Journal of Empirical Finance 31 (2015) 36–53

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Market proxies as factors in linear asset pricing models: Still living with the roll critique☆

Todd Prono *

Department of Finance and Real Estate, Kogod School of Business at American University, 4400 Massachusetts Ave, N.W., Washington, DC 20016, United States

ARTICLE INFO

Article history:
Received 12 November 2011
Received in revised form 29 August 2013
Accepted 2 February 2015
Available online 7 February 2015

JEL classification:
G11
G12

Keywords:
Asset pricing
CAPM
Conditional CAPM
ICAPM
Measurement error
Momentum
Time-series testing
Value

ABSTRACT

A new model misspecification measure for linear asset pricing models is proposed for the case where misspecification maps to latency of one of the pricing factors; in this case, the market return. This measure is suited both for testing models that include the market return as a pricing factor in a traditional sense (i.e., whether the chosen model does or does not price a collection of risky assets) and ranking those models (i.e., determining which model performs best). The proposed measure is used in pricing portfolios reflecting the size, value, and momentum premia. The conditional CAPM of Jagannathan and Wang (1996) is found to best the performance of both the simple CAPM and the ICAPM of Petkova (2006). Moreover, it is discovered that winner stocks in a momentum portfolio may have higher market betas than loser stocks.

1. Introduction

The recent literature on empirical asset pricing has focused some attention on the effects of model misspecification in the testing (and ranking) of various pricing models (see, as examples, Kan and Robotti, 2008, 2009; Kan et al., 2013). Since any model is but an abstraction of reality, it seems likely that, at least, some (and, perhaps, most) of these models are misspecified; for example, either in terms of the factors that they select for pricing risky assets or how (some of) those selected factors are measured. Standard practice is to ask whether a given model either prices a collection of test assets or explains the cross-sectional variation in the expected returns of those assets.[1] While the answer to this question is, certainly, of interest, it is only partially satisfying in light of the (likely) potential for model misspecification. In addition, one might like to have a means for comparing the performance of different models; for example,

☆ I owe thanks to my thesis advisors Christopher Baum, Arthur Lewbel, and Zhijie Xiao, to Celso Brunetti, Ethan Cohen-Cole, Alex Horenstein, James Moser, David Reiffen, Michel Robe, Theo Vermaelen, an associate editor, and two anonymous referees, as well as to seminar participants at Acadian Asset Management, the Federal Reserve Bank of Boston, the Boston College Econometrics Seminar, the Commodity Futures Trading Commission, and the 2010 meetings of the Eastern, Midwest, and Southwestern Finance Associations for helpful comments and discussions. Nicholas Kraninger provided excellent research assistance on an earlier version of this paper. All errors are my own.

* Tel.: +1 202 885 2745.
E-mail address: tprono@american.edu.

[1] See Kan et al. (2013) for a distinction between pricing returns in a stochastic discount factor sense and explaining the variation in expected returns, either in a cross-sectional or time-series context.

http://dx.doi.org/10.1016/j.jempfin.2015.02.001

as in the case when none of the aforementioned models completely explains the observed behavior of the test assets (i.e., all of the models are 'rejected' by the data). Doing so, of course, requires a scalar measure of model misspecification.

In terms of explaining prices via a stochastic discount factor, the Hansen and Jagannathan (1997), hereafter HJ, distance measure has become quite popular. Kan and Robotti (2008, 2009) investigate how model misspecification affects this measure. In a cross-sectional context, Kan et al. (2013) discuss how misspecification affects the two-pass regression methodology. In this paper, I examine misspecification effects in the context of time-series tests of linear asset pricing models (see as a seminal example of time-series testing, Gibbons et al., 1989). Specifically, latency of the true market return as argued by Roll (1977) is examined, where an all-stock index serves as a proxy for this unobservable variable.[2] The scalar measure of model misspecification in this case, rather than being a quadratic-form of model pricing errors as in HJ or the cross-sectional R^2 as in Kan et al. (2013), is the maximum (or highest) correlation between the true market return and an observable proxy return that supports the given pricing model's ability to explain expected returns. Clearly, this misspecification measure is limited in its application to pricing models that include the market return as a factor. Fortunately, however, the set of models to which this measure applies includes such popular alternatives as the simple CAPM, the conditional CAPM, the ICAPM, and the Fama and French (1993), hereafter FF, three-factor model.[3] Moreover, this misspecification measure has intuitive appeal in that one can expect (see Roll, 1977) that an all-stock proxy should have a high correlation with the true market return. A low value of the upper bound to this correlation, therefore, reflects a correspondingly low likelihood (in an intuitive sense) that the model to which this value attaches serves as a descent approximation to reality.

In developing this misspecification measure, I begin with the works of Shanken (1987) and Kandel and Stambaugh (1987, 1995). Specifically, I generalize the findings of Shanken to allow for (1) multiple sources of priced risk and (2) a complete treatment for the effects of misspecifying the market return factor. Regarding (2), I demonstrate that not only should the latency of the market return be reflected through an imperfect correlation between it and its selected proxy, but also through what the use of that proxy return means in the estimation of the familiar market model in the case of the simple CAPM, and (more generally) the linear multi-factor model generating factor betas. The key realization here is that the proxy return is endogenous to these linear models by nature of the measurement error associated with it. As a consequence, these linear models cannot be treated as projection equations and estimated by OLS without loss of generality. In particular, I find that ignoring this measurement error in the evaluation of the given pricing model biases the associated test results towards rejection. Moreover, this measurement error also impacts the relative ranking of competing models that are based (in whole, or in part) on the true market return.

The misspecification measure I propose is used to test the simple CAPM, the conditional CAPM of Jagannathan and Wang (1996), and the ICAPM of Petkova (2006) on three portfolios representing the size, value, and momentum premium, respectively. This measure is based on a set of second-order covariance estimators developed in Prono (2014) that are capable of controlling for latency in the market return factor. Previewing the results of these tests, while all three models are rejected by the data, the conditional CAPM is found to perform the best, offering an improvement over the simple CAPM. No such improvement is found for the ICAPM. Moreover, illustrating the importance of misspecification effects in the market return factor, all three models are found to perform meaningfully better in tests that control for these effects relative to ones that do not. Helping to explain this improved performance is the finding that winning stocks (those that account for the long positions in a momentum portfolio) may actually have higher market betas than losing stocks (those that account for the short positions in a momentum portfolio). Use of imperfect proxies for the true market return may, therefore, help explain (a portion of) the empirically observed momentum premium.

2. Misspecification

Assume there exists an observable risk-free rate. Let r_t be an N-vector of observable risky asset returns in excess of that risk-free rate. Define $r_{m,t}$ as the unobservable excess return on a portfolio of all risky assets, including those in r_t. Let $r_{p,t}$ be the excess return on an observable proxy for $r_{m,t}$, and assume that the $N+1$ components of r_t and $r_{p,t}$ are linearly independent. In addition, let the K-vector X_t contain additional, observable factors that price risky asset returns. Define $E(\cdot)$, $\sigma^2(\cdot)$, and $cov(\cdot)$ as the expectation, variance, and covariance operators, respectively. Consider the following linear model for r_t and $r_{p,t}$

$$r_t = \alpha + \beta_X X_t + \beta_m r_{m,t} + U_t, \tag{1}$$

$$r_{p,t} = \alpha_p + X_t' \beta_{X,p} + \beta_m r_{m,t} + U_{p,t}, \tag{2}$$

where U_t and $U_{p,t}$ are both mean zero innovations, β_X is $(N \times K)$ and $\beta_{X,p}$ is $(K \times 1)$.

Assumption A1. (i) $X_t, r_{m,t} \perp U_t, U_{p,t}$; (ii) $X_t \perp r_{m,t}$; (iii) α and α_p are both equal to zero.

Given A1, the model of Eqs. (1) and (2) is consistent with the APT of Ross (1976). It is also consistent with the intertemporal CAPM of Merton (1973), as well as the simple CAPM of Sharpe (1964) and Lintner (1965).[4] If $r_{m,t} \equiv r_{c,t}$, where $r_{c,t}$ is the excess return on a

[2] This practice is common in both academic work as well as industry practice, with the latter sourcing to the need of locating tradable instruments for hedging purposes.

[3] Note that works such as Petkova (2006) attribute the empirical success of the FF three-factor model to an ICAPM explanation.

[4] In the simple CAPM case, β_X and $\beta_{X,p}$ are both equal to zero. In the ICAPM case, the vector X_t contains elements that summarize the investment opportunity set.