Accepted Manuscript

Market timing over the business cycle

Magnus Sander

PII: S0927-5398(17)30120-2
DOI: https://doi.org/10.1016/j.jempfin.2017.12.002
Reference: EMPFIN 1024

To appear in: *Journal of Empirical Finance*

Received date : 14 November 2016
Revised date : 16 November 2017
Accepted date : 12 December 2017

Please cite this article as: Sander M., Market timing over the business cycle. *Journal of Empirical Finance* (2017), https://doi.org/10.1016/j.jempfin.2017.12.002

*Highlights (for review)

Highlights for: Market Timing over the Business Cycle

- Stock and bond return predictability switches distinctly across the business cycle
- I adjust standard forecasting models to take these switches into account
- The adjustments lead to significant economic gains for a mean-variance investor
- The risk-adjusted returns are robust to using real-time recession indicators

Market Timing over the Business Cycle*

Magnus Sander†

Aarhus University and CREATES

*I thank Daniel Ferreira (the Editor), an anonymous referee, participants at the 2015 MFA conference, Tom Engsted, Chang Lee, Stig Vinther Møller, Thomas Quistgaard Pedersen, and Allan Timmermann for helpful comments. I acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

†CREATES, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus, Denmark. Phone: (47) 40101930. E-mail: mdsj@econ.au.dk.

Market Timing over the Business Cycle

Abstract

This paper analyzes the economic value of linking return predictability to the business cycle. Recent studies show that stock returns are predictable in recessions while bond returns are predictable in expansions. I examine whether these findings can be exploited in real-time trading by letting the coefficients of popular return regressions switch across states of the economy. The switching models I propose are easy to implement and provide meaningful economic gains relative to their constant coefficient versions. However, choosing a good recession signal is important as inaccurate business cycle turning points corrupt the switching extensions.

Keywords: Portfolio choice, business cycles, return predictability.

JEL: C53, E44, G11.

1. Introduction

This paper measures the economic value of conditioning stock and bond return forecasts on recession dummies. Motivated by recent studies, I restrict stock returns to only be predictable in recessions and bond returns to only be predictable in expansions. These restrictions generate notable increases in risk-adjusted performance in a standard setup with mean-variance utility and a simple bivariate regression model. I also let regression coefficients switch across recessions and expansions without restricting them. This extension performs particularly well when combined with forecast averaging. The risk-adjusted returns are robust over time and prevail using real-time recession signals. However, high accuracy of business cycle turning points is essential and seemingly small tweaks to the classification rule can have large economic implications.

Goyal and Welch (2008) forcefully argue that simple regression models cannot forecast stock returns out of sample. Thornton and Valente (2012) reach similar conclusions for bonds. However, these studies do not condition on the business cycle. Henkel et al. (2011) let coefficients be regime-dependent and document that stock returns are predictable but only in economic recessions.[1] Andreasen et al. (2016) find the mirror image for bond returns as they are only forecastable in expansions. Thus, return predictability seems to be highly dependent on the business cycle and, moreover, it is asymmetric across stocks and bonds. This motivates why I extend standard regression models to take macroeconomic conditions into account.

Markov switching models are one of the most popular approaches to capturing state-dependence in returns. These models let the return distribution depend on unobservable realizations of a Markov chain without imposing an economic interpretation. Examples of papers taking this approach are Guidolin and Timmermann (2007), Guidolin and Hyde (2012), and Henkel et al. (2011). Henkel et al. (2011) find a strong connection between the inferred states of their return forecasting model and the business cycle. On one hand, Markov switching models provide flexibility by not imposing that the regimes are related to specific economic events. On the other hand, they are highly nonlinear and rely on potentially unstable numerical estimation methods which may hamper their usefulness

[1] See also Rapach et al. (2010), Dangl and Halling (2012), Rapach and Zhou (2013), and Neely et al. (2014). Further, Kacperczyk et al. (2014) examine the time-varying skill of mutual fund managers and conclude that they are only good at market timing in contractions.

in out-of-sample forecasting. If the underlying states are related to recessions, which are short-lived and infrequent, more robust techniques could be useful. I therefore instead use observable recession dummies which allows me to estimate the models using simple linear regressions. This choice is natural since most of the studies that identify differences in predictability across states of the economy have done so using recession dummies (see citations above). I do also show results for Markov switching models but find that, in contrast to the dummy switching strategies, they do not improve on constant coefficient models for real-time trading. Further, tying financial markets explicitly to recessions is in the spirit of the macro-finance asset pricing literature as summarized in Cochrane (2017).

Several other papers allow predictability to change over time. Timmermann (2008) argues that profit seeking traders cause predictability to only be present in pockets of time and proposes an adaptive forecast combination approach. Pesaran and Timmermann (2002) suggest a two-step procedure to forecasting in the presence of structural breaks. Dangl and Halling (2012) find that time-varying parameter models improve stock market timing over models with constant coefficients. In contrast to these studies, I specifically impose that predictability is linked to the business cycle.

The rest of the paper is organized as follows. Section 2 presents data on returns, the main predictors, and recession indicators. Section 3 motivates the switching strategies. Section 4 implements single-predictor switching strategies in a cross-asset setup. Section 5 extends the analysis to multivariate forecasting models. Section 6 implements Markov switching models. Section 7 concludes.

2. Data

I use S&P 500 stock returns and long-term U.S. government bond returns. The bond returns are from Ibbotson's Stocks, Bonds, Bills, and Inflation Yearbook. I compute excess returns on stocks and bonds by subtracting the risk-free return, which is the lagged three-month T-bill rate. The main predictors are the log dividend-price ratio of the S&P 500 for stocks and the term spread for bonds. The term spread is computed as the difference between the yield on Ibbotson's long-term bond and the three-month T-bill rate. The dividend-price ratio and the term spread are among the most popular predictors in the literature, see, e.g., Campbell and Shiller (1988) and Fama and French (1989). All returns and predictors are from the updated Goyal and Welch (2008) data set.[2] Unless otherwise specified, all data are collected on a monthly frequency from 1927:1 to 2013:12.

To identify recessions, I first consider the recession dates from the National Bureau of Economic Research (NBER) as they are the standard choice in the literature. While the NBER dates constitute the most popular business cycle classification, they are only available with a significant publication lag. Therefore, I also rely on two real-time dummies. The first is based on the Aruoba-Diebold-Scotti Business Conditions Index (ADS), see Aruoba et al. (2009). ADS is a daily index based on a dynamic factor model of economic variables and is updated continuously as new data are released, which is at least once a week. It has a mean of zero. The second dummy is based on the Purchasing Managers' Index (PMI), which Christiansen et al. (2014) show is a strong predictor of U.S. recessions. PMI is released on the first business day of the month and is based on a survey of the manufacturing sector. It is an index from 0 to 100 with values below 50 indicating a recession in the manufacturing economy.[3]

2.1. Real-time turning points

To translate ADS and PMI into dummy variables, I follow Berge and Jordà (2011) and use Receiver Operating Characteristic (ROC) curves to identify the thresholds that maximize each signal's ability to identify NBER turning points. Let $NBER_t$ be a dummy variable that is one if month t is classified as a recession by the NBER and zero otherwise. A signal Y_t indicates an expansion when $Y_t \geq c$ and a recession when $Y_t < c$, where c is a threshold. The true positive rate ($TP(c)$) and the false positive rate ($FP(c)$) are given by

$$TP(c) = P[Y_t \geq c \,|NBER_t = 0] \tag{1}$$

$$FP(c) = P[Y_t \geq c \,|NBER_t = 1]. \tag{2}$$

[2] I thank Amit Goyal for making it available on his website.
[3] www.instituteforsupplymanagement.org.

Further, the unconditional probability of an expansion is denoted by $\pi = P\left[NBER_t = 0\right]$. To find the optimal threshold, I solve the following maximization problem:

$$\max_{c}\left(2\widehat{\pi}\widehat{TP}\left(c\right)-\widehat{\pi}\right)-\left(2\left(1-\widehat{\pi}\right)\widehat{FP}\left(c\right)-\left(1-\widehat{\pi}\right)\right), \tag{3}$$

where the hats reflect that all three probabilities, $\widehat{TP}\left(c\right)$, $\widehat{FP}\left(c\right)$, and $\widehat{\pi}$ are sample estimates. Berge and Jordà (2011) find thresholds of -0.80 for ADS and 44.5 for PMI using full-sample analysis. I instead estimate real-time thresholds for my evaluation period using an expanding window of data going back to the first ADS observation in 1960:3 and the first PMI observation in 1948:1. Each period, I simply recalculate the optimal thresholds by solving (3). The maximization problem relies on NBER data and so the publication lag needs to be addressed. Ng (2012) reports an average delay of nine months for the period 1980 to 2008. However, Ng also argues that due to the availability of real-time economic data, it is not very realistic that people first know that they are in a recession nine months ex-post and instead assumes a three-month publication lag, see also Kauppi (2008) and Christiansen et al. (2014). I follow this approach and use data up to $t-3$ to estimate the thresholds for my ADS and PMI real-time dummies at time t.

I collect NBER dates and the Purchasing Managers' Index from St. Louis Fed's FRED database and the Aruoba-Diebold-Scotti Business Conditions Index from the Philadelphia Fed's website.[4]

3. Return predictability across the business cycle

I create real-time forecasts of stock and bond returns from the following bivariate regression

$$R_{t+1} = \alpha + \beta X_t + \varepsilon_{t+1}, \tag{4}$$

where R_{t+1} is the excess return on a risky asset from time t to $t+1$, and X_t is the associated predictor observed at time t. I use data from 1927:1 to 1970:12 to estimate (4) using OLS and

[4] I use the latest vintage (collected March 2016) of all data as real-time vintages are not available for my sample period. Berge and Jordà (2011) argue that data revisions should affect indices much less than single series, and Chauvet and Piger (2008) find that business cycle turning points are quite robust to data revisions.

create the first return forecast for 1971:1. My benchmark is to impose no predictability, $\beta = 0$, and compute forecasts from

$$R_{t+1} = \alpha_0 + \varepsilon_{0,t+1}. \tag{5}$$

I repeat this exercise every month until 2013:12 using an expanding estimation window. This choice of evaluation period ensures that I have at least two recessions to train the real-time recession indicators for the first forecast.

To measure forecasting performance from a mean squared error perspective, I compute out-of-sample R^2:

$$R^2_{oos} = 1 - \frac{\sum_{t=T_0}^{T-1} \widehat{\varepsilon}^2_{t+1}}{\sum_{t=T_0}^{T-1} \widehat{\varepsilon}^2_{0,t+1}}, \tag{6}$$

where $\widehat{\varepsilon}_{t+1}$ and $\widehat{\varepsilon}_{0,t+1}$ are the recursive residuals from the candidate model and the benchmark, respectively, and T_0 is the time of the first forecast.[5] I evaluate H_0: $R^2_{oos} \leq 0$ against H_A: $R^2_{oos} > 0$ using the Clark and West (2007) test for equal predictive power which takes into account that the benchmark is nested in the candidate model.

To also assess forecasting performance from an economic perspective, I follow Campbell and Thompson (2008) and calculate the change in Certainty Equivalent Return (ΔCER) for a mean-variance utility investor. The utility function is

$$U_{t+1}\left(R_{p,t+1}\right) = E\left(R_{p,t+1}\right) - \frac{1}{2}\gamma Var\left(R_{p,t+1}\right), \tag{7}$$

where

$$R_{p,t+1} = \omega_t R_{t+1} + R_{f,t+1},$$

and γ is the coefficient of relative risk aversion, ω_t is the chosen risky asset weight, and $R_{f,t+1}$ is the risk-free return. ΔCER is the fixed fee that equates expected utility from using the no-predictability benchmark (5) with expected utility from using the regression model (4):

$$\Delta CER = CER - CER_0, \tag{8}$$

[5] I compute recession R^2_{oos} as $R^2_{oos,rec} = 1 - \left(\sum_{t=T_0}^{T-1} \widehat{\varepsilon}^2_{t+1} \times NBER_t\right) / \left(\sum_{t=T_0}^{T-1} \widehat{\varepsilon}^2_{0,t+1} \times NBER_t\right)$. The expansion values are computed similarly.

where CER is computed as $\widehat{\mu}_p - \frac{1}{2}\gamma\widehat{\sigma}_p^2$ using the sample mean ($\widehat{\mu}_p$) and variance ($\widehat{\sigma}_p^2$) of portfolio returns derived from the candidate model, and CER_0 is computed similarly for the benchmark.[6] I set the coefficient of relative risk aversion γ to 5 and the investment horizon to one month. The optimal real-time weights maximize the utility function in (7):

$$\omega_t = \frac{1}{\gamma}\frac{\widehat{R}_{t+1}}{\widehat{\sigma}_{t+1}^2}, \tag{9}$$

where $\widehat{R}_{t+1}$ is generated from the relevant regression model and $\widehat{\sigma}_{t+1}^2$ is the estimated variance. I follow Campbell and Thompson (2008) and use a five-year rolling window for the variance estimate. Through winsorization the investor is constrained to vary his weight in the single risky asset between 0% and 100% with the rest placed in the risk-free asset. To evaluate H_0: $\Delta CER \leq 0$ against H_A: $\Delta CER > 0$, I follow McCracken and Valente (2016) and use a stationary bootstrap to create artificial samples of returns and predictors (including recession dummies). In each sample, I recompute optimal portfolio weights and resulting performance gains to map out a bootstrapped distribution of ΔCER estimates that captures time-series dependencies as well as estimation uncertainty in the weights. As in McCracken and Valente (2016) the number of samples is 999 and the average block length is $T^{0.6}$, where T is the total number of observations. The performance gains are recentered to reflect the null hypothesis.

Table 1 shows forecasting performance of the bivariate regression versus the no-predictability benchmark in NBER recessions and expansions. Panel A presents results for stock returns using the log dividend-price ratio as a predictor and panel B presents results for bond returns using the term spread as a predictor. P-values corresponding to the percentages of recentered bootstrap estimates larger than the sample ΔCERs are in parentheses. We see evidence that stock returns are only predictable in recessions, while bond returns are only predictable in expansions. This is true both when using statistical and economic evaluation criteria. In recessions, timing the stock market beats the no-predictability benchmark reflected by a positive R_{oos}^2 and ΔCER, whereas timing the bond market produces a negative R_{oos}^2 and ΔCER. In expansions, the opposite is true.

[6] To compute recession ΔCERs, I use $CER_{rec} = \widehat{\mu}_{p,rec} - \frac{1}{2}\gamma\widehat{\sigma}_{p,rec}^2$, where the mean and variance are calculated using recession observations only. Expansion values are computed similarly.

These findings echo those of Henkel et al. (2011) for stocks and Andreasen et al. (2016) for bonds. Figure 1 provides more intuition. I modify (4) by conditioning the right-hand side on the NBER dummy:

$$R_{t+1} = \left(\alpha^{REC} + \beta^{REC} X_t\right) NBER_t + \left(\alpha^{EXP} + \beta^{EXP} X_t\right)\left(1 - NBER_t\right) + \widetilde{\varepsilon}_{t+1}. \tag{10}$$

I plot recursive OLS estimates of (10) and (4) showing β^{REC} in recessions (grey bars) and β^{EXP} in expansions. For stocks, the coefficient jumps up in recessions reflecting stronger predictability of returns in bad states of the economy. The increase in stock slope coefficients is consistent with countercyclical equity risk premia, which is a feature of several established asset pricing models such as Campbell and Cochrane (1999) and Bansal and Yaron (2004). However, generating the strong nonlinearity shown in Figure 1 could pose a challenge to these models. Gargano (2013) investigates the issue and indeed finds that adjustments are needed. He proposes a new long-run risk model which has the ability to match the time-varying return predictability in the data.

For bonds, the coefficients in Figure 1 even switch sign being positive in expansions but negative in recessions. Andreasen et al. (2016) show how this pattern is consistent with bond risk premia decreasing during recessions and suggest that it could be fueled by accommodating monetary policy in these periods. The dramatic switches in bond and stock return coefficients we see in Figure 1 explain why the standard forecasting models without switching tend to be unreliable and motivate why I propose new models which allow for asymmetric predictability over the business cycle.

4. Gains from switching predictability

Based on the distinctive patterns in return predictability documented in Section 3, I propose two adjustments of the standard regression model. Motivated by Table 1, the restricted coefficient strategy (REST) uses the standard model in (4) but shuts down market timing when it does not work. For stocks, the strategy is

$$R_{t+1} = \begin{cases} \alpha + \beta X_t + \varepsilon_{t+1}, & \text{if } I_t = 1 \\ \alpha_0 + \varepsilon_{0,t+1}, & \text{if } I_t = 0, \end{cases} \tag{11}$$

where I_t is a recession dummy. For bonds, market timing is only activated in expansions. REST exploits how predictability is only there in pockets of time and disappears for stocks in expansions and bonds in recessions.

Motivated by Figure 1, the unrestricted coefficient strategy (UNREST) is more ambitious by allowing for market timing in both states of the economy. UNREST lets the coefficients switch freely based on (10):

$$R_{t+1} = \begin{cases} \alpha^{REC} + \beta^{REC} X_t + \widetilde{\varepsilon}_{t+1}, & \text{if } I_t = 1 \\ \alpha^{EXP} + \beta^{EXP} X_t + \widetilde{\varepsilon}_{t+1}, & \text{if } I_t = 0. \end{cases} \tag{12}$$

Both strategies require a recession dummy, I_t. The two real-time indicators based on ADS and PMI described in the Data Section above are only available in the later part of the sample, but recursive estimation of UNREST requires a full history of dummy observations. Consistent with the assumed three-month publication lag of NBER turning points (see the Data Section), I therefore use the real-time dummies from $t-2$ to t for the estimation at time t, and the NBER dates for the remaining history. In this way I utilize the most popular recession classification for estimation but real-time signals for portfolio construction. In summary, to create a forecast at time t, I go through the following steps:

- Estimate the recession threshold using (3) based on $\{NBER_1, NBER_2, ..., NBER_{t-3}\}$ and $\{Y_1, Y_2, ..., Y_{t-3}\}$.
- Construct a dummy I_t for the most recent period by comparing Y_t with the threshold from the previous step.
- Estimate the forecasting model coefficients in (11)-(12) using $\{NBER_1, NBER_2, ..., NBER_{t-3}, I_{t-2}, I_{t-1}\}$ $\{X_1, X_2, ..., X_{t-3}, X_{t-2}, X_{t-1}\}$, and $\{R_2, R_3, ..., R_{t-2}, R_{t-1}, R_t\}$.[7]

[7] The lagged real-time dummies $\{I_{t-2}, I_{t-1}\}$ needed in this step are available from previous periods.

- Construct return forecasts using X_t, I_t, and the model coefficients.

To identify the impact of using real-time turning points I also show results assuming no publication lag for the NBER dates by using $\{NBER_{t-2}, NBER_{t-1}, NBER_t\}$ for $\{I_{t-2}, I_{t-1}, I_t\}$.

Both of the above strategies are motivated by the recent evidence of asymmetric predictability across the business cycle cited in the introduction. One could worry whether investors back in time would have known that standard models should be adjusted to take switching into account. On the other hand, the early recursive estimates for both bonds and stocks in Figure 1 show the same pattern as the coefficients using the full sample more than 40 years later. UNREST follows directly from this pattern, and REST can be seen as a more conservative implementation of it with less parameters.

4.1. Certainty Equivalent Returns

Table 2 presents ΔCER values for a mean-variance utility investor comparing the REST and UNREST strategies with the no-predictability benchmark. I also show results for the constant coefficient model in (4). To get an aggregate measure of the economic value of predictability, I consider a joint setup with both stocks, bonds, and the risk-free asset. Portfolio returns are therefore now calculated as

$$R_{p,t+1} = \boldsymbol{\omega}_t^{\prime} \boldsymbol{R}_{t+1} + R_{f,t+1}, \tag{13}$$

where $\boldsymbol{R}_{t+1} = (R_{s,t+1}, R_{b,t+1})^{\prime}$ is the vector of excess returns on stocks and bonds, and $\boldsymbol{\omega}_t = (\omega_{s,t}, \omega_{b,t})^{\prime}$ is the vector of portfolio weights.

Panel A presents results with weights on the two risky assets restricted to the interval $[0, 1]$, and panel B presents results for the interval $[-0.5, 1.5]$. The weight restrictions apply to each asset individually and to the sum of weights. At each point in time, the weights are computed through numerical maximization of (7) with the short selling constraints imposed. Expected portfolio return is computed using the vector of stock and bond return forecasts, $\widehat{\boldsymbol{R}}_{t+1}$, from a given regression model, and portfolio variance is computed using the 60-month rolling covariance matrix estimate, $\widehat{\boldsymbol{\Sigma}}_{t+1}$. Using an Exponentially Weighted Moving Average (EWMA) covariance estimator gives similar results. All models are estimated recursively to simulate a real-time trading environment. The coefficient of relative risk aversion is 5, but setting it to other reasonable values such as 2 or 10 shows a similar picture.

In parentheses are p-values from the McCracken and Valente (2016) test. Since ADS and PMI are not available for the full sample, I proxy their missing observations with the NBER dummy to be able to apply the stationary bootstrap procedure. First, this choice has no effect on the ΔCER point estimates but only affects the bootstrap distribution. Second, as demonstrated below in Section 4.3, the results are robust to alternative ways of handling these missing observations.

The ΔCER figures in Table 2 indicate that a mean-variance utility investor who cannot short sell would be willing to pay a fixed annual fee of 0.8% to market time using the standard regression model rather than assume zero predictability. The utility gain is statistically insignificant with a p-value of around 0.20. However, the REST strategy using NBER dates increases the fee to 1.5% and lowers the p-value to 0.03. When allowing for short selling the fee goes from 0.3% to 2.4% and the p-value decreases from 0.40 to 0.02 by using REST rather than the standard model. Further, the performance improvement is not limited to the NBER turning points. Both ADS and PMI seem to be good real-time indicators of recessions.

The UNREST strategy also provides economic value, however, the utility gains are generally not statistically significant using this simple specification with one predictor. UNREST introduces extra parameter uncertainty by partitioning the observations according to I_t when estimating the model. Further, the choice of recession indicator seems quite important. The strategy even loses versus the no-predictability benchmark for the PMI dummy. Figure 2 provides an explanation. It shows NBER dates versus the raw ADS and PMI indices with stars indicating values below the real-time thresholds. The plot reveals that PMI misses the early part of some recessions such as the recent financial crisis. Table 3 quantifies the effect. For 6% of the months in the evaluation period, PMI signals an expansion while, at the same time, the NBER committee signals a recession. This is a large effect given that NBER recessions only constitute 14% of months in the evaluation period. Imprecise turning points can spoil business cycle strategies such as REST and UNREST.

4.2. Portfolio weights

Figure 3 shows stock and bond weights over time for the REST strategy (using NBER dates), the constant coefficient model, and the no-predictability benchmark. Weights for the UNREST strategy are in the Appendix. As expected, the REST weights bounce between the constant coefficient weights and the no-predictability weights depending on the state of the economy. One interesting period in time is the bull-market of the 1990s during which REST places a higher weight on stocks than the constant coefficient model. This is because the low level of dividend-price ratios after 1990 pulls down the stock weight in the standard regression model through a low forecast of the equity premium, see also Campbell and Thompson (2008). Meanwhile, since the U.S. economy is expanding for most of this period, REST adheres to the no-predictability benchmark. The no-predictability forecasts are detached from the dividend-price ratio and therefore more optimistic about equity returns, which pulls up the weight on stocks.

Overall, the weights of the two market timing strategies display more variability than the no-predictability benchmark. The impact of transaction costs on relative performance could therefore be a concern. However, the Appendix shows how transaction costs would need to be unrealistically high given the portfolio turnover reflected in Figure 3 to eliminate the gains associated with market timing for the more reliable REST strategy.

4.3. Robustness

The results presented above are robust to alternative ways of handling the missing real-time recession observations and also hold up when using alternative business cycle dates. Further, the portfolio returns are stable over time. I now go through each of these robustness checks in detail.

4.3.1. Handling missing observations

In the main specification ADS and PMI are proxied with the NBER dummy in the early part of the sample to be able to apply the stationary bootstrap of McCracken and Valente (2016).

While this choice has no impact on the ΔCER estimates, it does potentially affect their simulated distribution and consequently the p-values. To investigate whether my conclusions are sensitive to the treatment of missing observations, I recompute the p-values in Table 2 using two alternative approaches. In the first called Overlap, I only draw observations for the bootstrap from the part of the sample where the given recession signal is available without using a proxy. For ADS this is from 1960:3 and for PMI this is from 1948:1. In each simulated sample I repeat the original construction of real-time dummies using recursive ROC analysis of the bootstrapped raw recession signals and NBER dates. In the second alternative I estimate a simple Markov chain for each real-time recession dummy and simulate from it independently of the other variables.[8] This approach explicitly imposes orthogonality between the real-time signal and the financial markets under the null. From Table 4 it is evident how the significance shown above is robust to the two variations of the bootstrap. If anything, the proxy approach seems conservative when comparing the p-values in Table 2 and Table 4.

4.3.2. Other turning points

In the above tables I benchmark the real-time recession indicators against the NBER dates as they are the standard choice in the literature. The NBER turning points are decided periodically by the members of the Business Cycle Dating Committee based on a range of indicators. Hamilton (2011) argues that the committee decisions should be supplemented with more mechanical algorithms. Two examples of such approaches are Hamilton's GDP-based recession probabilities and the Chauvet and Piger (2008) dynamic factor model recession probabilities. I collect Hamilton's turning points and the Chauvet and Piger probabilities from the FRED database and proxy the early part of the sample with NBER dates. The Hamilton series start in 1967:10 and the Chauvet and Piger series start in 1967:6, both before the beginning of the evaluation period in 1971:1.[9] Consistent with the

[8] The estimates of the diagonal elements in the transistion probability matrix are $\widehat{p}_{ii} = \frac{n_{ii}}{\sum_{j=1}^{2} n_{ij}}$ for $i = 1, 2$, where n_{ij} is the number of transistions from state i to state j in the sample. The off-diagonal elements follow directly given that each row must sum to one.

[9] Hamilton's dates are technically not available in real-time since the probabilities are only available on a quarterly basis, are lagged by one quarter, and his dating algorithm uses smoothed probabilities to identify beginnings of recessions. I assume the quarterly turning points are effective from the first month of each quarter. The Chauvet and Piger probabilities are also not available in real-time as only smoothed probabilities are publicly available, and the probabilities are only released with a two-month lag. I use a simple cutoff of 50% to convert their probabilities into

assumed three-month NBER publication lag, I also include results lagging the NBER dummy three months as an alternative to the above ROC based real-time dummies. Table 5 presents results using these alternative recession dates. All three approaches produce meaningful increases in ΔCER for REST relative to the constant coefficient model, whereas UNREST gives less reliable improvements for these specifications.

4.3.3. Performance over time

Figure 4 sheds more light on the robustness of the REST and UNREST strategies over time. In the left side of the plot, I show the cumulative log ΔCER computed relative to the no-predictability benchmark for the case of short selling. The cumulative ΔCER measures the compounded risk-adjusted gain from using REST or UNREST rather than the benchmark. In the right side of the plot I show cumulative log returns also relative to the no-predictability benchmark. From the upper part of the plot, we see that the improved performance from (occasionally) timing the market with REST versus assuming zero predictability clearly is not confined to any period. Both the cumulative ΔCER and the cumulative return gain are increasing rather steadily over time. From the lower part of the plot it is evident how UNREST loses on market timing during the 1990s. The REST model ignores the low dividend-price ratio during this period and therefore does not suffer from this loss. Further, we see how the choice of recession signal is important for UNREST as the real-time PMI dummy performs consistently worse than the no-predictability benchmark over time.

5. More predictors

So far, I have only presented results using a single predictor per risky asset. One could be concerned that switching only matters if you start out with a poor forecasting model. To address this worry, I expand the information set to $i = 1, ..., N$ predictors. When using many predictors the risk of in-sample overfitting is higher, and I deal with this issue using two popular approaches from the recent literature on return forecasting. For each approach, I implement multivariate versions of

dummy variables. Chauvet and Piger (2008) show how real-time probabilities from their model line up with NBER turning points.

REST and UNREST as well as a standard constant coefficient model.

The first is the mean forecast combination (FC):

$$R_{t+1} = \frac{1}{N}\sum_{i=1}^{N}\left(\alpha_i + \beta_i X_{i,t} + \varepsilon_{i,t+1}\right), \tag{14}$$

where $(\alpha_i, \beta_i)'$ are the OLS estimates for predictor i. See, for instance, Rapach et al. (2010).

Secondly, I use diffusion indices (DI):

$$R_{t+1} = \alpha + \boldsymbol{\beta}'\mathbf{F}_t + \varepsilon_{t+1}, \tag{15}$$

where $\mathbf{F}_t$ holds common factors from principal component analysis of the N predictors. See Ludvigson and Ng (2007, 2009).

I follow Neely et al. (2014) and use a broad panel of 14 economic (econ) and 14 technical predictors (tech). Rapach and Zhou (2013) and Neely et al. (2014) provide evidence of switching return predictability for several of these predictors for the case of stocks. The group of economic predictors is very popular in the literature and is comprised of: the log dividend-price ratio, term spread, log dividend yield, log earnings-price ratio, log payout ratio, realized volatility, book-to-market ratio, net equity expansion, T-bill rate, long-term bond yield, long-term bond return, default yield spread, default return spread, and lagged inflation. The first six technical predictors I use are trend-following strategies based on moving averages of the S&P 500 index, P_t. Let $S_{i,t} = 1$ denote a buy signal and $S_{i,t} = 0$ a sell signal. The signals are generated as follows:

$$S_{i,t} = \begin{cases} 1, & \text{if } MA_{s,t} \geq MA_{l,t} \\ 0, & \text{if } MA_{s,t} < MA_{l,t}, \end{cases} \tag{16}$$

where

$$MA_{j,t} = \frac{1}{j}\sum_{i=0}^{j-1} P_{t-i} \text{ for } j = s, l, \tag{17}$$

with $s = 1, 2, 3$ and $l = 9, 12$. The next two predictors capture momentum:

$$S_{i,t} = \begin{cases} 1, & \text{if } P_t \geq P_{t-m} \\ 0, & \text{if } P_t < P_{t-m}, \end{cases} \tag{18}$$

where $m = 9, 12$. For further details about these two variables, see Neely et al. (2014). Finally, I compute six bond moving average signals using the long term yield instead of the S&P 500 index in (16)-(17). All predictors are based on the updated Goyal and Welch (2008) data set.

For DI, I use the first common factor when only including the 14 economic predictors. When including both the economic and technical predictors, I follow Neely et al. (2014) and use the first four common factors to fully take advantage of the two different information sets. Both REST and UNREST strategies for FC and DI are straightforward extensions of the single predictor case. REST only activates the given forecasting model in recessions for stocks and in expansions for bonds. It uses the no-predictability benchmark for stocks in expansions and bonds in recessions. UNREST interacts predictors and intercepts with recession and expansion dummies. All models are estimated recursively.

5.1. Statistical and economic significance

It is important to address the accumulation of size that arises due to analyzing a large number of predictors combined with several model specifications. I therefore employ the fixed-regressor bootstrap procedure developed by Clark and McCracken (2012) to address multiple testing. Their test is based on the maximum mean squared forecast error F statistic (MSE$-F$) across the M models under consideration:

$$\max_{j=1,\ldots,M} [\text{MSE} - F_j] = \max_{j=1,\ldots,M} \left[(T - T_0) \times \left(\frac{\text{MSE}_0 - \text{MSE}_j}{\text{MSE}_j} \right) \right], \tag{19}$$

where $\text{MSE}_j = (T - T_0)^{-1} \sum_{t=T_0}^{T-1} \widehat{\varepsilon}^2_{j,t+1}$, and MSE_0 is computed similarly for the benchmark. The bootstrap is designed to take heteroskedasticity into account and explicitly imposes the null by simulating from the no-predictability model. For more details see Clark and McCracken (2012). In each of the 9,999 bootstrap samples I search for the max MSE$-F$ across all combinations of forecasting models (bivariate, FC, DI), switching strategies (CONST, REST, UNREST), and main recession signals (NBER, ADS, PMI) considered so far.

In Table 6 I show the top five models sorted by MSE$-F$ for stocks and bonds separately. Next to the max MSE$-F$ is the bootstrap p-value computed as the fraction of simulated max MSE$-F$ higher than the original sample value. Next to R^2_{oos}s are Clark and West (2007) p-values. With max MSE$-F$ p-values of 0.021 for stocks and 0.009 for bonds, it does not seem likely that I find return predictability due to searching across many models. Interestingly, for bonds the bivariate model performs the best while for stocks DI and FC models comprise the top five. For both stocks and bonds, however, UNREST and REST strategies are the top performers illustrating the importance of conditioning on the state of the economy in return forecasting.

Turning to economic significance, Table 7 shows how the UNREST strategy seems to interact particularly well the mean combination model, which is designed to reduce forecast variability.[10] The REST strategy provides good results for the diffusion index. Looking at econ forecast combinations, the performance fee without short selling increases from roughly 1% to 2.4% using the UNREST model with NBER dates rather than the standard model. With short selling, the fee increases from around 1.2% to 2.8%. While meaningful ΔCERs are also realized using the most precise turning points from ADS, we again see how using popular but less accurate recession signals can turn the gains into losses. As an example, econ forecast combinations with the UNREST strategy and short selling provides a risk-adjusted return of roughly 2.4% using ADS but -0.1% using PMI.

Finally, looking at models that only rely on economic predictors (econ) versus both economic and technical predictors (econ, tech), using an expanded information set gives some performance improvements. However, taking switching into account seems to have a larger effect on risk-adjusted returns than the number of predictors. In addition, when comparing Table 2 and 7, the multivariate strategies do not seem to improve reliably on the strategies with one predictor, which suggests that the more parsimonious choice could be sufficient to capture economic gains from switching

[10] For the UNREST FC strategies I exclude predictors if they have zero variation in an economic state in a given bootstrap sample. This happens for a few of the technical indicators but only rarely.

predictability.

6. Markov switching

The Markov switching vector autoregression (MS-VAR) is an alternative approach to capturing regime switching. Notable examples from finance include Henkel et al. (2011), Guidolin and Timmermann (2007), and Guidolin and Hyde (2012). MS-VARs are an appealing way of modeling business cycles, because the states are treated as unobservable. In this way, MS-VARs naturally incorporate uncertainty (even ex-post) about which economic regime each period belongs to. On the other hand, the models are highly nonlinear and sometimes produce wild forecasts. The MS-VAR is

$$\mathbf{Y}_{t+1} = \boldsymbol{A}_{s_t} + \boldsymbol{B}_{s_t}\mathbf{Y}_t + \boldsymbol{\eta}_{t+1}, \tag{20}$$

where $\mathbf{Y}_{t+1}$ is a vector of returns and predictors, $\boldsymbol{\eta}_{t+1} \sim N(\mathbf{0}, \boldsymbol{\Omega}_{s_t})$, and the latent state variable s_t follows a two-state Markov chain with a fixed transition probability matrix:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \tag{21}$$

where p_{ij} refers to the probability of switching from state i at time t to state j at time $t+1$. I impose restrictions on $\boldsymbol{B}_{s_t}$ to reduce the parameter space. The first specification of returns is a bivariate model in which stock returns are predicted only by the log dividend-price ratio and bond returns are predicted only by the term spread. This is directly related to the dummy regression in (10), only that the states are now unobserved. The two predictors only load on lagged values of themselves. The second specification I consider uses the diffusion index of economic predictors from Section 5 for both stock and bond returns. The diffusion index only loads on its lagged value. I estimate the models using an expectation maximization algorithm from Krolzig (1997).

I report full-sample estimation results in the Appendix. One regime has lower persistence and much higher volatility than the other. I identify this as the recession regime. Figure 5 shows smoothed state probabilities revealing that the recession regime has a relatively large degree of overlap with NBER dates. Table 8 reports ΔCERs versus the no-predictability benchmark.[11] For comparison, I also report results for the constant coefficient versions of the bivariate and diffusion index models in (4) and (15). All models are estimated recursively. The MS-VAR models offer little to no improvement on the constant coefficient specifications. With short selling, the MS-VAR DI model even performs worse than the no-predictability benchmark. While the full-period estimates of the MS-VARs seemed to produce regimes similar to the NBER classification, the out-of-sample performance of the MS-VARs is underwhelming.[12] The added flexibility from not imposing an economic structure on the regimes seems to come at the cost of poor forecasting performance. The limited amount of recession observations makes it especially challenging to achieve accurate parameter estimates of a heavily parameterized and strongly nonlinear model like the MS-VAR.[13] The simpler REST and UNREST strategies using dummy variables and linear regressions appear to be a more attractive way of capturing regime switches in predictability.

7. Conclusion

I measure the economic value of allowing return predictability to switch across the business cycle. I do this for a mean-variance utility investor choosing between stocks, bonds, and a risk-free asset. First, I let stocks only be predictable in recessions and bonds only be predictable in expansions. Second, I let the regression intercept and slope coefficients change freely across recessions and expansions. Both strategies combine recession dummies with standard return forecasting methods such as bivariate regressions and diffusion indices. For the most accurate recession signals, these dummy based strategies considerably improve risk-adjusted returns when compared to constant coefficient forecasts. However, the gains depend strongly on the choice of recession indicator, and I show how choosing the wrong turning points can have large economic consequences.

[11] See Section 4 for details on the calculation of optimal weights. I compute return forecasts by weighting conditional means across states using estimated state probabilities. Further, the covariance matrix is the same 60-month rolling sample estimate I use for the other models to isolate the effect of risk premium forecasts.

[12] Guidolin and Timmermann (2007) and Guidolin and Hyde (2012) find economic gains to using Markov switching models with four and three states, respectively.

[13] Henkel et al. (2011) use Bayesian techniques to get stable performance by conditioning on prior information.

References

Andreasen, M.M., Engsted, T., Møller, S.V., Sander, M., 2016. Bond market asymmetries across recessions and expansions: new evidence on risk premia. Unpublished working paper. Aarhus University.

Aruoba, S.B., Diebold, F.X., Scotti, C., 2009. Real-time measurement of business conditions. Journal of Business and Economic Statistics 27 (4), 417-427.

Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: some utility cost calculations. Journal of Financial Economics 52, 47–78.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Berge, T., Jordà, Ò., 2011. Evaluating the classification of economic activity into expansions and recessions. American Economic Journal: Macroeconomics 3, 246–277.

Binsbergen, J.H.v., Brandt, M.W., 2007. Solving dynamic portfolio choice problems by recursing on optimized portfolio weights or on the value function. Computational Economics 29, 355–367.

Brandt, M.W., Goyal, A., Santa-Clara, P., Stroud, J.R., 2005. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. Review of Financial Studies 18 (3), 831–873.

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

Campbell, J.Y., and R.J. Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1 (3), 195–227.

Campbell, J.Y., Thompson, S.B., 2008. Predicting the equity premium out of sample: can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Cochrane, J.H., 2017. Macro-finance. Review of Finance 21 (3), 945-985.

Chauvet, M., Piger, J., 2008. A comparison of the real-time performance of business cycle dating methods. Journal of Business and Economic Statistics 26 (1), 42–49.

Christiansen, C., Eriksen, J.N., Møller, S.V., 2014. Forecasting US recessions: the role of sentiment. Journal of Banking and Finance 49, 459-468.

Clark, T.E., McCracken, M.W., 2012. Reality checks and comparisons of nested predictive models. Journal of Business and Economic Statistics 30 (1), 53–66.

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

Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. Journal of Financial Economics 106, 157-181.

Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23-49.

Gargano, A., 2013. Return predictability in recessions: an asset pricing perspective. Unpublished working paper. University of California, San Diego.

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

Guidolin, M., Hyde, S., 2012. Can VAR models capture regime shifts in asset returns? A long-horizon strategic asset allocation perspective. Journal of Banking and Finance 36, 695-716.

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

Hamilton, J.D., 2011. Calling recessions in real time. International Journal of Forecasting 27, 1006-1026.

Han, Y., 2006. Asset allocation with a high dimensional latent factor stochastic volatility model. Review of Financial Studies 19 (1), 237–271.

Henkel, S.J., Martin, J.S., Nadari, F., 2011. Time-varying short-horizon predictability. Journal of Financial Economics 99, 560–580.

Kacperczyk, M., Van Nieuwerburgh S., Veldkamp L., 2014. Time-varying fund manager skill. Journal of Finance 69, 1455-1484.

Kauppi, H., 2008. Yield-curve based probit models for forecasting U.S. recessions: stability and dynamics. Helsinki Center of Economic Research Discussion Paper 221, 1–36.

Krolzig, H.-M., 1997. Markov-switching vector autoregressions: modelling, statistical inference, and application to business cycle analysis. Springer-Verlag, Berlin.

Ludvigson, S.C., Ng, S., 2007. The empirical risk-return relation: a factor analysis approach. Journal of Financial Economics 83, 171–222.

Ludvigson, S.C., Ng, S., 2009. Macro factors in bond risk premia. Review of Financial Studies 22 (12), 5027–5067.

McCracken, M.W., Valente, G., 2016. Asymptotic inference for performance fees and the predictability of asset returns. Journal of Business and Economic Statistics, forthcoming.

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

Ng, E.C., 2012. Forecasting US recessions with various risk factors and dynamic probit models. Journal of Macroeconomics 34, 112–125.

Pesaran, M.H., Timmermann, A., 2002. Market timing and return prediction under model instability. Journal of Empirical Finance 9, 495–510.

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

Rapach, D.E., Zhou, G., 2013. Forecasting stock returns. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting 2. Elsevier, Amsterdam.

Thornton, D.L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: an asset allocation perspective. Review of Financial Studies 25 (10), 3141-3168.

Timmermann, A., 2008. Elusive return predictability. International Journal of Forecasting 24, 1–18.

A. Appendix

A.1. Estimation results

Table A.1 shows OLS estimates and White standard errors for the bivariate and diffusion index models in panels A and B. In panels C and D are estimates of the corresponding Markov switching models. I estimate the Markov switching models using the expectation maximization algorithm of Krolzig (1997) and compute standard errors using the outer product method.

Table A.1: Full period estimates

This table reports estimation results for forecasting models of stock and bond returns. I show results for a bivariate regression (Bivariate) and a diffusion index (DI) regression. Bivariate uses the log dividend-price ratio to predict stock returns and the term spread to predict bond returns. DI uses the first common factor from principal component analysis of 14 economic predictors to forecast both returns. In panels A and B, the parameters switch according to the NBER dummy. These models are estimated using OLS, and White standard errors are in parentheses. The MS Bivariate and MS DI models in panels C and D are Markov switching versions of the models in panels A and B. The MS models are estimated using an expectation maximization algorithm, and the standard errors in parentheses are computed using the outer product method. Slope is the estimated slope coefficient, Variance is the estimated error variance, and Stay Probability (Stay Prob.) is the estimated probability of staying in a given state from month to month. The sample period is 1927:01-2013:12.

	Panel A: Bivariate		Panel B: DI	
	Recessions	Expansions	Recessions	Expansions
Stock Slope	0.026	0.007	0.543×10^{-2}	0.215×10^{-2}
	(0.021)	(0.004)	(0.507×10^{-2})	(0.106×10^{-2})
Bond Slope	-0.217	0.345	0.029×10^{-2}	-0.056×10^{-2}
	(0.280)	(0.068)	(0.077×10^{-2})	(0.041×10^{-2})

	Panel C: MS Bivariate		Panel D: MS DI	
	Recessions	Expansions	Recessions	Expansions
Stock Slope	0.025	0.005	0.837×10^{-2}	0.255×10^{-2}
	(0.011)	(0.003)	(0.309×10^{-2})	(0.075×10^{-2})
Bond Slope	0.240	0.249	-0.102×10^{-2}	-0.007×10^{-2}
	(0.105)	(0.051)	(0.183×10^{-2})	(0.046×10^{-2})
Stock Variance	0.722×10^{-2}	0.141×10^{-2}	1.005×10^{-2}	0.135×10^{-2}
	(0.044×10^{-2})	(0.008×10^{-2})	(0.117×10^{-2})	(0.007×10^{-2})
Bond Variance	0.117×10^{-2}	0.035×10^{-2}	0.155×10^{-2}	0.037×10^{-2}
	(0.007×10^{-2})	(0.002×10^{-2})	(0.014×10^{-2})	(0.002×10^{-2})
Stay Prob.	0.82	0.93	0.81	0.96
	(0.05)	(0.04)	(0.07)	(0.03)

A.2. UNREST weights

Figure A.1 shows stock and bond weights for the no-predictability benchmark, the bivariate constant coefficient model, and the bivariate UNREST model.

Figure A.1: Stock and bond weights of the UNREST strategy

The plot shows recursive stock and bond weights for the CONST and UNREST models in Table 2. UNREST uses the NBER recession dates (grey bars). I also show weights from the no-predictability benchmark (NOPRED).

A.3. Transaction cost

Table A.2 reports break-even transaction cost measuring the unit fee that would make the investor indifferent between using a candidate forecasting model and the benchmark. Let $W_t = 1 + R_{p,t}$ be portfolio wealth at time t. Following Han (2006) and Thornton and Valente (2012), I calculate average traded value (turnover) as

$$V = \frac{1}{T - T_0 - 1} \sum_{t=T_0+1}^{T-1} \left(\left| \omega_{s,t} - \omega_{s,t-1} \frac{(1 + R_{s,t} + R_{f,t})}{W_t} \right| + \left| \omega_{b,t} - \omega_{b,t-1} \frac{(1 + R_{b,t} + R_{f,t})}{W_t} \right| \right).$$

I then measure break-even cost using

$$\tau^{BE} = \frac{\frac{1}{T-T_0} \sum_{t=T_0}^{T-1} (W_{t+1} - W_{0,t+1})}{V - V_0},$$

where V and V_0 are the portfolio and benchmark average traded value, respectively. Table A.2 both reports break-even cost in percent and annualized traded value in percent. The break-even costs for REST are all much higher than the normal range considered in the literature. As an example, Balduzzi and Lynch (1999) use 0.50% for individual stocks and 0.01% for S&P 500 futures.

Table A.2: Break-even transaction cost

This table reports unit transaction cost, τ^{BE}, that would make a mean-variance utility investor indifferent between the candidate model and the benchmark. $12 \times V$ measures the (annualized) traded value for each strategy. Both are in percent. The numbers correspond to the CONST, REST, and UNREST models in Table 2.

	Panel A: $LB = 0,\ UB = 1$				Panel B: $LB = -0.5,\ UB = 1.5$			
	CONST				CONST			
	τ^{BE}		$12 \times V$		τ^{BE}		$12 \times V$	
	0.01		109		-0.26		181	
	REST		UNREST		REST		UNREST	
	τ^{BE}	$12 \times V$	τ^{BE}	$12 \times V$	τ^{BE}	$12 \times V$	τ^{BE}	$12 \times V$
NBER	2.35	91	0.48	155	2.52	171	1.17	283
ADS	1.66	105	0.13	210	1.86	198	0.54	389
PMI	1.76	94	-0.75	155	2.18	176	-0.29	286

Table 1: Out-of-sample return predictability

The table reports out-of-sample R^2 (R^2_{oos}) and annualized certainty equivalent returns (ΔCERs) relative to the no-predictability benchmark. Both are in percent. The metrics are based on recursive estimates of $R_{t+1} = \alpha + \beta X_t + \varepsilon_{t+1}$, where R_{t+1} is either excess stock returns (panel A) or excess bond returns (panel B). The predictor X_t is the log dividend-price ratio for stocks and the term spread for bonds. The evaluation period is 1971:1 to 2013:12, and I report results for both the full period and conditioning on NBER dated recessions and expansions. ΔCER is computed for a one-month mean-variance utility investor with a coefficient of relative risk aversion of 5. The investor has access to a single risky asset (either stocks or bonds) and T-bills. The weight on the risky asset can vary between 0% and 100%. The p-values in parentheses are from the Clark and West (2007) test for R^2_{oos} and the McCracken and Valente (2016) test for ΔCER.

	Full period	Recessions	Expansions
		Panel A: Stocks	
R^2_{oos}	-0.68 (0.13)	2.57 (0.02)	-1.91 (0.33)
ΔCER	-0.62 (0.78)	5.37 (0.06)	-1.63 (0.97)
		Panel B: Bonds	
R^2_{oos}	1.42 (0.01)	-1.68 (0.65)	2.69 (0.00)
ΔCER	1.14 (0.03)	-1.37 (0.86)	1.55 (0.03)

Table 2: Economic value of switching predictability

This table reports annualized certainty equivalent returns (ΔCERs) relative to the no-predictability benchmark for a mean-variance utility investor with access to stocks, bonds, and T-bills. The ΔCERs are in percent. CONST uses a bivariate regression: $R_{t+1} = \alpha + \beta X_t + \varepsilon_{t+1}$, where R_{t+1} is either excess stock returns or excess bond returns. The predictor X_t is the log dividend-price ratio for stocks and the term spread for bonds. REST only uses the bivariate regression for stocks in recessions and bonds in expansions. It shifts to the no-predictability benchmark for stocks in expansions and bonds in recessions. UNREST uses $R_{t+1} = (\alpha^{REC} + \beta^{REC} X_t) I_t + (\alpha^{EXP} + \beta^{EXP} X_t)(1 - I_t) + \widetilde{\varepsilon}_{t+1}$, where I_t is a recession dummy. To identify recessions, I use either NBER dates (assuming no publication lag), the Aruoba-Diebold-Scotti index (ADS), or the Purchasing Managers' Index (PMI). For ADS and PMI, I recursively estimate optimal thresholds using ROC analysis. The weights on stocks and bonds are restricted to be between 0% and 100% (panel A) or -50% and 150% (panel B). The restrictions apply to both risky assets individually and to the sum of weights. The models are estimated recursively and the evaluation period is 1971:1-2013:12. The investment horizon is one month and the coefficient of relative risk aversion is 5. The p-values in parentheses are from the McCracken and Valente (2016) test of equal average utility versus the no-predictability benchmark.

	Panel A: $LB = 0,\ UB = 1$		Panel B: $LB = -0.5,\ UB = 1.5$	
	CONST		CONST	
	0.84 (0.20)		0.33 (0.40)	
	REST	UNREST	REST	UNREST
NBER	1.52 (0.03)	1.66 (0.08)	2.37 (0.02)	3.35 (0.03)
ADS	1.34 (0.05)	1.23 (0.15)	2.23 (0.02)	2.37 (0.09)
PMI	1.33 (0.04)	-0.29 (0.61)	2.28 (0.02)	-0.37 (0.62)

Table 3: Recession indicators' overlap with NBER dates

The table reports the overlap between the NBER dummy and other recession dummies as a percentage of number of observations in the evaluation period: 1971:1-2013:12. The other dummy is either the Aruoba-Diebold-Scotti index (ADS) or the Purchasing Managers' Index (PMI) using recursively estimated cutoffs from ROC analysis. The dummies are all lagged to be consistent with the regression models in other tables.

NBER	Expansion	Recession	Expansion	Recession
Other Indicator	Expansion	Recession	Recession	Expansion
NBER	86.0	14.0		
ADS	83.5	13.2	2.5	0.8
PMI	84.3	7.8	1.7	6.2

Table 4: P-values for alternative bootstrap choices

The table reports alternative p-values for the $\Delta CERs$ in Table 2. The Overlap columns correspond to using the McCracken and Valente (2016) bootstrap but only drawing observations from the part of the sample where the given recession indicator is available without using the NBER as a proxy. For ADS this is from 1960:3 and for PMI this is from 1948:1. The Markov Chain columns correspond to simulating recession dummies using ML estimates of a Markov chain where each draw is independent of the bootstrapped data. For further details see Table 2.

	Panel A: $LB = 0,\ UB = 1$			
	REST		UNREST	
	Overlap	Markov Chain	Overlap	Markov Chain
NBER	0.03	0.03	0.08	0.08
ADS	0.01	0.02	0.08	0.18
PMI	0.02	0.02	0.62	0.62

	Panel B: $LB = -0.5,\ UB = 1.5$			
	REST		UNREST	
	Overlap	Markov Chain	Overlap	Markov Chain
NBER	0.02	0.02	0.03	0.03
ADS	0.01	0.01	0.05	0.11
PMI	0.01	0.01	0.60	0.61

Table 5: Alternative turning points

The table reports ΔCERs for alternative business cycle turning points. The NBER lagged row presents results using the NBER dummy lagged three months. The Hamilton row shows results using the Hamilton GDP based turning points. The Chauvet and Piger row shows results using the Chauvet and Piger (2008) dynamic factor model recession probabilities with a threshold of 50%. For further details see Table 2.

	Panel A: $LB = 0,\ UB = 1$		Panel B: $LB = -0.5,\ UB = 1.5$	
	CONST		CONST	
	0.84 (0.20)		0.33 (0.40)	
	REST	UNREST	REST	UNREST
NBER lagged	1.34 (0.05)	0.85 (0.26)	2.26 (0.02)	1.05 (0.31)
Hamilton	2.17 (0.01)	2.37 (0.03)	3.16 (0.01)	3.98 (0.03)
Chauvet and Piger	1.22 (0.06)	0.62 (0.31)	2.11 (0.04)	1.50 (0.22)

Table 6: Reality check

The table reports out-of-sample R^2 (R^2_{oos}) in percent and mean squared error F statistics (MSE-F) for the top five MSE-F models considered. The benchmark is always the no-predictability model. Next to R^2_{oos}s are p-values from the Clark and West (2007) test for equal predictive power and next to the maximum MSE-F is the Clark and McCracken (2012) reality check p-value. The reality check bootstrap uses residuals from an unrestricted model including all economic and technical predictors, as well as the NBER dummy interacted with a constant and the main predictors: the dividend-price ratio and the term spread. CONST uses either a bivariate regression, forecast combination (FC), or diffusion indices (DI) to forecast returns. REST only uses the constant coefficient model for stocks in recessions and bonds in expansions. It shifts to the no-predictability benchmark for stocks in expansions and bonds in recessions. UNREST interacts predictors and intercepts with recession and expansion dummies. econ includes 14 economic predictors. econ, tech includes the 14 economic predictors along with 14 technical indicators. DI econ uses one factor and DI econ, tech uses four factors.

Model	Indicator	Strategy	R^2_{oos}	MSE$-F$
Panel A: Stocks				
DI econ, tech	NBER	REST	2.11 (0.01)	11.15 [0.021]
DI econ, tech	ADS	REST	1.89 (0.01)	9.96
FC econ, tech	NBER	UNREST	1.65 (0.02)	8.63
DI econ	NBER	REST	1.46 (0.00)	7.66
DI econ	NBER	UNREST	1.41 (0.01)	7.38
Panel B: Bonds				
Bivariate	ADS	UNREST	2.39 (0.00)	12.64 [0.009]
Bivariate	NBER	UNREST	2.23 (0.00)	11.75
Bivariate	NBER	REST	1.91 (0.00)	10.05
Bivariate	ADS	REST	1.89 (0.00)	9.94
Bivariate	PMI	UNREST	1.54 (0.01)	8.10

Table 7: More predictors

The table reports annualized certainty equivalent returns (ΔCERs) for multivariate forecasting models. The benchmark is always the no-predictability model. CONST uses forecast combination (FC) or diffusion indices (DI) to forecast returns. REST only uses the constant coefficient model for stocks in recessions and bonds in expansions. It shifts to the no-predictability benchmark for stocks in expansions and bonds in recessions. UNREST interacts predictors and intercepts with recession and expansion dummies. econ includes 14 economic predictors. econ, tech includes the 14 economic predictors along with 14 technical indicators. DI econ uses one factor and DI econ, tech uses four factors. For further details see Table 2.

Panel A: $LB = 0,\ UB = 1$								
	FC econ		FC econ, tech		DI econ		DI econ, tech	
	CONST		CONST		CONST		CONST	
	1.04 (0.04)		1.74 (0.01)		0.51 (0.27)		0.87 (0.28)	
	REST	UNREST	REST	UNREST	REST	UNREST	REST	UNREST
NBER	0.57 (0.08)	2.40 (0.02)	0.96 (0.02)	2.57 (0.00)	1.69 (0.02)	1.14 (0.16)	1.77 (0.04)	2.01 (0.12)
ADS	0.46 (0.10)	2.26 (0.02)	0.88 (0.02)	2.45 (0.01)	1.34 (0.04)	1.18 (0.15)	1.70 (0.03)	2.38 (0.08)
PMI	0.08 (0.38)	0.51 (0.28)	0.30 (0.15)	0.53 (0.28)	-0.28 (0.74)	0.26 (0.40)	0.14 (0.46)	2.45 (0.08)

Panel B: $LB = -0.5,\ UB = 1.5$								
	FC econ		FC econ, tech		DI econ		DI econ, tech	
	CONST		CONST		CONST		CONST	
	1.22 (0.05)		1.66 (0.02)		0.04 (0.47)		0.20 (0.46)	
	REST	UNREST	REST	UNREST	REST	UNREST	REST	UNREST
NBER	0.76 (0.09)	2.76 (0.03)	1.10 (0.03)	3.33 (0.01)	2.57 (0.02)	1.32 (0.22)	2.54 (0.06)	2.16 (0.22)
ADS	0.65 (0.12)	2.42 (0.05)	1.04 (0.03)	3.15 (0.02)	2.11 (0.03)	1.10 (0.26)	2.32 (0.06)	2.80 (0.17)
PMI	0.18 (0.34)	-0.14 (0.53)	0.37 (0.16)	-0.10 (0.53)	-0.57 (0.77)	-1.17 (0.77)	-0.02 (0.55)	2.44 (0.19)

Table 8: Markov switching

This table reports annualized certainty equivalent returns (ΔCERs) relative to the no-predictability benchmark for a mean-variance utility investor with access to stocks, bonds, and T-bills. The ΔCERs are in percent. CONST uses either a bivariate regression (Bivariate) or a diffusion index (DI) to forecast returns. The MS-VARs are Markov switching versions of these models. The weights on stocks and bonds are restricted to be between 0% and 100% (panel A) or -50% and 150% (panel B). The restrictions apply to both risky assets individually and to the sum of weights. See Section 4 for details on how the weights are computed. The models are estimated recursively and the evaluation period is 1971:1-2013:12. The investment horizon is one month and the coefficient of relative risk aversion is 5. The p-values in parentheses are from the McCracken and Valente (2016) test of equal average utility versus the no-predictability benchmark.

	Panel A: $LB = 0,\ UB = 1$		Panel B: $LB = -0.5,\ UB = 1.5$	
	CONST	MS-VAR	CONST	MS-VAR
Bivariate	0.84 (0.20)	0.87 (0.20)	0.33 (0.40)	0.49 (0.40)
DI	0.51 (0.27)	0.10 (0.51)	0.04 (0.47)	-1.55 (0.84)

Figure 1: Slope coefficients

The plot shows recursive estimates of the constant coefficient model $R_{t+1} = \alpha + \beta X_t + \varepsilon_{t+1}$ and the switching coefficient model $R_{t+1} = (\alpha^{REC} + \beta^{REC} X_t) NBER_t + (\alpha^{EXP} + \beta^{EXP} X_t)(1 - NBER_t) + \widetilde{\varepsilon}_{t+1}$. R_{t+1} is either excess stock returns (upper part of the plot) or excess bond returns (lower part of the plot). The predictor X_t is the log dividend-price ratio for stocks and the term spread for bonds. $NBER_t$ is the NBER dummy. The first estimates use data from 1927:1 to 1970:12 and the window is expanding thereafter. I show β^{REC} in NBER recessions (grey bars) and β^{EXP} in expansions for the switching coefficient model.

Figure 2: Recession indicators

The plot shows the Aruoba-Diebold-Scotti index (ADS) and the Purchasing Managers' Index (PMI). NBER dated recessions are marked with grey bars. Stars indicate that the given indicator is below the recursively estimated thresholds using ROC analysis. For details see Section 2.

Figure 3: Stock and bond weights

The plot shows recursive stock and bond weights for the CONST and REST models in Table 2. REST uses the NBER recession dates (grey bars). I also show weights from the no-predictability benchmark (NOPRED). Short selling is allowed.

Figure 4: Cumulative ΔCER and return

The plot shows the cumulative increase in certainty equivalent returns (ΔCERs) and cumulative increase in portfolio returns of the bivariate REST and UNREST strategies in Table 2. I compute both measures relative to the no-predictability benchmark (NOPRED) and measure the returns in logs. I compute the monthly log ΔCERs using an expanding window and then sum them over time. The cumulative returns are the monthly differences in realized log portfolio returns summed over time. Grey bars are NBER recession dates.

Figure 5: Smoothed MS-VAR recession probabilities
The plot shows smoothed recession probabilities (%) using full-sample estimates of the MS-VARs in Section 6. The grey bars are NBER recessions.

Figure1

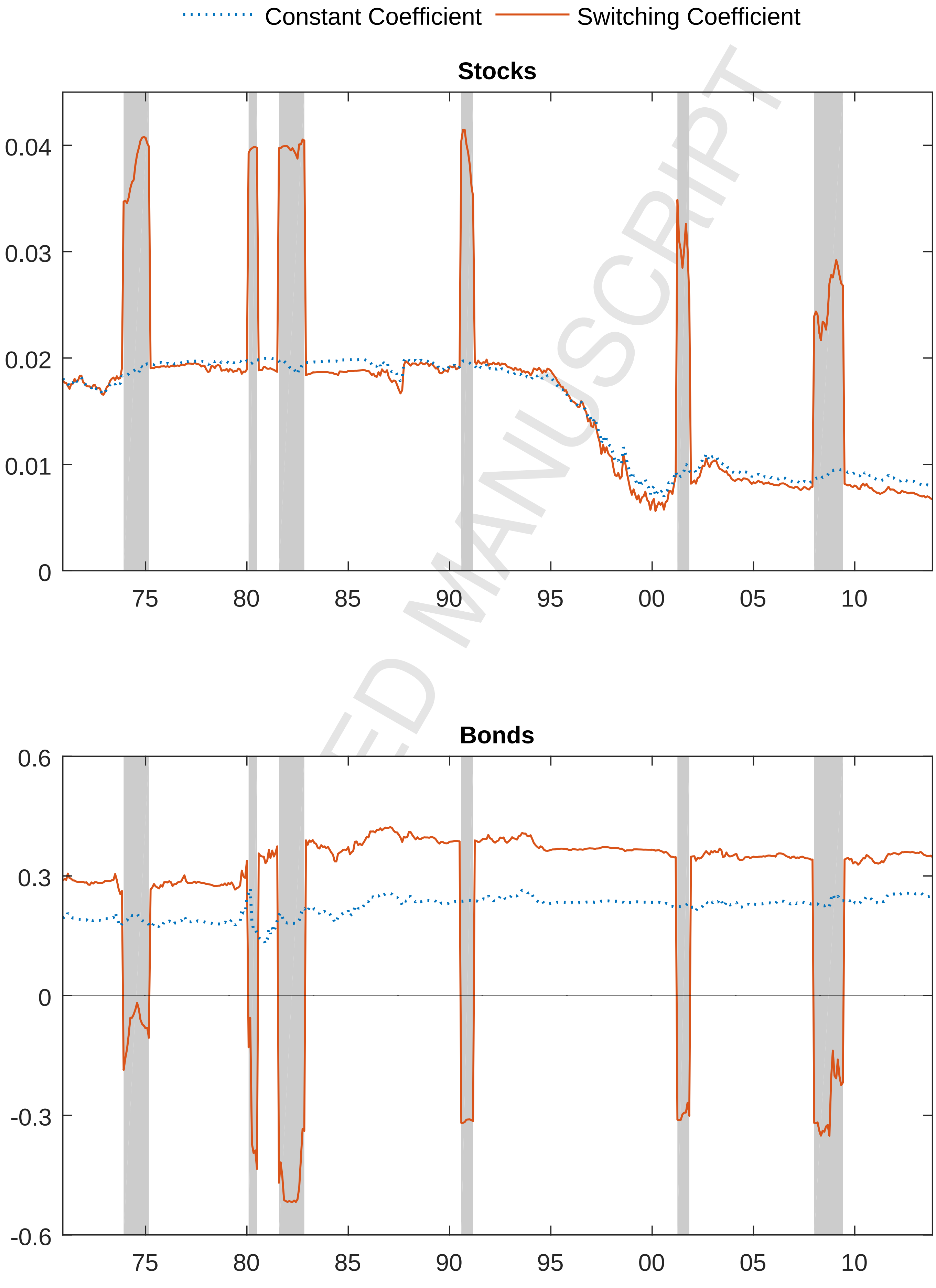

Constant Coefficient
Switching Coefficient
Stocks
0.04
0.03
0.02
0.01
0
75
80
85
90
95
00
05
10
Bonds
0.6
0.3
0
-0.3
-0.6
75
80
85
90
95
00
05
10

Figure2

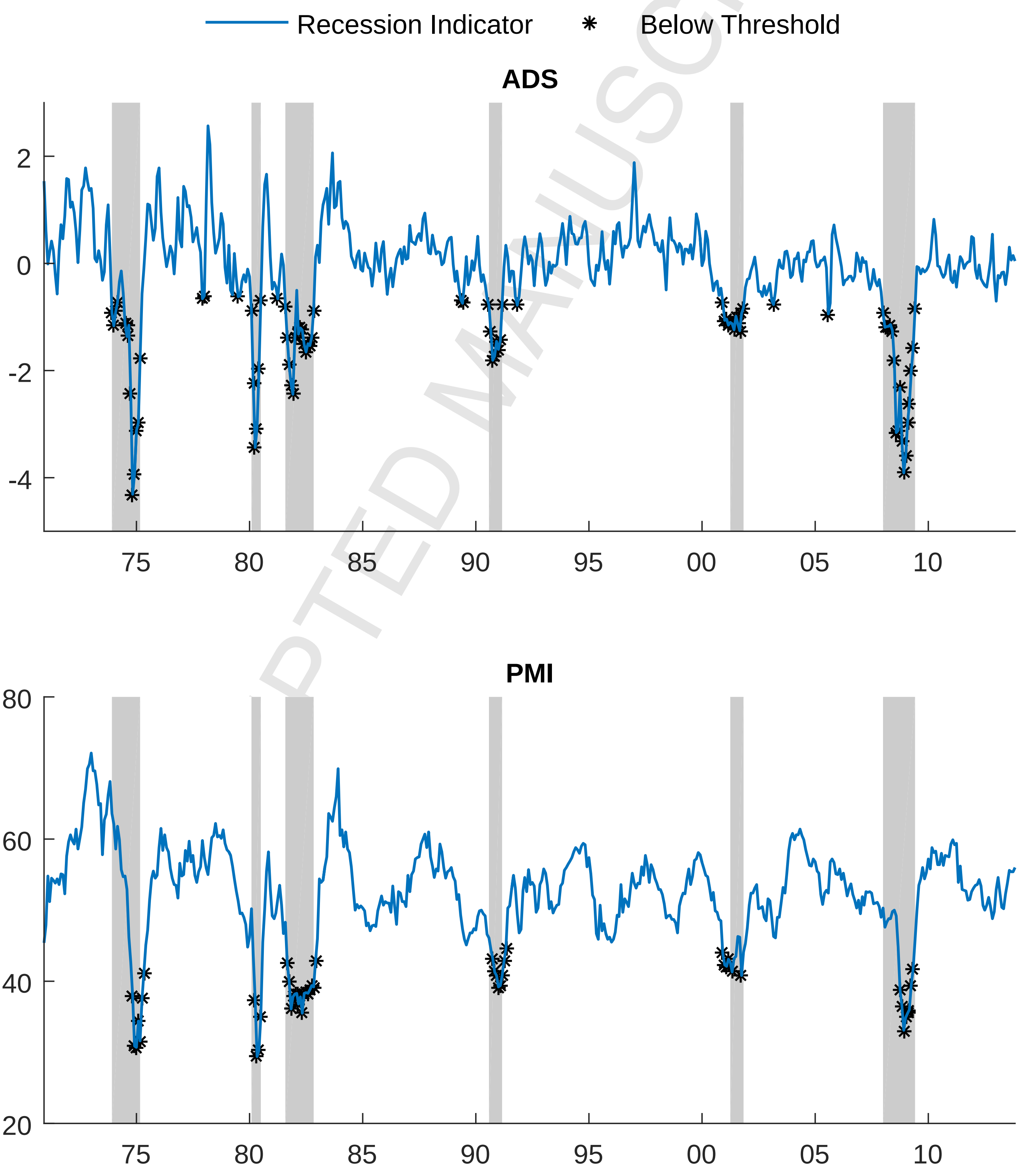

Recession Indicator
Below Threshold
ADS
2
0
-2
-4
75
80
85
90
95
00
05
10
PMI
80
60
40
20
75
80
85
90
95
00
05
10

Figure3

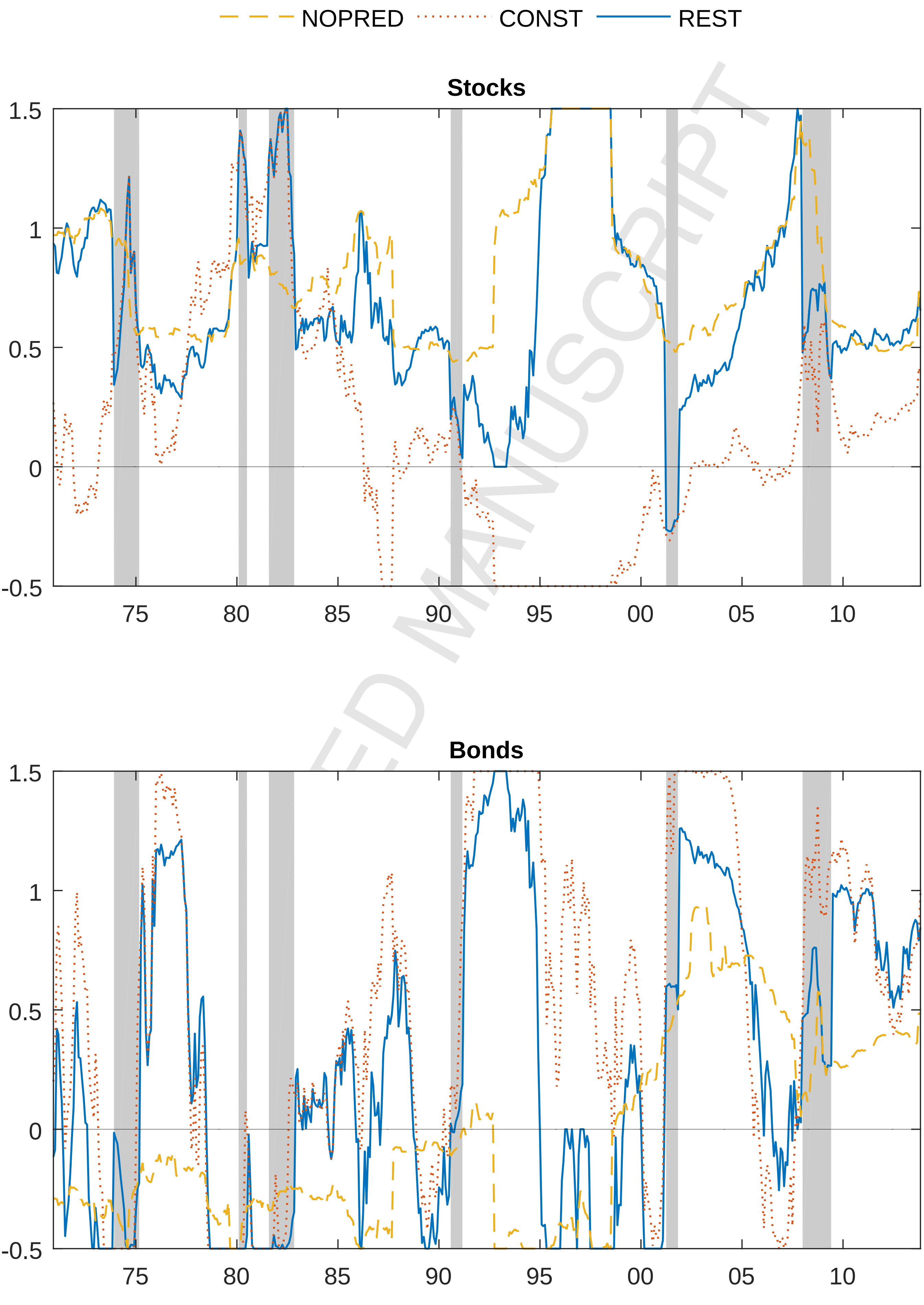

NOPRED
CONST
REST
Stocks
1.5
1
0.5
0
-0.5
75
80
85
90
95
00
05
10
Bonds
1.5
1
0.5
0
-0.5
75
80
85
90
95
00
05
10

Figure4

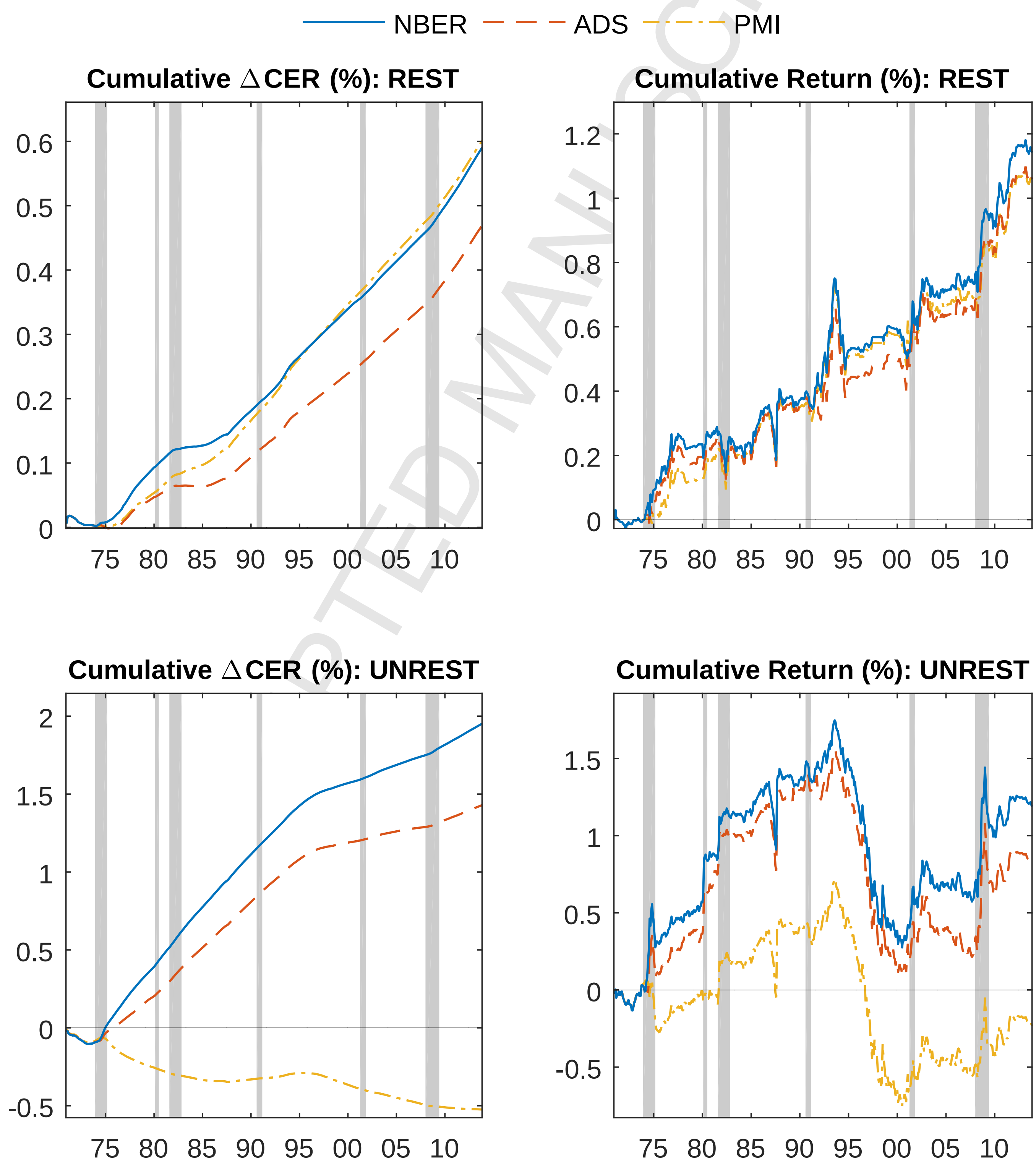
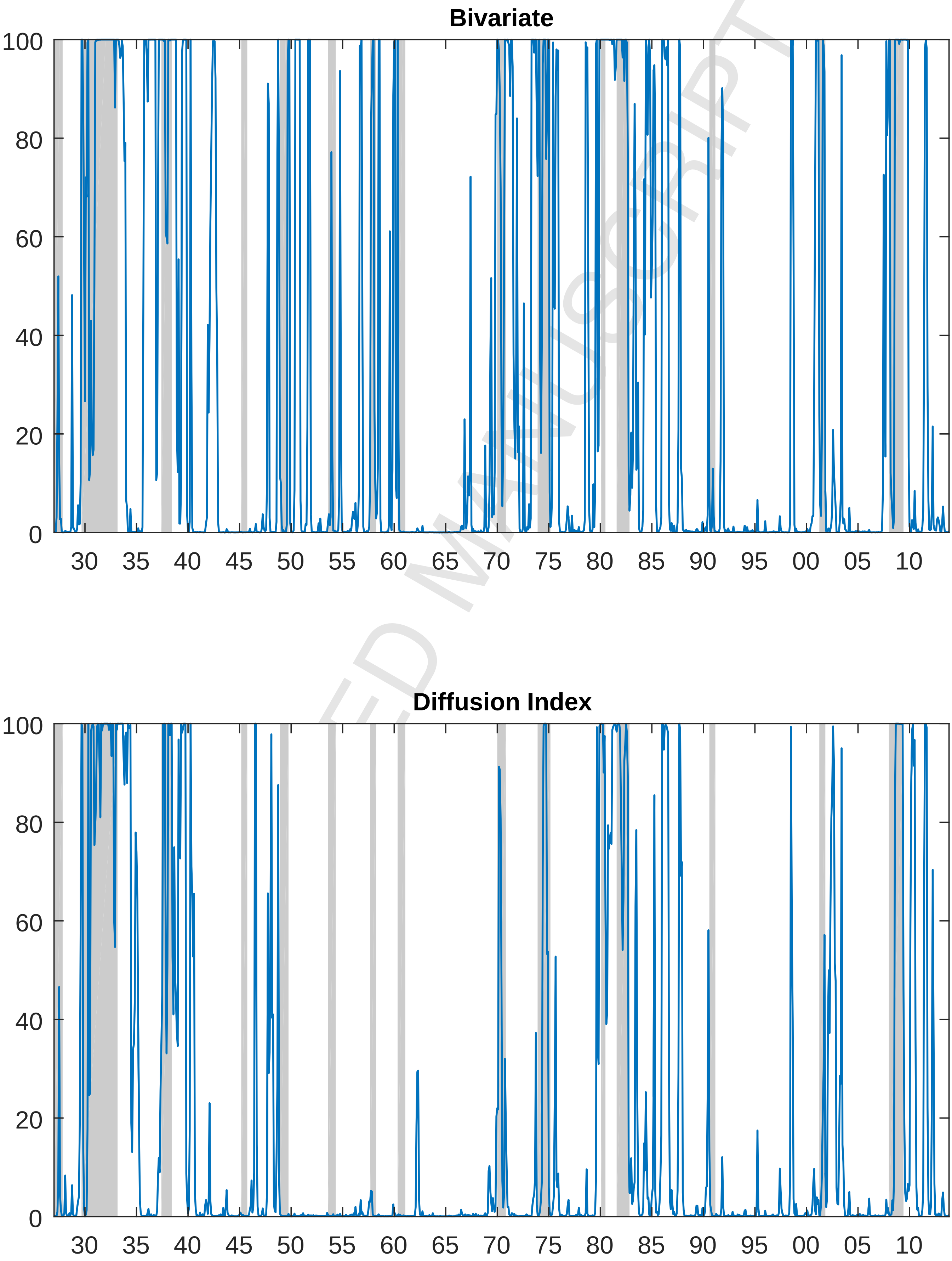
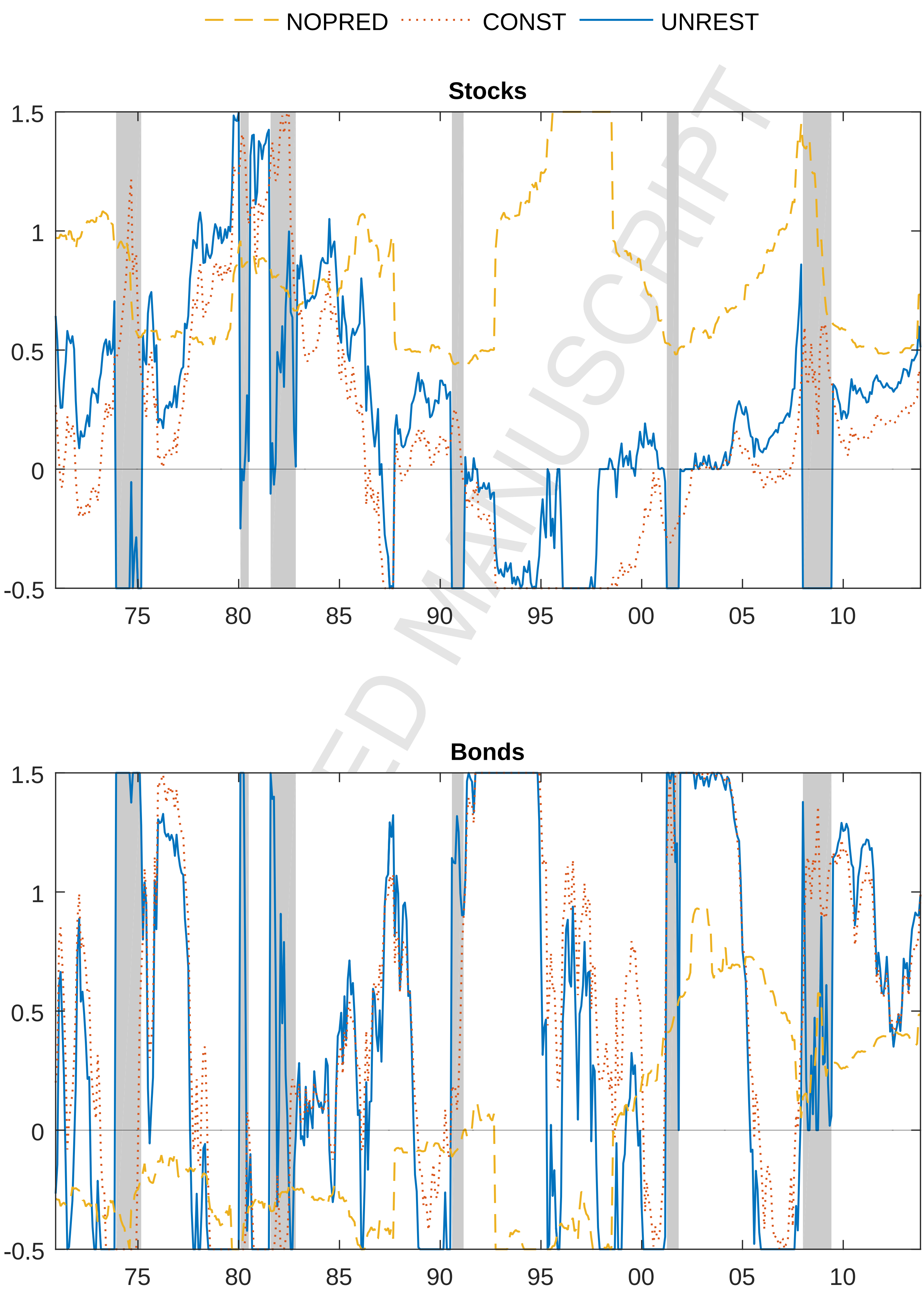

NBER
ADS
PMI
Cumulative △CER (%): REST
Cumulative Return (%): REST
Cumulative △CER (%): UNREST
Cumulative Return (%): UNREST
75 80 85 90 95 00 05 10

Figure5


Bivariate
100
80
60
40
20
0
30
35
40
45
50
55
60
65
70
75
80
85
90
95
00
05
10
Diffusion Index
100
80
60
40
20
0
30
35
40
45
50
55
60
65
70
75
80
85
90
95
00
05
10

FigureA1


NOPRED
CONST
UNREST
Stocks
1.5
1
0.5
0
-0.5
75
80
85
90
95
00
05
10
Bonds
1.5
1
0.5
0
-0.5
75
80
85
90
95
00
05
10