,t+1} - r_{i,t+1})(\hat{r}_{i,t+1} - f_{i,j,t}) = \rho_{j,t}, \tag{19}$$

which is exactly the quantity used to update the weights. Summing over t,

$$\sum_t \big(L_t(w_t) - L_t(q)\big) \le \sum_j q_j \sum_t \rho_{j,t} \le \max_j \sum_t \rho_{j,t}, \tag{20}$$

which is the first inequality in (16).

Step 3 (BOA on linear losses). Define the linear loss of expert j at date t as $\ell_{j,t} = \nabla L_t(w_t)' e_j$ and that of the forecaster as $\hat{\ell}_t = \nabla L_t(w_t)' w_t$, so that $\rho_{j,t} = \hat{\ell}_t - \ell_{j,t}$. Algorithm 2 is BOA with adaptive learning rates applied to the sequence of linear losses $(\ell_{1,t}, \dots, \ell_{N,t})_{t\ge 1}$: the weights at t depend only on $\rho_{j,s}$, $s < t$. These losses are bounded by $4M^2$ in absolute value, so the instantaneous regrets are bounded by $8M^2$. The second-order regret bound of Wintenberger (2017) for BOA with adaptive learning rates holds for any bounded sequence of losses, whatever the process that generates it; it gives the second inequality in (16). The fact that each linear loss is itself a cross-sectional average plays no role in that bound, which only uses its boundedness. □

Remarks. Three features of the guarantee should be kept in mind. (i) The benchmark is the best *common* convex combination of experts, the same for all securities, chosen in hindsight; the bound says nothing about a security-specific combination. (ii) The loss that is bounded weights every date equally, whatever the number of securities quoted at that date; the pooled R^2_{OOS} of Section 3.5 instead weights dates by their number of securities, whereas the Newey–West test on monthly cross-sectional averages is closer to the quantity that

the bound controls. (iii) Returns that are not known by the next calendar date, because the security was not quoted, do not enter the update and are therefore not covered by the bound, although they are included in the evaluation.

C Tests

C.1 Pooled test on the cross-sectional loss differential

Let $d_{i,t} = (r_{i,t+1} - b_{i,t})^2 - (r_{i,t+1} - \hat{r}_{i,t+1})^2$ be the loss differential of security i at date t against benchmark b. For each month m, $\bar{d}_m$ is the average of $d_{i,t}$ over all securities and the two half-months of the month. The null of equal accuracy is $\mathbb{E}[\bar{d}_m] = 0$ against $\mathbb{E}[\bar{d}_m] > 0$; the statistic is the mean of $\bar{d}_m$ divided by its Newey–West standard error with 12 lags (Newey and West, 1987; Gu et al., 2020).

C.2 Diebold–Mariano and Clark–West tests by security

For a security, with one-step-ahead non-overlapping forecasts, the Diebold–Mariano statistic is the mean of $d_{i,t}$ divided by its Newey–West standard error with $\lfloor n^{1/3} \rfloor$ lags, asymptotically standard normal under the null (Diebold and Mariano, 1995). When the model nests the benchmark, as the aggregate nests the historical mean, the squared error of the larger model is biased upward under the null by the estimation noise of the additional parameters; Clark and West (2007) propose the adjusted differential $d^{CW}_{i,t} = d_{i,t} + (b_{i,t} - \hat{r}_{i,t+1})^2$, tested in the same way. With several thousand securities, the share of rejections at 5% is reported together with the share after the Benjamini and Hochberg (1995) procedure, which controls the false-discovery rate.

D Market-structure variables

The five variables used in Table I.34 are annual, 1870–2025, and coded as follows; an event dated before 30 June counts from that year, an event dated after 30 June from the following year. *Transaction tax*: 1 when a tax on securities transactions is in force, the *impôt sur les opérations de bourse* created by the law of 28 April 1893 and repealed by article 11 of the 2008 finance law, and the *taxe sur les transactions financières* in force since 1 August 2012 (0.2% until 2016, 0.3% from 2017, 0.4% from April 2025, on firms with a capitalisation above one billion euros). *Forward market*: fraction of the year during which the forward market (*marché à terme*) of the Paris Bourse was open; it closed on 27 July 1914 and reopened in January 1920, closed on 10 May 1940 and reopened in 1949 (coded 0.5 for 1949, the exact date being unknown); after 2000 the deferred-settlement service is treated as its continuation. *Dual market*: 1 while the official market of the

agents de change and the *Coulisse* coexist, the latter being unofficial until the law of 14 February 1942 and official until the merger under the law of 29 July 1961. *Capital controls*: 1 from 1940 to 1989; exchange controls were lifted on 1 January 1990 [confirm the start date, probably September 1939]. *Electronic trading*: 1 from 1986, when the CAC continuous quotation system was introduced and the MATIF opened; trading was fully electronic by 1989. Sources: the chronology of French financial markets of the Service des archives économiques et financières (Ministry of the Economy); the *Grande Encyclopédie Larousse*; Hautcœur and Riva (2012); Hautcœur et al. (2010); and the legal texts cited. A fuller set of indicators (the 1898 reform of the *agents de change*, the creation of the COB in 1967, the end of the brokers' monopoly in 1988, Euronext in 2000, MiFID in 2007 and 2018) is available but too collinear with the post-1977 indicator to be used jointly.

E Additional tables and figures

Table E.5: Ex-post best expert across securities

Best expert ex post	*N*	Share
bivar_R_m	2299	0.311
zero	1065	0.144
char	649	0.088
bivar_NASDAQNQFR10T_logdiff	538	0.073
diffusion	478	0.065
bivar_VIXCLS_logdiff	449	0.061
cs_mean	360	0.049
bivar_bond_yield_logdiff	203	0.027
bivar_rb_logdiff	188	0.025
bivar_non_bank_stock_index_logdiff	118	0.016
combination	102	0.014
bivar_SPASTT01FRM661N_logdiff	99	0.013
bivar_IHLIDXFRTPBAFI_logdiff	98	0.013
bivar_IRLTLT01FRM156N_logdiff	91	0.012
bivar_esc_bdf_logdiff	82	0.011

Expert with the lowest root mean squared error for the security, among experts available for at least half of the security's forecasts; fifteen most frequent. bivar_R_m is the regression on the lagged market return; the other bivariate experts are regressions on the log difference of the named predictor. Source: paper_analysis.R, table T7a.

F Is the Paris Bourse predictable? Evidence from the index (Appendix)

F.1 The stale-price artefact

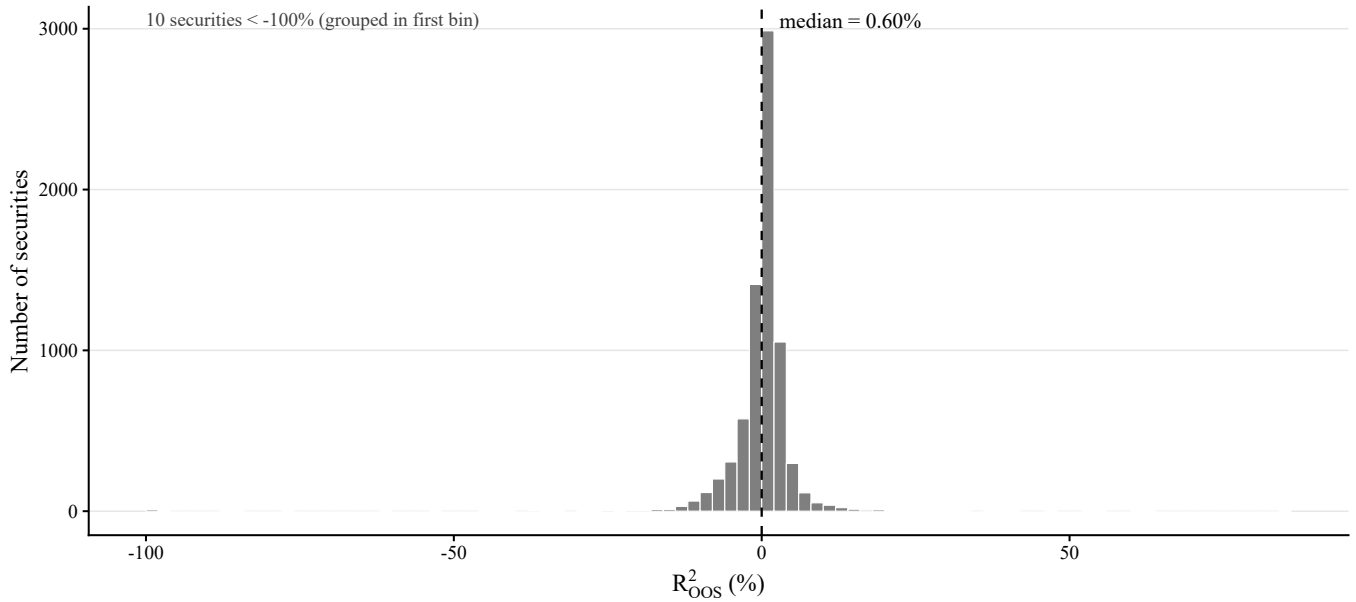


Figure E.4: Distribution of the security-level R^2_{OOS} against the historical mean (%), all securities; values below -100% are grouped in the first bin.

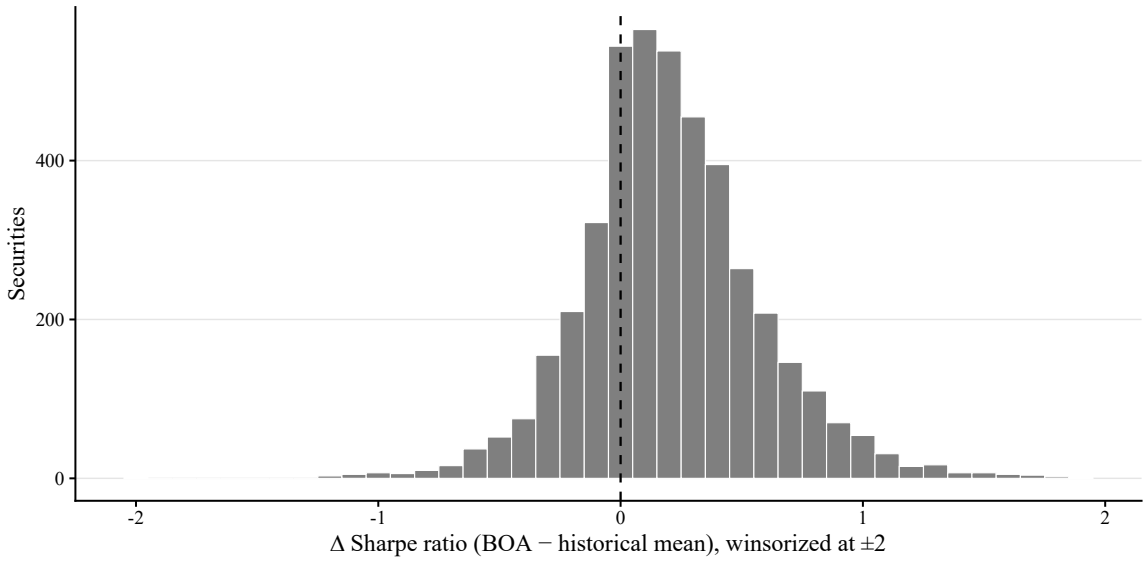


Figure E.5: Distribution of the security-level difference in annualised Sharpe ratio, BOA timing minus historical-mean timing (winsorised at ±2).

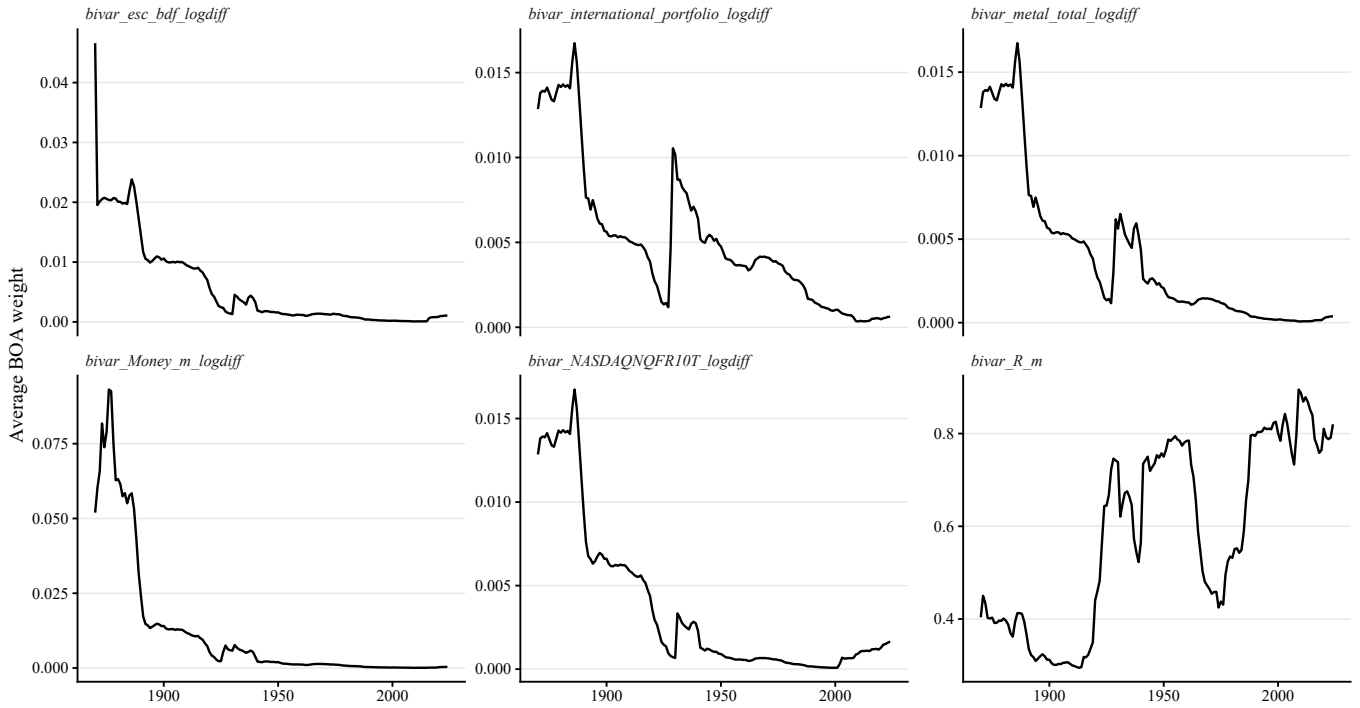


Figure E.6: Average BOA weight of the six bivariate predictors with the largest cumulative weight, by year (note the different vertical scales).

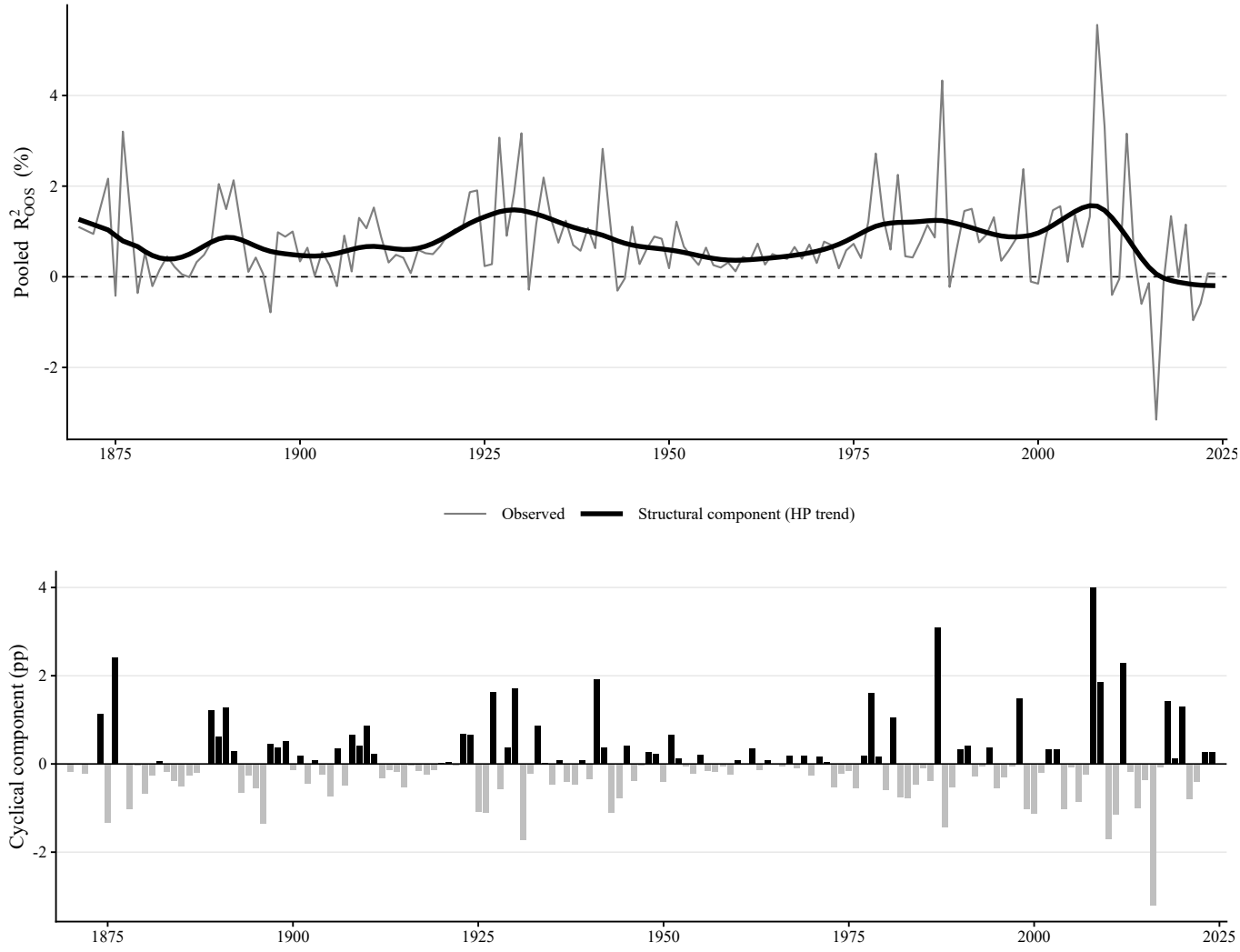


Figure E.7: Annual pooled R^2_{OOS} against the historical mean: observed series and Hodrick–Prescott trend (top), cyclical component (bottom).

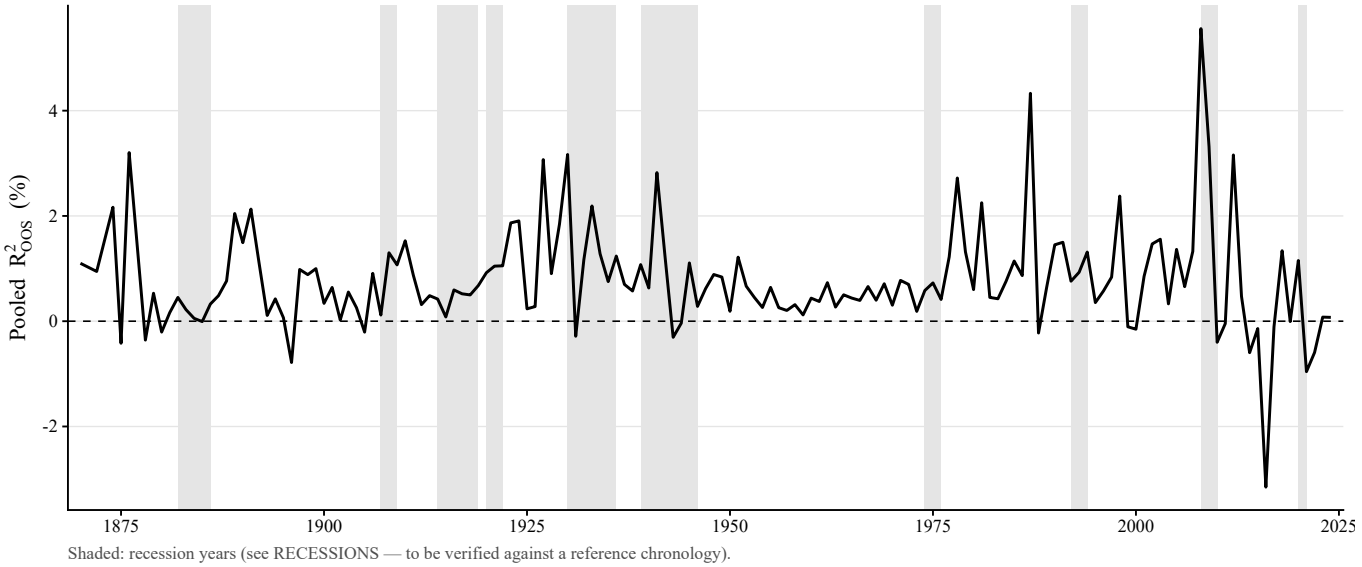


Figure E.8: Annual pooled R^2_{OOS} against the historical mean; shaded areas are recession years (provisional chronology).

Table F.6: Autocorrelation of the half-monthly equal-weighted index return

Sample	Obs.	ac1	ac2	ac3	r2 ar1
All	3694	0.319	0.193	0.042	10.192
DFIH (1870–1976)	2550	0.299	0.211	0.026	8.967
EUROFIDAI (1977–2025)	1144	0.342	0.151	0.052	11.698

Autocorrelations of the half-monthly log return on the equal-weighted market index at lags 1 to 3, by source, 1870–2025. $R^2_{AR(1)} = \rho_1^2$ is the in-sample R^2 of a first-order autoregression, that is, the predictability generated by stale prices alone. Source: index_forecast.R, table TI0.

Table F.7: Market index with all predictors, including the lagged index: the stale-price artefact

Rule	Sample	Obs.	R^2_{OOS} vs. HM	R^2_{OOS} vs. 0	DM t	DM p	CW t	CW p
BOA	All	3693	9.59	10.35	5.78	0	8.90	0
BOA	DFIH (1870–1976)	2550	8.33	8.51	4.22	0	6.46	0
BOA	EUROFIDAI (1977–2025)	1143	11.63	13.27	4.06	0	6.50	0
BOA (CT)	All	3693	7.59	8.36	7.10	0	8.68	0
BOA (CT)	DFIH (1870–1976)	2550	5.46	5.64	4.57	0	5.70	0
BOA (CT)	EUROFIDAI (1977–2025)	1143	11.02	12.68	6.04	0	7.53	0
MLpol	All	3693	10.07	10.82	6.39	0	8.70	0
MLpol	DFIH (1870–1976)	2550	8.62	8.79	5.18	0	6.46	0
MLpol	EUROFIDAI (1977–2025)	1143	12.40	14.03	4.09	0	6.30	0
Fixed Share	All	3693	9.96	10.71	6.48	0	8.86	0
Fixed Share	DFIH (1870–1976)	2550	8.78	8.95	5.23	0	6.91	0
Fixed Share	EUROFIDAI (1977–2025)	1143	11.85	13.49	4.11	0	5.79	0
Equal weights	All	3693	1.96	2.78	6.05	0	6.52	0
Equal weights	DFIH (1870–1976)	2550	1.71	1.90	3.85	0	4.05	0
Equal weights	EUROFIDAI (1977–2025)	1143	2.37	4.18	5.52	0	6.36	0

As Table 1, with all predictors, including the lagged equal-weighted index return and the price-based indices. Source: index_forecast.R with SPEC = “all”, table TI1.

F.2 Out-of-sample accuracy

Table F.8: Market index: individual experts and the value of learning the weights

A. Macro-financial predictors		B. All predictors	
Expert	R^2_{OOS}	Expert	R^2_{OOS}
VIX	0.35	Lagged equal-weighted index	9.95
3-month interbank rate	0.32	OECD share price index	0.88
Total deposits	0.19	Nasdaq France technology index	0.70
Livret A rate	0.19	Non-bank stock index	0.63
Monetary-policy operations	0.18	VIX	0.35
Experts with $R^2_{\text{OOS}} > 0$	12	Experts with $R^2_{\text{OOS}} > 0$	17
Equal-weighted average	0.44	Equal-weighted average	1.96
BOA	0.48	BOA	9.59
DM t, BOA vs. equal weights	0.14	DM t, BOA vs. equal weights	5.39

Out-of-sample R^2 (%) against the historical mean, 1870–2025, of the five best experts taken alone; an expert is replaced by the historical mean before it becomes available, so that its R^2 reflects only the years in which it exists. Equal-weighted average: uniform combination of all experts with fixed weights. DM t: Diebold–Mariano statistic of the squared-error differential between the equal-weighted average and BOA (positive when BOA is more accurate), Newey–West with 24 lags. Source: index_forecast.R, table TI1b, specifications “macro” and “all”.

Table F.9: Market index: economic value for a mean-variance investor

Rule	Sample	CER gain (% p.a.)	Sharpe	Sharpe (HM)	Avg. equity weight
BOA	All	1.88	0.66	0.51	0.72
BOA	DFIH (1870–1976)	0.70	0.33	0.24	0.58
BOA	EUROFIDAI (1977–2025)	4.55	1.12	0.90	1.03
BOA (CT)	All	1.38	0.59	0.51	0.87
BOA (CT)	DFIH (1870–1976)	0.54	0.30	0.24	0.76
BOA (CT)	EUROFIDAI (1977–2025)	3.30	0.99	0.90	1.12
MLpol	All	1.61	0.62	0.51	0.79
MLpol	DFIH (1870–1976)	0.52	0.30	0.24	0.66
MLpol	EUROFIDAI (1977–2025)	4.10	1.06	0.90	1.10
Fixed Share	All	7.45	1.16	0.51	0.67
Fixed Share	DFIH (1870–1976)	5.55	0.86	0.24	0.60
Fixed Share	EUROFIDAI (1977–2025)	11.77	1.67	0.90	0.82
Equal weights	All	1.26	0.60	0.51	0.77
Equal weights	DFIH (1870–1976)	0.63	0.32	0.24	0.64
Equal weights	EUROFIDAI (1977–2025)	2.70	1.00	0.90	1.05

The investor allocates $w_t = \hat{r}_{t+1}/(\gamma\hat{\sigma}_t^2)$ to the index, $\gamma = 3$, $\hat{\sigma}_t^2$ the return variance of the past five years, $w_t \in [0, 1.5]$; the rest is held in cash with a zero return. CER gain: annualised certainty-equivalent return of the strategy minus that of the historical-mean allocation, in percent per year. Sharpe ratios are annualised. Source: index_forecast.R, table TI2.

Table F.10: Market index: out-of-sample R^2 by market state

Rule	Market state	Obs.	R^2_{OOS} vs. HM	R^2_{OOS} vs. 0	DM t	DM p	CW t	CW p
BOA	High volatility (real time)	697	1.31	1.98	1.54	0.06	2.56	0.01
BOA	Normal volatility	2996	0.00	0.92	-0.01	0.50	1.63	0.05
BOA	Crisis	756	1.01	0.51	1.23	0.11	1.73	0.04
BOA	No crisis	2937	0.06	1.91	0.10	0.46	2.61	0.00
BOA (CT)	High volatility (real time)	697	0.60	1.27	1.85	0.03	2.20	0.01
BOA (CT)	Normal volatility	2996	0.38	1.31	1.62	0.05	2.34	0.01
BOA (CT)	Crisis	756	-0.21	-0.72	-0.84	0.80	-0.56	0.71
BOA (CT)	No crisis	2937	0.98	2.81	3.85	0.00	4.54	0.00
MLpol	High volatility (real time)	697	1.22	1.89	2.47	0.01	2.70	0.00
MLpol	Normal volatility	2996	0.16	1.09	0.46	0.32	1.72	0.04
MLpol	Crisis	756	0.03	-0.48	0.09	0.47	0.50	0.31
MLpol	No crisis	2937	0.95	2.78	2.73	0.00	3.80	0.00
Fixed Share	High volatility (real time)	697	1.59	2.25	0.56	0.29	2.53	0.01
Fixed Share	Normal volatility	2996	1.43	2.35	1.06	0.14	5.28	0.00
Fixed Share	Crisis	756	4.16	3.67	2.86	0.00	2.97	0.00
Fixed Share	No crisis	2937	-0.56	1.30	-0.24	0.59	4.19	0.00
Equal weights	High volatility (real time)	697	0.82	1.49	2.24	0.01	2.41	0.01
Equal weights	Normal volatility	2996	0.22	1.14	1.04	0.15	1.75	0.04
Equal weights	Crisis	756	0.23	-0.28	0.72	0.24	0.87	0.19
Equal weights	No crisis	2937	0.60	2.43	2.65	0.00	3.43	0.00

High volatility: the variance of index returns over the past six months exceeds the 80th percentile of its history up to that date (known in real time). Crisis: the half-month belongs to one of the episodes listed in the text. Statistics as in Table 1. Source: index_forecast.R, table TI3.

F.3 Economic value and market states

G Predictability of individual securities

G.1 Can the aggregate beat the historical mean?

Table G.11: Distribution of security-level out-of-sample R^2

Sample	N	Median	Q1	Q3	Share > 0
DFIH (1870–1976)	1505	0.80	0.20	1.39	0.81
EUROFIDAI (1977–2025)	5143	0.51	-1.31	1.91	0.59
All	7382	0.60	-0.99	1.78	0.62

Security-level R^2_{OOS} (%) of the aggregate forecast against the security's historical mean: median, quartiles and share of securities with a positive value, by period. Source: paper_analysis.R, table T2b.

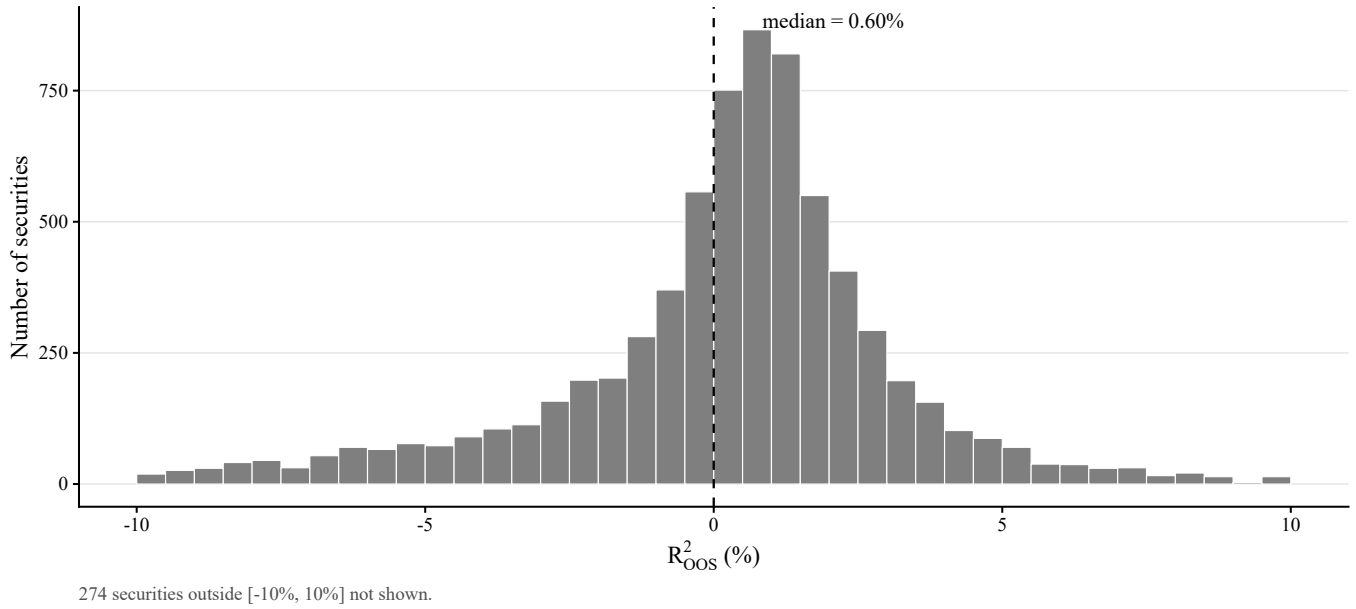


Figure G.9: Distribution of the security-level out-of-sample R^2 against the historical mean (%), securities within $[-10, 10]$%; bins of 0.5 point.

Table G.12: Share of securities with significant predictability

Test	Securities	Signif. 5% (raw)	Signif. 5% (BH)	Share (raw)	Share (BH)
Diebold–Mariano, BOA vs historical mean	6513	1403	247	0.215	0.038
Clark–West (nested), BOA vs historical mean	6513	2258	1134	0.347	0.174
Diebold–Mariano, historical mean vs BOA (reverse)	6513	432	1	0.066	0.000

Number and share of securities for which the test rejects at 5%, with raw p-values and with the Benjamini–Hochberg false-discovery-rate correction, among securities with at least 30 forecasts. Diebold–Mariano: one-sided test that the aggregate has a lower squared error than the historical mean; Clark–West: the same with the correction for nested models; reverse: one-sided test that the historical mean is more accurate. Newey–West standard errors with $n^{1/3}$ lags. Source: paper_analysis.R, table T3.

Table G.13: Significance tests by period

Period	Securities	Signif. 5% (raw)	Signif. 5% (BH)	Share (raw)	Share (BH)	Median R^2_{OOS} (%)
DFIH (1870–1976)	1505	539	188	0.358	0.125	0.801
EUROFIDAI (1977–2025)	5143	878	107	0.171	0.021	0.515

Diebold–Mariano test, aggregate against historical mean, computed within each period for securities with at least 30 forecasts in the period; median security-level R^2_{OOS} (%). Source: paper_analysis.R, table T4.

G.2 Which experts predict?

Table G.14: Bivariate predictors with the largest average BOA weight, by period

Period	Expert	Avg. BOA weight
DFIH (1870–1976)	bivar_R_m	0.598
DFIH (1870–1976)	bivar_international_portfolio_logdiff	0.005
DFIH (1870–1976)	bivar_Money_m_logdiff	0.004
DFIH (1870–1976)	bivar_metal_total_logdiff	0.003
DFIH (1870–1976)	bivar_esc_bdf_logdiff	0.003
EUROFIDAI (1977–2025)	bivar_R_m	0.798
EUROFIDAI (1977–2025)	bivar_NASDAQNQFR10T_logdiff	0.001
EUROFIDAI (1977–2025)	bivar_international_portfolio_logdiff	0.001
EUROFIDAI (1977–2025)	bivar_esc_bdf_logdiff	0.000
EUROFIDAI (1977–2025)	bivar_tx_long_logdiff	0.000

Top five predictors of the bivariate-regression family by average weight. R_m is the lagged return of the equal-weighted market; the other predictors are in log differences. Source: paper_analysis.R, table T7c.

Table G.15: Timing portfolios by period

Period	Strategy	Return (% p.a.)	Volatility (% p.a.)	Sharpe	Max. drawdown (%)	Utility (%)	Months	Utility gain vs. B&H (%)
DFIH (1870–1976)	Buy & hold	2.46	13.27	0.19	79.98	-0.18	1192	0.00
DFIH (1870–1976)	BOA timing	4.32	8.74	0.49	34.91	3.17	1192	3.36
DFIH (1870–1976)	Hist.-mean timing (recomputed)	0.74	7.94	0.09	63.42	-0.21	1192	-0.02
DFIH (1870–1976)	Hist.-mean timing	0.19	8.42	0.02	69.82	-0.87	1192	-0.69
EUROFIDAI (1977–2025)	Buy & hold	-0.21	15.96	-0.01	78.41	-4.03	576	0.00
EUROFIDAI (1977–2025)	BOA timing	7.16	8.25	0.87	19.13	6.14	576	10.17
EUROFIDAI (1977–2025)	Hist.-mean timing (recomputed)	0.68	8.67	0.08	56.07	-0.45	576	3.58
EUROFIDAI (1977–2025)	Hist.-mean timing	0.86	8.17	0.11	53.54	-0.14	576	3.89
All	Buy & hold	1.59	14.20	0.11	79.98	-1.44	1768	0.00
All	BOA timing	5.24	8.59	0.61	34.91	4.14	1768	5.57
All	Hist.-mean timing (recomputed)	0.72	8.18	0.09	66.64	-0.29	1768	1.15
All	Hist.-mean timing	0.41	8.34	0.05	71.12	-0.63	1768	0.80

Equal-weighted monthly portfolios. BOA timing: each security is held when its aggregate forecast is positive and in cash otherwise; hist.-mean timing: the same rule with the historical mean as forecast ("recomputed": with the historical mean recomputed after the burn-in); buy and hold: all securities. Annualised return and volatility (%), Sharpe ratio, maximum drawdown (%), and the utility of a mean-variance investor with $\gamma = 3$ (% per year) and its gain over buy and hold. Source: paper_analysis.R, table T5.

Table G.16: Security-level Sharpe ratio differences

Sample	N	Median	Share > 0	Wilcoxon p
DFIH (1870–1976)	1258	0.224	0.838	0
EUROFIDAI (1977–2025)	3157	0.152	0.695	0
All	4337	0.181	0.735	0
All, vs buy & hold	5899	0.106	0.639	0

Difference between the annualised Sharpe ratio of the BOA timing portfolio of each security and that of its historical-mean timing portfolio (and, in the last row, of buy and hold), for securities with at least 24 months: median, share of positive differences and one-sided Wilcoxon p-value. Source: paper_analysis.R, table T6.

G.3 Economic value

G.4 One hundred and fifty-five years of predictability

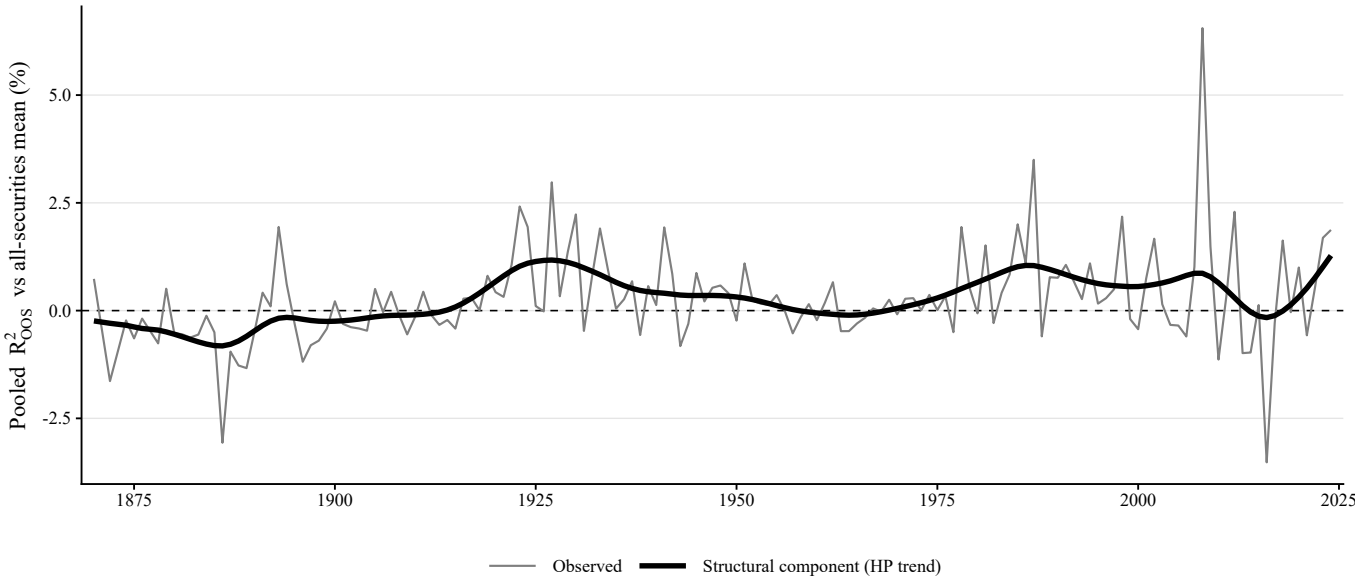


Figure G.10: Annual pooled R^2_{OOS} against the all-securities mean (grey) and its structural component, a Hodrick–Prescott trend with $\lambda = 100$ (black).

Table G.17: Structural breaks in annual predictability

Series	Sample	Break year	95% CI low	95% CI high	Nearest event
R2	1870–2024	NA	NA	NA	no break (BIC)
P	1870–2024	NA	NA	NA	no break (BIC)
P10	1872–2024	NA	NA	NA	no break (BIC)

Bai–Perron breaks in the mean of the annual series: pooled R^2_{OOS} against the historical mean (R2), share of securities beaten in the year (P) and over the trailing ten years (P10). Number of breaks selected by BIC, at most five, 15% trimming. Source: paper_analysis.R, table T8.

Table G.18: Predictability in recession and expansion years

Measure	Recession	Expansion	N rec.	N exp.	Wilcoxon p
R2	1.031	0.713	33	119	0.256
P	61.571	59.510	33	119	0.305
P10	73.235	69.627	33	118	0.267
cycle	0.161	-0.045	33	119	0.671

Means of the annual measures over recession and expansion years and two-sided Wilcoxon p-value. R2: pooled R^2_{OOS} against the historical mean; P and P10: share of securities beaten in the year and over ten years; cycle: cyclical component of R2. Source: paper_analysis.R, table T9b.

Table G.19: Predictability around institutional events and crises

Year	Event	Measure	Before	After	Difference	Permutation p
1882	Union Générale crash	R2	0.032	0.215	0.183	0.722
1914	WWI – Bourse closure	R2	0.854	0.474	-0.380	0.494
1929	1929 crash	R2	1.279	1.504	0.225	0.676
1945	Bank nationalizations	R2	0.868	0.563	-0.304	0.588
1967	COB created	R2	0.467	0.578	0.111	0.812
1986	Financial deregulation	R2	1.006	1.539	0.533	0.358
1988	Insider-trading law	R2	1.505	1.056	-0.448	0.449
2007	MiFID / algorithmic trading	R2	1.075	2.317	1.242	0.076
2008	Global financial crisis	R2	1.047	1.300	0.253	0.641
1882	Union Générale crash	P	59.676	65.833	6.157	0.408
1914	WWI – Bourse closure	P	53.523	58.637	5.114	0.525
1929	1929 crash	P	65.918	67.255	1.337	0.839
1945	Bank nationalizations	P	55.957	60.790	4.833	0.594
1967	COB created	P	60.817	62.803	1.986	0.810
1986	Financial deregulation	P	64.995	67.050	2.055	0.805
1988	Insider-trading law	P	67.594	68.392	0.798	0.919
2007	MiFID / algorithmic trading	P	58.822	63.860	5.038	0.549
2008	Global financial crisis	P	61.879	57.564	-4.315	0.659

Mean of the annual measure over the five years before and the five years after the event, their difference, and the permutation p-value of the difference (2,000 random event years). Source: paper_analysis.R, table T10.

Table G.20: Determinants of annual predictability

Dependent variable	Variable	Estimate	NW s.e.	p	
Pooled R²OOS	(Intercept)	0.799	0.400	0.048	**
Pooled R²OOS	recessionTRUE	0.111	0.226	0.623	
Pooled R²OOS	mkt_vol	0.046	0.016	0.004	***
Pooled R²OOS	log_n	-0.108	0.074	0.149	
Share BOA > RW	(Intercept)	65.024	5.419	0.000	***
Share BOA > RW	recessionTRUE	-0.526	2.442	0.830	
Share BOA > RW	mkt_vol	0.573	0.173	0.001	***
Share BOA > RW	log_n	-2.139	1.048	0.043	**

OLS regressions of the annual pooled R^2_{OOS} and of the annual share of securities for which the aggregate beats the historical mean on a recession indicator, the annualised volatility of the equal-weighted market (%), and the log number of securities. Newey–West standard errors with two lags. ***, **, *: significant at 1, 5 and 10%. Source: paper_analysis.R, table T11.

H The cross-section: characteristics and momentum

H.1 Real-time Fama–MacBeth regressions

Table H.21: Fama–MacBeth regressions of relative returns on characteristics

Sample	mom b	mom t	strev b	strev t	ltrev b	ltrev t	vol b	vol t	beta b	beta t	zero b	zero t	dp b	dp t	nsec b	nsec t	cross sections
All	0.28	9.26	0.03	0.87	-0.01	-0.56	-0.06	-1.60	0.00	-0.09	0.11	3.99	0.06	1.24	0.01	0.33	3221
DFIH (1870–1976)	0.20	5.33	0.04	0.69	-0.05	-1.42	0.02	0.38	0.04	1.01	0.06	1.94	0.06	0.93	0.02	0.56	2070
EUROFIDAI (1977–2025)	0.42	9.66	0.02	0.71	0.05	1.32	-0.19	-3.14	-0.09	-2.17	0.20	3.89	0.04	1.12	-0.01	-0.20	1151

At each half-month, the next relative return $r_{i,t+1} - \bar{r}_{t+1}$ (winsorised at the 1st and 99th percentiles of the date) is regressed on the cross-sectional ranks in $[-1, 1]$ of the characteristics known at t: momentum (return from $t-23$ to $t-2$), short-term reversal (return at $t-1$), long-term reversal (return from $t-71$ to $t-24$), 12-month volatility, 24-month beta on the equal-weighted index, share of zero returns over 12 months, dividend–price ratio lagged one period, and the number of listed securities of the firm. Average slope in percent per half-month and Newey–West t-statistic with 24 lags. Source: char_experts.R, table TC1.

Table H.22: Out-of-sample R^2 of real-time relative-return forecasts

Sample	Securities	Obs.	R^2_{OOS} relative (%)	Correlation
All	14839	2703738	0.08	0.03
DFIH (1870–1976)	1879	465947	0.01	0.01
EUROFIDAI (1977–2025)	13485	2237791	0.09	0.04

The forecast at T uses the average Fama–MacBeth slopes over all dates before T, after at least five years of cross-sections. R^2_{OOS} against zero, in percent, and correlation between forecast and realised relative return. Source: char_experts.R, table TC2.

H.2 Long–short portfolios

Table H.23: Long–short decile portfolios on the characteristics forecast

Holding (half-months)	Sample	Periods	stocks	Turnover (%)	Gross (% p.a.)	gross vol ann	Sharpe gross	t_{NW} gross	Net (% p.a.)	net vol ann	Sharpe net	t_{NW} net
0.5 month(s)	All	2672	903.93	42.66	10.43	16.43	0.63	6.78	-10.05	16.63	-0.60	-5.15
0.5 month(s)	DFIH (1870–1976)	1521	263.51	57.37	8.26	15.36	0.54	4.38	-19.28	15.48	-1.25	-8.86
0.5 month(s)	EUROFIDAI (1977–2025)	1151	1750.23	23.21	13.29	17.73	0.75	5.33	2.15	17.74	0.12	0.84
6 month(s)	All	2172	1048.52	7.78	10.49	15.93	0.66	5.94	6.75	16.03	0.42	3.79
6 month(s)	DFIH (1870–1976)	1026	282.42	10.07	9.95	15.93	0.62	3.71	5.12	16.13	0.32	1.90
6 month(s)	EUROFIDAI (1977–2025)	1146	1734.40	5.73	10.97	15.93	0.69	4.72	8.22	15.93	0.52	3.52

Top minus bottom decile of the real-time characteristics forecast, equal-weighted, bottom decile of price excluded, held one half-month or six months. Annualised mean return (%), volatility (%), Sharpe ratio and Newey–West t-statistic, gross and net of a 0.5% cost per transaction. Turnover: average share of each leg replaced per half-month. Source: char_experts.R, table TC3.

H.3 One characteristic at a time

Table H.24: Long–short decile portfolios sorted on one characteristic

Characteristic	Sample	Periods	Turnover (%)	Gross (% p.a.)	Sharpe gross	t_{NW} gross	Net (% p.a.)	Sharpe net	t_{NW} net
mom	All	2098	7.61	10.77	0.63	5.54	7.12	0.41	3.62
mom	DFIH (1870–1976)	951	9.86	11.61	0.65	3.86	6.88	0.38	2.23
mom	EUROFIDAI (1977–2025)	1147	5.74	10.07	0.61	4.02	7.31	0.44	2.91
strev	All	2120	9.45	6.35	0.40	3.67	1.81	0.11	1.06
strev	DFIH (1870–1976)	978	11.88	8.27	0.53	3.29	2.57	0.16	1.04
strev	EUROFIDAI (1977–2025)	1142	7.38	4.71	0.29	1.99	1.17	0.07	0.49
ltrev	All	2083	5.97	-5.76	-0.40	-3.39	-8.63	-0.59	-5.01
ltrev	DFIH (1870–1976)	939	8.00	-7.57	-0.47	-3.06	-11.42	-0.70	-4.52
ltrev	EUROFIDAI (1977–2025)	1144	4.30	-4.28	-0.33	-1.85	-6.34	-0.49	-2.75
vol	All	2122	5.85	7.85	0.40	3.02	5.04	0.25	1.96
vol	DFIH (1870–1976)	981	7.58	16.61	0.76	4.38	12.98	0.59	3.45
vol	EUROFIDAI (1977–2025)	1141	4.36	0.31	0.02	0.10	-1.78	-0.10	-0.56
beta	All	2089	5.41	5.22	0.25	2.02	2.62	0.12	1.03
beta	DFIH (1870–1976)	952	7.36	14.25	0.69	3.74	10.72	0.52	2.85
beta	EUROFIDAI (1977–2025)	1137	3.79	-2.34	-0.11	-0.75	-4.16	-0.19	-1.33
zero	All	2199	4.37	2.91	0.32	2.81	0.81	0.09	0.76
zero	DFIH (1870–1976)	1067	6.71	1.55	0.15	1.08	-1.67	-0.17	-1.15
zero	EUROFIDAI (1977–2025)	1132	2.15	4.19	0.51	2.88	3.16	0.38	2.18
dp	All	1302	2.80	1.82	0.13	0.82	0.48	0.03	0.21
dp	DFIH (1870–1976)	260	3.35	-3.16	-0.20	-0.85	-4.76	-0.30	-1.30
dp	EUROFIDAI (1977–2025)	1042	2.66	3.06	0.24	1.18	1.78	0.14	0.69
nsec	All	2295	2.51	-0.87	-0.07	-0.72	-2.07	-0.17	-1.74
nsec	DFIH (1870–1976)	1149	3.63	0.79	0.07	0.46	-0.95	-0.08	-0.55
nsec	EUROFIDAI (1977–2025)	1146	1.39	-2.53	-0.20	-1.55	-3.20	-0.25	-1.97

For each characteristic, top minus bottom decile, equal-weighted, bottom decile of price excluded, six-month holding. Annualised mean return (%), Sharpe ratio and Newey–West t-statistic, gross and net of a 0.5% cost per transaction. Turnover: average share of each leg replaced per half-month. Source: char_experts.R, table TC4.

Table H.25: Sensitivity of net long–short returns to transaction costs

Strategy	Sample	Cost (%)	Net (% p.a.)	t_{NW} net
Combined model	All	0.25	8.62	4.87
Combined model	All	0.50	6.75	3.79
Combined model	All	1.00	3.02	1.67
Combined model	All	2.00	-4.45	-2.31
Combined model	DFIH (1870–1976)	0.25	7.53	2.81
Combined model	DFIH (1870–1976)	0.50	5.12	1.90
Combined model	DFIH (1870–1976)	1.00	0.28	0.10
Combined model	DFIH (1870–1976)	2.00	-9.39	-3.23
Combined model	EUROFIDAI (1977–2025)	0.25	9.59	4.12
Combined model	EUROFIDAI (1977–2025)	0.50	8.22	3.52
Combined model	EUROFIDAI (1977–2025)	1.00	5.47	2.33
Combined model	EUROFIDAI (1977–2025)	2.00	-0.03	-0.01
Momentum	All	0.25	8.94	4.58
Momentum	All	0.50	7.12	3.62
Momentum	All	1.00	3.46	1.73
Momentum	All	2.00	-3.84	-1.81
Momentum	DFIH (1870–1976)	0.25	9.24	3.04
Momentum	DFIH (1870–1976)	0.50	6.88	2.23
Momentum	DFIH (1870–1976)	1.00	2.14	0.68
Momentum	DFIH (1870–1976)	2.00	-7.32	-2.15
Momentum	EUROFIDAI (1977–2025)	0.25	8.69	3.46
Momentum	EUROFIDAI (1977–2025)	0.50	7.31	2.91
Momentum	EUROFIDAI (1977–2025)	1.00	4.56	1.80
Momentum	EUROFIDAI (1977–2025)	2.00	-0.95	-0.37

Net annualised return (%) and Newey–West t-statistic of the momentum and combined-model long–short portfolios, six-month holding, for alternative costs per transaction; the cost charged is four times the cost per transaction times the average turnover of the two legs. Source: char_experts.R, table TC5.

Table H.26: Minimum coverage of each leg

Strategy	Min. coverage	Sample	Periods	Gross (% p.a.)	t_{NW} gross	Net (% p.a.)	t_{NW} net
Combined model	0.33	All	2232	10.46	6.03	6.82	3.91
Combined model	0.33	DFIH (1870–1976)	1086	9.92	3.85	5.35	2.06
Combined model	0.33	EUROFIDAI (1977–2025)	1146	10.97	4.72	8.22	3.52
Combined model	0.50	All	2172	10.49	5.94	6.75	3.79
Combined model	0.50	DFIH (1870–1976)	1026	9.95	3.71	5.12	1.90
Combined model	0.50	EUROFIDAI (1977–2025)	1146	10.97	4.72	8.22	3.52
Momentum	0.33	All	2174	10.19	5.41	6.67	3.51
Momentum	0.33	DFIH (1870–1976)	1026	10.29	3.64	5.91	2.06
Momentum	0.33	EUROFIDAI (1977–2025)	1148	10.10	4.04	7.34	2.93
Momentum	0.50	All	2098	10.77	5.54	7.12	3.62
Momentum	0.50	DFIH (1870–1976)	951	11.61	3.86	6.88	2.23
Momentum	0.50	EUROFIDAI (1977–2025)	1147	10.07	4.02	7.31	2.91

A date is used only if at least the indicated share of the securities of each leg has an observed return: one half (baseline) or one third. Six-month holding. Source: char_experts.R, table TC6.

Table H.27: Size-neutral sorts

Characteristic	Sample	Periods	Gross (% p.a.)	Sharpe gross	t_{NW} gross	Net (% p.a.)	t_{NW} net
beta	All	2138	2.12	0.13	1.18	-0.06	-0.03
beta	DFIH (1870–1976)	998	5.53	0.37	2.25	2.60	1.07
beta	EUROFIDAI (1977–2025)	1140	-0.86	-0.05	-0.35	-2.38	-0.95
ltrev	All	2111	-3.44	-0.34	-2.74	-5.89	-4.66
ltrev	DFIH (1870–1976)	977	-3.46	-0.31	-1.84	-6.62	-3.49
ltrev	EUROFIDAI (1977–2025)	1134	-3.42	-0.36	-2.03	-5.25	-3.12
mom	All	2180	8.17	0.62	5.49	5.10	3.39
mom	DFIH (1870–1976)	1034	8.05	0.61	3.70	4.26	1.93
mom	EUROFIDAI (1977–2025)	1146	8.28	0.64	4.07	5.87	2.87
vol	All	2145	3.67	0.25	1.90	1.19	0.62
vol	DFIH (1870–1976)	1001	5.75	0.40	2.20	2.49	0.95
vol	EUROFIDAI (1977–2025)	1144	1.85	0.13	0.67	0.05	0.02
zero	All	2193	2.35	0.41	3.73	0.56	0.89
zero	DFIH (1870–1976)	1049	2.98	0.42	2.93	0.14	0.14
zero	EUROFIDAI (1977–2025)	1144	1.76	0.42	2.39	0.94	1.29

Top minus bottom quintile of the characteristic within each size group, six-month holding. Size groups: market-capitalisation terciles when capitalisation is available at the date; otherwise the number of listed securities of the firm (1, 2, 3 or more). Source: char_experts.R, table TC7.

Table H.28: Long–short portfolios with market capitalisations and own spreads, 1977–2025

			Gross			Net, 0.5% per trade		Net, own half-spreads		
Strategy	Holding	Periods	Mean	Sharpe	t_{NW}	Mean	t_{NW}	Mean	t_{NW}	Cost
Momentum, equal-weighted	6m	1149	11.0	0.68	4.45	8.2	3.32	10.4	4.22	0.59
Momentum, equal-weighted	1m	1151	12.9	0.71	5.22	5.4	2.15	11.4	4.62	1.52
Momentum, value-weighted	6m	1137	1.7	0.06	0.44	-1.5	-0.39	0.9	0.24	0.73
Momentum, size-neutral (cap. terciles)	6m	1137	7.4	0.57	3.87	5.0	2.57	6.9	3.59	0.52
Momentum, excluding smallest size quintile	6m	1136	6.0	0.34	2.12	3.2	1.13	5.4	1.88	0.66
Size (small minus big)	1m	1149	9.1	0.52	3.08	6.2	2.11	8.7	2.94	0.42
Illiquidity, Amihud (high minus low)	1m	887	5.5	0.41	2.57	-6.6	-2.95	3.5	1.61	2.00
Illiquidity, zero returns (high minus low)	1m	680	9.5	0.85	3.77	5.3	2.10	9.3	3.73	0.14
Volatility (high minus low)	1m	1151	3.3	0.18	1.04	-1.4	-0.45	2.5	0.78	0.80
Volatility, size-neutral	1m	1148	2.0	0.12	0.77	-2.3	-0.90	1.2	0.46	0.77
Beta (high minus low)	1m	1150	-4.5	-0.20	-1.34	-9.1	-2.70	-5.2	-1.56	0.77

Decile long–short portfolios on EUROFIDAI data, 1977–2025, with the holding period indicated (six months: rebalanced every twelve half-months; one month: every two). Equal-weighted unless indicated; value-weighted: weights proportional to market capitalisation at formation; size-neutral: top minus bottom quintile within each capitalisation tercile; excluding smallest quintile: securities in the lowest capitalisation quintile at formation are dropped. Size: small minus big capitalisation; Amihud: average absolute daily return per million euros traded; zero returns: share of days without a price change. Annualised mean (%), Sharpe ratio and Newey–West t-statistic, gross; net of a uniform 0.5% cost per transaction; and net of each security's own half-spread (quoted when available, Corwin–Schultz otherwise) charged on the turnover of each rebalancing, with the resulting average annual cost. Source: eurofidai_extra.R, table TE4.

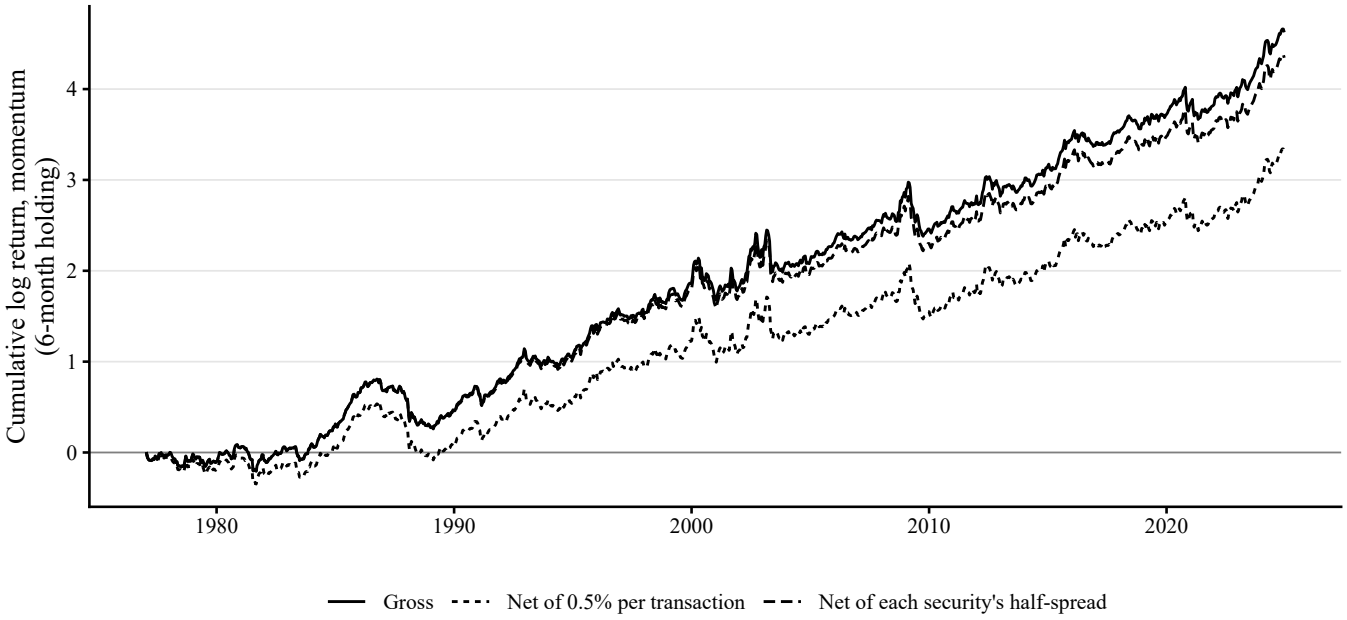


Figure H.11: Equal-weighted momentum, 1977–2025, six-month holding: cumulative log return gross, net of a uniform 0.5% cost per transaction, and net of each security's own half-spread. Source: eurofidai_extra.R, figure FE3.

Table I.29: French equity factors, 1870–2025

Factor	Period	Months	Mean (% p.a.)	Vol. (% p.a.)	Sharpe	t_{NW}
MKT_EW	All	1852	7.85	15.71	0.50	5.47
MKT_EW	DFIH (1870–1976)	1276	7.76	15.25	0.51	4.82
MKT_EW	EUROFIDAI (1977–2025)	576	8.05	16.68	0.48	2.72
BETA	All	1285	4.53	23.43	0.19	1.94
BETA	DFIH (1870–1976)	709	11.57	22.75	0.51	3.97
BETA	EUROFIDAI (1977–2025)	576	-4.14	24.03	-0.17	-1.24
ILLIQ	All	1316	3.31	9.63	0.34	3.34
ILLIQ	DFIH (1870–1976)	740	1.32	9.57	0.14	1.16
ILLIQ	EUROFIDAI (1977–2025)	576	5.87	9.67	0.61	3.55
LTREV	All	1276	-4.25	15.00	-0.28	-2.46
LTREV	DFIH (1870–1976)	701	-5.60	16.23	-0.35	-2.16
LTREV	EUROFIDAI (1977–2025)	575	-2.60	13.35	-0.19	-1.21
MOM	All	1313	10.74	18.12	0.59	6.30
MOM	DFIH (1870–1976)	737	8.77	16.89	0.52	3.81
MOM	EUROFIDAI (1977–2025)	576	13.26	19.56	0.68	5.30
VOL	All	1309	10.94	21.17	0.52	4.59
VOL	DFIH (1870–1976)	733	16.71	21.78	0.77	5.15
VOL	EUROFIDAI (1977–2025)	576	3.61	20.18	0.18	1.13

Monthly returns of equal-weighted decile long–short portfolios rebalanced monthly (two half-months of holding), bottom decile of price excluded: MOM (momentum), LTREV (long-term reversal), VOL (volatility), BETA, ILLIQ (share of zero returns); MKT_EW is the equal-weighted market. Annualised mean (%), volatility (%), Sharpe ratio and Newey–West t-statistic (12 lags). Source: history_factors.R, table TH1.

Table I.30: Equal- and value-weighted market returns, 1977–2025

Factor	Periods	Mean (% p.a.)	Vol. (% p.a.)	Sharpe	t_{NW}
MKT_EW	1151	7.76	14.75	0.53	2.61
MKT_VW	1150	6.41	21.47	0.30	1.92

Half-monthly simple returns of all EUROFIDAI securities with a return at the date, equal-weighted and weighted by market capitalisation at the previous date; annualised mean and volatility (%), Sharpe ratio and Newey–West t-statistic (12 lags). Source: eurofidai_extra.R, table TE5.

Table I.31: Momentum crashes: the Daniel–Moskowitz regression

Variable	Estimate	t_{NW}	p
Intercept	0.97	6.24	0.00
Bear(t-1)	0.14	0.41	0.68
Bear(t-1) x MKT(t)	-56.49	-3.31	0.00
MKT(t)	1.36	0.23	0.82

Monthly momentum return (%) regressed on a bear-market indicator (24-month market return below zero, known at $t-1$), its interaction with the contemporaneous equal-weighted market return, and the market return. Coefficients are multiplied by 100; the interaction and market coefficients are therefore in percentage points of momentum return per unit (100%) of market return. Newey–West t-statistics with 12 lags. Source: history_factors.R, table TH2.

Table I.32: The fifteen worst momentum months

Month	MOM (%)	Market (%)	Bear	24-month market (%)
2003-05-01	-29.4	6.5	1	-34.1
2020-11-01	-29.1	24.5	1	11.1
2009-04-01	-24.4	16.5	1	-47.5
2002-10-01	-23.6	11.6	1	-48.1
1988-02-01	-23.3	18.7	0	15.3
2002-11-01	-22.2	10.9	1	-38.9
2001-10-01	-19.1	7.6	1	-20.2
1957-09-01	-18.2	-8.2	0	-9.6
1931-12-01	-14.9	15.6	1	-41.1
2001-11-01	-14.6	6.0	1	-20.3
1935-03-01	-14.5	3.0	1	-4.7
2000-03-01	-13.8	-2.1	0	25.0
1931-01-01	-13.6	16.0	1	-17.8
1932-02-01	-13.1	12.0	1	-36.9
1920-05-01	-12.9	-12.3	0	27.4

Monthly momentum return, equal-weighted market return and 24-month cumulative market return, in percent; bear = 1 if the 24-month market return at $t-1$ is negative. Source: history_factors.R, table TH3.

Table I.33: Momentum in France, the United States and Europe

Period	Comparison	Months	Correlation	France (% p.a.)	Other (% p.a.)	t_{NW} (diff.)
All	MOM_US	1067	0.32	10.30	7.28	1.49
DFIH (1870–1976)	MOM_US	491	0.14	6.83	7.69	-0.26
EUROFIDAI (1977–2025)	MOM_US	576	0.49	13.26	6.93	2.59
All	MOM_EU	410	0.70	14.69	10.15	1.96
EUROFIDAI (1977–2025)	MOM_EU	410	0.70	14.69	10.15	1.96

Monthly momentum returns: France (this paper), United States (Kenneth French data library, from 1927) and Europe (from 1990). Correlation, annualised means (%) and Newey–West t-statistic of the mean difference, on common months. Source: history_factors.R, table TH4.

H.4 Robustness

I A long-run view: factors, crashes, other markets and mechanisms

I.1 A French factor dataset, 1870–2025

I.2 France, the United States and Europe

I.3 Mechanisms: predictability and market structure

Table I.34: Annual predictability and market structure

Dependent variable	Variable	Estimate	t_{NW}	p	N
mom	(Intercept)	0.028	0.084	0.933	127
mom	log_n	0.031	0.802	0.424	127
mom	illiq	-2.348	-1.583	0.116	127
mom	disp	0.408	0.600	0.549	127
mom	mkt_vol_l	-0.467	-2.523	0.013	127
mom	mkt_ret_l	0.069	0.607	0.545	127
mom	I(year >= 1977)TRUE	0.082	1.089	0.278	127
mom	stt_in_force	0.079	1.089	0.279	127
mom	forward_market	-0.085	-1.248	0.215	127
mom	dual_market	0.021	0.361	0.719	127
mom	capital_controls	-0.074	-1.750	0.083	127
mom	electronic_trading	-0.116	-1.234	0.220	127
r2_cs	(Intercept)	1.396	0.831	0.408	151
r2_cs	log_n	-0.079	-0.392	0.696	151
r2_cs	illiq	-16.274	-1.936	0.055	151
r2_cs	disp	5.049	1.044	0.298	151
r2_cs	mkt_vol_l	-0.500	-0.328	0.743	151
r2_cs	mkt_ret_l	-0.011	-0.020	0.984	151
r2_cs	I(year >= 1977)TRUE	0.401	1.392	0.166	151
r2_cs	stt_in_force	-0.488	-0.674	0.501	151
r2_cs	forward_market	0.102	0.475	0.635	151
r2_cs	dual_market	0.162	0.815	0.416	151
r2_cs	capital_controls	0.011	0.046	0.963	151
r2_cs	electronic_trading	-0.369	-0.852	0.396	151

Annual regressions, 1870–2025. Dependent variables: annualised momentum return (decimal) and pooled R^2_{OOS} of the aggregate forecast against the all-securities mean (%). Regressors: log number of listed securities, average share of zero returns, cross-sectional standard deviation of returns, lagged annualised market volatility and return (decimals), a post-1977 indicator, and the market-structure variables described in the text and in Appendix D. Newey–West t-statistics with two lags. Source: history_factors.R, table TH5.

J Robustness: the lagged market return and stale prices

Section 5.2 shows that the aggregate gives most of its weight to the regression of a security's return on the lagged market return, and that part of this lead-lag effect could be the stale-price mechanism of Section 2.7 seen from the side of the security: a security that is not traded on a quotation date carries its last price, and its next recorded return mechanically "catches up" with the market. Four exercises bound the contribution of this mechanism; the last two use the trading data of Section 2.6 and cover 1977–2025 only.

(i) Securities quoted on consecutive dates. The evaluation is restricted to forecasts whose return is realised on the next calendar date (the 15th for a forecast made on the 1st, the 1st of the next month for a forecast made on the 15th). A security that is not quoted on the next date has a return that spans several periods and is realised only when the next quote appears; these are the observations for which the stale-price mechanism is strongest. The restriction keeps 1.40 of the 1.55 million forecasts (91%; 331,000 of 443,000 in the historical period, 1.07 of 1.10 million in the modern one) and 7,338 of the 7,382 securities. Table J.35 reports the pooled R^2_{OOS} on this sample. Against the historical mean it is 0.77% ($t = 6.2$) instead of 0.69%; against zero 0.72% ($t = 3.3$) instead of 0.63%; against the all-securities mean 0.71% ($t = 3.4$) instead of 0.62%, with 0.47% ($t = 1.7$) in the historical period instead of 0.29% and 0.77% ($t = 3.1$) in the modern one instead of 0.75%. Removing the observations most exposed to stale pricing raises the measured predictability, because those observations are the ones for which every forecast, including the lead-lag expert, is least accurate. The weights of the aggregation are unchanged (the bivariate family, almost entirely the lead-lag expert, receives 66% and 80% of the weight in the two periods, the characteristics expert 14% and 7%), and the timing portfolio of Section 5 keeps a Sharpe ratio of 0.56 against 0.11 for buy and hold. The lead-lag effect is a property of traded prices.

(ii) Without the lead-lag expert. The aggregation is re-run with the regression on the lagged market return removed from the set of experts, all other experts and the evaluation being unchanged.[22] Table J.37 reports the result, which is the decomposition of Table 2. Without the lead-lag expert, the aggregate still beats the historical mean, by 0.18% ($t = 7.0$; 0.41% in the historical period, 0.09% in the modern one), but this is shrinkage: against zero and against the all-securities mean the gain is 0.11% and not significant over the full sample ($t = -0.2$), −0.01% in the historical period ($t = -1.7$) and 0.15% in the modern period ($t = 3.2$). The weights move to the characteristics expert (70% and 69% in the two periods), the zero forecast (12% and

[22] The experts are not re-estimated: the lead-lag expert is dropped from the stored forecasts before the pooled aggregation. The combination and diffusion experts, available from 2008, still use the lagged market return among their inputs; their weight is below 0.3%.

13%) and the all-securities mean (11% in both); the forty macro-financial regressions together receive 6 to 7%. The timing portfolio has a Sharpe ratio of 0.21 against 0.11 for buy and hold, and the share of securities with a significant Clark–West test after the Benjamini–Hochberg correction falls from 17.4% to 5.9%. The lead-lag expert therefore accounts for about four fifths of the information content of the security-level forecasts, 0.51 of the 0.62 percentage points against the all-securities mean; the characteristics expert accounts for the rest, and only in the modern period. This is consistent with Section 6: at a half-month horizon the characteristics explain 0.08% of the variance of relative returns, and their value is realised through six-month long–short portfolios, not through half-monthly forecasts.

Table J.35: Pooled out-of-sample R^2 on consecutive-date forecasts

Sample	Securities	Obs.	R^2_{OOS} vs. HM	t_{NW}	p	R^2_{OOS} vs. 0	t_{NW}	p	R^2_{OOS} vs. CS mean	t_{NW}	p	R^2_{OOS} (1/var. weighted)	R^2_{OOS} (trimmed)
DFIH (1870–1976)	1535	330536	0.91	5.98	0	0.44	1.58	0.06	0.47	1.67	0.05	0.97	0.93
EUROFIDAI (1977–2025)	5981	1072738	0.74	4.01	0	0.79	3.00	0.00	0.77	3.12	0.00	0.51	0.83
All	7338	1403274	0.77	6.25	0	0.72	3.26	0.00	0.71	3.40	0.00	0.62	0.85

As Table 2, evaluation restricted to forecasts whose return is realised on the next calendar date. Source: paper_analysis.R with CONSECUTIVE_ONLY = TRUE, table T2.

Table J.36: Timing portfolios on consecutive-date forecasts

Period	Strategy	Return (% p.a.)	Volatility (% p.a.)	Sharpe	Max. drawdown (%)	Utility (%)	Months	Utility gain vs. B&H (%)
DFIH (1870–1976)	Buy & hold	2.38	13.75	0.17	74.16	-0.45	1089	0.00
DFIH (1870–1976)	BOA timing	4.07	9.64	0.42	62.36	2.68	1089	3.13
DFIH (1870–1976)	Hist.-mean timing (recomputed)	1.11	8.16	0.14	54.84	0.11	1089	0.56
DFIH (1870–1976)	Hist.-mean timing	0.41	9.40	0.04	59.31	-0.92	1089	-0.47
EUROFIDAI (1977–2025)	Buy & hold	0.05	16.20	0.00	76.69	-3.89	576	0.00
EUROFIDAI (1977–2025)	BOA timing	7.33	8.38	0.87	18.86	6.28	576	10.17
EUROFIDAI (1977–2025)	Hist.-mean timing (recomputed)	0.78	8.75	0.09	54.34	-0.37	576	3.52
EUROFIDAI (1977–2025)	Hist.-mean timing	0.89	8.25	0.11	53.24	-0.14	576	3.75
All	Buy & hold	1.57	14.64	0.11	76.69	-1.64	1665	0.00
All	BOA timing	5.20	9.23	0.56	62.36	3.92	1665	5.56
All	Hist.-mean timing (recomputed)	0.99	8.37	0.12	54.84	-0.06	1665	1.59
All	Hist.-mean timing	0.57	9.02	0.06	59.31	-0.65	1665	0.99

As Table G.15, on the consecutive-date sample. Source: paper_analysis.R with CONSECUTIVE_ONLY = TRUE, table T5.

Table J.37: Pooled out-of-sample R^2 without the lead-lag expert

Sample	Securities	Obs.	R^2_{OOS} vs. HM	t_{NW}	p	R^2_{OOS} vs. 0	t_{NW}	p	R^2_{OOS} vs. CS mean	t_{NW}	p	R^2_{OOS} (1/var. weighted)	R^2_{OOS} (trimmed)
DFIH (1870–1976)	1545	443353	0.41	7.80	0	-0.04	-1.67	0.95	-0.01	-1.75	0.96	0.39	0.41
EUROFIDAI (1977–2025)	6034	1104415	0.09	2.58	0	0.17	2.13	0.02	0.15	3.15	0.00	0.53	0.33
All	7382	1547768	0.18	6.97	0	0.11	-0.24	0.60	0.11	-0.18	0.57	0.49	0.35

As Table 2, with the regression on the lagged market return removed from the experts before the pooled aggregation. Source: pooled_boa.R with DROP_EXPERTS = “bivar_R_m”, then paper_analysis.R, table T2.

(iii) Trading on the sampling day. For 1977–2025, the daily volumes say whether the sampled price of a security is a transaction price. Table J.38 splits the modern-period forecasts by this indicator, by tercile of turnover, of Amihud illiquidity and of quoted spread, all formed in real time at each date, and Figure J.12 plots the R^2_{OOS} against the all-securities mean of the aggregate and of the lead-lag expert alone. Volumes are observed for 813,000 of the 1.10 million modern-period forecasts. Among the 670,000 whose price at t is a transaction price, the aggregate attains 0.89% ($t = 3.6$) and the lead-lag expert 0.89%; among the 604,000 traded on both sampling dates, 0.90% ($t = 3.4$) and 0.90%; among the 143,000 whose sampled price is a

carried-forward quote, 0.13% (t = 0.6) and −0.11%. The predictability is not in the stale quotes; it is in the prices set by trades, and it is larger there than in the full sample. By turnover, the aggregate attains 0.06% (t = 1.1) in the lowest tercile, where the lead-lag expert is at −0.14%, 0.86% (t = 4.4) in the middle tercile and 1.16% (t = 3.5) in the highest, where the lead-lag expert reaches 1.24%. By the Amihud ratio, which requires the value of volume and covers 357,000 forecasts, it is 0.60% (t = 0.4) in the most liquid tercile, 1.11% (t = 2.2) in the middle one and 0.69% (t = 2.4) in the least liquid. Together with the capitalisation quintiles of Table 5, the picture is consistent: the lagged market return predicts the returns of small securities that trade actively, and neither those of the largest securities, which react first, nor those of securities that hardly trade, whose recorded prices do not move. The quoted-spread terciles cover 18,000 forecasts and are reported for completeness only.

(iv) Mid-quote returns. Where bid and ask prices are recorded on two consecutive sampling dates, the realised return can be computed from the mid-quotes instead of the closing prices, which removes the bid–ask bounce. Table J.39 compares the two evaluations on the same forecasts. The sample is small, 369 forecasts on 129 securities, almost all after 2005, and the test is correspondingly weak: the R^2_{OOS} against the all-securities mean is 2.06% on closing prices and 1.67% on mid-quotes (t = 0.6 and 0.5), 1.82% and 1.44% among those traded on both dates (t = 1.3 for both). The predictability does not vanish at mid-quotes; whether it is attenuated cannot be decided with the bid and ask data that EUROFIDAI provides.

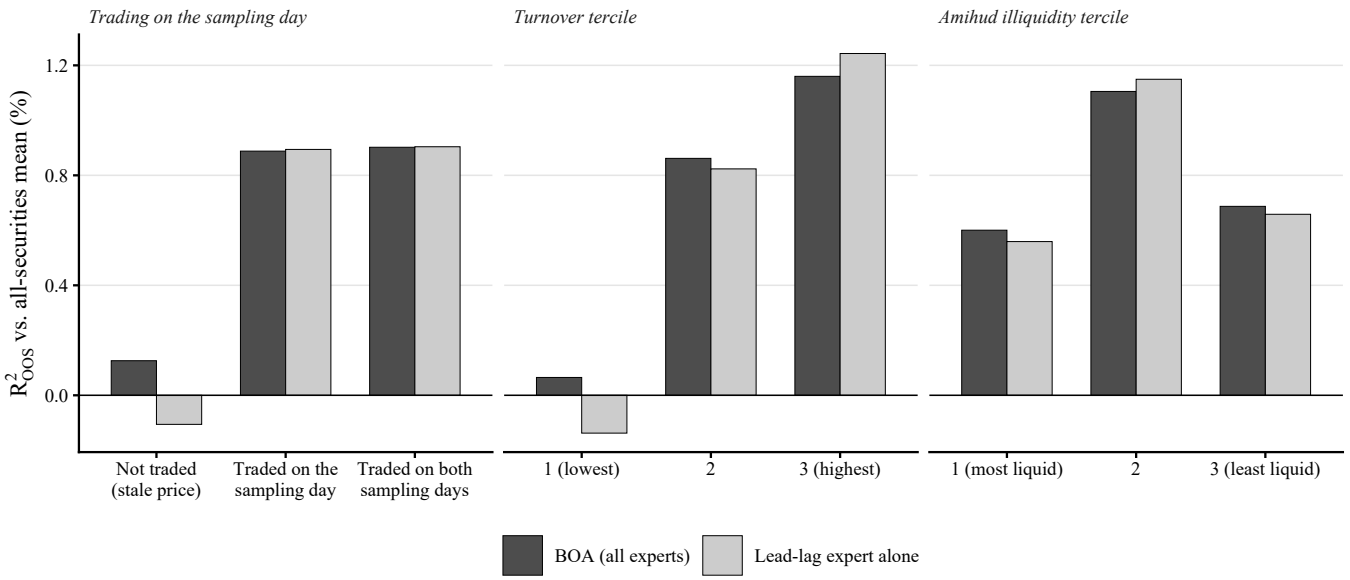


Figure J.12: Pooled R^2_{OOS} against the all-securities mean (%), 1977–2024, of the aggregate and of the lead-lag expert alone, by trading on the sampling day, by turnover tercile and by Amihud illiquidity tercile (Table J.38). Source: eurofidai_extra.R, figure FE2.

Table J.38: Predictability and trading, 1977–2024

			R^2_{OOS} (%) vs.				Expert alone, vs. CS mean	
	Obs.	Securities	HM	Zero	CS mean	t_{NW}	Lead-lag	Characteristics
All (1977-2024)	1,104,440	6,059	0.69	0.77	0.75	3.24	0.70	0.07
Price at t								
Not traded (stale price)	142,739	3,883	0.14	0.08	0.13	0.57	-0.11	0.10
Traded on the sampling day	669,846	4,765	0.81	0.90	0.89	3.60	0.89	0.05
Prices at t and t+1								
Traded on both sampling days	604,488	4,166	0.90	0.91	0.90	3.44	0.90	0.05
Turnover tercile								
1 (lowest)	266,760	3,996	0.50	0.05	0.06	1.08	-0.14	0.01
2	267,171	4,118	0.87	0.85	0.86	4.43	0.82	0.09
3 (highest)	267,551	2,818	0.72	1.19	1.16	3.52	1.24	0.07
Amihud illiquidity tercile								
1 (most liquid)	118,932	1,982	0.98	0.58	0.60	0.37	0.56	-0.01
2	119,224	1,992	1.41	1.10	1.11	2.20	1.15	0.05
3 (least liquid)	119,538	1,980	0.87	0.67	0.69	2.41	0.66	0.07
Quoted spread tercile								
2	17,774	912	1.16	0.80	0.91	0.39	0.82	0.14
3 (widest)	217	80	1.90	1.50	1.46	-0.12	1.78	-0.06

Forecasts of 1977–2024 split by whether the security traded on the day its price at t is sampled (volume strictly positive), on both sampling days t and $t + 1$ (consecutive dates only), and by real-time terciles of turnover, Amihud illiquidity and quoted spread over the half-month ending at t. Pooled R^2_{OOS} (%) of the aggregate against the historical mean, zero and the all-securities mean; t_{NW}: Newey–West t-statistic (12 lags) on the monthly cross-sectional loss differential against the all-securities mean; last two columns: R^2_{OOS} against the all-securities mean of the lead-lag expert and of the characteristics expert, each taken alone. Forecasts without volume data are omitted from the splits. Source: eurofidai_extra.R, table TE2.

Table J.39: Closing-price and mid-quote returns

Sample	Obs.	Securities	R^2_{OOS} vs. CS mean (closing prices)	t_{NW}	R^2_{OOS} vs. CS mean (mid-quotes)	t_{NW} (mid-quotes)	Lead-lag expert alone (mid-quotes)
Consecutive dates with quotes at t and t+1	369	129	2.06	0.59	1.67	0.53	1.88
of which traded on both sampling days	335	123	1.82	1.35	1.44	1.31	1.58

Forecasts for which bid and ask prices are recorded at t and at $t + 1$ on consecutive sampling dates; the realised return is computed from closing prices and from bid–ask mid-quotes, and the same forecasts are evaluated against the all-securities mean. Pooled R^2_{OOS} (%), Newey–West t-statistic, and R^2_{OOS} of the lead-lag expert alone on mid-quote returns. Source: eurofidai_extra.R, table TE3.