

Research Article

War Discourse and Financial Markets: Some Concerns

Avanidhar Subrahmanyam* 

Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, United States

Abstract

One of the key issues in finance is whether sentiment is related to stock prices. Sentiment can influence prices via its effect on investors' moods. A recent innovation in this space has been to perform text analysis on newspaper articles. A recent technique that is applied for this purpose is supervised Latent Dirichlet Allocation (sLDA), which allows articles to be scored based on topics such as War, Conflict, Financial Crisis, and so on. This work clarifies some issues surrounding two research papers published and forthcoming on the subject of whether war discourse in the New York Times predicts returns. One paper indicates that war discourse positively predicts next-month ahead market returns. The rationale is that the sentiment surrounding war discourse suppresses current prices and thus increases future expected returns. The second paper argues that loadings (betas) on the war factor command negative premia in the cross-section of stock returns. The rationale is that people require lower returns on stocks that form a hedge against the war factor. In this paper, I take a deeper look at the papers' findings. I argue that authors are making judgment calls that need to be discussed and disclosed, and that the results are not robust.

Keywords

Factors, Risk Premia, Stock Returns

1. Introduction

This paper is a comment on two papers: Hirshleifer, Mai, and Pukthuanthong [5] and Hirshleifer, Mai, and Pukthuanthong [6] (henceforth the "first" and "second" papers, respectively). In the first paper, the authors use a text analysis method called supervised Latent Dirichlet Allocation (sLDA) to isolate topics related to War and related "cluster" words (such as "conflict" and "terrorism") from articles in the New York Times. They then create a monthly War factor culled from words in the Times. They show that such a factor positively predicts one-month-ahead market (Standard and Poor's 500) returns.

In the second paper, the authors construct a monthly War factor to specifically focus on War (not related cluster words). The factor is constructed as AR(1) innovations of the sLDA

vector. They show that loadings on the War factor command a negative premium in the cross-section of several sorted portfolios.

The application of sLDA to financial markets is creative, and I have no issues with the incremental contribution of that aspect of the paper. Also, to their credit the authors post the R code to replicate their paper, and post the War factors. However, a practitioner or academic might be interested in knowing how robust are the results on such a key topic of return predictability, and my note addresses this issue.

Specifically, I demonstrate using their posted factors that the results are not robust to reasonable perturbations. This is not a note about replication. Given their posted data, the code appears to produce the results exactly. Instead, this is about

*Corresponding author: subra@anderson.ucla.edu (Avanidhar Subrahmanyam)

Received: 21 September 2025; **Accepted:** 16 October 2025; **Published:** 27 October 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

robustness.

2. The First Paper

The obvious issue in whether War discourse predicts next month's market returns is a comprehensive control for macroeconomic variables. We note from their Table 5, column (6) that the War variable's t -statistic is down to 1.98 (p -value 0.048) with very few macro variables: the dividend and earnings to price ratios, and market variance. 1.98 is quite borderline for 150 plus years (1800 monthly observations).

The argument for including so few macro variables is that these are the only ones available throughout the sample period from the 1870s to 2010s. Although this is true, inflation is available from the 1920s to the present - a significant part of their sample period. Further, the authors themselves present results elsewhere for the 2000-2016 period (see their Table 6), and the options volatility index VIX is available throughout that sample.

Inflation and VIX are such key controls (inextricably linked with War) that it is hardly necessary to justify their inclusion. Accordingly, the attached Table 1 includes Inflation (INFL) and VIX in their Table 5, column (6), for the periods for which these variables are available. I present Newey and West [9] (NW)-corrected t -statistics with lags of 1 and 12 (the paper uses a lag of 1). The reader can easily discern that their variable loses significance in either case (and its coefficient drops by at least 50% in magnitude).

On a parenthetical note, their primary market risk control in their Table 5, column (6) is SV AR (market variance). The market variance is very susceptible to outliers, and many of us might advocate for market standard deviation instead.

3. The Second Paper

For the second paper, I begin by stating a standard asset pricing equation. Let the subscript i denote a stock, j denote one of L factors, and t denote time. Let R_f denote the risk-free rate, R the return, β the loading on a factor, and λ the risk-premium on the factor. The conditional form of an asset pricing equation of the APT/ICAPM type, which holds period-by-period, is

$$E(R_{it}) = R_f t + \sum_{j=1}^L \beta_{ijk} \lambda_{jk} t - 1 \quad \forall k, t.$$

This is an equation rooted in economics, and says that the risk premium demanded by agents to hold a stock is higher if its factor betas (i.e., covariance risks) are higher. Obviously, the agent has to know the betas to decide what premium to charge when pricing assets; hence the $t-1$ subscript on the betas and the t subscript on the expected return. If the betas and premia are not time-varying, then the unconditional form of this equation can also hold. However, betas on portfolios sorted by characteristics such as book/market likely do

time-vary (see, e.g., Engle [1], and McLean and Pontiff [8] show that the premia on such portfolios also vary materially over time.

Their Table 2 of the second paper, which documents that loadings on War are priced for various off-the-shelf portfolios, uses full sample betas, as per the legend and code. That is, both betas and premia are estimated over the full sample once, starting from 1972. Further, the premia are estimated not via Fama and MacBeth [4] (FM), but via a regression of average portfolio returns on full sample betas, and Shanken [10]-corrected t -statistics are presented. With time-varying betas, the former aspect means that their Table 2 has significant lookahead problems. For example, in effect, the premium of the Fama and French (1993) HML in early months is allowed to depend on the covariance risk of HML in later months within the regressions they run [2]. Thus, this is not a predictive regression. In addition, their full sample beta uses information that the agent does not have at the time prices are set in all but the last month of the sample.

Next, their Table 5 with the Fama and French [3] industry portfolios from 1926 onwards switches method. It uses rolling betas over the past sixty months, as indicated in the code, thus conforming to the conditional asset pricing equation, and uses Fama-MacBeth to estimate risk premia. However, the premium on War does not clear the 5% hurdle in the first regression ($t=-1.89$) of their Table 5, with rolling betas, and only the War factor included. The method switch from their Tables 2 to 5 is not explained or motivated, as far as I can see.

Table 2 of the present paper uses their own Table 5 method (Fama-MacBeth with rolling 60-month betas) on the first column of their Table 2 (which uses the Hou, Xue, and Zhang [7] portfolios). I present the Shanken [10]-corrected FM t -statistics, though simple ones leave the result virtually unchanged. As can be seen, the beta on War completely loses significance relative to the full sample method. Further, the absolute magnitude of the premium drops from 1.33 to 0.10. It could be argued that betas are more precisely estimated in the full sample, justifying their method, but this is unlikely to be the crux of the issue for well-diversified portfolios. Again, the method I use in my Table 2 is one they themselves use later in the paper, and accords with the conditional asset pricing equation.

A simple way to ascertain if look-forward issues affect their paper's Table 2 result is to cross-sectionally regress, in turn, average returns from the first 50% of the sample period on betas from the second 50% of their sample period (with betas estimated once per stock over the entire half-sample), and vice versa. If betas are stable, and look-ahead is not an issue, the results should be comparable.

Yet, regressing the second half mean return on the first-half betas yields a coefficient of just -0.26% ($t=-0.72$), whereas regressing the first-half mean return on the second-half betas yields a coefficient of as high as -1.26% ($t=-2.44$). Since agents do *not* know the second-half beta when they set the premium (average return) in the first half, the second coeffi-

cient does not accord with conditional asset pricing. Instead, this is strongly suggestive of look-ahead issues contaminating the inference that War betas command premia in asset returns. I will leave the state of the research there.

There are a couple of other parenthetical notes on the second paper.

First, in their Tables 1 and 5, they apply Newey and West [9] with a lag of 6. The lag is not motivated or mentioned in the paper.

Second, since War has a systematic component with the market, there is no apparent reason that their Table 1, which documents the betas on the War factor, should be restricted to non-traded factors. It is of critical practical importance

whether War is supplanted by the market. The answer to the “whether” is not known because the betas in their Table 1 do not control for the market.

4. Conclusion

Basic robustness checks, and full description of and motivations for such checks, are desirable in research. Issues such as choice of asset pricing methods, lags used in applications of Newey-West, regression controls, should all be clearly described in published papers.

Table 1. Predicting S&P 500 returns one month ahead with the dividend-price ratio (DP), earnings-price ratio (EP), market return variance (SVAR), the Treasury Bill rate (TBL), inflation (INFL), and the option-based volatility index (VIX). NW stands for the Newey-West [9] error correction.

	NW lags = 1			NW lags = 12		
War	2.53 (1.99)	1.25 (0.77)	0.69 (0.22)	2.53 (1.70)	1.25 (0.72)	0.69 (0.19)
DP	-0.87 (-0.33)	-0.28 (-0.07)	6.00 (1.79)	-0.87 (-0.30)	-0.28 (-0.06)	6.00 (2.09)
EP	3.88 (1.44)	4.54 (1.21)	5.86 (1.47)	3.88 (1.18)	4.54 (0.96)	5.86 (1.29)
SVAR	-0.28 (-0.07)	-1.10 (-0.22)	-20.51 (-4.33)	-0.28 (-0.07)	-1.10 (-0.22)	-20.51 (-6.80)
TBL	-4.39 (-2.87)	-4.22 (-2.40)	-3.40 (-1.07)	-4.39 (-3.03)	-4.22 (-2.40)	-3.40 (-1.02)
INFL		-3.70 (-1.89)	-1.48 (-0.57)		-3.70 (-1.85)	-1.48 (-0.67)
VIX			19.03 (3.74)			19.03 (4.36)
R2	0.80	0.68	5.95	0.80	0.68	5.95
Beg	187101	191302	199001	187101	191302	199001
End	201909	201909	201909	201909	201909	201909

Table 2. Cross-sectional return prediction using betas (loadings) on the War factor, 1972-2019.

Regression	Y	X	Coefficient	tstat
1	Full-sample mean return	Full-sample War beta	-1.33	(-2.89)
2	FM next month return	Rolling 60-month lagged War beta	-0.10	(-0.95)

Abbreviations

APT	Arbitrage Pricing Theory
CAPM	Capital Asset Pricing Model
NW	Newey-West
FM	Fama-MacBeth
HML	High Minus Low Book/Market Factor
DP	Dividend-price Ratio
EP	Earnings-price Ratio
SVAR	Market Return Variance
TBL	Treasury Bill Rate
INFL	Inflation
VIX	Option-based Volatility Index

Author Contributions

Avanidhar Subrahmanyam is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Engle, Robert F., 2016, Dynamic conditional beta, *Journal of Financial Econometrics* 14, 643-667. <https://doi.org/10.1093/jjfinec/nbw006>
- [2] Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- [3] Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153-193. [https://doi.org/10.1016/S0304-405X\(96\)00896-3](https://doi.org/10.1016/S0304-405X(96)00896-3)
- [4] Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636. <https://doi.org/10.1086/260061>
- [5] Hirshleifer, David, Dat Mai, and Kuntara Pukthuanthong, 2025a, War discourse and disaster premium: 160 years of evidence from the stock market, *Review of Financial Studies* 38, 457-506. <https://doi.org/10.1093/rfs/hhae081>
- [6] Hirshleifer, David, Dat Mai, and Kuntara Pukthuanthong, 2025b, War discourse and the cross section of expected stock returns, *Journal of Finance*, forthcoming. <https://doi.org/10.3386/w31348>
- [7] Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019-2133. <https://doi.org/10.1093/rfs/hhy131>
- [8] McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5-32. <https://doi.org/10.1111/jofi.12365>
- [9] Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708. <https://doi.org/10.2307/1913610>
- [10] Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1-33. <https://doi.org/10.1093/rfs/5.1.1>