[Unveiling Macro Momentum.]

In partnership with Lombard Odier
Master 203 Financial Markets, Université Paris-Dauphine
Abstract

In the multi-asset realm, risk factors are extensively used by portfolio managers as generators of risk premia. In [5], AQR Capital Management reveals its usage as a strong indicator to generate returns. In this context, the following paper focuses on a preponderant factor — momentum — applied to macroeconomic indicators. The resulting investment strategy, referred to as macro momentum, involves the implementation of a macro signals analysis, and its translation into tactical allocations within a multi-asset portfolio framework.

Keywords: Momentum, macroeconomics, business cycle, portfolio management
Acknowledgements

This report has been conducted with the support and guidance of Lombard Odier. It represents the conclusion of our Cross Cutting Project, realized in the context of the Master 203 Financial Markets from Université Paris-Dauphine, and is the fruit of research and analyses carried out with the help of Thierry Béchu and Benjamin Meyer from the Lombard Odier portfolio management team based in Geneva, Switzerland and Paris, France. We would like to thank them for their continuous and precious advice and counsels.
Contents

1 Introduction 5
2 Literature review 7
3 Signals construction 9
   3.1 Methodology used in our reference paper 9
   3.2 Recovering the signals/asset classes’ relationship 10
      3.2.1 Rebuilding the indicators 10
      3.2.2 Regression analysis on full and sub-periods 12
   3.3 Extension of the signals pool through the use of our own signals 14
4 Portfolio construction and implementation 18
   4.1 Methodology 18
   4.2 Portfolio construction and back-testing 21
   4.3 Performance analysis 23
5 Results 25
   5.1 Optimal choice of indicators 25
   5.2 Empirical findings 27
   5.3 Interpretations 29
   5.4 Integrating the macro momentum strategy 30
   5.5 Further research 33
6 Conclusion 34
1 Introduction

“Now that you’ve seen the evidence and know where to find it, those repeating the [pejorative] myths above regarding momentum should have a harder time maintaining credibility”, Clifford Asness, founder of the AQR investment group in New York.

The theorization of financial markets’ behavior has led to major breakthroughs in our perception and analysis of risk and return. From the efficient market hypothesis to asset pricing models, the search for rational explanations behind market-linked phenomenons is still very much alive today. However, conclusions are often divergent between professionals and academics, and even between academic researchers themselves. Momentum represents an interesting example of such a controversy, as it is the subject of hot debates between “believers” and “non-believers”. With myths slowly emerging around the concept of momentum, some professionals (cf. above quote) have made attempts at quantitatively demonstrating its concrete existence. In the following report, we will make no “a priori” arbitrary assumption on the conceptual validity of momentum, but will rather focus on practical issues linked to indicators construction and portfolio implementation, in order to generate new sources of return.

Let’s first provide a few words of definition on the concept of momentum strategy: it can be described as an investment strategy aiming at benefiting from the perpetuation of existing trends in the market. Hence, it implies buying already performing assets, and selling poorly performing ones over the same period of time. In [11], Jegadeesh and Titman proved that momentum investing between 1965 and 1989 would generate an extra return with a compound annual growth rate of 12.01%, while the S&P500 (SPX Index) yielded an annual rate of 10%. As an intuitive way of implementation of trend-following methodologies, the “momentum strategy” thus rapidly raised a lot of interest.

Considered as the father of momentum investing, Richard Driehaus asserted this strategy. In [15], Schwager reported Driehaus’ words: “Far more money is made buying high and selling at even higher prices. That means buying stocks that have already had good moves and have high relative strength—that is, stocks in demand by investors. I would much rather invest in a stock that’s increasing in price and take the risk that it may begin to decline than invest in a stock that’s already in a decline and try to guess when it will turn around.”

The most crucial topic in our study is the embedding of a macro component inside a momentum strategy. A macro-based strategy aims at capturing and taking advantage of macroeconomic variables moves in order to generate profit through tactical asset allocation. When talking about macroeconomic events, one may for instance think of major interest rates moves, drastic changes in risk sentiment, or shocks on international trade competitiveness.

Let’s now delve deeper into what would be the consequences of combining both concepts, through what is called “macro momentum” investing. It consists in using recent macroeconomic releases from a variety of sources including central banks, governmental or non-governmental institutions as well as private investors’ publications, in order to find fruitful investment opportunities and generate value. This requires a careful scrutiny of contemporaneous and upcoming macroeconomic events, and a thorough analysis of their potential implications on securities prices. Based on intuition, it seems for example consistent to think that any interest rate hike announcement from the US Federal Reserve (Fed) or the European Central Bank (ECB) would have a negative impact on the Government Bonds’ prices. As another trivial example, on the equity market side, our intuition would lead us to think that an improving consumer confidence would have a positive impact on equity returns.

Therefore, the combination of tactical asset allocation momentum strategies with macroeconomic indicators analysis seems to represent a potential substantial source of profit: this is what we are going to study in this paper. After carrying out an investigation on the academic literature linked to macro momentum strategies, we will try to construct an aggregated indicator using various relevant macroeconomic signals in order to shape our allocation strategy. This will then lead us to define a portfolio construction methodology, translating the signals analysis into concrete

\[1\text{Cf. [5]}\]
strategic investment decisions. Finally, this paper will provide some important empirical results and interpretations arising from our study.
2 Literature review

Discussions on the implementation of a profitable “macro momentum” strategy in the context of asset allocation are far from new: while our reference paper is recent (March 2017), many other asset managers and money market players had already investigated this category of strategies. Evli Fund Management Company has for example demonstrated the added value of such an approach in signaling favorable periods for tactical asset allocation decisions in an environment characterized by increasing returns and falling draw-downs and volatility. This idea of a superior signaling capacity is also highlighted by the existence of specific investment horizons where an increase in risky exposure (e.g. equity) and a decrease in fixed income positions could generate positive returns.

Macro momentum, and momentum in a more general fashion, also appear extensively in the factor investing space. In the realm of investing, momentum stands high as a strong excess return generator alongside quality, value and size factors. Typically, combined with the value factor, momentum has historically proved to bring diversification benefits because of the surprisingly low to negative correlation it exhibits with this factor. This advantageous characteristics have incentivized not only asset managers to enter this type of strategy, but also indices providers to embed them in all types of structures (e.g. MSCI World Momentum Index).

Intuitively, one could claim that macro momentum goes against the concept of diversification because of its intrinsic pro-cyclical nature. In [12], Moskowitz et al. have nonetheless shown that across all asset classes, momentum strategies generate abnormal returns with little exposure to standard asset pricing factors and with best performance during stress periods. This last point is crucial for the development of this project. Indeed, building on the idea that macro momentum could bring diversification to an already existing portfolio, one could consider this approach as an overlay strategy - a point that will be investigated and tested further in this report.

Applying the concept of factor investing to macro momentum would imply the opportunity for investors to earn potentially significant returns in exchange for taking on a corresponding “macro risk” (characterized by the exposure to a generalized macroeconomic factor). This is the hypothesis developed by Horsting in [10], who was able to demonstrate that momentum does exist and is significant in a sufficiently large macroeconomic context and outside homogeneous markets only (e.g. stocks). However, the author expressed several reservations regarding the characteristics of such a factor. First, the risk-reward relationship of macro momentum has not been empirically demonstrated, with no direct link exhibited between macro momentum returns and pro-cyclical risk. With this result, it would thus appear that macro momentum does not directly and proportionally reward investors for their increased exposure to the current market cycle. Additionally, “winners” and “losers” of macro momentum strategies did not exhibit consistently different risk exposures to predetermined factors: it weakens even more the definition of macro momentum as a risk factor (considering that no other preponderant risk factors have been omitted in the study). Finally, the attempt made at using economic signals as predictive variables yielded poor results in the forecasting of the reference indices price moves. Nevertheless, and despite these restrictions emitted by the author, it remains true that macro momentum can be exploitable as a tool to unlock new sources of value creation.

Apart from conventional asset pricing models (e.g. CAPM) that took time to acknowledge the existence of a momentum factor, the research field of behavioral finance have also questioned the presence of such a phenomenon. Barberis, Shleifer, and Vishny in [3]; Daniel, Hirshleifer, and Subrahmanyam in [4] and; Hong and Stein in [9] have thus suggested the idea that momentum is caused by an underreaction of agents to private and public information. Momentum, due to a delayed reaction to new pieces of information, should then lead to a slow price increase that will overshoot its fundamental value, this overreaction ultimately causing a reversal of the price. According to the paper, the horizon of such a phenomenon would be variable across asset classes and continuously changing, from 6 to 12 months.

Chordia and Shivakumar show in [6] that profits made through momentum strategies can be explained by a set of lagged macroeconomic variables. This implies that once they are adjusted for these macroeconomic variables, these strategies’ payoffs are null : these authors therefore support the possibility of exploiting macroeconomic information in order to set in place a profitable tactical
asset allocation strategy. As they use lagged macroeconomic variables in order to prove their results, the main challenge remains to find consistent indicators that prove to have a substantial predictive power.

All in all, while the intuition behind macro momentum strategies is clear if we are referring to fundamental factor investing theories, researchers have highlighted many gray areas around the characteristics of such returns. The questions of knowing whether it is diversifying or not, and whether it is performing best in periods of growth or during crisis, are not definitely answered.
3 Signals construction

3.1 Methodology used in our reference paper

Our work finds its roots in a paper of AQR Capital Management [5], published in 2017. Using historical data from a number of sources, the paper focuses on building a global macro momentum strategy based on several macroeconomic indicators, supposed to act as signals emphasizing the improvement or deterioration of macroeconomic trends. Hence, according to the signal given by the predetermined indicators, the macro momentum strategy implies going “long assets for which fundamental macroeconomic trends are improving, and short assets for which fundamental macroeconomic trends are deteriorating”. It uses four different macroeconomic indicators, each one underlining a specific aspect of macroeconomic trends:

- Business Cycle
- International Trade
- Monetary Policy
- Risk Sentiment

The construction of the indicators is the following:

The Business Cycle indicator is built using a 50/50 combination of data on GDP growth and Inflation. In order to be able to capture the business cycle momentum and future implications on the markets, the data used is based on economic forecasts. Nonetheless, the paper differentiates the period from 1990 onward, from the prior 20 years period running from 1970 to 1990. For the most recent period, the indicator is built using one-year differentiations on GDP growth and CPI inflation forecasts, whereas for the previous period, the methodology used is based on one-year differentiations in realized year-on-year real GDP growth and CPI inflation. The paper further assumes that lagging this data by one quarter is “equivalent to changes in forecasts assuming that real GDP growth and CPI inflation follow random walks”. The previous statement seems rather naive and may be called into question, as the proxy used here to derive a forecast value from realized GDP growth and inflation may have simply been used to deal with a lack of forecast data over the period 1970-1990. Indeed, the paper does not further explain the link between realized data on GDP growth or inflation “lagged one quarter”, and forecasts. The methodology is far from being straightforward, as getting data “lagged one quarter” could be used in order to backtest the strategy on a past period, but cannot be used for the real-time implementation of the strategy since data on future realizations of GDP growth and inflation are not contemporaneous. Using this type of proxy involves a strong bias in the backtest methodology, and would not have been effectively implementable by a fund over this period.

The International Trade indicator is constructed by AQR based on “one-year changes in spot exchange rates against an export-weighted basket”. This choice of indicator seems a priori consistent as fluctuations in exchange rates have an altering effect on relative prices of imported and exported goods between countries, and therefore on the trade balance.

The Monetary Policy indicator is built from the front end of the yield curve, therefore on short-term interest rates levels. As for the Business Cycle indicator, the data treatment is split between two periods separated by the year 1992. Prior to 1992, AQR’s methodology for the Monetary Policy indicator is based on one-year changes of the Libor, whereas for the most recent period, it is based on one-year changes in the two-year yields. No more precisions are given neither for the choice of this time split nor for the fact that two different time series are used. Moreover, Libor’s inception dates back to 1984, and the paper does not give precisions on the data chosen for this indicator for the period ranging between 1970 and 1984. We believe that AQR’s recurrent use of heterogeneous indicators in order to represent one single aspect of the economy would have required a more in-depth justification, or at least additional details regarding the smoothing and scaling methodologies applied.
The *Risk Sentiment* indicator is built from the equity market excess returns, using one-year changes as for the other signals. This choice of signal is intrinsically linked to the invalidation of the efficient market hypothesis: once having empirically acknowledged that investors don’t act as pure rational agents, it can be observed that sentiment beliefs have a significant impact on the formation of stock prices. This is the reason why the choice of equity market excess returns as an indicator of risk sentiment seems consistent, both variables being linked through a reciprocal influence dynamic.

Once each indicator has been built, the strategy built by AQR uses their historical distribution in order to allocate resources between four different asset classes:

- Equity
- Government Bonds
- Currencies
- Interest Rates

Our work will focus on the study of the two first asset classes (Equity and Government Bonds), for which we aim to build a macro momentum overlay strategy used to bring diversification to a conventional multi-asset portfolio composed of equity and bonds.

The paper’s strategy relies on qualitative relationships between the chosen indicators and the different asset classes. The summary of these relationship for Equity and Bonds is given in Table 1:

<table>
<thead>
<tr>
<th>Indicators / Asset Classes</th>
<th>Business Cycle</th>
<th>International Trade</th>
<th>Monetary policy</th>
<th>Risk Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth (increasing)</td>
<td>Inflation (increasing)</td>
<td>Increasing Competitiveness</td>
<td>Policy Tightening</td>
</tr>
<tr>
<td>Equity</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Bonds</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Prior to building our own macro momentum portfolio, it was therefore necessary to first try to retrieve these qualitative relationships, while fitting as best as possible the original data used by AQR. Indeed, the construction of both the portfolio in the reference paper and our overlay strategy using equity and bonds are based on the relative intensity of the chosen signals and their relationships with the different asset classes. Hence, the first step was to try to gather the same data as the one used in the paper, and fit regressions between the indicators values and the two asset classes we chose to study.

### 3.2 Recovering the signals/asset classes’ relationship

#### 3.2.1 Rebuilding the indicators

As explained above, the paper’s strategic allocation is based on qualitative relationships between the chosen signals and the investment environment, separated between asset classes.

We first tried to rebuild Table 1 using data on the U.S. However, as we were not able to retrieve the full set of data used by AQR, we tried to build the closest possible indicators with available data.

The set of data we used was the following:
— Business Cycle indicator

We were not able to retrieve the exact data used by the paper. Without an access to the Consensus Economics Database (used by AQR to get forecasts on GDP Growth and Inflation from 1990 onwards), we could not find sufficiently granular data on GDP Growth Forecasts. Nonetheless, we were able to find quarterly data on CPI Inflation forecasts dating back to 1970 (beginning of the paper studied period). Hence, we used one year changes of this time series over the whole period (1970-2018), instead of following the contested methodology “lagged one quarter”, described by the paper for the period prior to 1990. Regarding the GDP Growth, we chose to use year-on-year changes in realized value, thinking that this would create less bias than lagging the data, and could be further used to build our portfolio until the current period.

— International Trade indicator

As presented above, the paper uses one-year changes in spot exchange rates against an export weighted basket, without further description of the methodology used. As our investment universe is firstly based on the US, we chose to use the U.S. Dollar Index (DXY), as a proxy for this indicator. This index represents a measure of the value of the USD compared to a weighted basket of the most traded currencies. The index is maintained, and published by the Intercontinental Exchange, with the name “U.S. Dollar Index” a registered trademark, and is build as the weighted geometric mean of the dollar value relative to this basket of currencies. The Index goes up when the USD gains strength compared to other currencies. Thus, it corresponds to the description made by AQR of their International Trade indicator, and we used one-year changes in the index value in order to build our own indicator.

— Monetary Policy indicator

AQR builds its indicator on a split between the period 1970-1992 and 1992-onwards, using one year changes in two year-yields for the most recent period and “Libor and its international equivalents”, for the period prior to 1992. We reproduced this methodology in our US case using one-year changes in the US 2-year government bond yield for the period running back to 1992, and one-year changes in the US FED Funds Rate for the previous period.

— Risk Sentiment indicator

The original Risk Sentiment indicator built by AQR relies on one-year changes in equity market returns. We were able to rebuild an indicator according to this methodology, and relative to our US framework. We built a risk-free rate security over the period, starting at a base of 100 on January 1st 1970, and further invested at the daily Fed Funds Rate until 2018. We then calculated one year changes of both the S&P500 Total Return and of the previously described security invested at the Fed Funds Rate. Our “one-year equity market excess return” was then calculated as the difference of the two.

As described in Table 1, AQR’s macro momentum strategy is based on the following assumptions regarding the relationships between indicators and asset classes:

— An increasing growth is assumed to be bullish for equities and bearish for fixed income. Increasing inflation is assumed to be bearish for equities and bullish for fixed income.

— A depreciating currency, leading to an improvement of the International Trade indicator through an increased competitiveness, is assumed to be bullish for equities, and bearish for fixed income.

— An expansionary monetary policy is assumed to be bullish for equities and bearish for fixed income, whereas a tightening policy is assumed to be bearish for equities and bullish for fixed income.

— An increasing risk sentiment is assumed to be bullish for equities and bearish for fixed income.
3.2.2 Regression analysis on full and sub-periods

We first attempted to prove back AQR paper’s results, running regressions between the reconstructed signals and our preferred asset classes: Equity and Bond.

In order to represent both the US Equity and Fixed Income markets, we used the following indices:

— An aggregated index for the Bond Market: the Barclays Aggregate Bond Index. Created in 1973 under the name “Lehman Aggregate Bond Index”, it is currently maintained by Bloomberg. The index is market-cap weighted, according to the market size of each constituent bond. It contains a large range of US Bonds (including Treasury securities, Government agency Bonds, Mortgage-backed Bonds, and Corporate Bonds) and thus represents well the overall US Bond Market.

The aforementioned regressions gave us very poor results over the whole period, leading us to think that the relationships described in the AQR paper were more based on a theoretical and qualitative aspect than on thoroughly studied quantitative relations.

Figures 2 & 3 allow to take a look at the regression results obtained for each asset class using the five reconstructed indicators as explanatory variables over the period 1980-2018. Figures 4 & 5 represent the equivalent regression results over the most recent period 2000-2018.

In all four cases, we observed very low to null R-squared, implying that the linear regressions did not yield statistically relevant relationships. Moreover, all coefficients were also close to 0: our regressions did not achieve to prove AQR paper’s assumptions on the relationships between asset classes and indicators. Nonetheless, this seems rather odd since most of the relationships make a lot of economic sense. We should obviously get some sort of negative relationship between the Monetary Policy indicator and Bond returns, based on the intrinsic construction of the indicator. An increase in the Monetary policy indicator should reflect a tightening policy, leading to a rate increase and thus to a depreciation in the current Bond market.

The poor regressions results might have been at first glance a consequence of excessive data manipulation. For instance, using time series with different periodicity and merging them on the same monthly basis could have generated non-negligible biases in the regressions. We tried similar regressions on various time periods in order to test the existence of such relationships in different regimes (periods of prosperity vs. extreme stress periods), but it didn’t yield any improvement in the results. We then thought about using logistic regressions, instead of linear ones, in order to display the qualitative relationships between the indicators and the chosen asset classes. But once again we didn’t obtain significantly more relevant results. Our last attempt was to fit polynomial regressions, but following another failure, we finally believed that only the introduction of more complex relationship characteristics would be able to yield sufficiently exploitable results (for example the use of Sieve Reduced Rank Regressions in order to display highly non-linear relationships as used in [1]).

In order to go further in our study, we assumed AQR’s stated relationships to be verified in order to start constructing our portfolio allocation methodology. We decided to go forward on this part because, as the AQR paper itself shows, indicators did not individually drive the performance of the strategy in a linear way over the studied investment period. The main interest of the methodology relied in the aggregation of the various signals, we therefore decided not to spend too much time on the analysis of individual predictive powers. Finally, the asset allocations resulting from the aggregated signals proved to be appropriate relative to the respective time periods. Thus the indicators, despite not yielding satisfying individual regression results, were able to provide appropriate investment timings for long and short positions in Equities and Bonds (which remains the most important in the context of our study).
### Figure 2: Equity Regression 1980-2018

| Dep. Variable: | Equities | R-squared: | 0.009 |
| Model: | OLS Adj. R-squared: | 0.002 |
| Method: | Least Squares | F-statistic: | 0.7839 |
| Date: | Sat, 28 Apr 2018 | Prob (F-statistic): | 0.562 |
| Time: | 09:58:54 | Log-Likelihood: | 793.48 |
| DF Residuals: | 450 | BIC: | -1549. |
| DF Model: | 5 | |
| Covariance Type: | nonrobust | |

| coef | std err | t | P>|t| | [0.025 | 0.975 |
|------|---------|---|---------|---------|---------|
| const | 0.0004 | 0.002 | 3.813 | 0.000 | 0.004 | 0.013 |
| Monetary Policy | 0.0001 | 0.001 | 0.110 | 0.913 | -0.002 | 0.002 |
| International Trade | 0.0199 | 0.022 | 0.917 | 0.368 | -0.023 | 0.002 |
| Risk Sentiment | 0.0001 | 0.012 | 0.656 | 0.512 | -0.016 | 0.033 |
| Growth | -0.0008 | 0.001 | -0.938 | 0.349 | -0.002 | 0.001 |
| Inflation | -0.0017 | 0.001 | -3.861 | 0.167 | -0.004 | 0.001 |

Omnibus: 57.491 Durbin-Watson: 1.943
Prob(Omnibus): 0.000 Jarque-Bera (JB): 135.424
Skew: -0.915 Prob(JB): 3.92e-30
Kurtosis: 5.323 Cond. No. 29.4

### Figure 3: Bond Regression 1980-2018

| Dep. Variable: | Bonds | R-squared: | 0.033 |
| Model: | OLS Adj. R-squared: | 0.022 |
| Method: | Least Squares | F-statistic: | 3.085 |
| Date: | Sat, 28 Apr 2018 | Prob (F-statistic): | 0.08948 |
| Time: | 09:58:54 | Log-Likelihood: | 1261.1 |
| DF Residuals: | 450 | BIC: | -2485. |
| DF Model: | 5 | |
| Covariance Type: | nonrobust | |

| coef | std err | t | P>|t| | [0.025 | 0.975 |
|------|---------|---|---------|---------|---------|
| const | 0.0073 | 0.001 | 1.633 | 0.000 | 0.006 | 0.009 |
| Monetary Policy | 0.0003 | 0.000 | 0.068 | 0.493 | -0.006 | 0.008 |
| International Trade | -0.0021 | 0.000 | -0.270 | 0.787 | -0.017 | 0.013 |
| Risk Sentiment | -0.0148 | 0.004 | -3.325 | 0.001 | -0.024 | 0.006 |
| Growth | -0.0002 | 0.000 | -0.520 | 0.590 | -0.001 | 0.000 |
| Inflation | -0.0004 | 0.000 | -0.871 | 0.384 | -0.001 | 0.000 |

Omnibus: 108.482 Durbin-Watson: 1.610
Prob(Omnibus): 0.000 Jarque-Bera (JB): 776.473
Skew: 0.785 Prob(JB): 9.45e-170
Kurtosis: 9.244 Cond. No. 29.4

### Figure 4: Equity Regression 2000-2018

| Dep. Variable: | Equities | R-squared: | 0.845 |
| Model: | OLS Adj. R-squared: | 0.822 |
| Method: | Least Squares | F-statistic: | 1.965 |
| Date: | Sat, 28 Apr 2018 | Prob (F-statistic): | 0.0051 |
| No. Observations: | 216 | AIC: | -756.3 |
| DF Residuals: | 210 | BIC: | -736.0 |
| DF Model: | 5 | |
| Covariance Type: | nonrobust | |

| coef | std err | t | P>|t| | [0.025 | 0.975 |
|------|---------|---|---------|---------|---------|
| const | 0.0067 | 0.003 | 1.977 | 0.049 | 1.0e-05 | 0.003 |
| Monetary Policy | 0.0009 | 0.003 | 0.317 | 0.007 | -0.001 | 0.013 |
| International Trade | 0.8343 | 0.836 | 0.942 | 0.347 | -0.037 | 0.186 |
| Risk Sentiment | -0.0020 | 0.024 | -0.084 | 0.933 | -0.050 | 0.046 |
| Growth | 0.0018 | 0.002 | 0.857 | 0.292 | -0.002 | 0.005 |
| Inflation | -0.0045 | 0.002 | -2.497 | 0.013 | -0.008 | 0.001 |

Omnibus: 8.268 Durbin-Watson: 1.886
Prob(Omnibus): 0.016 Jarque-Bera (JB): 10.236
Skew: -0.207 Prob(JB): 0.80999
Kurtosis: 3.886 Cond. No. 32.6
3.3 Extension of the signals pool through the use of our own signals

The indicators extracted from [5] may, at first sight, seem a little restrictive. Moreover, as the regressions we have performed revealed not to be fruitful, we decided to perform a more thorough research to determine what would be the most accurate macroeconomic indicators in order to generate significant and sustainable returns.

In table 2, we referenced a list of alternative macroeconomic indicators we decided to try to use as signals in order to build our macro momentum indicator. In order to stay consistent with the initial purpose of this paper, we have based our study on the general categories proposed by AQR Research, but tried to replace the proposed indicators by consistent alternative ones of our choice. We believed that putting into question the indicators already provided by AQR, and proposing different ones through both a qualitative analysis and a trial and error process would be a crucial step towards the enhancement of our own portfolio building strategy.

Table 2: Other Macroeconomic Indicators Used As Signals

<table>
<thead>
<tr>
<th>Business Cycle</th>
<th>International Trade</th>
<th>Monetary Policy</th>
<th>Risk Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth rate of US Industrial Production (USIP)</td>
<td>US Expected Inflation (University of Michigan)</td>
<td>Trade Weighted U.S. Dollar Index (DTEXM) (Federal Reserve Bank of St. Louis)</td>
<td>TED Spread (difference between 3m T-Bill and LIBOR)</td>
</tr>
<tr>
<td>US Consumer Confidence Index (CCI)</td>
<td>3-Month T-Bill</td>
<td>VIX Index (S&amp;P500 options volatility)</td>
<td></td>
</tr>
<tr>
<td>Purchasing Managers’ Index (PMI)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Let’s now have a closer look at the theoretical and empirical foundations of our chosen alternative indicators:
— Business Cycle indicators

Thanks to the research of the The Conference Board published in [13], we have been able to identify three different categories and thus analyze their meaning and relevance to our own research. Here are the three aforementioned categories of indicator:

- **Leading**
- **Coincident**
- **Lagging**

A *leading* indicator is defined as a tool with predictive power. In the context of macro-momentum, a leading indicator would tend to give an insight on what would potentially be the next phase of the economic cycle - we can then define it as a “forward looking indicator”. That is, its direction shifts in advance of the business cycle. At first sight, one may use manufacturing average weekly hours or manufacturers’ new orders, but collecting data on this kind of indicator may prove to be difficult or at least incomplete and potentially biased. Thus, we chose to use as leading indicator the Consumer Confidence Index, for which we were able to collect precise data from the Federal Reserve of St. Louis. Moreover, the main reason that orientated this choice was that it is one of the only leading indicator that is at the same time fully expectation-based and reflecting change in consumer attitudes concerning the future environment.

A *coincident* indicator, such as production or employment actual levels, define the contemporaneous state of the business cycle. Their behavior is supposed to be in synchronicity with the economic activity. We first chose the PMI as a coincident Business cycle indicator to test as it had been used by several research papers and practitioners. For instance, Evli Fund Management [2] has built a strategy based on a macro momentum tactical allocation, using PMI as the main measure of macroeconomic momentum. They tried to demonstrate that global macro momentum added significant value and diversification in the context of a tactical asset allocation, and thus could be used as an overlay for a broader multi asset strategy. Therefore, they compared the performance of a classical 50/50 Equity/Bond portfolio on European Markets — using the STOXX 600 Europe Total Return as equity reference, and the iBoxx Euro Sovereigns for bond allocation — to a portfolio that over/underweights allocations on each assets based on the estimation of global macro momentum. This estimation is based on the JP Morgan’s Global manufacturing PMI, and uses boundaries at 30%/70% levels on each side, depending on the relative overweight or underweight of Equity over Fixed Income, respectively in a state of positive or negative global macro momentum. Moreover, the PMI is often referred as a relevant indicator to predict growth in the economy, thus a positive business cycle trend. In [14], Rossiter shows for example using bridge equations, that the PMI is a useful tool to forecast improvements in global growth. Beyond academics and practitioners, PMI is also considered by international institutions as one of the main business cycle indicators. For instance, the European Commission in a its July 2017 technical paper on European Business Cycle Indicators [8], defined the PMI as one of the “two most commonly used indicators for tracking euro-area GDP growth”.

We also chose to select, as a coincident indicator to test, the growth rate of US industrial production. This figure is based on the monthly raw volume of goods produced by industrial or manufacturing firms (mines, electric utilities, newspaper, book publishing etc.). This is a major leading indicator and is watched very closely by the Fed, first because inflation logically appears at the industrial level before being translated to individual consumers. But most importantly because the relatively high volatility observable in the nominal output during peak periods, it has been empirically verified that important variations in growth rate of industrial production can be a robust indicator of business cycle shift.

Finally, a *lagging* indicator depicts the economic activity with a lag period. They tend to shift direction once the business cycle has already been impacted. The fact that they are lagging may seem useless in our study, which goal is to predict the economic cycle in order to make profit out of it. However, it appears that the lagging indicators can represent the cost of doing business. For the Conference Board, “an accelerated rise in the lagging indicators, which often occurs late in an expansion, provides a warning that an imbalance in rising costs may be developing”. But even with the knowledge of this potential explanatory power, we decided not to implement lagging indicators in the context of our work, putting the emphasis on leading and coincident ones.
— International Trade indicators

We believed that the choice made by AQR of using an export-weighted foreign exchange rate index was the most consistent. Indeed, several academic studies supported the existence of an inverse relationship between exchange rate and trade balance, given that a certain number of conditions on macroeconomic adjustment mechanisms are met (especially concerning flexibility of imported goods border prices, elasticity of end consumer’s demand to such price variations, and translation into rational consumption choices between local or imported goods). Therefore, even if exchange rate levels do not entirely explain trade imbalances between countries, they have a non-negligible effect on trade: an export-weighted spot exchange rate index for the USA appeared to be the one with the most robust explanatory power. We thus tried to gather data representing the same figure, but coming from a different source than the one used by AQR in order to see if a slight variation in the index construction methodology would yield more interesting results.

— Monetary Policy indicators

We first chose to test the 3-Month T-Bill as it is one of the best indicator of variations in short-term yields. This focus on short-term interest rates as indicators of Monetary Policy is widely justified by two empirical phenomenon linking the US central banking system and short-term yields. First, the Fed is regularly acting on the Treasury bonds market in order to regulate the monetary aggregates, therefore influencing short-term interest rates through supply and demand dynamics. Second, in an even more straightforward fashion, the central bank can affect the Fed funds rate through signaling mechanisms. The Fed is therefore in a position to have a preponderant and direct influence on these rates, knowing moreover that the very short duration and important liquidity characterizing these securities minimize their risk-premium component. In second place, we chose to test the US Expected Inflation. This is tightly linked to the fact that in order to carry out its monetary policy, the Fed looks very closely to forward looking inflation indicators - price stability being one of its two core objectives.

— Risk Sentiment indicators

We first chose to test an indicator closely related to stock market’s returns volatility, acknowledging the relationship between price movements and investors’ economic sentiment.

This tight link between both figures can be explained by several investors’ behavioral biases such as overreaction (excessively emotional reactions to newly available information), overconfidence (overestimation of one’s capacity to successfully perform a task), herding (tendency of individuals to mimic actions of a broader group), anchoring (attachment of thoughts to an unfounded reference point), prospect asymmetry (different valuation of gains and losses), confirmation bias (overweighting of information confirming the investor’s original idea), hindsight bias (past events considered as obvious and predictable) or availability bias (decisions oriented towards interpretation of more recent information). These cognitive biases lead to a translation of agents’ subjective perceptions in stock market price movements: investor sentiment can therefore be considered as a predominant factor explaining stock volatilities.

We first considered using a realized volatility index in order to express the risk sentiment, but as our work aims at finding signal with exploitable predictive power, we orientated our choice towards a forward looking volatility indicator. Therefore, we chose the VIX Index as it provides a clear view on the forward looking implied volatility. It estimates expected volatility by averaging the weighted prices of SPX puts and calls over a wide range of strike prices. An important precision to give is that the relationship between VIX and Equity or Bond returns is non-linear: above a certain threshold (around 80%), the sensitivity is inversed due to the shift towards an extreme stress period in the economy (this phenomenon is partly explained by funding constraints endogenously linked to market volatility levels, translating into flight-to-safety reallocations). In this case, the lagged VIX starts to have a negative impact on stock returns, and a positive impact on Bond returns. As a second alternative Risk Sentiment indicator, we chose the Ted Spread, exhibiting the difference between 3-month Treasury Bill and 3-month LIBOR based on US dollars. This indicator seemed complementary to the previous one, as it is based on fixed income and not equity markets: it gives a proper view on the perceived difference in credit risk between government and banking industry.
related securities. An increase in the Ted Spread implies that the default risk on interbank loans is considered to be increasing.

Since the regressions implemented on the new indicators did not yield better results than on the initial ones used by AQR, we will further on assume the following qualitative relationships between our alternative indicators and the studied asset classes, in order to build our optimized portfolio (summed up in Table 3):

Table 3: Impact of the alternative indicators on asset classes

<table>
<thead>
<tr>
<th>Indicators / Asset Classes</th>
<th>Business Cycle</th>
<th>International Trade</th>
<th>Monetary policy</th>
<th>Risk Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PMI</td>
<td>US CCI</td>
<td>USIP Growth</td>
<td>DTEXM</td>
</tr>
<tr>
<td>Equity</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bonds</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
4 Portfolio construction and implementation

4.1 Methodology

As mentioned in the Literature review, a simple implementation of a momentum strategy consists in buying assets which have recently performed well, while selling assets which have performed poorly. Holding this portfolio of assets for a period ranging from 3 to 12 months should then yield positive returns.

Going back to our reference paper, it appears that AQR had put in place quite a different and more complex structure for its portfolio construction. The method used within each asset class and each theme (category of indicator) can be summarized and divided into two parts:

— Long-short portfolios: those portfolios have long (resp. short) positions in assets with beneficial (resp. detrimental) macroeconomic trends relative to the cross-sectional average. Additionally, they are market neutral.

— Directional portfolios: those portfolios have long (resp. short) positions in assets with beneficial (resp. detrimental) macroeconomic trends regardless of other markets. Additionally, they are market neutral on average.

Using intermediary “Asset Class” and “Thematic” portfolios, they were then able to build an “Aggregate Macro Momentum” portfolio using an equally weighted average across all their portfolios (16 in total). Indeed, the paper insists on the fact that “the performance of the hypothetical composite macro momentum strategy is neither driven by a single asset class (i.e. trading on all macro momentum themes, but only in equity markets) nor by a single theme (i.e. trading on macro momentum in all markets, but only using monetary policy trends)”.

The methodology described above seemed really appealing, as it created on the one hand easily aggregable and balanced long-short portfolios with similar risk levels, while on the other hand providing diversification and limiting individual markets risk for the directional portfolios. Nevertheless, considering the disappointing results from our first analysis,—which intended to quantitatively demonstrate the relationships between the macro momentum signals and the selected assets—it was not possible to replicate such a procedure. We thus implemented a more intuitive and straightforward approach.

In our first attempt to construct our own macro momentum portfolios, we developed the following approach. At each optimization date, we considered each macro momentum signal independently with the objective of constructing four resulting portfolios (one by macro signal) that would then be aggregated.

— According to each signal’s sign, we determined if we would have a long or short position in each asset class using the qualitative relationships given by AQR (see Table 1).

— According to each signal’s magnitude, we then determined a corresponding intensity, choosing from two different possible intensities as shown in Figure 6.

Figure 6: Signal intensities using medians
— If the signal was positive, we compared it to the median of all the signal’s past positive values:

- If the value was above the computed median, we allocated a high intensity (2) to this signal observation.
- If the value was below the computed median, we allocated a small intensity (1) to this signal observation.

— A symmetric approach was followed in the case of a negative signal.

— Having determined these different investment allocation intensities, we then optimized the allocation between the two asset classes using 10-year rolling returns and a target volatility level of 10%.

— Finally, we linearly aggregated the four portfolios obtained, rescaling the volatility of the aggregated output. We thus obtained an aggregated portfolio coming from the best signal-driven optimized allocation between our two assets at each optimization date.

The next step of our work was to improve this rather simple portfolio construction methodology. We wanted to make the signal treatment more granular, \textit{i.e.} to have a more precise view on the relative signal intensities. A first way to do so was to introduce a whole set of quantiles, as shown in the graph displayed in Figure 7.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{signal_intensities.pdf}
\caption{Signal intensities using four quantiles}
\end{figure}

However, it seemed that a too large granularity was not beneficiary. One of the drawbacks of using too many quantiles was the instability of the optimization (\textit{i.e.} a few additional data points might change the boundaries). Such an instability was contrary to our search for an improved robustness. Moreover, too many quantiles would cause the optimization to be too restrictive (\textit{i.e.} the quantile rank of the indicator forces the exact asset allocation, letting no space for optimization). Through trial and error, we finally found that the optimal number of quantiles would vary from 1 (median only) to 3 (quartiles), depending on the method used.

Another way of improving the portfolio construction methodology was to use Z-score with the same window of data (from 0 to \(t\)). This method seemed easier to interpret. It allowed us to calibrate the model by finding a “threshold” Z-score (or even several thresholds to add granularity) that could better discriminate the indicators values, looking at the relative magnitude of the signal compared to the predetermined thresholds (Figure 8). One first limit to this approach was that the volatility metric (and therefore Z-scores) changed at each optimization date, taking into consideration the fact that new values were being sequentially added to the data. On the other hand, the choice of a constant threshold for the 40 years of optimization seemed limited: we therefore chose to develop a more dynamic approach.
In the Z-score methodology, we did not distinguish positive and negative values. We computed the Z-score of the indicator at time $t$ based on all the data available from 0 to $t$. In our first simple implementation, we determined one negative threshold and one positive threshold. If the Z-score was in-between them, the signal was discriminated and assigned a value of 0. Hence it had no influence on the portfolio construction. If it was above the positive threshold or below the negative threshold, it was considered as a strong signal and therefore impacted the optimization. The choice of the thresholds was delicate, but we made it dynamic and dependent on the signal.

We found out that a discriminating rule was too restrictive and yielded poor results. We thus preferred a rule closer to our first intuition linked to quantiles, as shown in Figure 9.

- If the Z-score threshold was too restrictive, the optimization would rely on only one or even no signal at each date, therefore either yielding an investment based on one signal only (the others being discriminated), or no optimization at all (every signal being ignored).
- On the contrary, if the Z-score signal was too lax, the portfolio had very poor results and we decided not to go further in this direction.

About the additivity of signals, the AQR paper insists on the fact that macro momentum’s interest does not rely in one signal in particular, but in the combination of all the four. Hence, in our view, we should not combine the strength of the signals before doing portfolio optimizations, as it might lead us to miss a major part of the value-added. Let’s take an example. If two signals have negative effects on the same asset class (e.g. positive growth and restrictive monetary policy on equities), this would cancel out and the signal would be not to invest in equities at all. However if we build a portfolio for each signal, and then aggregate both portfolios, the aggregated portfolio will combine the value of each signal. This is not so obvious when our investment universe is composed of only two assets (equities and fixed income). But it becomes very important as we increase the investment universe. The cross-effects of indicators are captured better by doing so. We will discuss how we aggregate the portfolios in following part.

Finally, this methodology allows to add assets and increase the investment universe as much as we want. We simply have to assess the qualitative relationships of the new asset with each of the macroeconomic signal. If it is easy for a macro asset (e.g. a currency pair), it becomes infeasible for a more granular investment universe (e.g. sub-sectors of the SPX). In that case, one can simply regroup the assets by family (equity, fixed income) and apply the initial qualitative assessments. However we also believe that sticking to a restrained investment universe is not handicapping in
our framework, as we want to capture macroeconomic movements, rather than making precise allocations in sub asset classes.

4.2 Portfolio construction and back-testing

In order to build our portfolio, we implemented a Python optimizer based on the methodology described above. We first chose a frequency for the data and the optimization (typically 1 month), as well as different optimization periods. The back-test was ran from 1980, but splitting the performance into different periods (like in AQR’s paper) facilitated the analysis. We did the split by decade, to avoid any hindsight bias. The basic four periods we chose are the following:

— Jan. 1, 1990 to Dec. 31, 1999
— Jan. 1, 2000 to Dec. 31, 2009
— Jan. 1, 2010 to Dec. 31, 2017

The optimization was done at each month, and the portfolio was then rebalanced for the next month accordingly. We chose a Markovian mean-variance optimization for its intuitive aspect and its relative simplicity. The macro momentum aspect of our methodology should prevent our portfolio from large drawdowns (assuming that the strategy is reactive enough) and allows to put in place a volatility constraint. It is a simple and well-known optimization framework, and simplicity is key to our approach.

Two “hyperparameters” can quite influence the results of our optimization:

— The first is the target volatility. Following a period of high volatility, a too low volatility target will create a strong bias towards fixed income. On the contrary, in a low volatility regime such as recently, a too high volatility target will force the optimizer to overweight the absolute investment in equities, even if it is not in line with the macro momentum signals. We first set the target volatility at 10%, but will discuss its values later.

— The second is the trailing period for computing expected returns and volatility of the investable assets. A too short period will not be robust enough and the resulting optimization will be too sensitive to recent past data. On the contrary, a too long period will incorporate data that became irrelevant. Eventually we chose 10 years (120 months), which seemed to yield consistent and stable results in all of our back-tests.

At each optimization date \( t \), we look at all of the historical data for the macroeconomic signals (from 0 to \( t \)). Then, based on our decision rules, and for each signal taken separately, we attribute a value to it (characterized by both its sign and its intensity). This value will directly influence the optimization boundaries for the asset classes. For example, if there is a strong increasing growth, it is a bullish signal for equities. So we initialize the equity weight at a high value, and we put high boundaries to force the portfolio to be highly invested into equities for the next month. Then, the Markovian optimization will adjust the position to match the volatility constraint while maximizing historical returns. A visual way to see our procedure is depicted in Figure 10.
This is done for each signal, and gives us a certain number of optimized portfolios at a given optimization date. The next step is to aggregate them. We should indeed aggregate all of them because the macro momentum value-added relies inside all of the signals and not one in particular. However, doing so can break our volatility constraint, because we aggregate uncorrelated portfolios. The method we implemented in order to deal with this overall volatility constraint is pretty simple:

— We average the weights of all the portfolios. This can be a weighted average in order to account for the possibility of having multiple signals for a single macroeconomic theme (e.g., growth and inflation for the business cycle).

— We compute the aggregated portfolio’s in-sample volatility (typically under the volatility target due to the diversification effect).

— We then adjust the relative risk-free rate position in order to scale the portfolio’s volatility, so that we are closer to the initial target.

This eventually gives us the global optimized portfolio, resulting from an aggregation all sub-portfolios, individually based on all macroeconomic signals. We store its weights and its out-of-sample performance (for the month after), and proceed to the next optimization date.

Let’s now give a few precisions on the volatility target. This parameter is crucial in a Markovian optimization framework, and by looking at the backtests results, it seems to work very well (out-of-sample volatility is below target). But this could be due to backtesting biases (e.g., volatility has been generally decreasing). Furthermore, a target volatility should not lead to the inclusion of volatile assets just to meet the target. We should keep only “good volatility” and keep a diversified portfolio that is the optimization output of the signals. One way to do so, was to calibrate the target volatility for each period, and based on the volatility of each asset class on this period. However, this creates a hindsight bias: using the equity volatilities on 1990-2000 to do our optimization in 1990. Another way to get around the problem is to use a rolling window like we did for the returns. We implemented this solution by averaging the trailing historical volatility of each asset classes (including the risk-free rate in order to lower the final target). Our target volatility oscillates between 10% (before 2000) and 6% (after the crisis).

• Combining portfolios

We developed several methods for aggregating the portfolios of each signals into the final portfolio of the strategy. The most basic one, though robust, simply consists in averaging the weights of each portfolio. By doing so, different indicators’ directions compensate each-others, and our final portfolio allocation depends on the general state of the economy, i.e. taking into account
all the indicators at the same time. A variation of this method consists in using a weighted average of the indicators. This was especially useful when we had several indicators for the same macroeconomic data (e.g. Growth and Inflation for the Business Cycle).

A third way of combining the portfolios that we implemented was to consider the recent performance of each indicators, and over-weight the indicators that performed the best recently. This would introduce a kind of momentum in the signals, which are momentum signals themselves. The implementation we used was an Exponentially Weighted Moving Average:

- At each period, we look at the performance of all the signals on the previous period. We sort them based on their performance. The best-performing signal is attributed the higher score, etc.

- We incorporate this score in the weighting of the current period: the better performing signals will see their weights increase, while the worse performers will see their weights decrease, relative to the weights of the previous period.

The parameter of this EMWA method is the speed of incorporation of the new information (i.e. the past period performance). A coefficient of 0 means we do not incorporate past period performance, which corresponds to the initial linear aggregation method (weighted or not). A coefficient of 1 means the weighting is only based on past period performance, ignoring the previous weights. In practice, the coefficient is between 0 and 1, and we prefer small values (between 0.2 and 0.5). Otherwise, the weighting of the new information becomes too important. This method is called EMWA as the parameter corresponds to the rate of decay of recent information (i.e. past period’s performance).

\[
\text{Weight}(t) = \text{Coefficient} \times \text{Performance}(t) + (1 - \text{Coefficient}) \times \text{Weight}(t - 1)
\]

• Volatility limit

We also developed several different methods for the Markowitz optimization. We will try to give an overview of the most relevant ones below (we will also sum up some of the methodologies described above):

- Using a rolling window for computing the target volatility. We implemented this solution by averaging the trailing historical volatility of each asset classes (including the risk-free rate so to lower the final target). Our target volatility oscillates between 10% (before 2000) and 6% (after the crisis). This is a robust way of constraining the out-of-sample volatility of our portfolio.

- Maximizing the Sharpe Ratio of the in-sample portfolio, regardless of the volatility. This method worked surprisingly well, and yielded among the best results for the optimization. The out-of-sample Sharpe ratio was strong and consistent over the 4 decades. This was the preferred method together with the rolling window target.

- A last method we implemented was to maximize a linear utility function of the form \( E(r) - \frac{\lambda}{2} V(r) \), where \( \lambda \) is the risk-aversion. This is actually quite similar to the Sharpe Ratio maximization, but causes a problem: the parametrization of \( \lambda \). As the optimization also depends on other factors, we did not think of a good enough heuristic for \( \lambda \), hence we preferred putting this method aside and continuing with one of the previous two.

4.3 Performance analysis

Running the optimizations on the four different 10-year sub-periods gives us the weights found by the optimizer, as well as the monthly returns of the strategy. We then built different performance analysis indicators and visualization tools.
— The basic returns / volatility and Sharpe ratio analysis.
— The correlation of our strategy returns to all of the investable assets.
— The maximum drawdown for the period.
— Visualizing the weights of the different asset classes through the investment period.
— We also compare the results to “naive” portfolios such as a 60/40 static weighting.

• Indicators cross-correlation

The choice of our indicators is key: the level of their pair-wise correlations should be low, or inexistent, in order to capture the purest signals and incorporate them in our portfolio construction methodology. Hence, we computed the correlation matrix of the indicators, over the periods. Both AQR’s indicators and ours appear not to be correlated one to another, or if they are, the correlation level is very low.

In the Figures 11 to 17 we display the correlation matrices.

Figure 11: AQR Indicators for 1980-1989.
Figure 12: AQR Indicators for 1990-1999.
Figure 13: AQR Indicators for 2000-2009.
Figure 14: AQR Indicators for 2010-2017.
Figure 15: Our Indicators for 1990-1999.
Figure 16: Our Indicators for 2000-2010.
Figure 17: Our Indicators for 2010-2017.
5 Results

5.1 Optimal choice of indicators

Let’s now describe our choice of indicators and the process we used in order to determine it. First, let’s have a look at the individual performances of the indicators over each sub-period:

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Returns over the period</th>
<th>Returns over the period</th>
<th>Returns over the period</th>
<th>Returns over the period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth (From AQR)</td>
<td>4.85%</td>
<td>9.16%</td>
<td>5.79%</td>
<td>2.43%</td>
</tr>
<tr>
<td>Inflation (From AQR)</td>
<td>10.79%</td>
<td>8.13%</td>
<td>7.85%</td>
<td>0.79%</td>
</tr>
<tr>
<td>International Trade (From AQR)</td>
<td>14.21%</td>
<td>6.55%</td>
<td>5.64%</td>
<td>1.88%</td>
</tr>
<tr>
<td>Monetary Policy (From AQR)</td>
<td>12.65%</td>
<td>8.72%</td>
<td>3.24%</td>
<td>4.30%</td>
</tr>
<tr>
<td>Risk Sentiment (From AQR)</td>
<td>13.20%</td>
<td>10.90%</td>
<td>6.43%</td>
<td>4.93%</td>
</tr>
<tr>
<td>US PMI</td>
<td>10.10%</td>
<td>8.05%</td>
<td>4.80%</td>
<td>6.92%</td>
</tr>
<tr>
<td>US CCI</td>
<td>7.87%</td>
<td>9.51%</td>
<td>7.20%</td>
<td>5.05%</td>
</tr>
<tr>
<td>USIP Growth</td>
<td>9.03%</td>
<td>10.88%</td>
<td>4.23%</td>
<td>6.17%</td>
</tr>
<tr>
<td>DTEXM</td>
<td>14.96%</td>
<td>8.30%</td>
<td>5.06%</td>
<td>2.47%</td>
</tr>
<tr>
<td>3M T-Bill</td>
<td>13.21%</td>
<td>7.79%</td>
<td>2.78%</td>
<td>4.10%</td>
</tr>
<tr>
<td>US Expected Inflation</td>
<td>9.68%</td>
<td>6.55%</td>
<td>8.12%</td>
<td>3.25%</td>
</tr>
<tr>
<td>TED Spread</td>
<td>9.94%</td>
<td>3.79%</td>
<td>3.95%</td>
<td>2.11%</td>
</tr>
<tr>
<td>VIX Index</td>
<td>11.41%</td>
<td>5.82%</td>
<td>6.36%</td>
<td>3.81%</td>
</tr>
</tbody>
</table>

The first observation to be made with regards to this individual performance table is the confirmation that the overall performance of the strategy is not driven by a single indicator. Each indicator contributes to diversify the returns, and the stability of the strategy over time relies on the use of multiple indicators and themes, as exposed by AQR: "The performance of the hypothetical composite macro momentum strategy is neither driven by a single asset class, nor by a single theme [...] each asset class and theme has contributed over the full sample and has exhibited broadly comparable performances during that time period”. Therefore, our choice of optimal strategy will be a combination of several indicators, thus using diversification in order to increase the stability of the returns over the various time periods.

— Business Cycle

Regarding the Business Cycle theme, in order to determine our best set of signals, we tried to implement several combinations of indicators, everything else remaining equal (thus keeping the same AQR indicators for the 3 other themes). Using a global Business Cycle theme by equally weighting each Business Cycle indicator, as well as dropping successively each one of the indicators in the overall combination, yielded quite similar results. Globally, dropping one of the indicators from the theme led to a trade-off between an improvement of the most recent period’s Return & Sharpe ratio, and an improvement of the former periods’ results. Nonetheless, we observed that dropping the Growth indicator (the one we constructed in 3.2 reproducing AQR’s methodology), improved our overall stability on the full sample period. The explanation for this result may be found in the difference between AQR’s indicator and our attempt at replicating it. Indeed, AQS had built its indicator using forecasts for the most recent periods and the ambiguous "lagged one
“quarter” method on realized GDP data for the former periods. On our side, we were not able to get enough granularity on GDP forecasts data, and since we chose not to lag the data from one quarter (see 3.2), our indicator was only built using one-year changes in realized GDP, proving to be relatively inefficient in predicting business cycle momentum. Overall, we therefore chose to drop the Growth indicator we had previously built, and we kept an equally-weighted combination of the other remaining indicators within our Business Cycle theme.

— International Trade:

Regarding the International Trade theme, we compared the individual performances and Sharpe Ratios of our two sets of export-weighted spot foreign exchange rate indices. Among the two indices (the one provided by the ICE - DXY - and the one provided by the Federal Reserve of St.Louis - DTEXM -), we observed that the most relevant, both in terms of performance and Sharpe Ratios was the index provided by the Federal Reserve of St.Louis (See Table 4 for performance, and Figures 18 & 19 for Sharpe Ratios analysis). Thus, we kept the DTEXM Index as a reference for our International Trade Indicator.

Figure 18: Performance using only the Export Weighted Dollar Index From ICE (DXY).

Figure 19: Performance using only the Export Weighted Dollar Index From Federal Reserve of St.Louis (DTEXM).
— Monetary policy:

Just as for the Business Cycle theme, we have implemented several different combinations of indicators within the Monetary Policy theme, in order to determine the best set of indicators to finally use in our optimization, everything else remaining equal for the remaining 3 themes. Our results showed that dropping either the initial indicator built by AQR or the 3M T-Bill led to a deterioration of the performance over the more recent period. As a reminder, AQR’s estimator is built using the Fed Funds rate for the period prior to 1992, and the 2Y Treasury yield for the period between 1992 and 2017. Thus, for contemporaneous strategies, we chose to build our Monetary Policy theme using a combination of one-year changes on the 3M T-Bill and on the 2Y Treasury yield; hence getting a diversified exposure to variations in the front end of the yield curve.

— Risk Sentiment:

We followed the same back-testing methodology for the Risk Sentiment theme as for the Business Cycle and Monetary Policy themes. Our results showed that both AQR’s indicator (Market Excess Return), and the VIX index were very important performance drivers over the most recent period. Dropping either of them led to an important decrease in returns over the period 2010-2017, and a slight decrease over the period 2000-2010, whereas it did not really impact our results for the studied periods prior to 2000. On the other hand, dropping the TED Spread from our Risk Sentiment theme led to a non-negligible improvement of returns since 1990. Over the period 2010-17, the returns of the strategy without the Ted Spread indicator were 63% higher, and 12% higher over the period 2000-10. Therefore, we decided to keep only a combination of AQR’s initial indicator and one-year change in the VIX index, in order to build our Risk Sentiment theme.

5.2 Empirical findings

• Relatively poor regression results

In the process of checking the relevance of the signals/asset classes relationships provided by AQR (see Table 1. Impact of the indicators on asset classes), OLS regressions were conducted over different time horizons, using sequentially equities and bonds as our dependent variables and including all of AQR’s signals as explanatory variables (see Section 3.2). As previously mentioned, the results of these operations were rather disappointing, with R-Squared hardly significant and coefficients close to zero for all regressors. The output of this procedure was however still not null. Indeed, it has first of all highlighted the need of not taking any theoretical or logical macroeconomic cause-consequence relationships for granted. Real data does not always verify what would appear as elementary and straightforward chains of events in financial markets and it has emphasized the need to remain cautious in formulating our initial assumptions. Moreover, it has pushed us to look for additional indicators, which eventually led to better results. This has increased the depth of our study and provided additional insights for the construction and comparison of signals’ return generation potential.

• Portfolio metrics: consistent but mitigated performance

Having implemented our own macro momentum strategy, we will comment the portfolio performances in the following paragraph. Coming back to AQR’s research and performance metrics, three main periods could be differentiated:

— The decade 1980-1990, during which a macro momentum strategy would have performed very well (Annualized Sharpe ratio of 1.7)

— The decades 1990-2010 which displayed Sharpe ratios in-between 1.3 and 1.4, still highlighting very advantageous outcomes from the strategy

— The period since 2010 where the portfolio performance decreased, with a Sharpe ratio at 0.7.
We could then compare these initial results to our own macro momentum strategy which resulted in the following trends:

— Our indicators performed quite poorly for the decade spanning over 1980-1990
— For the period 1990-2010, our portfolio performances were slightly better
— Since 2010, our macro momentum strategy showed encouraging results even if the obtained Sharpe ratio is lower than the ones for equities and bonds over this period. Indeed, it highlights a certain resilience ability in value creation even after a large crisis episode - The Great Financial Crisis in 2009.

Looking more in details into the origins of these performances, we observe that whatever the selection of indicators is, the strategy remains consistent with market trends; the portfolio holds a bigger proportion of equities during periods of smooth and trendy markets and operates a flight-to-quality towards bonds in stress and highly volatile periods. We thus believe that our methodology covering both the indicators construction, and the portfolio construction and optimization has enabled to capture efficiently macroeconomic trends.

While running this strategy, we still put effort into acknowledging and working on remaining limitations to our methodology. For example, the strategic portfolio allocation is highly sensitive to the parametrization of the optimization. This potentially implies instability and big changes in the optimal portfolio allocations with small signals movements leading to radical shifts in the repartition. Additionally, for both our strategy and the one from AQR, we witnessed a degradation of the performance since 2010. A potential explanation of this phenomenon could flow from the major changes seen in fiscal and monetary policies after 2008. Exceptional quantitative easing has significantly impacted the interconnectedness of diverse macroeconomic indicators, ultimately reducing the predictability of their behaviors in the current cycle. Another point highlighted relates to the quite high turnover implied by our strategy, which comes back to the idea of allocation instability. Thus, at observation points corresponding to momentum turnabouts, positions need to be drastically readjusted. The transaction costs engendered would then most probably surpass the generated returns of the strategy. Finally, the frequency at which the signals macroeconomic data is published should be looked at closely. Indeed, information about most of our macroeconomic indicators is published on a monthly or even quarterly basis, leading to almost “lagging” signals that are potentially slow to react to turning points. One could then imagine that especially during crashes such as in February, our portfolio would perform poorly as it would not anticipate the market correction. This remark is then of course not applicable to ”market” indicators (e.g. VIX).

![Figure 20: Macro momentum strategy returns in 2017](image-url)
5.3 Interpretations

• Do macroeconomic data and momentum strategies represent a pertinent combination?

Looking at the results of our optimization, we can unreservedly support the fact that macroeconomic data represents a strong basis for the building of a macro strategy. Indeed we have been able to observe that, through the aggregation of variety of signals' combined effects, it is possible to obtain a consistent view of the broad macroeconomic trends, and to derive an affiliated optimal allocation. This finding reinforces both the idea of the importance of fundamentals behavior in portfolio management, and the belief in a predominant link between macroeconomic variables and investment opportunities.

At first sight, we had emitted doubts on the feasibility of applying a momentum strategy to macroeconomic data: characteristics like constant revisions, low frequencies, and delayed publications seemed hardly compatible with a robust portfolio construction. But with careful cross-checking of different data sources and meticulous choices of reliable indicators, it is possible to overcome this difficulty, and even to take advantage of the reduced presence of noise and impurities in time series compared for instance to stock prices.

• What is the relative contribution of each asset class?

Throughout this study, we made the choice of focusing on two central asset classes, Bonds and Equities. This decision has been taken in order to provide the most intuitive and interpretable results, as we knew that extending our methodology to broader asset classes would be easier after having constructed a robust and transparent tool.

According to our results, we can state that macro momentum strategies fit particularly well the Equity and Bond investment universe, as both asset classes exhibit interesting diversification features while being very sensitive to different macroeconomic data. Another advantage of choosing a constrained investment universe as we did is that our strategy can be applied as an overlay to an already existing Equity and Bond portfolio without radically changing its exposures in terms of asset classes.

Should one would like to extend our methodology to other asset classes, it would be interesting to consider foreign exchange rates (for instance major currency pairs) as they tend to be very sensitive to the global macro environment, but then a more international framework should be envisioned in order to take into account relative strengthening or weakening effects of macro trends between countries. Moreover, adding further granularity in the asset classes, as mentioned before, could yield a more precise allocation, but could lead to a decreased value-added: macro momentum should not be captured through very narrow channels as it represents global trends.
5.4 Integrating the macro momentum strategy

- Macro momentum as a strategy in itself or as an overlay?

At the very beginning of our project, we considered the macro momentum strategy as an independent portfolio strategy in itself. Indeed, we believed in the idea of building a whole cross-asset portfolio from scratch based on such approach, considering it would be very much correlated to conventional asset classes and portfolio structures.

Nevertheless, throughout our work, we progressively realized that besides being an investment tool in itself allowing to generate substantial profits, it could also be used as an overlay, which is perfectly in line with adding the strategy on top of various pre-existing portfolios. As we mentioned briefly above, this comes from the fact that our strategy can work as a diversifier. Indeed, we were surprised by the marginal contribution a macro momentum strategy could bring: at the beginning, we expected the returns to be strongly and positively correlated for instance to a conventional trend-following strategy on Equities and Bonds (as fundamentals and prices trends tend to adjust coordinately on the long run). But we found that such a strategy differed in many points from a classical trend-following one, notably in one central aspect: our strategy provided positive returns during each of the major drawdown periods during which a simple trend-following strategy would have performed badly. We can therefore say that applying our momentum strategy to a macroeconomic framework provides a kind of integrated “tail hedge” displaying interesting option-like payoffs, complementary to a conventional trend-following strategy or any other cross-asset portfolio structure.

By looking at the correlation of the strategy with the different assets on the different periods (Figure 21), we see that our strategy is generally uncorrelated from the other asset classes (regardless of the optimization method and the choice of the indicators). Hence we have strong reasons to believe that adding the macro-momentum portfolio on top of a traditional asset allocation would help diversifying risk away.

![Figure 21: Correlation of the strategy with asset classes for the 4 periods.](image)

In that case, implementing the strategy through Futures seems to be a good alternative: it is easier to rebalance and to go short, which is of interest due to the strategy turnover.

- Returns & Sharpe Ratios of the Different Strategies

Table 5 summarizes the performance and Sharpe Ratios of the three implemented Strategies.
Table 5: Returns and Sharpe Ratios of each Strategy

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Replication of AQR’s Indicators (3.2)</td>
<td>13.44%</td>
<td>10.43%</td>
<td>7.92%</td>
<td>4.85%</td>
</tr>
<tr>
<td>Implementation of Alternative Indicators (3.3)</td>
<td>14.57%</td>
<td>8.77%</td>
<td>7.92%</td>
<td>5.01%</td>
</tr>
<tr>
<td>Mixed Strategy (Best Choice of Indicators: 5.3)</td>
<td>14.49%</td>
<td>10.49%</td>
<td>7.08%</td>
<td>5.55%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Replication of AQR’s Indicators (3.2)</td>
<td>0.40</td>
<td>0.88</td>
<td>0.82</td>
<td>0.93</td>
</tr>
<tr>
<td>Implementation of Alternative Indicators (3.3)</td>
<td>0.56</td>
<td>0.61</td>
<td>0.81</td>
<td>0.89</td>
</tr>
<tr>
<td>Mixed Strategy (Best Choice of Indicators: 5.3)</td>
<td>0.51</td>
<td>0.88</td>
<td>0.68</td>
<td>1.04</td>
</tr>
</tbody>
</table>

We first tried to reproduce AQR’s results, using the closest possible indicators to the ones built in the paper (3.2). The first observation that can be made looking at this first line of results is that we did not manage to get AQR’s performance nor Sharpe Ratio over the studied period, using our indicator’s replication. In fact the performance we obtain is always lower than the one displayed by AQR (They show performances of 16.7% over the period 1980-89, 14.1% for 1990-99, 15.4% for 2000-09 and 6.5% for 2010-16). We only obtain a better Sharpe Ratio for the most recent period (0.93 over 2010-17, whereas the paper shows a Sharpe Ratio of 0.7 over the period 2010-16). This difference in the results is straightforward. On the first hand, we did not manage to get the same data as AQR, nor did we follow their exact indicator construction methodology, sometimes opaque (see for instance the “lagged one quarter” methodology to get forecast proxies). On the other hand, the methodology used by AQR to build its portfolio optimization is not fully explained in the paper, thus we used our own methodology to construct the optimization. This difference in methodology may be the main driver of the gap between both results.

We further tried to implement several alternative indicators (3.3), and calculated the performance and Sharpe Ratio of this new portfolio, equally weighting each set of indicators within a particular theme, thus having the same weight of each theme within our aggregated portfolio. Looking at the results, we first observed that the performance of the portfolio using those alternative indicators was close to the initial performance obtained when replicating AQR’s methodology. Thus, we may say that the new set of indicators was still relevant to describe macro-momentum, and that our choice of decomposition of the indicators within each theme was accurate. Moreover, looking at the performance on each sub-period, we saw that several of them displayed improvements compared to the first portfolio built using AQR’s indicator replication. Therefore, we assumed that it was possible to improve our overall performance by diversification, using a combination of the first used indicators and our alternative ones.

The choice of this ”Mixed indicators Strategy” is described and explained above in part 5.3. We chose to combine the indicators in a certain way that provided the best results over the most recent period (thus would potentially be more accurate in the near future), and that improved the overall performance when looking at the sum of each period. Our mixed Strategy loses some performance over the period 2000-09, but improves the performance of the other time studied periods.
• Adding Granularity to the Signal Intensity Interpretation?

Table 6: Implementing various granularity levels to study the signal intensity

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Quantiles: 1</th>
<th>Quantiles: 2</th>
<th>Quantiles: 3</th>
<th>Quantiles: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sharpe Ratio</td>
<td>Returns</td>
<td>Sharpe Ratio</td>
<td>Returns</td>
</tr>
<tr>
<td>Jan 1980 - Dec 1989</td>
<td>0.51</td>
<td>14.49%</td>
<td>0.40</td>
<td>13.68%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
<td>14.24%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.49</td>
<td>14.40%</td>
</tr>
<tr>
<td>Jan 1990 - Dec 1999</td>
<td>0.88</td>
<td>10.49%</td>
<td>0.73</td>
<td>9.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
<td>8.82%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
<td>8.42%</td>
</tr>
<tr>
<td>Jan 2000 - Dec 2009</td>
<td>0.68</td>
<td>7.08%</td>
<td>0.52</td>
<td>6.27%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
<td>6.60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.63</td>
<td>6.94%</td>
</tr>
<tr>
<td>Jan 2010 - Dec 2017</td>
<td>1.04</td>
<td>5.55%</td>
<td>0.53</td>
<td>2.99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.39</td>
<td>2.29%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.36</td>
<td>2.10%</td>
</tr>
</tbody>
</table>

We finally tried to implement our optimal indicators’ portfolio through different levels of granularity regarding the treatment of the signal intensity. As exposed in part 4.1, it seems that adding too much granularity do not yield any improvement in the results. Indeed, the best performance was obtained when dealing with the median as the only discriminating factor for signal intensity (see Table 6).
5.5 Further research

Having thoroughly investigated and applied the concept of macro momentum to real-life data, one element stood out as particularly crucial in the implementation of such a strategy; the quality of the signals used. Indeed, we realized very early while carrying out this project that intuitive ideas and relationships between asset classes and macroeconomic variables, that are often taken for granted, are not always quantitatively verified (cf. Section 5.1 Empirical Findings). Moreover, signals have been seen to sometimes yield inconsistent results depending on the state of the economy (stable markets vs. crisis) which is not easily integrated in an allocation decision strategy. This issue of signals instability could thus be looked at in more details.

While implementing the above macro momentum strategy, we observed two different types signals potential instability in the portfolio optimization problem:

- Indicators with non linear monotonic behaviors
- Indicators with non linear non monotonic behaviors

In the first case, we are referring here to signals whose relationships with our two asset classes of interest - equities and bonds - are always positive or negative (same sign) but with varying intensities (e.g. PMI). In the second case, we look at indicators whose behavior in relation with equities and bonds is dependent on the intensity of the deviation (e.g. VIX and TED Spread).

Looking at the tail of the distributions, we should therefore acknowledge that a unique and simple positive or negative relationship is not representative of real data. Moreover, coming back to the idea of a macro momentum portfolio as an overlay product, this point is highly problematic as it implies that the constructed portfolios would be optimized with wrong assumptions during crisis periods, harming performances even more.

This point adds another layer of complexity to the modeling and implementation of these macro momentum signals in a simple Markowitz optimization framework and would require the subdivision of signals into different states following the status of the economy. Further research could thus be conducted with the purpose of unveiling and breaking down the changes in behaviors of the selected indicators. This would require extensive analysis of historical data and past distributions of our signals in relation with equities and bonds. While this has not been directly implemented in this report, we acknowledge the need of looking more closely at these tail risks.

![Figure 22: Relationship between the VIX and both Market and 1Y Treasury returns](image-url)


6 Conclusion

In this study, we finally achieved to underline the potential value-added of a macro momentum tactical allocation overlay in a multi-asset portfolio framework. On the basis of the available academic literature and of our own analysis, we have first defined the essential characteristics of such a strategy in a risk factor framework, and discussed its potential sources and remaining gray areas.

From there we were able to have a more critical view on our subject, and provided a thorough analysis of the AQR paper, from its core assumptions to the practical implementation of their strategy. As we have found several inconsistencies in the process described in the paper, the reproduction of its results has revealed to be rather arduous. Eventually, we achieved to construct a tool that yielded satisfying results, but it required a strong questioning and deviation from the reference paper’s methodology.

Indeed, we have carried out an in-depth signal study in order to determine the optimal set of macroeconomic indicators and the most favorable treatment associated to it (using statistical tools ranging from distribution quantiles to Z-scores analysis). We finally chose the Z-score analysis using dynamic thresholds and relatively low granularity as it proved to yield the most profitable results. Our portfolio construction procedure has been elaborated starting from a Markovian mean-variance approach, and has been enhanced by fine-tuning the approach through trial and error (our final choice being the maximization of the in-sample portfolio Sharpe Ratio regardless of the volatility). After the optimization of each individual portfolio using Equities and Bonds (one portfolio per macroeconomic indicator) through a boundary rule linked to signals’ signs and intensities, we were able to perform the final aggregation.

As stated in the AQR paper, the core value-added of such macro momentum strategies doesn’t rely in the separated interpretation of individual signals, but in their total aggregation: there can be found a real predictive power encompassing the global macro environment trends. We have implemented different alternative combination methodologies, from a simple weights averaging to an Exponentially Weighted Moving Average approach in order to account for momentum in the signals themselves. Using risk-free rate positions in order to scale the overall volatility to the initial target, we finally obtained our finite product: a tactically allocated cross-asset portfolio using macro momentum signals.

Moreover, our investment tool can be easily expanded to other indicators and asset classes, the only requirement being the assessment of the relationships linking the additional element with the original ones. With our final strategy returns being relatively uncorrelated to the original investable assets, and displaying a strong Sharpe Ratio across the majority of time periods, we can say that the macro momentum strategy yields robust and implementable results. It is nonetheless important to state that some improvements can be incorporated in the methodology described in our paper, including a more detailed and quantitative analysis of the relationships linking signals and asset classes (taking into account non-linearity and dynamic evolutions) or a the division of the signals into subsets corresponding to different macro economic regimes (taking into account the relative overperformance of certain indicators in stress periods).
References


