Dissecting market expectations in the cross-section of book-to-market ratios: A Comment*

Bryan Kelly
Yale University and
AQR Capital Management

Seth Pruitt
Arizona State University

June 16, 2021

Abstract

Souza (2021) replicates the findings of Kelly and Pruitt (2013), then critiques their findings with alternative empirical choices. We challenge this critique, and argue that the choices of Kelly and Pruitt (2013, 2015) are the natural economic and statistical choices. The attenuation of predictability in Souza's (2021) empirical analysis is due primarily to a number of ill-advised implementation choices. We conclude that the results of Kelly and Pruitt (2013) are notably robust in the sample following publication of their paper.

Overview

Throughout we refer to Kelly and Pruitt (2013) as "KP." We also refer to a companion paper, Kelly and Pruitt (2015) or "KP15", which analyzes the econometric properties of the PLS method and is referenced in KP. We refer to Souza (2021) as "Souza21."

Souza21 first shows that the findings of KP are exactly replicable using the data and methodology described in their paper. Souza21 then claims that

> "the evidence of market premium predictability, in particular, essentially disappears by making any one of the following changes: (i) Updating the sample to June 1926 – December 2019; (ii) not taking logs of the book-to-markets used as regressors; (iii) not dividing book-to-markets by their time-series standard deviations; or (iv) not taking one extra book-to-market lag (for monthly forecasts)."

We disagree with each of Souza21's claims, and discuss the flaws in each of these criticisms in turn.

*We thank Thiago de Oliveira Souza for providing code and data.

Table 1: KP vs Souza21 Replication

	KP Code	Souza21 Code		
	KP Data	KP Data	Souza21 Data_{2010}	Souza21 Data_{2019}
		Panel A: Annual R^2		
6 Portfolios				
In-sample	7.718	7.712	8.527	7.298
OOS	5.811	5.791	6.672	4.356
25 Portfolios				
In-sample	13.502	13.506	12.430	9.758
OOS	3.493	3.510	8.241	1.669
100 Portfolios				
In-sample	18.054	18.047	19.925	17.356
OOS	13.068	13.030	14.332	9.664
		Panel B: Monthly R^2		
6 Portfolios				
In-sample	0.559	0.598	0.574	0.479
OOS	0.654	0.651	0.654	0.491
25 Portfolios				
In-sample	1.115	1.112	0.875	0.654
OOS	0.769	0.759	0.649	0.223
100 Portfolios				
In-sample	2.380	2.337	2.544	2.297
OOS	0.925	0.924	0.611	0.483

Change (i): Extended Sample

Souza21's first critique is that extending KP's sample (which ends in 2010) to include post-publication data through 2019 weakens the evidence of market return predictability.

First, Souza21's Tables 1 and 2 show that in-sample analyses for the 2010 vs. 2019 samples are nearly identical. This is a valuable and affirmative robustness check of KP that Souza21 fails to mention.

Second, Souza21 concludes from its analysis that there is a decline in the out-of-sample R^2 with the inclusion of post-publication data. However, Souza21's own analysis shows that this decline is *not* due to the inclusion of post-publication data, but is instead due to the *changes in variable construction* used by Souza21.

We demonstrate this point in Table 1, which replicates the findings of KP on the original sample ending in 2010 and the Souza21 sample ending in 2019. It does this using two sets of

code. One code is from KP and has been available on the authors' websites since publication. The other code is from Souza21 and provided to us by the author.

The second and third columns of Table 1 show that in the sample ending in 2010, Souza21's results match those of KP, despite the different methodological choices made by Souza21. This is another indication of the robustness of KP's original findings.

The fourth column of Table 1 uses the extended 2019 sample period, uses the code of Souza21 with the exact data construction and model specification choices of KP. We find that the out-of-sample R^2's are all positive and are generally consistent with KP findings for the 2010 data. In fact, this analysis closely matches the top three rows of Table 2 in Souza21, which show that the out-of-sample predictive power of PLS remains statistically significant in the post-publication sample! This tracks down the source of attenuation in predictive power from Souza21—it comes entirely through its changes in variable construction. In the subsequent sections, we discuss the flaws in Souza21's variable construction.

In summary, Souza21 continues to find significant out-of-sample return predictability in the full 1926-2019 sample, yet again demonstrating the robustness of KP's original findings. Taken at face value, Souza21 promotes a misleading conclusion from its own empirical findings. Souza21's abstract states "the evidence of market premium predictability, in particular, essentially disappears by making any one of the following changes: (i) Updating the sample to June 1926–December 2019; ...." This is a direct contradiction of Souza21's own Table 2. When updating the sample to 1926–2019 and making no changes to variable construction, there is *no* disappearance in return predictability. This is confirmed by Souza21's own out-of-sample significance tests.

In addition, KP reports out-of-sample R^2 for a wide range of sample splits in order to demonstrate the robustness of their results to sample split. See KP Figures 1, 2, 3, 4, and 5 for sample split robustness in a variety of analyses. It is also worth noting that with a 1980 sample split, the out-of-sample test period is 33% longer than the test period beginning in 1990, which obviously leads to more precise inference about the out-of-sample reliability of KP's findings.

Change (ii): Not Taking Logs of B/M

The derivation of the KP model uses the log-linearization of book-to-market ratios and price-dividend ratios proposed by Vuolteenaho (2002) and Campbell and Shiller (1988). In light of this, we are puzzled by Souza21's criticism of our choice to use logs of valuation ratios as predictors—log ratios are precisely the form of predictor variables dictated by our model. Souza21 favors a model with levels of valuation ratios, an odd choice that lacks the theoretical motivation of KP's log specification.

A related criticism levied by Souza21, though not emphasized in the abstract, regards our "choice of forecasting the return on the market, as opposed to the market premium." First, Souza21's nomenclature is unusual: what is meant by "market premium" is in fact excess return; i.e., the realized return on the market in excess of the risk-free rate (as opposed to some notion of expected excess return, the object for which the term "premium" is typically

reserved). The present value system we analyze is rooted in the Campbell and Shiller (1988) identity, which is an identity involving the total return. Furthermore, the total return is the forecast target in a long literature studying present-value relations that serves as the basis of comparison and discussion in KP, including Campbell and Shiller (1988), Cochrane (1992, 2008), Lettau and Van Nieuwerburgh (2008), van Binsbergen and Koijen (2010), among many others. This total return objective is clearly stated as the first line of KP's abstract and repeated throughout the paper.

Change (iii): Not Variance-standardizing Predictors

In this criticism, Souza21 takes issue with our choice to variance-standardize each of the predictor variables that go into the PLS algorithm. It states that "From a strict asset pricing perspective, this adjustment implies that the empirical results of KP rely on a possible theoretical linear relation between the latent factor and ratios between BMs and standard deviations, not exactly BMs."

The grounds for Souza21's objection are not at all clear. Regardless, this criticism is flawed on basic statistical grounds. It fails to understand the distinction between a model and an estimator. Equation 7 in KP is a model. PLS is an estimator. PLS delivers a consistent estimate of the expected return regardless of the predictor scaling; this point is carefully analyzed in KP15. Rescaling puts evidence from all predictors on equal footing and improves the statistical efficiency of the estimator. More plainly, Souza21's claim is analogous to criticizing a researcher for using GLS rather than OLS to estimate a linear model.

Furthermore, the standard and widely accepted approach for applying PLS is to work with variance-standardized predictors. Our paper explicitly states that we follow the methodology laid out in KP15 (KP reference the 2012 version, which was eventually published in 2015), stating that we use a "matrix of predictors that have been standardized to have unit time series variance" and discusses the sensibility of this choice. In particular, it is well known that PLS lacks invariance to volatility of the predictors, prompting our recommendation to variance standardize.

But don't take our word for it. The reigning textbook authority Hastie et al. (2001) states "like principal component regression, partial least squares (PLS) is not scale invariant, so we assume that each x_i [predictor] is standardized to have mean 0 and variance 1." Just like in PCA (e.g. Stock and Watson, 2002), a variable with an unusually large variance will dominate PLS's dimension reduction and generally lead to unstable forecasts.[1]

This instability is exactly what Souza accomplishes by its avoidance of variance standardization. Ill-advised methodological choices can kill any positive empirical result. As shown in Souza21, by far the biggest driver of its poor PLS results is the choice to *not* variance-standardize predictors.

[1]Note that when KP consider pc1 or pc123 forecasts in Table II, the PCs were taken from the same standardized predictors – because it is the standard procedure given PCA's lack of scale invariance.

Change (iv): Timing of Predictors

The timing of book-to-market predictors in Souza21 differs from that used by KP. KP constructed book-to-market ratios from Ken French's website in 2011. These book-to-market ratios divide a portfolio's previous fiscal year book equity by its market equity starting in July.[2] We settled on this timing through a series of exchanges with Ken French regarding the details of his data construction and based on feedback from referees during the review process, particularly in response to a referee request that we make absolutely sure the book equity data would have been known to the market. Souza21 describes this timing as "taking one extra lag," but this is not how we, Ken French, nor our referee conceived the data construction.

In any case, our data has no look-ahead bias and are, if anything, overly conservative in their information content. And while the divergence in Souza21's predictor timing contributes some discrepancy in its results versus KP, we do not consider it a primary driver of the differences.

Overlooked Evidence in Favor of KP's Conclusions

The first set of analyses reported by KP use 1) book-to-market ratios of 2) Fama-French size and value portfolios to forecast 3) the aggregate US equity market return. This empirical setting represents the full extent of Souza21's re-evaluation of KP's findings.

But this is only one empirical application of the much more general valuation ratio framework and PLS estimation procedure laid out by KP. The evidence supporting KP's conclusions is far more expansive than the evidence re-evaluated by Souza21. In particular, KP study

- A range of alternative predictor variables beyond book-to-market ratios, including
  - moving average price ratios (that are not subject to balance sheet data limitations)
  - price-dividend ratios
- A range of other predictor cross sections beyond Fama-French portfolios, including
  - valuation ratios for individual stocks
  - valuation ratios for country-level equity portfolios
- A range of other forecast targets beyond the aggregate US equity returns, including
  - returns on a wide variety of characteristic-sorted portfolios
  - returns on the global ex-US equity portfolio

[2]More recently, Ken French's data format has changed such that portfolios' book equity is no longer separately reported.

Collectively, these analyses amount to a preponderance of evidence that the framework and methodology outlined by KP robustly predicts future asset returns. In addition, our analyses above indicate that even the main results in KP's Table 1 are impressively robust in the extended sample through 2019. Returning to Souza21's abstract statement that "I find no evidence that the procedure generates a valid forecasting model of market premiums with persistently positive out-of-sample R^2 in the full 1926-2019 sample," we conclude that Souza21 misrepresents the totality of the empirical evidence, and draws flawed conclusions regarding return predictability using valuation ratio predictors and PLS estimation.

References

Campbell, J. Y., Shiller, R., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Cochrane, J. H., 1992. Explaining the variance of price-dividend ratios. Review of Financial Studies 5, 243–280.

Cochrane, J. H., 2008. The dog that did not bark: A defense of return predictability. Review of Financial Studies 21, 1533.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer, New York.

Kelly, B. T., Pruitt, S., 2013. Market expectations in the cross section of present values. Journal of Finance 68, 1721–1756.

Kelly, B. T., Pruitt, S., 2015. The three-pass regression filter: A new approach to forecasting with many predictors. Journal of Econometrics 186, 294–316.

Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. Review of Financial Studies 21, 1607–1652.

Souza, T. d. O., 2021. Dissecting market expectations in the cross-section of book-to-market ratios. Critical Finance Review .

Stock, J. H., Watson, M. W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.

van Binsbergen, J., Koijen, R., 2010. Predictive regressions: A present-value approach. Journal of Finance 65.

Vuolteenaho, T., 2002. What drives firm-level stock returns? Journal of Finance 57, 233–264.