Mispricing Proxies in Factor Models for Asset Returns *

Gabriele Confalonieri† **Carlo A. Favero**‡ **Ilaria Leoni**§

March 23, 2025

Abstract

This paper examines the influence of mispricing proxies on stock return dynamics within the framework of Fama-French five factor models. Specifically, we assess the role of mispricing proxies derived from cointegration between asset prices and factor prices, as well as sentiment indicators extracted from quarterly earnings conference calls. Using quarterly data from 1980 to 2023 for the cross-section of DJIA-listed firms, our empirical analysis shows that deviations from long-run trends, driven by factor prices, have predictive power for stock returns after controlling for the five Fama-French factors. Stock-specific sentiment further enhances predictability . The additional predictability generated by mispricing proxies is fully explained by a non-linear model where sentiment determines the speed of adjustment toward the long-run trend identified by cointegration analysis when stock prices are above it.

JEL codes: C38, G12, G14, G17.

Keywords: Factor Models, Mispricing, Cointegration, Sentiment Measures, Natural Language Processing.

*This paper is part of the research activities of the Baffi Centre unit on Macroeconomic Trends, Cycles, and Asset Prices. Dev Srivastava provided outstanding research assistance.

†Bocconi University. E-mail: gabriele.confalonieri@unibocconi.it. Website: https://sites.google.com/view/gabrieleconfalonieri/home.

‡Bocconi University, IGIER, Baffi Centre and CEPR. E-mail: carlo.favero@unibocconi.it. Website: https://mypage.unibocconi.eu/carloambrogiofavero/.

§Bocconi University. E-mail: ilaria.leoni@unibocconi.it. Website: https://sites.google.com/view/ilarialeoni/home.

1 Introduction

Dynamic predictive modeling of stock market returns is commonly implemented in the framework of factor models. Factor models of stock returns are usually interpreted in a no-arbitrage framework as decomposing risk in two parts: a common risk component, captured by the factors, and an idiosyncratic risk component captured by the residuals from the projection of excess returns on factors. In theory, common risk, being undiversifiable, needs a compensation while idiosyncratic risk, being diversifiable, does not need a compensation. Therefore, within a no-arbitrage framework, deviations of expected excess returns of stock returns from their exposure to expected factors are interpreted as mispricing. In this paper, we investigate the role of measures of mispricing specific to single stocks in the context of the Fama-French five factors model (Fama and French, 2015). In particular, we consider the potential role of mispricing proxies generated by cointegration analysis between asset prices and the five factor prices and by sentiment measures extracted from quarterly earnings conference-calls. Following (Christoffersen and Diebold, 1998) we experiment with the possibility of forecast enhancement for returns generated by using information in the current deviations from the cointegrating relationships. The empirical evidence documents that deviations of prices from their long-run(stochastic) trend driven by factor prices do predict stock returns, and that sentiment also features some additional predictability. Interestingly, the additional predictability generated by mispricing proxies is fully explained by a non-linear model where sentiment determines the speed of adjustment toward the long-run trend identified by cointegration analysis when stock prices are above it.

Asset prices feature trends, and financial markets are characterized by frequent deviations of asset prices from their long-run trends (Dong et al., 2022). If factors prices successfully capture the stochastic trend in asset prices, and deviations of asset prices from their long-run trend affect returns, then these deviations are naturally interpreted as mispricing in standard factor models. Mispricing can result for different reasons, such as overreaction to news (Gennaioli and Shleifer, 2018), non-rational expectations (Bordalo et al., 2019) or financial frictions (Duffie, 2010). We consider a general framework in which mis-pricing proxies are specified as asset-specific additional regressors in a standard factor model specification that projects asset returns (we focus on the companies in the Dow Jones Industrial Average (DJIA) index) on the 5 FF factors. A time-series approach to the relation between asset prices, factor prices, asset returns and factor returns derives our first proxy. We start from the intuition that asset prices are driven by a permanent and a temporary component, Fama and French (1988) and we build a model in which factor prices, i.e. the values of buy-and-hold portfolios in the 5 Fama-French factors, are used to capture the permanent component in asset prices. If this model is successful, then its residuals are stationary and mean-reverting. We consider them as a proxy of mispricing, as they are an additional regressor with a significant coefficient in the standard factor model.

A second measure of idiosyncratic mispricing is then constructed by using a granular measure of sentiment for each firm in our sample. This measure is obtained by analyzing firm-by-firm quarterly earnings conference-calls. Even though the earnings calls are available at a lower frequency than business news, they let us treat each company consistently, as some firms are sistematically overrepresented in the news (e.g. Apple). Earnings calls transcripts are available for each company in our sample, and they are widely followed and covered by financial markets and the economic literature (Hassan et al., 2019). We compute the sentiment for each document using a state-of-the-art natural language processing (NLP) model, FinBERT (Araci, 2019).[1]

These two measures of mispricing are first compared by looking at their time series behaviour for the same assets and at their correlation across different assets. We then proceed to evaluate their significance in standard empirical factor models to find that most

[1] More details on the class of BERT models is provided in the Appendix A

of the estimated coefficients on the cointegration based mis-pricing proxies are significant and negative (implying that price above the equilibrium in the current quarter have a negative impact on the next quarter return) while the coefficients on sentiment based mis-pricing proxies are positive and also statistically different from zero. On the basis of this evidence we proceed to investigate if the information from the two different proxies can be combined by allowing the speed of adjustment with respect to the cointegration based disequilibrium to be a function of sentiment.

Heuristically, if an asset is overpriced but investor sentiment remains high, it will take longer to revert to equilibrium. This intuition is consistent with a model of diagnostic expectations in which investors' beliefs are confirmed by the sentiment signal (Bordalo et al., 2021). We regress the quarterly returns of each company against the lagged deviation of ther prices from their stochastic trend (the Equilibrium correction Term, ECT), and we impose that the speed of adjustment, i.e. the coefficient on the ECT, is a linear function of sentiment. We account for an asymmetric reaction to positive and negative mispricings. Our findings show that, when the price is above its equilibrium level, a larger sentiment lowers the speed of adjustment, so that it will take more periods to correct the overpricing shocks. We only find weak, and not statistically significant, effects in the case of negative ECTs.

The remainder of the paper is organized as follows. Section 2 places our contribution in the literature. Section 3 describes the methodology used in our analysis. Section 4 describes our data, presents the empirical results and discusses their implications. Section 5 concludes and provides directions for future research.

2 Placing our Contribution in the Literature

Our work contributes to the two related strands of the literature on mispricing in factor models and sentiment indicators.

Mispricing in a factor model has been traditionally evaluated by the standard test of efficiency for a given portfolio. this test concentrates on the null hypothesis that all the constants (alphas) in the system projecting excess returns on factors are not statistically different from zero (Gibbons et al., 1989). This specification strategy allows to identify "average" mis-pricing without explicitly allowing for time-varying mispricing. (Stambaugh and Yuan, 2017) consider explicitly the possibility of time-varying mispricing by exploring parsimonious factor models that include factors combining information from a range of anomalies. Their baseline evidence on mispricing for factor models is obtained by constructing a four factor model that includes two mispricing factors along with market and size factors. They also relate mis-pricing to investor sentiment by showing that investor sentiment predicts their mispricing factors, particularly their short (overpriced) legs and that, unlike the size factor constructed by Fama and French (Fama and French, 1993), their size factor- constructed to be less contaminated by mispricing—is not predicted by sentiment. Interestingly, they interpret the evidence on correlation between mispricing and investor sentiment as consistent with the arbitrage asymmetry in buying versus shorting. When mispricing is present, stocks that are more difficult to short should also be those for which overpricing is less easily corrected. The key innovation in our approach to assessing mispricing in factor models is that our measure derives from information inherently embedded within the chosen factor model, rather than relying on anomalies arising from external information excluded from the factor set.

Our second measure of mispricing is constructed by using a granular measure of sentiment for each of the 30 components of the DJIA index in 2023:1, built with a state-of-the-art natural language processing (NLP) model, FinBERT (Araci, 2019). This measure is obtained by analyzing firm-by-firm quarterly earnings conference-calls.

The study of sentiment analysis has seen considerable advancements with the evolution of (NLP) natural language processing technologies. FinBERT, developed by (Araci, 2019), represents a significant leap in this field. By tailoring the BERT model specifically to financial texts, FinBERT has demonstrated superior performance in extracting sentiment from financial documents compared to earlier methods.

The underlying BERT model, by (Devlin et al., 2018), introduced a novel approach with deep bidirectional transformers that have set a new standard in language understanding. Although BERT was not initially aimed at financial applications, its capacity to understand context and semantics has made it a powerful tool for analyzing financial sentiment.

In their comprehensive review, (Mishev et al., 2021) explore various sentiment analysis techniques within the finance sector. They compare traditional methods that rely on lexicons with newer transformer-based models like FinBERT. Their analysis highlights how modern approaches provide more accurate sentiment insights, reflecting a broader shift toward using advanced models to interpret financial data effectively.

3 Mispricing Proxies in Factor Models for Asset Returns

Factor models are commonly used to characterize parsimoniously the predictive distribution of asset returns. Specifically, multi-factor models in which k factors characterize in a lower parametric dimension the distribution of n asset (excess) returns, have the following general form:

$$\mathbf{r}_{t+1} = \alpha + \beta \mathbf{f}_{t+1} + \mathbf{v}_{t+1}, \qquad \text{with } \mathbf{v}_{t+1} \sim \mathcal{D}\left(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{v}}\right) \tag{1}$$

$$\mathbf{f}_{t+1} = E\left(\mathbf{f}_{t+1} \mid I_t\right) + \epsilon_{t+1} \qquad \text{with } \epsilon_{t+1} \sim \mathcal{D}\left(\mathbf{0}, \mathbf{\Sigma}\right) \tag{2}$$

where the $(n \times n)$ matrix $\mathbf{\Sigma}_v$ is diagonal, the $(k \times k)$ matrix $\mathbf{\Sigma}$ is full, and $Cov\left(v_{i,t+1}, \epsilon_{j,t+1}\right) = 0$ for each i, j. $\mathbf{f}_{t+1}$ is a vector of k factors at time $t+1$, $\mathbf{r}_{t+1}$ is a vector of excess returns on the n assets at time $t+1$, and the matrix β contains the loadings for the n assets on the k factors. Equation (1) specifies the conditional distribution of returns on factors, while equation (2) specifies the predictive distribution for factors at time $t+1$ conditioning on information available at time t. A baseline specification for this system assumes away factors predictability thus implying that conditional expectations of factors have no variance (i.e., $E\left(\mathbf{f}_{t+1} \mid I_t\right) = \mu$). In this case the predictive distribution of returns coincides with the unconditional one and it takes the following specification: $\left(\mathbf{r}_{t+1} \mid I_t\right) \sim \mathcal{D}\left(\left(\alpha + \beta' \mu\right), \left(\beta \mathbf{\Sigma} \beta' + \mathbf{\Sigma}_v\right)\right)$.

If a factor model is interpreted in a no-arbitrage context, then it is evident that the total compensation for risk is decomposed two parts: a compensation for a common risk component, captured by the factors, and a compensation for an idiosyncratic risk component captured by the residuals from the projection of excess returns on factors. Common risk, being undiversifiable, needs a compensation while idiosyncratic risk, being diversifiable, does not need a compensation, therefore a common test for the validity of a factor model is performed by evaluating the null $\alpha = 0$. The reduction in dimensionality of the parameters in the variance-covariance matrix of returns occurs as a consequence of orthogonality between the common and the idiosyncratic risk components for all assets and of the orthogonality between the idiosyncratic risk components of all assets.

To empirically evaluate the potential role of mis-spricing in factor models consider the following general specification for each of the available n test-assets :

$$r_{i,t+1} = \alpha_i + \beta_i \mathbf{f}_{t+1} + \delta_i(L) u_{i,t+1}^{MP} + v_{i,t+1}, \quad \text{with } \mathbf{v}_{t+1} \sim \mathcal{D}\left(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{v}}\right) \tag{3}$$

where $u_{i,t+1}^{MP}$ represents the idiosyncratic variables that are candidates to capture mispricing in the standard factor model. Given the general specification and different variables candidate to capture mispricing, we shall also test the hypothesis that the response to mis-pricing is common across different assets, i.e. $\delta_i(L) = \delta(L)$.

3.1 Cointegration-based Mispricing Proxies

Mispricing proxies can be naturally derived by taking a cointegration-based approach to factor models Favero et al. (2020)

Consider factor models specified on log excess returns and define log cumulative excess returns on asset i $\ln P_{i,t}$ as:

$$\ln P_{i,t} = \ln P_{i,t-1} + r_{i,t} \ , \tag{4}$$

i.e., (log)prices of any asset are cumulative (log) returns. The analogous of the (log) price for an asset can be constructed for any given factor.

Specifically, the generic factor prices associated with the log factor return $\mathbf{f}_t$ evolves according to the following process:

$$\ln \mathbf{F}_t = \ln \mathbf{F}_{t-1} + \mathbf{f}_t \ . \tag{5}$$

Test assets returns and factors are stationary, but asset prices and factor prices are non-stationary.

If factor prices are the non-stationary variables that drive the non-stationary dynamics of asset prices, then a linear combination of asset prices and factor prices should be stationary; i.e., asset and factor prices should be cointegrated.

Consider the following model describing the exposure of a given asset price $P_{i,t}$ to a given set of factor prices $\mathbf{F}_t$:

$$\begin{aligned} \ln P_{i,t+1} &= \alpha_{0,i} + \alpha_{1,i} t + \beta_i' \ln \mathbf{F}_{t+1} + u_{i,t+1} \ , \qquad i = 1, \ldots, n \\ u_{i,t+1} &= \rho_i u_{i,t} + v^r_{i,t+1} \end{aligned} \tag{6}$$

If the residuals are stationary, i.e. $|\rho_i| < 1$, then we have cointegration between factors' and asset prices. In this case, $u_{i,t}$ captures temporary deviations of prices from the long-run value determined by the factor prices. Thus, it is natural to refer to $u_{i,t}$ as the "Equilibrium Correction Term" (henceforth, ECT) associated with asset i at time t.

$$ECT^P_{i,t} \equiv \ln P_{i,t} - \hat{\alpha}_{0,i} - \hat{\alpha}_{1,i}t - \hat{\beta}'_i \ln \mathbf{F}_t \ .$$

Interestingly, the existence of cointegration and therefore the stationarity of the ECT is naturally interpreted as an indication of validity of a factor models. The validity of a factor model can be judged by the capability of factor prices to capture the stochastic trend in asset prices. In fact, this argument could be further refined by imposing two requirements for a valid factor models: no-cointegration between factor prices (the presence of common trends among factors prices would be an indication of "redundancy of some of the factors") and cointegration between factor prices and asset prices.

3.1.1 The Augmented Factor Model

The presence of cointegration generates a mispricing proxy in factor models.

By differencing (6) we see that cointegration between asset and factors prices implies the presence of a new additional predictive term in the standard factor model projection of returns on factors

$$r_{i,t+1} = \alpha_{1,i} + \beta'_i \mathbf{f}_{t+1} + \underbrace{(\rho_i - 1)}_{\delta_i} \underbrace{u^{CB}_{i,t}}_{\equiv ECT^P_{i,t}} + v^r_{i,t+1}, \tag{7}$$

Note that the presence of the mispricing term in standard factor models generated by the cointegration approach is ruled out only in the case of the presence of a unit root in the ECT term i.e. when there is no cointegration between asset prices and factor prices. By construction, the CB-based mispricing proxy affects asset returns with a one-period lag.

3.2 Sentiment-based Mispricing Proxies

An alternative mispricing proxy to the cointegration-based one has been developed, using a sentiment indicator constructed from the analysis of earnings call transcripts with natural language processing (NLP) techniques, specifically employing the FinBERT model. Earnings calls, where corporate executives discuss financial performance and future outlook, are valuable sources of information for investors. To distill insights from these complex and lengthy conversations, FinBERT, a model fine-tuned on financial data, was leveraged to extract sentiment from these texts and quantify the overall tone of the discussion. The process of constructing this sentiment indicator involved several steps, including data extraction, sentence-level sentiment scoring, and the final calculation of an overall sentiment score for each earnings call.

3.2.1 Data Extraction and Preprocessing

The first step involved obtaining the earnings call transcript in a format suitable for sentiment analysis. The text was extracted and broken down into individual sentences, with each sentence serving as a unit of analysis. This granular approach allowed for a more detailed and accurate sentiment assessment. Once segmented, the data was converted into a structured CSV file. Each row contained a single sentence from the earnings call, which was then analyzed independently.

3.2.2 Applying FinBERT for Sentiment Analysis

To evaluate the sentiment of each sentence, the FinBERT model was employed. FinBERT is a pre-trained transformer model specifically tailored for sentiment analysis in financial texts. Unlike general-purpose NLP models, FinBERT is fine-tuned on financial data, making it particularly effective in capturing the nuances of language used in the finance sector. Its training involved a large annotated dataset from the Financial Phrasebank, ensuring its ability to identify positive, negative, and neutral sentiments in the context of financial markets.

The Financial Phrasebank, developed as part of this research, is a collection of over 4,800 sentences annotated by financial experts. Each sentence is classified as positive, negative, or neutral based on its likely impact on stock prices from an investor's perspective. Sentences with annotations from at least 16 contributors form the basis for training FinBERT. This dataset is essential in ensuring that the model is fine-tuned specifically for the financial domain, where certain phrases or statements may carry sentiment that could easily be missed or misinterpreted by models trained on non-financial data.

Using FinBERT, each sentence in the earnings call transcript was processed, and a sentiment score was assigned based on FinBERT's analysis. The model categorizes sentences into three categories: positive, negative, or neutral. Positive sentences typically highlight favorable financial performance or optimistic future outlooks, while negative sentences focus on challenges, losses, or unfavorable trends. Neutral sentences, on the other hand, do not present any clear sentiment relevant to financial outcomes.

3.2.3 Calculating the Overall Sentiment Indicator

Once each sentence was associated with a sentiment score, the overall sentiment for the entire earnings call was calculated. This step involves aggregating the sentence-level scores to provide a single summary metric representing the tone of the call. To compute this, the average of the individual sentiment scores was taken, giving equal weight to each sentence. The resulting outcome serves as the final sentiment indicator for the earnings call, offering a snapshot of the call's overall sentiment—whether positive, negative, or neutral.

3.2.4 The Financial Phrasebank

The accuracy of FinBERT's sentiment scoring can be attributed to the Financial Phrasebank, which plays a critical role in ensuring that the model is finely attuned to the language used in financial communications. The Phrasebank contains sentences that were manually annotated by finance students and researchers with relevant domain knowledge, ensuring that the resulting model captures the specific nuances of financial language. The dataset offers multiple agreement levels for annotations—ranging from 100% consensus to 50% majority voting—providing flexibility in building sentiment models. For this sentiment indicator, using sentences with high agreement ensures a more reliable and consistent interpretation of sentiment.

The Financial Phrasebank focuses specifically on sentences that influence stock prices from an investor's perspective. This domain-specific annotation enables FinBERT to accurately assess how a given sentence, and by extension, the entire earnings call, might affect investor sentiment and market behaviour. Sentences that may appear neutral or irrelevant to non-financial models are carefully evaluated in terms of their potential impact on stock prices, significantly enhancing the model's applicability to real-world financial analysis.

3.2.5 Example: JPM Sentiment Indicator

This section presents an example for JPMorgan Chase (JPM). The following time series illustrates the sentiment scores for multiple earnings calls, highlighting the minimum and maximum sentiment scores.

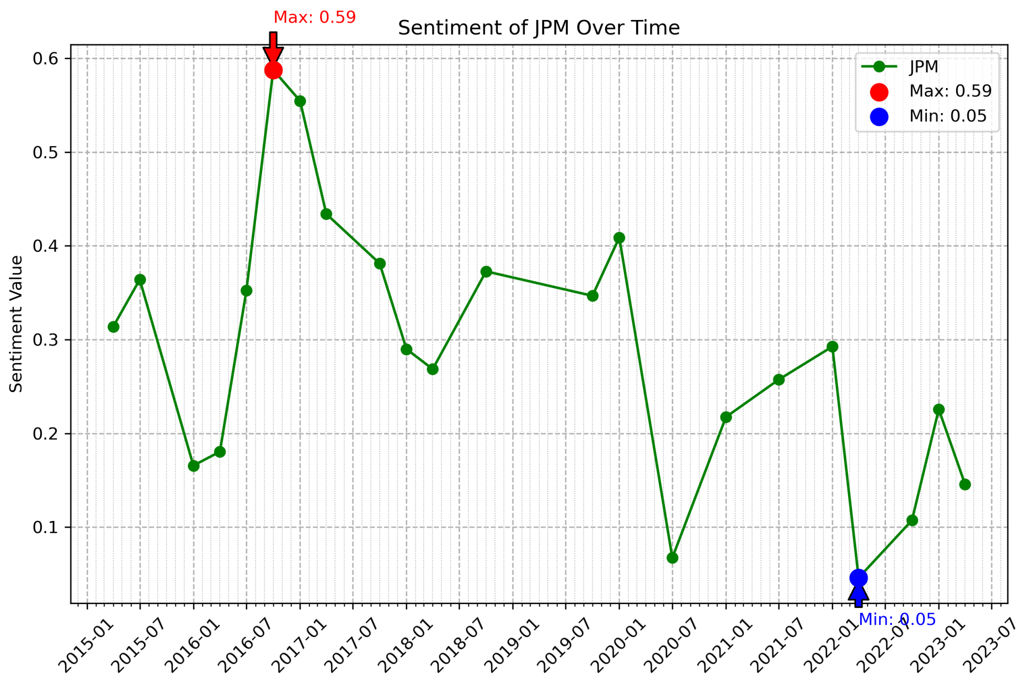


Figure 1: JPM Sentiment Scores for Earnings Calls

Below are the two earnings call transcripts corresponding to the minimum and maximum sentiment scores for JPM. The left image shows the earnings call with the minimum sentiment score, while the right image shows the call with the maximum sentiment score.

16	Regarding loan growth, we're continuing to see positive trends with loans up 8% year-on-year and 1% quarter-on-quarter ex PPP with the sequential growth driven by a continued pickup in demand in our Wholesale businesses, including ongoing strength in AWM.	1
17	On Page 2, we have some more detail on our results	-1.40E-06
18	**Revenue** of $31.6 billion was **down $1.5 billion or 5% year-on-year**	-0.99999744
19	NII ex Markets was up $1 billion or 9% on balance sheet growth and higher rates, partially offset by lower NII from PPP loans	0.99994177
20	**NIR ex Markets was down $2.2 billion or 17%** predominantly driven by lower IB fees, lower Home Lending production revenue, **losses in Credit Adjustments** & Other and CIB as well as **investment securities losses** in corporate	-0.99998677
21	And **Markets revenue was down $300 million or 3%** against a record first quarter last year.	-0.9999984
22	Expenses of $19.2 billion were up approximately $500 million or 2% predominantly on higher investments and structural expenses, largely offset by lower volume and revenue-related expenses	0.8322505
23	Credit costs were $1.5 billion for the quarter	-0.000568593
24	We built $902 million in reserves driven by increasing the probability of downside risks due to high inflation and the war in Ukraine as well as builds for Russia-associated exposures in CIB and AWM.	0.004683026
25	Net charge-offs of $582 million were down year-on-year and comparable to last quarter and remain historically low across our portfolios.	-0.99884903
26	On to balance sheet and capital on Page 3	-3.66E-06
27	Our **CET1 ratio** ended at 11.9%, **down 120 basis points from the prior quarter**	-0.9987935
28	As a reminder, we exited the fourth quarter with an elevated buffer to absorb anticipated changes this quarter, the largest being SA-CCR adoption as well as some pickup in seasonal activity.	0.3082016
29	In addition to those anticipated items, there were a couple of other drivers	-0.012411644
30	The rate sell-off led to AOCI drawdowns in our AFS portfolio	-0.07326184
31	But keep in mind, all else equal, these mark-to-market losses accrete back to capital through time and as securities mature	-0.006182162
32	And price increases across commodities resulted in higher counterparty credit and market risk RWA.	0.62100106
33	While, of course, the **environment is uncertain**, many of these effects are now in the rearview mirror	-0.8923381
34	And as a result, we believe that our current capital and future earnings profile position us well to continue supporting business growth while meeting increasing capital requirements as we look ahead.	1
35	With that, let's go to our businesses, starting with Consumer & Community Banking on Page 4	-1.71E-06
36	CCB reported **net income** of $2.9 billion on revenue of $12.2 billion, which was **down 2% year-on-year**	-0.9999996
37	In Consumer & Business Banking, revenue was up 8% predominantly driven by growth in deposit balances and client investment assets, partially offset by deposit margin compression.	0.9999978
38	Deposits were up 18% year-on-year and 4% quarter-on-quarter, consistent with last quarter	0.99992645
39	And client investment assets were up 9% year-on-year largely driven by flows in addition to market performance.	0.9999981
40	In Home Lending, **revenue was down 20% year-on-year** on **lower production revenue** from both **lower margins and volumes** against a very strong quarter last year, largely offset by higher net servicing revenue	-0.99999887
41	**Originations** of $24.7 billion **declined 37% with the rise in rates**	-0.999426
42	And as a result, **mortgage loans were down 3%.**	-0.9985784

(a) Earnings Call with Min Sentiment

15	Starting on page 1 and taking a look at the quarter we had **strong performance in each of our businesses** despite the continuation of reasonably challenging conditions.	0.9999998
16	Bringing it all together this quarter's result was clean with no significant items and with the firm reporting net income of $6.3 billion, EPS of $1.58 and the return on tangible common equity of 13% on $25.5 billion of revenue	0.9730507
17	Highlights of the quarter include the **highest reported revenue for a third quarter in CIB** with **IBCs up 15%** and **markets revenues up 33%** with **strong performance across the board**.	1
18	**Robust core loan growth** for the Company of **15%** on the back of **sustained demand across businesses**	1
19	And the **continuation of strong credit performance** including a net release for Oil & Gas.	1
20	**Card sales are back to double-digit growth** year on year and we saw a strong positive market reaction to new proprietary products	1
21	Finally, we had record **consumer deposit growth up 11%.**	1
22	Before I move on we recently submitted our 2016 resolution filing	-1.28E-05
23	The Board and management believes that we submitted a credible plan and **more than met the requirements** for the October submission	0.99812907
24	It was a tremendous effort across the Company involving all businesses and functions and we took many significant actions, perhaps most notably **improving the firm's overall liquidity** and pre-positioning our material legal entities for both liquidity and capital.	0.99999994
25	We determined this is in the best interest of the Company, albeit at some cost	-0.018679436
26	We took many other important actions which hopefully you've had the chance to review in our public filings.	0.000510423
27	Moving back to the quarter and moving on to page 2 **revenue** of $25.5 billion was up $2 billion year on year or **up 8%**	0.99937934
28	On the back of **continued strong growth in core loans net interest income was up** $700 million and is trending for the full year to be above the $2.5 billion guided last quarter.	1
29	**Non-interest revenue was up** $1.3 billion driven by **strong performance in the CIB**	1
30	Adjusted expense of $14.5 billion was up $500 million both year on year and quarter on quarter, largely driven by two notable expense items in Consumer which I'll talk about later as well as the increase in FDIC surcharge which took effect this quarter and some higher marketing expense.	0.77875733
31	Credit cost of $1.3 billion in the quarter includes Consumer reserve builds of $225 million primarily Card	-0.000193629
32	But against that we have a net reserve release in wholesale for Oil & Gas of about $50 million.	-0.001153902
33	So as I said net income was $6.3 billion	-0.000119323
34	And while down 8% year on year you will recall that there were a number of significant items in last year's results, most notably significant tax benefits	0.29547194
35	If you adjust for tax, legal expense and **credit reserves net income is up** over $800 million year on year.	0.99999905
36	Dealing with Oil & Gas here, we are encouraged by how quickly investor sentiment and risk appetite for the sector returned as the **outlook for both oil and gas prices continued to improve**	1

(b) Earnings Call with Max Sentiment

Figure 2: Earnings Calls for JPM with Min and Max Sentiment Scores

In conclusion, using FinBERT, the Financial Phrasebank and by breaking down the text into individual sentences and applying a financial-domain-specific NLP model, an accurate sentiment score was generated that reflects the overall tone of the call. This sentiment indicator gauges the mood of corporate executives during earnings calls, helping to inform investment decisions. By leveraging FinBERT's specialized training and the comprehensive annotation of the Financial Phrasebank, the resulting sentiment scores offers a measure of financial sentiment in earnings calls.

3.2.6 The augmented factor model specification

In the case of sentiment based mis-pricing proxies, the relevant augmented factor model will be the following:

$$r_{i,t+1} = \alpha_{1,i} + \beta_i' \mathbf{f}_{t+1} + \delta_i u_{i,t+1}^{SB} + \ v_{i,t+1}^{r}, \tag{8}$$

where $u_{i,t+1}^{SB}$ are the asset specific sentiment-based mis-pricing proxies. Note that, while the econometric specification strategy leads to the inclusion in the specification of the cointegration based proxy with a lag, the sentiment based proxy is simply an additional variable that affects returns contemporaneously as the factors. Both the CB-based and the SB-based proxy are idiosyncratic to test assets and are meant to capture fluctuations in the idiosyncratic risk components, but the timing of the CB based proxy is different because of the link between cointegration between asset prices and factor prices and the implied Equilibrium Correction Term for asset returns.

4 The Empirical Evidence

We describe our empirical evidence by illustrating first the sources for our data and the construction of our indicators, by moving then to preliminary data analysis, to eventually illustrate our estimation results.

4.1 Data

We consider as a baseline model a standard factor model for a representative sample of large and liquid stocks, in particular we consider the quarterly data over the sample $1980:01-2023:01$ of the 30 constituents of the Dow Jones Industrial Average index (DJIA) at the end of the sample, omitting from our analysis Dow Inc., which was included in the index only in 2019. This gives us an initial cross-section of 29 assets. The data on the Fama-French 5 factors (Fama and French, 2015) are retrieved from Ken French data library (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html), where also data on the time-series of the riskfree rate are made available. Data on factor prices and asset prices are constructed by cumulating log(returns).

Our measure of sentiment is constructed by analyzing the text of quarterly earnings call, which we obtain via Refinitiv and Bloomberg. We analyze them using a FinBERT model. Earnings calls are conducted by companies with their board members, investors, analysts and the press. These calls typically occur once every quarter and are used to discuss the company's financial results and performance. Earnings calls serve as a platform for the company to communicate its financial performance and outlook to stakeholders. They also provide an opportunity for investors and analysts to ask questions and gain insights into the company's operations and prospects.

Earnings calls usually follow a structured format, consisting of two parts: the main presentation and a question-and-answer (Q&A) session. During the main presentation portion, the company's management, typically the CEO and CFO, provide a detailed overview of the financial results for the quarter or year. Management may also discuss strategic initiatives, market trends, and other factors that have influenced the company's performance. The main presentation is scripted and prepared in advance to ensure that important information is communicated clearly and accurately. A Q&A session follows the main presentation. In this session, analysts, investors, and the press have the opportunity to interact with the company's management. These questions can cover a wide range of topics, including specific financial results, future guidance, industry trends, and more. The Q&A session allows for direct engagement between the company and its stakeholders, providing additional insights beyond the prepared remarks.

Additionally, earnings calls can offer insights into the company's sentiment and risk perceptions. This information is derived through textual analysis of transcripts from these calls. Earnings calls are considered valuable because they provide timely information on a company's financial performance and offer insights into its management's perspective on the business. Furthermore, they are used to monitor sentiment and risk trends, which can offer valuable insights into the overall health of a company and its industry.

4.2 Building mis-pricing proxies

Several steps are necessary to build the cointegration-based mis-pricing proxies. First, asset prices and factor prices are constructed by cumulating (log) returns:

$$\ln P_{i,t} = \ln P_{i,t-1} + r_{i,t}, \tag{9}$$

$$\ln F_{i,t} = \ln F_{i,t-1} + f_{i,t} \tag{10}$$

Second, the long-run properties of the factor prices are investigated by checking the absence of any co-integrating relationship among them. If factor prices were to share some stochastic trends, then some of the factors would be redundant for pricing the assets. To test for no-cointegration among factor prices, we estimate the following VAR on them to apply the Johansen cointegration analysis (Johansen, 1995)

$$\begin{aligned}\ln \mathbf{F}_t &= A_0 + A_1 t + A_2(L) \ln \mathbf{F}_{t-1} + \mathbf{v}_t \ . \qquad (11)\\ \ln \mathbf{F}_t &= (lnF_{mkt,t} - lnF_{rf,t}, lnF_{smb,t}, lnF_{hml,t}, lnF_{rmw,t}, lnF_{cma,t})\end{aligned}$$

Table 1 reports the empirical results and shows that the null of at most zero cointegrating vectors is not rejected at 1 percent critical level by the relevant test (the one for the specification that allows for the presence of a deterministic trend in the cointegrating vector).

Table I

Cointegration Test on Factor Prices

This table reports the statistic and critical values for the Johansen maximum eigenvalue test on factor prices. The factor prices are obtained as the cumulative returns of a buy-and-hold strategy that invests in the 5 Fama French factors. r denotes the number of cointegrating relations. Quarterly data from $1980 : Q1$ to $2023 : Q1$.

	test	10pct	5pct	1 pct
$r \leq 4$	3.47	10.49	12.25	16.26
$r \leq 3$	5.09	16.85	18.96	23.65
$r \leq 2$	10.49	23.11	25.54	30.34
$r \leq 1$	18.11	29.12	31.46	36.65
$r = 0$	39.70	34.75	37.52	42.36

We then move to the cointegration analysis between asset prices and the factor prices. Again we apply the Johansen procedure that requires the specification of 29 VARs, on the vector of variables obtained by augmenting in turn the (log-)factor prices with the (log-)asset prices of the constituents of the DJIA at the end of the sample:

$$\begin{aligned}\ln \mathbf{Z}_t &= A_0 + A_1 t + A_2(L) \ln \mathbf{Z}_{t-1} + \mathbf{v}_t \ . \qquad (12)\\ \ln \mathbf{Z}_t &= (lnP_{i,t}, lnF_{mkt,t} - lnF_{rf,t}, lnF_{smb,t}, lnF_{hml,t}, lnF_{rmw,t} lnF_{cma,t})\end{aligned}$$

Table II reports the result of the Johansen analysis for the thirty constituent of the DJIA, showing that for 26 of the 29 assets the null of at most zero cointegrating vectors can be rejected at the 10pct critical level, if the 5pct critical value is adopted the number reduces to 16 assets, and if the strictest 1pct critical value is adopted then null is rejected only for 5 assets.

Table II

Cointegration Test on Asset and Factor Prices

This table reports the statistic and critical values for the Johansen maximum eigenvalue test on asset and factor prices. Each row shows the test result considering the price of asset i and the price of the 5 Fama French factors. Each row shows tests for the null of at most zero cointegration relation among the series. The data is at the quarterly frequency, from $1980:Q1$ to $2023:Q1$ (unbalanced panel).

	test	10pct	5pct	1pct
AAPL	40.37	40.91	43.97	49.51
AMGN	43.01	40.91	43.97	49.51
AXP	47.36	40.91	43.97	49.51
BA	50.29	40.91	43.97	49.51
CAT	46.54	40.91	43.97	49.51
CRM	62.65	40.91	43.97	49.51
CSCO	59.13	40.91	43.97	49.51
CVX	49.65	40.91	43.97	49.51
DIS	46.82	40.91	43.97	49.51
GS	47.54	40.91	43.97	49.51
HD	45.33	40.91	43.97	49.51
HON	42.92	40.91	43.97	49.51
IBM	41.73	40.91	43.97	49.51
INTC	40.25	40.91	43.97	49.51
JNJ	41.02	40.91	43.97	49.51
JPM	48.21	40.91	43.97	49.51
KO	49.30	40.91	43.97	49.51
MCD	42.75	40.91	43.97	49.51
MMM	43.93	40.91	43.97	49.51
MRK	47.89	40.91	43.97	49.51
MSFT	44.99	40.91	43.97	49.51
NKE	49.13	40.91	43.97	49.51
PG	41.14	40.91	43.97	49.51
TRV	41.33	40.91	43.97	49.51
UNH	47.77	40.91	43.97	49.51
V	66.29	40.91	43.97	49.51
VZ	40.08	40.91	43.97	49.51
WBA	42.74	40.91	43.97	49.51
WMT	48.75	40.91	43.97	49.51

In the light of these mixed results on the capability of the Fama-French factor prices to capture the long-run trend in asset prices, we construct cointegration based mis-pricing proxies by considering only the 16 companies when the null of at most zero cointegrating vectors between factor prices and asset prices was rejected at the 5pct level and we drop from the analysis of returns those companies for which the null of no cointegrating vectors could not be rejected. Cointegration based mis-pricing proxies are then constructed as follows:

$$
\begin{aligned}
u_{i,t}^{CB} &\equiv \ln P_{i,t} - \hat{\alpha}_{0,i} - \hat{\alpha}_{1,i} t - \hat{\beta}_i' \ln \mathbf{F}_t. \\
\ln \mathbf{F}_t &= (\ln P_{AXP,t}, \ln P_{BA,t}, \ln P_{CAT,t}, \ln P_{CRM,t}, \ln P_{CSCO,t}, \ln P_{CVX,t}, \ln P_{DIS,t}, \\
&\quad \ln P_{GS,t}, \ln P_{HD,t}, \ln P_{JPM,t}, \ln P_{MRK,t}, \ln P_{MSFT,t}, \ln P_{NKE,t}, \ln P_{UNH,t}, \\
&\quad \ln P_{V,t}, \ln P_{WMT,t})
\end{aligned}
$$

where the cointegrating parameters are estimated via static long-run regressions.[2]

Sentiment-based mispricing proxies for our selected assets are built using the detailed textual analysis approach described in the previus section. Specifically, we apply the FinBERT model to the same set of companies for which cointegration analysis between factor prices and stock prices has been implemented. We run FinBERT on the transcripts of earnings calls, where company leaders discuss their financial results and future plans. The model analyzes these transcripts to determine the sentiment of each report, assessing the overall tone and emotion expressed. This tone is computed as a weighted average of the sentiment of each sentence within the earnings calls. The sentiment score from each report provides us with a proxy for how the market might emotionally react to the information shared during the earnings call.

[2]Robustness of these estimates to the omission of the short-term dynamics has been checked via the ARDL approach, which gives negligible differences in the mispricing proxies

4.3 Preliminary analysis of mis-pricing proxies

We report the cointegration-based and the sentiment based measures of mispricing in Figure 3.

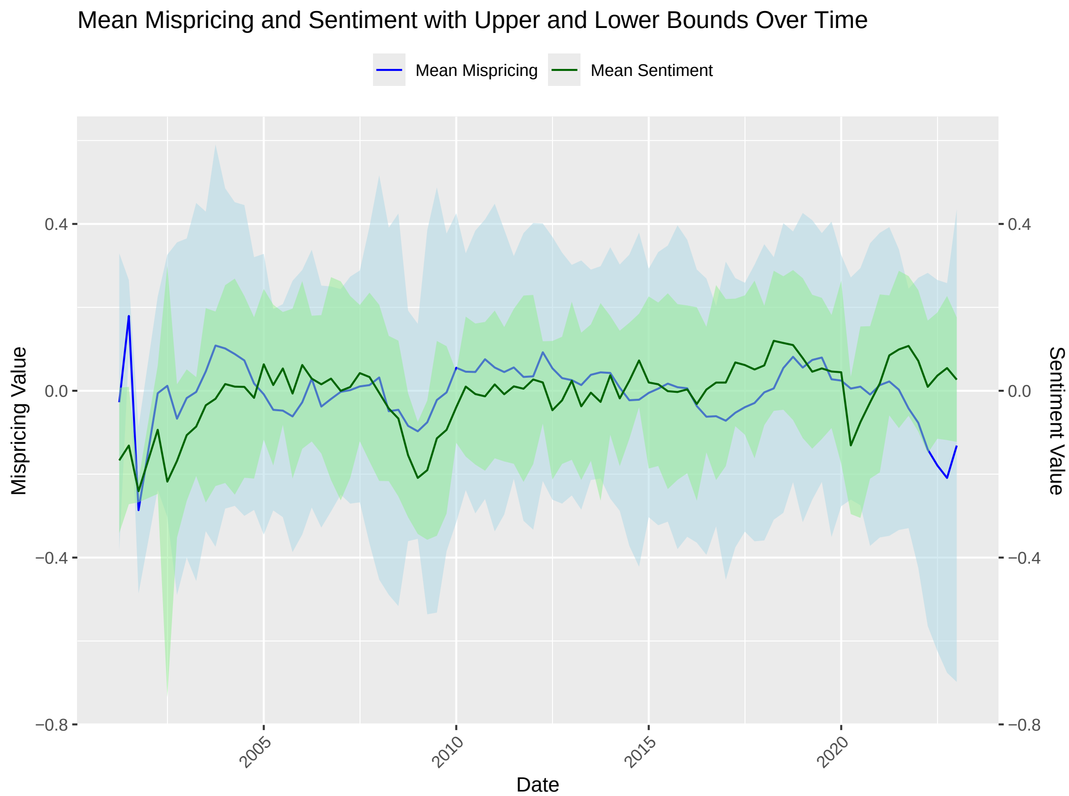


Figure 3: Sentiment based and Cointegration based measures of mispricing

Figure 3 concentrates on the subset of 16 stocks cointegrated with factor prices and reports the cross-sectional average and the upper and lower bounds of the two measures. Both measures feature stationarity with correlation that varies across the cross section of stocks. Figure 3 shows a small positive correlation between the cross-sectional averages of the two measures. In fact, there is rather remarkable heterogeneity in the cross-section as witnessed by Figure 4 that reports the correlation for the two different measures for each stocks. Correlation ranges from -0.4 for CSCO and V to 0.7 for BA, and it is positive for 12 out of the 16 considered stocks.

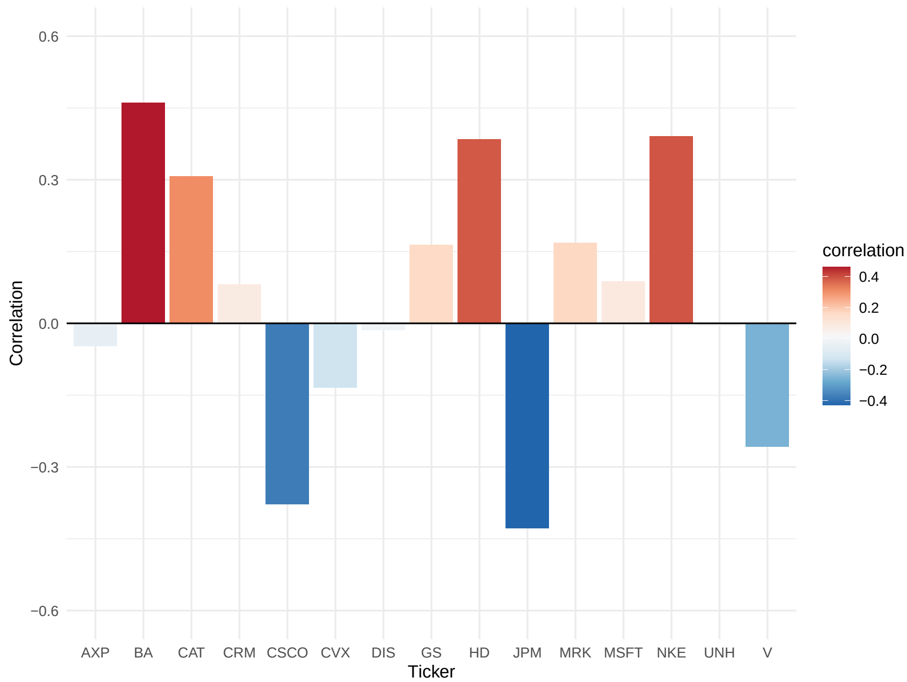
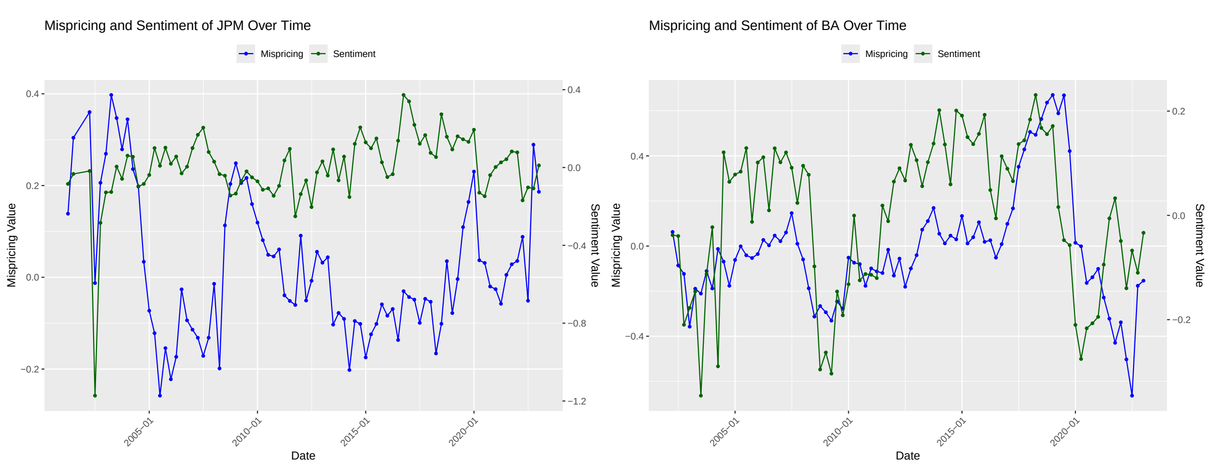


Figure 4: Correlation between sentiment based and cointegration based measures of mispricing

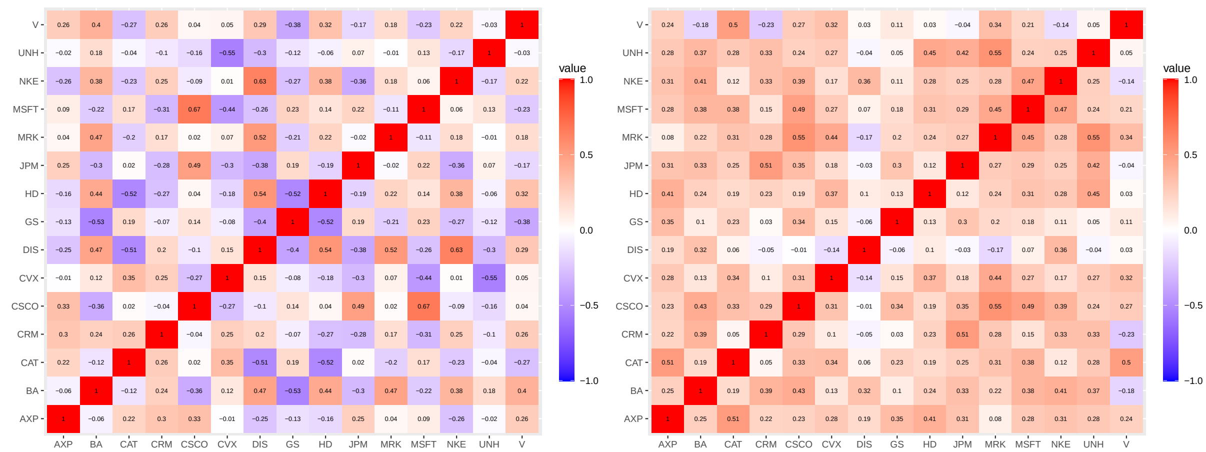
Figure 5 illustrates the pattern of correlation for the most negatively and the most positively correlated stocks.



(a) JPM **(b)** BA

Figure 5: Sentiment-based and cointegration based measures of mispricing for WBA and MSFT

It is also interesting to look at the correlation across tickers for the two measures to evaluate the idiosyncratic versus common nature of mispricing. Figure 6 reports correlation heatmaps across tickers for cointegration based and sentiment based measures of mispricing.



(a) Cointegration **(b)** Sentiment

Figure 6: Correlation across tickers of sentiment-based and cointegration based measures of mispricing

Correlation across tickers of the econometric based measure of mispricing is moderate and predominantly negative, while correlation across tickers of the sentiment-based measures for mispricing is also moderate but predominantly positive. Mispricing based by sentiment features positive comovement potentially related to a common sentiment, while the same does not hold for the econometric-based measure. Interestingly, the heatmaps reveal a different pattern of correlation across tickers for the two mispricing measures, suggesting different natures for the underlying phenomena captured by the two differently constructed measures.

On the basis of this initial evidence we proceed to unrestricted system estimation of 16 equations projecting the cointegration based mispricing on the sentiment based mis-pricing and to a restricted panel estimation. That is, we estimate

$$u_{i,t}^{SB} = \beta_i u_{i,t-1}^{CB} + \epsilon_{i,t}, \tag{13}$$

where $u_{i,t-1}^{CB}$ is the ECT for company i in quarter t, while $u_{i,t}^{SB}$ is the sentiment extracted with FinBERT from the earnings call transcript at time t. No constant is included in the regression as the ECT for each company has zero mean by construction, we also estimate a restricted version of the system imposing the restriction $\beta_i = \beta$. No constant is included as both variables have zero mean. We have an unbalanced panel, due to the availability of the sentiment measure. Our data is at the quarterly frequency, and our estimation window goes from $2001 : Q1$ to $2023 : Q2$ (although the sample is not balanced across all equations included in the system). We report our restricted and unrestricted system estimates in Tables III. The estimate of coefficient β from Equation (13) is 0.035, positive and statistically different from zero at the ten per cent level. Overall, the evidence from the estimation confirms that on average ECT and sentiment are mildly positively correlated, with wild variation across different tickers.

Table III

System and Panel Estimation

This table reports the coefficient estimates for the estimation of the system $u^{SB}_{i,t} = \beta_i u^{CB}_{i,t-1} + \epsilon_{i,t}$. $u^{CB}_{i,t}$ is the ECT for company i in quarter t, the residual from a regression of the log-price of asset i on risk-drivers. $u^{SB}_{i,t}$ is the FinBERT sentiment extracted from the earnings call transcripts released by company i at time t. No constant is included as both variables have zero mean. The table also reports the coefficient estimates for the panel regression $u^{SB}_{i,t} = \beta u^{CB}_{i,t-1} + \epsilon_{i,t}$ where the restrictions $\beta_i = \beta$, are imposed. Number of observations is not balanced across equations. Standard errors are shown in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels respectively.

	Dependent variable: $u^{SB}_{i,t}$							
	AXP	BA	CAT	CRM	CSCO	CVX	DIS	GS
$u^{CB}_{i,t-1}$	−0.049	0.259***	0.285***	0.027	−0.096***	−0.082	−0.008	0.158
	(0.111)	(0.055)	(0.099)	(0.039)	(0.027)	(0.067)	(0.058)	(0.106)
First Obs.	2002:2	2002:2	2001:2	2004:3	2003:1	2002:2	2001:2	2002:2
Last Obs.	2023:1	2023:1	2023:1	2023:1	2023:1	2023:1	2023:1	2023:1
R^2	0.002	0.211	0.090	0.007	0.143	0.018	0.0002	0.027
Adjusted R^2	−0.010	0.202	0.079	−0.007	0.131	0.006	−0.012	0.015
	HD	JPM	MRK	MSFT	NKE	UNH	V	Panel
$u^{CB}_{i,t-1}$	0.240***	−0.497***	0.149	0.044	0.193***	−0.002	−0.229**	0.035*
	(0.065)	(0.115)	(0.096)	(0.054)	(0.051)	(0.058)	(0.112)	(0.018)
First Obs.	2001:3	2001:2	2001:2	2001:2	2001:2	2001:2	2008:2	
Last Obs.	2023:1	2023:1	2023:1	2023:1	2023:1	2023:1	2023:1	
R^2	0.144	0.180	0.029	0.008	0.152	0.00002	0.066	0.03
Adjusted R^2	0.134	0.171	0.017	−0.004	0.141	−0.014	0.05	-0.001

Note: *p<0.1; **p<0.05; ***p<0.01

4.4 Mis-pricing proxies in the FF factor model

On the basis of the evidence of low correlation between the two mispricing measures we investigate their importance in FF factor model by estimating first a standard factor model for the 16 companies selected in the previous section:

$$r_{i,t+1} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{f}_{t+1} + \epsilon_{i,t+1}, \tag{14}$$

and then by augmenting the standard factor model with the two proposed mispricing proxies to estimate the following system of equations

$$r_{i,t+1} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{f}_{t+1} + \delta_i^{CB} u_{i,t}^{CB} + \delta_i^{SB} u_{i,t+1}^{SB} + \epsilon_{i,t+1}, \tag{15}$$

We report in Tables IV-V, the results from the estimation of the systems. The estimation of the standard Fama-French 5 factors models, in Table IV, gives evidence for α never significantly different from zero, for a dominant role of the market factor in determining the common risk and some role for the HML, RMW and CMA factors while SMB factor is never significant at the 1 per cent level. The evidence of α not significantly different from zero is not common. It would be intepreted in standard factor model as confirming the validity of the five factors model for these tickers. However, the 16 tickers considered have been selected on the basis of the evidence for cointegration between asset prices and factor prices and it is therefore important to consider the econometric specification with Equilibrium Correction Terms implied by cointegration before drawing inference on mispricing. Table V reports the results of augmenting the standard FF model with the Cointegration based and the Sentiment based mispricing proxies. Two versions of model (15) have been estimated: in the first version the speed of adjustment with respect to disequilibrium is allowed to be heterogenous across assets, while in the second version is restricted to be the same. In the first version of the model the inclusion of the mispricing proxies leaves estimated coefficients on factors substantially unaltered with respect to those obtained in the standard Fama-French specification, but both mispricing proxies carry some additional predictability. The coefficients on $u_{i,t}^{CB}$ are significant for 12 tickers, with a very similar magnitude. When the same speed of adjustment is imposed on all tickers, the estimated coefficient takes a value of -0.109 significant at the 1 per cent level, implying that about one-tenth of the mispricing is corrected over the one-quarter horizon. This evidence of a strongly significant partial adjustment coefficient estimated with panel restriction is consistent with the results in the recent literature estimating the impact of relative leverage in corporate finance (Ippolito et al., 2023). The coefficients on $u_{i,t}^{SB}$ are significant for 7 tickers, with a very similar magnitude. When the homogeneity restriction is imposed on all tickers, the estimated coefficient takes a value of 0.105 significnt at the 1 per cent level, implying that the returns are positively affected by sentiment after controlling for common factors and idiosyncratic fluctuations driven by the cointegration based-mispricing proxy.

Table IV

Fama-French Five Factors Model

This table reports the coefficient estimates for the estimation of system (14). Sample of quarterly data 2001:1-2023:1, number of observations not balanced across equations. Standard errors are shown in parentheses.

	Dependent variable: $r_{i,t+1}$					
	Intercept	exmkt	smb	hml	rmw	cma
AXP	0.003	1.323***	-0.428*	1.006***	-0.572**	-0.763**
	(0.009)	(0.117)	(0.251)	(0.202)	(0.25)	(0.315)
BA	-0.003	1.3***	0.218	0.868**	0.127	-0.419
	(0.015)	(0.193)	(0.413)	(0.333)	(0.414)	(0.518)
CAT	0.003	1.246***	0.211	0.185	0.023	0.972**
	(0.013)	(0.162)	(0.338)	(0.278)	(0.342)	(0.437)
CRM	0.029*	1.297***	-0.519	-0.517	-0.528	-0.698
	(0.015)	(0.186)	(0.388)	(0.333)	(0.425)	(0.553)
CSCO	0.004	1.022***	-0.223	0.021	-0.477	0.24
	(0.012)	(0.146)	(0.314)	(0.262)	(0.323)	(0.414)
CVX	0.001	0.847***	0.189	0.231	0.643**	0.784**
	(0.011)	(0.15)	(0.313)	(0.254)	(0.319)	(0.383)
DIS	0.003	1.064***	0.241	0.341	-0.4	-0.415
	(0.01)	(0.126)	(0.267)	(0.218)	(0.273)	(0.334)
GS	-0.008	1.336***	-0.141	0.996***	-0.316	-0.835**
	(0.012)	(0.155)	(0.326)	(0.268)	(0.324)	(0.41)
HD	0.004	0.953***	0.018	-0.202	0.35	0.542
	(0.011)	(0.139)	(0.303)	(0.244)	(0.305)	(0.38)
JPM	0.01	1.132***	-0.167	1.036***	-1.002***	-0.142
	(0.011)	(0.134)	(0.289)	(0.233)	(0.276)	(0.347)
MRK	0	0.596***	-0.426	-0.475**	-0.004	1.227***
	(0.01)	(0.128)	(0.308)	(0.225)	(0.274)	(0.342)
MSFT	0.012	1.138***	-0.594**	-0.113	-0.177	-0.469
	(0.01)	(0.12)	(0.253)	(0.207)	(0.245)	(0.306)
NKE	0.008	1.083***	-0.288	-0.106	0.369	0.302
	(0.014)	(0.171)	(0.352)	(0.29)	(0.393)	(0.468)
UNH	0.017	0.73***	0.164	-0.089	0.473*	0.284
	(0.011)	(0.128)	(0.29)	(0.239)	(0.277)	(0.374)
V	0.017	0.914***	-0.349	-0.373	-0.014	0.31
	(0.011)	(0.128)	(0.286)	(0.237)	(0.307)	(0.391)

Note: *p<0.1; **p<0.05; ***p<0.01

Table V

Fama-French Five Factors Model with CB and EB Mispricing proxies

This table reports the coefficient estimates for the estimation of system (15). Sample of quarterly data 2001:1-2023:1, number of observations not balanced across equations. Standard errors are shown in parentheses.

	Dependent variable: $r_{i,t+1}$							
	Intercept	exmkt	smb	hml	rmw	cma	$u_{i,t}^{CB}$	$u_{i,t+1}^{SB}$
AXP	0.004	1.363***	-0.473*	0.908***	-0.662**	-0.532	-0.156**	-0.069
	(0.009)	(0.12)	(0.261)	(0.214)	(0.257)	(0.337)	(0.068)	(0.063)
BA	-0.006	1.302***	0.298	0.611*	0.105	-0.304	-0.122*	0.217*
	(0.016)	(0.195)	(0.439)	(0.363)	(0.421)	(0.549)	(0.07)	(0.122)
CAT	0.01	1.142***	0.474	0.07	0.076	0.872**	-0.187***	0.144**
	(0.012)	(0.157)	(0.329)	(0.264)	(0.325)	(0.412)	(0.064)	(0.065)
CRM	0.031**	1.118***	-0.085	-0.678**	-0.507	-0.38	-0.241***	0.322
	(0.014)	(0.179)	(0.378)	(0.311)	(0.398)	(0.52)	(0.067)	(0.202)
CSCO	0.004	1.095***	-0.199	0.068	-0.492	0.368	-0.079**	0.024
	(0.012)	(0.141)	(0.303)	(0.255)	(0.304)	(0.394)	(0.031)	(0.122)
CVX	0.003	0.796***	0.193	0.257	0.629**	0.675*	-0.073	0.165*
	(0.011)	(0.149)	(0.31)	(0.251)	(0.314)	(0.385)	(0.061)	(0.096)
DIS	0.001	1.083***	0.223	0.33	-0.366	-0.449	-0.023	0.081
	(0.01)	(0.131)	(0.28)	(0.225)	(0.28)	(0.345)	(0.039)	(0.074)
GS	-0.009	1.259***	0.048	0.896***	-0.15	-0.66*	-0.184***	0.135*
	(0.012)	(0.148)	(0.313)	(0.255)	(0.313)	(0.392)	(0.068)	(0.068)
HD	0.003	0.983***	-0.061	-0.177	0.307	0.525	-0.034	-0.004
	(0.012)	(0.151)	(0.338)	(0.257)	(0.323)	(0.393)	(0.061)	(0.091)
JPM	0.011	1.075***	0.13	0.89***	-0.884***	-0.037	-0.168**	0.101*
	(0.01)	(0.128)	(0.284)	(0.228)	(0.261)	(0.328)	(0.073)	(0.059)
MRK	-0.004	0.695***	-0.499	-0.475**	-0.077	1.047***	-0.227***	0.123
	(0.011)	(0.135)	(0.333)	(0.228)	(0.277)	(0.352)	(0.069)	(0.086)
MSFT	0.012	1.152***	-0.576**	-0.06	-0.055	-0.425	-0.093**	0.095
	(0.009)	(0.117)	(0.251)	(0.202)	(0.243)	(0.298)	(0.039)	(0.074)
NKE	0.013	1.003***	-0.305	0.003	0.321	-0.016	-0.193**	0.438***
	(0.013)	(0.163)	(0.334)	(0.277)	(0.37)	(0.459)	(0.081)	(0.162)
UNH	0.002	0.775***	0.237	-0.147	0.551*	0.372	-0.14***	0.225**
	(0.012)	(0.125)	(0.293)	(0.234)	(0.276)	(0.365)	(0.045)	(0.092)
V	0.022*	0.894***	-0.45	-0.269	-0.329	0.304	-0.22*	-0.087
	(0.012)	(0.13)	(0.294)	(0.246)	(0.352)	(0.401)	(0.114)	(0.113)
$\delta_i^{CB} = \delta^{CB}$, $\delta_i^{SB} = \delta^{SB}$							-0.109^{***}	0.105^{***}
							(0.014)	(0.021)

Note: *p<0.1; **p<0.05; ***p<0.01

4.5 Can Sentiment Be Integrated in a Co-integrated Approach to Mispricing ?

The evidence in the previous section led us to conclude that there is evidence of mispricing of the standard five factor FF model driven both by a sentiment based measure and by a cointegration-based measure. Interestingly, both coefficients are rather homogenous in the cross-section of stocks and it can be validly restricted to be the same, with the coefficient on sentiment-based mispricing taking a positive value and the coefficient on cointegration-based mispricing taking a negative value. The specification most strongly supported by the data can therefore be written as:

$$r_{i,t+1} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{f}_{t+1} + \delta^{CB} u_{i,t}^{CB} + \delta^{SB} u_{i,t+1}^{SB} + \epsilon_{i,t+1}, \tag{16}$$

in system (16) the cointegration-based mispricing proxy captures the effect of deviations of asset prices from their long-run trend determined by factor prices. The (negative) coefficient on this regressor, δ^{CB}, determines the (common)speed with which returns on different stocks fluctuates to adjust in presence of disequilibrium. The larger the coefficient, the faster the price gets back to its long-run trend. Favero et al. (2019) assume that this coefficient is constant over time, and we followed this approach in Section 4.4.

The coefficient on the sentiment based measure, δ^{SB}, is instead positive: sentiments affects returns positively in addition to the standard factors and the cointegration-based proxy. In this section we consider an alternative specifications in which the speed of adjustment is asymmetric between situations of over and underpriced stocks and sentiment affects it.

We adopt the following alternative specification of the basic FF factors augmented factor model (16):

$$r_{i,t+1} = \alpha_i + \boldsymbol{\beta}'_i \boldsymbol{f}_{t+1} + \delta_0^{SB} u_{i,t+1}^{SB} + \left(\delta_0^{CB,+} + \delta_1^{CB,+} u_{i,t}^{SB}\right) u_{i,t}^{CB,+} + \left(\delta_0^{CB,-} + \delta_1^{CB,-} u_{i,t}^{SB}\right) u_{i,t}^{CB,-} + \epsilon_{i,t+1}, \tag{17}$$

in which $u_{i,t}^{+} = u_{i,t}, if\ (u_{i,t} > 0)$ and zero otherwise, while $u_{i,t}^{-} = u_{i,t}\ if\ (u_{i,t} < 0)$ and zero otherwise.

In the light of the results obtained by (Stambaugh and Yuan, 2017) on the asymmetric effect of overpricing versus underpricing related to the arbitrage asymmetry in buying versus shorting, the system (17) allows for a time-varying asymmetry, driven by sentiment, in the speed of adjustment when prices are above or below their long-run trend.

A behavioral model of diagnostic expectations as in Bordalo et al. (2019) and Bordalo et al. (2021) can help to understand why sentiment may affect the time to get back to equilibrium. When agents receive positive news about a company, they overreact and increase their demand for it, which drives the price upwards. Over time, agents learn about the true new fundamental value of the asset, and trade so that its price gets back to equilibrium. However, the speed with which they update their beliefs may be affected by sentiment. Receiving positive signals from an earnings call confirms agents' beliefs about the higher price, which implies that prices will take more time to get back to equilibrium.

The empirical results from the estimation of the extended system specifications are reported in Table (VI).

Table VI

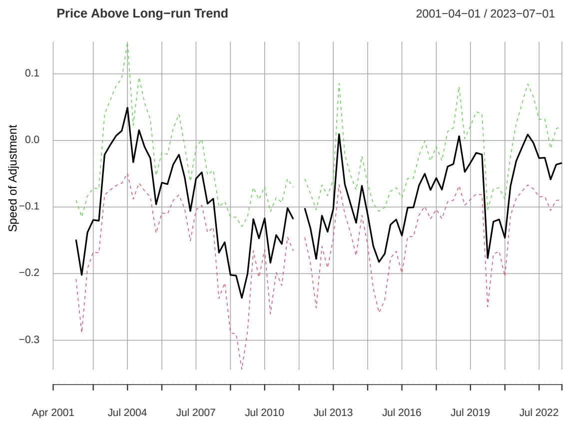
Speed of Adjustment

This table reports the coefficient estimates for the systems (16), and (17) . We include the companies in the Dow Jones Industrial Average index (DJIA), and the sample period is 2001 : $Q1$ to 2023 : $Q2$, at the quarterly frequency. Standard errors are shown in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels respectively.

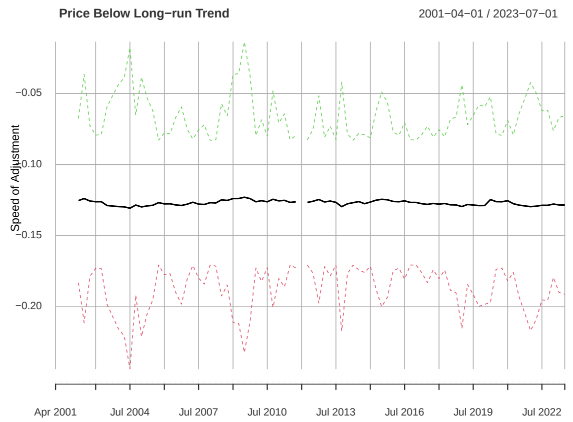
	δ^{CB}	δ^{SB}	δ_0^{SB}	$\delta_0^{CB,+}$	$\delta_1^{CB,+}$	$\delta_0^{CB,-}$	$\delta_1^{CB,-}$
$r_{i,t,t+1}$	-0.109^{***}	0.105^{***}	0.067^{*}	-0.087^{***}	0.444^{***}	-0.127^{***}	-0.012
	(0.014)	(0.021)	(0.035)	(0.026)	(0.177)	(0.027)	(0.193)

The results in Table VI show that there is significant asymmetry in the speed of adjustment which is constant and estimated at -0.127 when prices are below their long-run trend but it is time-varying and positively affected by sentiment when prices are above their long-run trend. A positive sentiment reduces the speed of adjustment when prices are above their long-run trend. Interestingly the independent effect of sentiment on asset returns becomes smaller and signicant only at the ten per cent level when the time-varying non-linearity is introduced. To give an example of how the speed of adjustment evolves over time we report in Figure 7 the behaviour over time of this variable for AAPL (Apple Inc). Figure 7 clearly shows that when prices are above their long-run trend (i.e. when $u_{i,t}^{CB} > 0$) positive sentiment can slow the speed of adjustment and in extreme cases can make it positive implying the prices will shift further away from their long-run trend rather than converging to it. On average the speed of adjustment fluctuates along the value estimated by the system with constant speed of adjustment. When prices are above their long-run trend their speed of reversion is very stable over time.

This non-linearity has relevant implications for predicting returns and their distribution and therefore for asset allocation, risk measurement and risk management.



(a) AXP-Speed of Adjustment



(b) AXP-Speed of Adjustment

Figure 7: The picture depicts the speed of adjustment for AXP, when the price is above the long-run trend (panel (a)) and when the price is below the long-run trend(panel (b). The speed of adjustments for the two types of mispricing are computed as $\left(\delta^{+}_{0,i} + \delta^{+}_{1,i} s_{i,t}\right), \left(\delta^{-}_{0,i} + \delta^{-}_{1,i} s_{i,t}\right)$. The dotted lines represent the 90% confidence intervals.

5 Conclusions

This paper explored mispricing in factor models by examining stock-specific deviations from long-run asset price trends and sentiment indicators derived from quarterly earnings conference calls. Our key empirical finding is that these measures capture significant mispricing, which can be modeled using a time-varying, asymmetric equilibrium correction model. In this model, the speed of adjustment toward the long-run trend, when prices exceed it, is influenced by sentiment.

Incorporating idiosyncratic mispricing measures into standard factor models has important implications for predicting returns and their distribution, and thus for asset allocation, risk measurement, and management. Factor-based strategies for asset allocation and risk assessment can be enhanced by monitoring mispricing metrics obtained from time-series analysis of non-stationary factor-prices and asset-prices, as well as sentiment indicators extracted using Natural Language Processing.

References

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Bordalo, P., Gennaioli, N., Kwon, S. Y., and Shleifer, A. (2021). Diagnostic bubbles. *Journal of Financial Economics*, 141(3):1060–1077.

Bordalo, P., Gennaioli, N., Porta, R. L., and Shleifer, A. (2019). Diagnostic expectations and stock returns. *The Journal of Finance*, 74(6):2839–2874.

Christoffersen, P. F. and Diebold, F. X. (1998). Cointegration and long-horizon forecasting. *Journal of Business & Economic Statistics*, 16(4):450–456.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dong, X., Li, Y., Rapach, D. E., and Zhou, G. (2022). Anomalies and the expected market return. *The Journal of Finance*, 77(1):639–681.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of finance*, 65(4):1237–1267.

Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2):246–273.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Favero, C. A., Melone, A., and Tamoni, A. (2019). Factor models with drifting prices. *Available at SSRN 3418352*.

Favero, C. A., Melone, A., and Tamoni, A. (2020). Factor models with drifting prices. Working Paper.

Gennaioli, N. and Shleifer, A. (2018). A crisis of beliefs. In *A Crisis of Beliefs*. Princeton University Press.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152.

Hassan, T. A., Hollander, S., Van Lent, L., and Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics*, 134(4):2135–2202.

Ippolito, F., Steri, R., Tebaldi, C., and Villa, A. (2023). Mind the gap: The market price of financial flexibility. *Available at SSRN 4553410*.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand.

Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., and Trajanov, D. (2021). Evaluation of sentiment analysis in finance: From lexicons to transformers. *IEEE Access*, 9:29583–29597.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing factors. *The review of financial studies*, 30(4):1270–1315.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

A BERT (Bidirectional Encoder Representations from Transformers)

An important advance in the class of Natural Language Processing (NLP) models has been the introduction of BERT (Bidirectional Encoder Representations from Transformers), a model developed by Devlin et al. (2018). BERT's architecture and training methodology represent a substantial advancement in the field, enabling improved and unmatched performances on a wide range of NLP tasks. BERT is built upon the Transformer model introduced by Vaswani et al. (2017). The Transformer model eschews conventional recurrent or convolutional layers, focusing instead on self-attention mechanisms to process text. Figure 8, which is take from Devlin et al. (2018) illustrates the architecture of a simple BERT model.

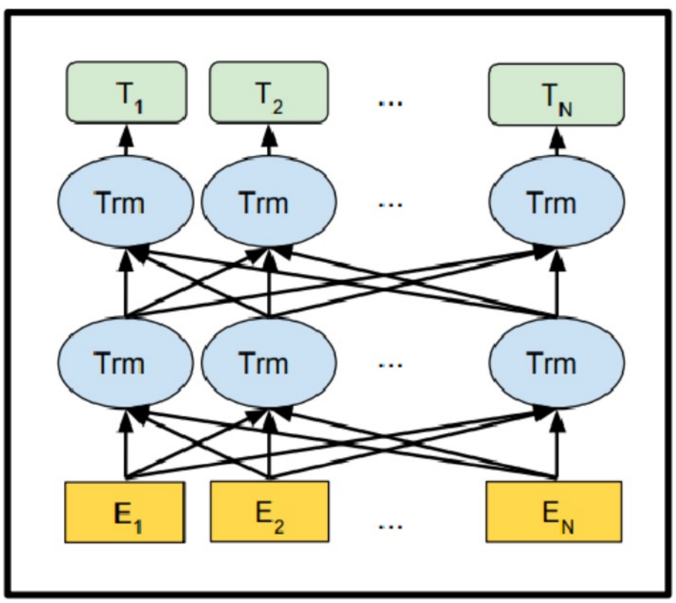


Figure 8: The picture depicts the BERT model architecture.

BERT's training involves two stages: pre-training and fine-tuning. The pre-training stage is unsupervised and utilizes two novel tasks: Masked Language Modeling (MLM) and Next Sentence Prediction (NSP). The fine-tuning stage is conducted after the pre-training, and fine-tunes BERT for specific tasks, wherein the entire model is slightly adjusted. This stage requires significantly less data compared to training a model from scratch. Figure 9, which is taken from Devlin et al. (2018) depicts the pre-training and fine-tuning stages.

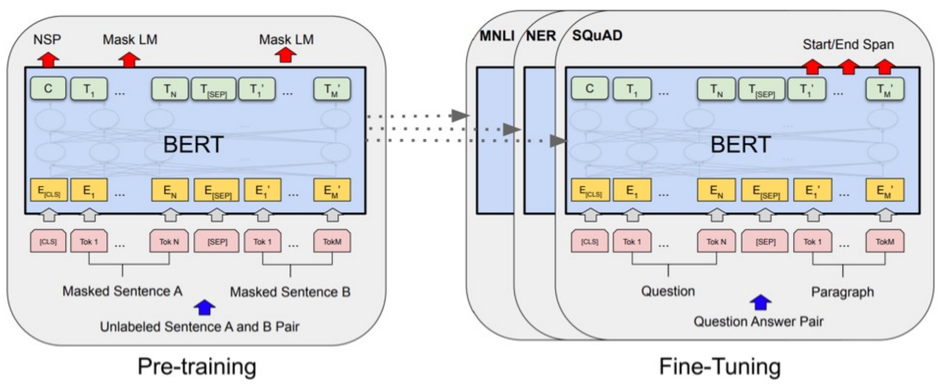


Figure 9: The picture depicts the pre-training and fine-tuning stages for BERT.

BERT's input representation is a blend of WordPiece token embeddings, positional embeddings, and segment embeddings. This approach allows BERT to handle out-of-vocabulary words effectively and provides the model with necessary positional and contextual information.

FinBERT's fine-tuning on specialized NLP tasks has resulted in performances that surpass those of traditional machine learning models, deep learning alternatives, and even fine-tuned versions of the original BERT model. Each fine-tuned variant of FinBERT is designed for a specific purpose and is readily accessible to the public via the Huggingface platform. FinBERT's pre-training encompasses a vast corpus of financial communication texts, amounting to 4.9 billion tokens. This corpus includes 2.5 billion tokens from Corporate Reports (10-K & 10-Q), 1.3 billion tokens from Earnings Call Transcripts, and 1.1 billion tokens from Analyst Reports.