Using Stocks or Portfolios in Tests of Factor Models*

Andrew Ang†
Columbia University and Blackrock and NBER

Jun Liu‡
UCSD

Krista Schwarz§
University of Pennsylvania

This Version: October 20, 2016

JEL Classification: G12
Keywords: Specifying Base Assets, Cross-Sectional Regression, Estimating Risk Premia, APT, Efficiency Loss

*We thank Rob Grauer, Cam Harvey, Bob Hodrick, Raymond Kan, Bob Kimmel, Georgios Skoulakis, Yuhang Xing, and Xiaoyan Zhang for helpful discussions and seminar participants at the American Finance Association, Columbia University, CRSP forum, Texas A&M University, and the Western Finance Association for comments. The usual disclaimer applies.

†Columbia Business School, 3022 Broadway 413 Uris, New York, NY 10027, ph: (212) 854-9154; email: aa610@columbia.edu; http://www.columbia.edu/∼aa610.

‡Rady School of Management, Otterson Hall, 4S148, 9500 Gilman Dr, #0553, La Jolla, CA 92093-0553; ph: (858) 534-2022; email: junliu@ucsd.edu; http://rady.ucsd.edu/faculty/directory/liu/.

§The Wharton School, University of Pennsylvania, 3620 Locust Walk, SH-DH 2300, Philadelphia, PA 19104; email: kschwarz@wharton.upenn.edu; https://fnce.wharton.upenn.edu/profile/979/.

Using Stocks or Portfolios in Tests of Factor Models

Abstract

We examine the efficiency of using individual stocks or portfolios as base assets to test asset pricing models using cross-sectional data. The literature has argued that creating portfolios reduces idiosyncratic volatility and allows more precise estimates of factor loadings, and consequently risk premia. We show analytically and empirically that smaller standard errors of portfolio beta estimates do not lead to smaller standard errors of cross-sectional coefficient estimates. Factor risk premia standard errors are determined by the cross-sectional distributions of factor loadings and residual risk. Portfolios destroy information by shrinking the dispersion of betas, leading to larger standard errors.

1 Introduction

Asset pricing models should hold for all assets, whether these assets are individual stocks or whether the assets are portfolios. The literature has taken two different approaches in specifying the universe of base assets in cross-sectional factor tests. First, researchers have followed Black, Jensen and Scholes (1972) and Fama and MacBeth (1973), among many others, to group stocks into portfolios and then run cross-sectional regressions using portfolios as base assets. An alternative approach is to estimate cross-sectional risk premia using the entire universe of stocks following Litzenberger and Ramaswamy (1979) and others. Perhaps due to the easy availability of portfolios constructed by Fama and French (1993) and others, the first method of using portfolios as test assets is the more popular approach in recent empirical work.

Blume (1970, p156) gave the original motivation for creating test portfolios of assets as a way to reduce the errors-in-variables problem of estimated betas as regressors:

> *...If an investor's assessments of α_i and β_i were unbiased and the errors in these assessments were independent among the different assets, his uncertainty attached to his assessments of $\bar{\alpha}$ and $\bar{\beta}$, merely weighted averages of the α_i's and β_i's, would tend to become smaller, the larger the number of assets in the portfolios and the smaller the proportion in each asset. Intuitively, the errors in the assessments of α_i and β_i would tend to offset each other. ... Thus, ...the empirical sections will only examine portfolios of twenty or more assets with an equal proportion invested in each.*

If the errors in the estimated betas are imperfectly correlated across assets, then the estimation errors would tend to offset each other when the assets are grouped into portfolios. Creating portfolios allows for more efficient estimates of factor loadings. Blume argues that since betas are placed on the right-hand side in cross-sectional regressions, the more precise estimates of factor loadings for portfolios enable factor risk premia to also be estimated more precisely. This intuition for using portfolios as base assets in cross-sectional tests is echoed by other papers in the early literature, including Black, Jensen and Scholes (1973) and Fama and MacBeth (1973). The majority of modern asset pricing papers testing expected return relations in the cross section now use portfolios.[1]

In this paper we study the relative efficiency of using individual stocks or portfolios in tests of cross-sectional factor models. We focus on theoretical results in a one-factor setting, but also consider multifactor models. We illustrate the intuition with analytical forms using maximum likelihood, but the intuition from these formulae are applicable to more general situations.[2] Maximum likelihood estimators achieve the Cramér-Rao lower bound and provide an optimal benchmark to measure efficiency. The Cramér-Rao lower bound can be computed with any set of consistent estimators.

Forming portfolios dramatically reduces the standard errors of factor loadings due to decreasing idiosyncratic risk. But, we show the more precise estimates of *factor loadings* do not lead to more efficient estimates of *factor risk premia*. In a setting where all stocks have the same idiosyncratic risk, the idiosyncratic variances of portfolios decline linearly with the number of stocks in each portfolio. The fewer portfolios used, the smaller the standard errors of the portfolio factor loadings. But, fewer portfolios also means that there is less cross-sectional variation in factor loadings to form factor risk premia estimates. Thus, the standard errors of the risk premia estimates *increase* when portfolios are used compared to the case when all stocks are used. The same result holds in richer settings where idiosyncratic volatilities differ across stocks, idiosyncratic shocks are cross-sectionally correlated, and there is stochastic entry and exit of firms in unbalanced panels. Creating portfolios to reduce estimation error in the factor loadings does not lead to smaller estimation errors of the factor risk premia.

The reason that creating portfolios leads to larger standard errors of cross-sectional risk premia estimates is that creating portfolios destroys information. A major determinant of the

[1] Fama and French (1992) use individual stocks but assign the stock beta to be a portfolio beta, claiming to be able to use the more efficient portfolio betas but simultaneously using all stocks. We show below that this procedure is equivalent to directly using portfolios.

[2] Jobson and Korkie (1982), Huberman and Kandel (1987), MacKinlay (1987), Zhou (1991), Velu and Zhou (1999), among others, derive small-sample or exact finite sample distributions of various maximum likelihood statistics but do not consider efficiency using different test assets.

standard errors of estimated risk premia is the cross-sectional distribution of risk factor loadings scaled by the inverse of idiosyncratic variance. Intuitively, the more disperse the cross section of betas, the more information the cross section contains to estimate risk premia. More weight is given to stocks with lower idiosyncratic volatility as these observations are less noisy. Aggregating stocks into portfolios shrinks the cross-sectional dispersion of betas. This causes estimates of factor risk premia to be less efficient when portfolios are created. We compute efficiency losses under several different assumptions, including cross-correlated idiosyncratic risk and betas, and the entry and exit of firms. The efficiency losses are large.

Finally, we empirically verify that using portfolios leads to wider standard error bounds in estimates of one-factor and three-factor models using the CRSP database of stock returns. We find that for both a one-factor market model and the Fama and French (1993) multifactor model estimated using the full universe of stocks, factor risk premia are highly significant. In contrast, using portfolios often produces insignificant estimates of factor risk premia in both the one-factor and three-factor specifications.

We stress that our results do not mean that portfolios should never be used to test factor models. In particular, many non-linear procedures can only be estimated using a small number of test assets. However, when firm-level regressions specify factor loadings as right-hand side variables, which are estimated in first stage regressions, creating portfolios for use in a second stage cross-sectional regression leads to less efficient estimates of risk premia. Second, our analysis is from an econometric, rather than from an investments, perspective. Finding investable strategies entails the construction of optimal portfolios. Finally, our setting also considers only efficiency and we do not examine power. A large literature discusses how to test for factors in the presence of spurious sources of risk (see, for example, Kan and Zhang, 1999; Kan and Robotti, 2006; Hou and Kimmel, 2006; Burnside, 2007) holding the number of test assets fixed. From our results, efficiency under a correct null will increase in all these settings when individual stocks are used. Other authors like Zhou (1991) and Shanken and Zhou (2007) examine the small-sample performance of various estimation approaches under both the null and alternative.[3] These studies do not discuss the relative efficiency of the test assets employed in cross-sectional factor model tests.

Our paper is related to Kan (2004), who compares the explanatory power of asset pricing models using stocks or portfolios. He defines explanatory power to be the squared cross-sectional correlation coefficient between the expected return and its counterpart specified by the model. Kan finds that the explanatory power can increase or decrease with the number of portfolios. From the viewpoint of Kan's definition of explanatory power, it is not obvious that asset pricing tests should favor using individual stocks. Unlike Kan, we consider the criterion of statistical efficiency in a standard cross-sectional linear regression setup. In contrast, Kan's explanatory power is not directly applicable to standard econometric settings. We also show that using portfolios versus individual stocks matters in actual data.

Two other related papers which examine the effect of different portfolio groupings in testing asset pricing models are Berk (2000) and Grauer and Janmaat (2004). Berk addresses the issue of grouping stocks on a characteristic known to be correlated with expected returns and then testing an asset pricing model on the stocks within each group. Rather than considering just a subset of stocks or portfolios within a group as Berk examines, we compute efficiency losses with portfolios of different groupings using all stocks, which is the usual case done in practice. Grauer and Janmaat do not consider efficiency, but show that portfolio grouping under the alternative when a factor model is false may cause the model to appear correct.

The rest of this paper is organized as follows. Section 2 presents the econometric theory and derives standard errors concentrating on the one-factor model. We describe the data and compute efficiency losses using portfolios as opposed to individual stocks in Section 3. Section 4 compares the performance of portfolios versus stocks in the CRSP database. Finally, Section 5 concludes.

---

[3] Other authors have presented alternative estimation approaches to maximum likelihood or the two-pass methodology such as Brennan, Chordia and Subrahmanyam (1998), who run cross-sectional regressions on all stocks using risk-adjusted returns as dependent variables, rather than excess returns, with the risk adjustments involving estimated factor loadings and traded risk factors. This approach cannot be used to estimate factor risk premia.

2 Econometric Setup

2.1 The Model and Hypothesis Tests

We work with the following one-factor model (and consider multifactor generalizations later):

$$R_{it} = \alpha + \beta_i \lambda + \beta_i F_t + \sigma_i \varepsilon_{it}, \tag{1}$$

where R_{it}, $i = 1, ..., N$ and $t = 1, ..., T$, is the excess (over the risk-free rate) return of stock i at time t, and F_t is the factor which has zero mean and variance σ_F^2. We specify the shocks ε_{it} to be IID $N(0,1)$ over time t but allow cross-sectional correlation across stocks i and j. We concentrate on the one-factor case as the intuition is easiest to see and present results for multiple factors in the Appendix. In the one-factor model, the risk premium of asset i is a linear function of stock i's beta:

$$\mathrm{E}(R_{it}) = \alpha + \beta_i \lambda. \tag{2}$$

This is the beta representation estimated by Black, Jensen and Scholes (1972) and Fama and MacBeth (1973). In vector notation we can write equation (1) as

$$R_t = \alpha 1 + \beta \lambda + \beta F_t + \Omega_\varepsilon^{1/2} \varepsilon_t, \tag{3}$$

where R_t is a $N \times 1$ vector of stock returns, α is a scalar, 1 is a $N \times 1$ vector of ones, $\beta = (\beta_1 \ \ldots \ \beta_N)'$ is an $N \times 1$ vector of betas, Ω_ε is an $N \times N$ invertible covariance matrix, and ε_t is an $N \times 1$ vector of idiosyncratic shocks where $\varepsilon_t \sim N(0, I_N)$.[4]

Asset pricing theories impose various restrictions on α and λ in equations (1)-(3). Under the Ross (1976) Arbitrage Pricing Theory (APT),

$$H_0^{\alpha=0}: \quad \alpha = 0. \tag{4}$$

[4] The majority of cross-sectional studies do not employ adjustments for cross-sectional correlation, such as Fama and French (2008). We account for cross-sectional correlation in our empirical work in Section 4.

This hypothesis implies that the zero-beta expected return should equal the risk-free rate. A rejection of $H_0^{\alpha=0}$ means that the factor cannot explain the average level of stock returns. This is often the case for factors based on consumption-based asset pricing models because of the Mehra-Prescott (1985) equity premium puzzle, where a very high implied risk aversion is necessary to match the overall equity premium.

However, even though a factor cannot price the overall market, it could still explain the relative prices of assets if it carries a non-zero price of risk. We say the factor F_t is priced with a risk premium if we can reject the hypothesis:

$$H_0^{\lambda=0}: \quad \lambda = 0. \tag{5}$$

A simultaneous rejection of both $H_0^{\alpha=0}$ and $H_0^{\lambda=0}$ economically implies that we cannot fully explain the overall level of returns (the rejection of $H_0^{\alpha=0}$), but exposure to F_t accounts for some of the expected returns of assets relative to each other (the rejection of $H_0^{\lambda=0}$). By far the majority of studies investigating determinants of the cross section of stock returns try to reject $H_0^{\lambda=0}$ by finding factors where differences in factor exposures lead to large cross-sectional differences in stock returns. Recent examples of such factors include aggregate volatility risk (Ang et al., 2006), liquidity (Pástor and Stambaugh, 2003), labor income (Santos and Veronesi, 2006), aggregate investment, and innovations in other state variables based on consumption dynamics (Lettau and Ludvigson, 2001b), among many others. All these authors reject the null $H_0^{\lambda=0}$, but do not test whether the set of factors is complete by testing $H_0^{\alpha=0}$.

In specific economic models such as the CAPM or if a factor is tradeable, then defining $\tilde{F}_t = F_t + \mu$, where $\tilde{F}_t$ is the non-zero mean factor with $\mu = \mathrm{E}(\tilde{F}_t)$, we can further test if

$$H_0^{\lambda=\mu}: \quad \lambda - \mu = 0. \tag{6}$$

This test is not usually done in the cross-sectional literature but can be done if the set of test assets includes the factor itself or a portfolio with a unit beta (see Lewellen, Nagel and Shanken, 2010). We show below, and provide details in the Appendix, that an efficient test for $H_0^{\lambda=\mu}$ is equivalent to the test for $H_0^{\lambda=0}$ and does not require the separate estimation of μ. If a factor is priced (so we reject $H_0^{\lambda=0}$) and in addition we reject $H_0^{\lambda=\mu}$, then we conclude that although the factor helps to determine expected stock returns in the cross section, the asset pricing theory requiring $\lambda = \mu$ is rejected. In this case, holding the traded factor F_t does not result in a long-run expected return of λ. Put another way, the estimated cross-sectional risk premium, λ, on a traded factor is not the same as the mean return, μ, on the factor portfolio.

We derive the statistical properties of the estimators of α, λ, and β_i in equations (1)-(2). We present results for maximum likelihood and consider a general setup with GMM, which nests the two-pass procedures developed by Fama and MacBeth (1973), in the Appendix. The maximum likelihood estimators are consistent, asymptotically efficient, and analytically tractable. We derive in closed-form the Cramér-Rao lower bound, which achieves the lowest standard errors of all consistent estimators. This is a natural benchmark to measure efficiency losses. An important part of our results is that we are able to derive explicit analytical formulas for the standard errors. Thus, we are able to trace where the losses in efficiency arise from using portfolios versus individual stocks. In sections 3 and 4, we take this intuition to the data and show empirically that in actual stock returns efficiency losses are greater with portfolios.

2.2 Likelihood Function

The constrained log-likelihood of equation (3) is given by

$$L = -\sum_t (R_t - \alpha - \beta(F_t + \lambda))'\Omega_\varepsilon^{-1}(R_t - \alpha - \beta(F_t + \lambda)) \quad (7)$$

ignoring the constant and the determinant of the covariance terms. For notational simplicity, we assume that σ_F and Ω_ε are known.[5] We are especially interested in the cross-sectional parameters $(\alpha\,\lambda)$, which can only be identified using the cross section of stock returns. The factor loadings, β, must be estimated and not taking the estimation error into account results in incorrect standard errors of the estimates of α and λ. Thus, our parameters of interest are $\Theta = (\alpha\,\lambda\,\beta)$. This setting permits tests of $H_0^{\alpha=0}$ and $H_0^{\lambda=0}$. In the Appendix, we state the maximum likelihood estimators, $\hat{\Theta}$, and discuss a test for $H_0^{\lambda=\mu}$.

2.3 Standard Errors

The standard errors of the maximum likelihood estimators $\hat{\alpha}$, $\hat{\lambda}$, and $\hat{\beta}$ are:

$$\begin{aligned}
\mathrm{var}(\hat{\alpha}) &= \frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\beta'\Omega_\varepsilon^{-1}\beta}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2} && (8)\\
\mathrm{var}(\hat{\lambda}) &= \frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{1'\Omega_\varepsilon^{-1}1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2} && (9)\\
\mathrm{var}(\hat{\beta}) &= \frac{1}{T}\frac{1}{\lambda^2+\sigma_F^2}\\
&\times \left[\Omega+\frac{\lambda^2}{\sigma_F^2}\frac{(\beta'\Omega_\varepsilon^{-1}\beta)11'-(1'\Omega_\varepsilon^{-1}\beta)\beta1'-(1'\Omega_\varepsilon^{-1}\beta)1\beta'+(1'\Omega_\varepsilon^{-1}1)\beta\beta'}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2}\right]. && (10)
\end{aligned}$$

We provide a full derivation in Appendix A.

To obtain some intuition, consider the case where idiosyncratic risk is uncorrelated across stocks so Ω_ε is diagonal with elements $\{\sigma_i^2\}$. We define the following cross-sectional sample moments, which we denote with a subscript c to emphasize they are cross-sectional moments

[5] Consistent estimators are given by the sample formulas

$$\begin{aligned}
\hat{\sigma}_F^2 &= \frac{1}{T}\sum_t F_t^2\\
\hat{\Omega}_\varepsilon &= \frac{1}{T}\sum_t (R_t-\hat{\alpha}-\hat{\beta}(F_t+\hat{\lambda}))(R_t-\hat{\alpha}-\hat{\beta}(F_t+\hat{\lambda}))'.
\end{aligned}$$

As argued by Merton (1980), variances are estimated very precisely at high frequencies and are estimated with more precision than means.

and the summations are across N stocks:

$$\begin{aligned}
\mathrm{E}_c(\beta/\sigma^2) &= \frac{1}{N}\sum_j \frac{\beta_j}{\sigma_j^2} \\
\mathrm{E}_c(\beta^2/\sigma^2) &= \frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2} \\
\mathrm{E}_c(1/\sigma^2) &= \frac{1}{N}\sum_j \frac{1}{\sigma_j^2} \\
\mathrm{var}_c(\beta/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^4}\right) - \left(\frac{1}{N}\sum_j \frac{\beta_j}{\sigma_j^2}\right)^2 \\
\mathrm{cov}_c(\beta^2/\sigma^2, 1/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^4}\right) - \left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2}\right)\left(\frac{1}{N}\sum_j \frac{1}{\sigma_j^2}\right). \qquad (11)
\end{aligned}$$

In the case of uncorrelated idiosyncratic risk across stocks, the standard errors of $\hat{\alpha}$, $\hat{\lambda}$, and $\hat{\beta}_i$ in equations (8)-(10) simplify to

$$\begin{aligned}
\mathrm{var}(\hat{\alpha}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\mathrm{E}_c(\beta^2/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)} \qquad (12) \\
\mathrm{var}(\hat{\lambda}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\mathrm{E}_c(1/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)} \qquad (13) \\
\mathrm{var}(\hat{\beta}_i) &= \frac{1}{T}\frac{\sigma_i^2}{(\sigma_F^2+\lambda^2)}\left(1+\frac{\lambda^2}{N\sigma_i^2\sigma_F^2}\frac{\mathrm{E}_c(\beta^2/\sigma^2)-2\beta_i\mathrm{E}_c(\beta/\sigma^2)+\beta_i^2\mathrm{E}_c(1/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)}\right). \qquad (14)
\end{aligned}$$

Comment 2.1 *The standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ depend on the cross-sectional distributions of betas and idiosyncratic volatility.*

In equations (12) and (13), the cross-sectional distribution of betas scaled by idiosyncratic variance determines the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$. Some intuition for these results can be gained from considering a panel OLS regression with independent observations exhibiting heteroskedasticity. In this case GLS is optimal, which can be implemented by dividing the regressor and regressand of each observation by residual standard deviation. This leads to the variances of $\hat{\alpha}$ and $\hat{\lambda}$ involving moments of $1/\sigma^2$. Intuitively, scaling by $1/\sigma^2$ places more weight on the asset betas estimated more precisely, corresponding to those stocks with lower idiosyncratic volatilities. Unlike standard GLS, the regressors are estimated and the parameters β_i and λ enter non-linearly in the data generating process (1). Thus, one benefit of using maximum likelihood to compute standard errors to measure efficiency losses of portfolios is that it takes into account the errors-in-variables of the estimated betas.

Comment 2.2 *Cross-sectional and time-series data are useful for estimating α and λ but primarily only time-series data is useful for estimating β_i.*

In equations (12) and (13), the variance of $\hat{\alpha}$ and $\hat{\lambda}$ depend on N and T. Under the IID error assumption, increasing the data by one time period yields another N cross-sectional observations to estimate α and λ. Thus, the standard errors follow the same convergence properties as a pooled regression with IID time-series observations, as noted by Cochrane (2001). In contrast, the variance of $\hat{\beta}_i$ in equation (14) depends primarily on the length of the data sample, T. The stock beta is specific to an individual stock, so the variance of $\hat{\beta}_i$ converges at rate $1/T$ and the convergence of $\hat{\beta}_i$ to its population value is not dependent on the size of the cross section. The standard error of $\hat{\beta}_i$ depends on a stock's idiosyncratic variance, σ_i^2, and intuitively stocks with smaller idiosyncratic variance have smaller standard errors for $\hat{\beta}_i$.

The cross-sectional distribution of betas and idiosyncratic variances enter the variance of $\hat{\beta}_i$, but the effect is second order. Equation (14) has two terms. The first term involves the idiosyncratic variance for a single stock i. The second term involves cross-sectional moments of betas and idiosyncratic volatilities. The second term arises because α and λ are estimated, and the sampling variation of $\hat{\alpha}$ and $\hat{\lambda}$ contributes to the standard error of $\hat{\beta}_i$. Note that the second term is of order $1/N$ and when the cross section is large enough it is approximately zero.[6]

Comment 2.3 *Sampling error of the factor loadings affects the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$.*

[6] The estimators are not N-consistent as emphasized by Jagannathan, Skoulakis and Wang (2002). That is, $\hat{\alpha} \nrightarrow \alpha$ and $\hat{\lambda} \nrightarrow \lambda$ as $N \to \infty$. The maximum likelihood estimators are only T-consistent in line with a standard Weak Law of Large Numbers. With T fixed, $\hat{\lambda}$ is estimated ex post, which Shanken (1992) terms an ex-post price of risk. As $N \to \infty$, $\hat{\lambda}$ converges to the ex-post price of risk. Only as $T \to \infty$ does $\hat{\alpha} \to \alpha$ and $\hat{\lambda} \to \lambda$.

Appendix A shows that the term $(\sigma_F^2 + \lambda^2)/\sigma_F^2$ in equations (12) and (13) arises through the estimation of the betas. This term is emphasized by Gibbons, Ross and Shanken (1989) and Shanken (1992) and takes account of the errors-in-variables of the estimated betas. If $H_0^{\lambda=\mu}$ holds and $\lambda = \mu$, then this term reduces to the squared Sharpe ratio, which is given a geometric interpretation in mean-variance spanning tests by Huberman and Kandel (1987).

2.4 Portfolios and Factor Loadings

From the properties of maximum likelihood, the estimators using all stocks are most efficient with standard errors given by equations (12)-(14). If we use only P portfolios as test assets, what is the efficiency loss? Let the portfolio weights be ϕ_{pi}, where $p = 1, \ldots, P$ and $i = 1, \ldots, N$. The returns for portfolio p are given by:

$$R_{pt} = \alpha + \beta_p \lambda + \beta_p F_t + \sigma_p \varepsilon_{pt}, \tag{15}$$

where we denote the portfolio returns with a superscript p to distinguish them from the underlying securities with subscripts i, $i = 1, \ldots, N$, and

$$\begin{aligned} \beta_p &= \sum_i \phi_{pi} \beta_i \\ \sigma_p &= \left(\sum_i \phi_{pi}^2 \sigma_i^2 \right)^{1/2} \end{aligned} \tag{16}$$

in the case of no cross-sectional correlation in the residuals.

The literature forming portfolios as test assets has predominantly used equal weights with each stock assigned to a single portfolio (see for example, Fama and French, 1993; Jagannathan and Wang, 1996). Typically, each portfolio contains an equal number of stocks. We follow this practice and form P portfolios, each containing N/P stocks with $\phi_{pi} = P/N$ for stock i belonging to portfolio p and zero otherwise. Each stock is assigned to only one portfolio usually based on an estimate of a factor loading or a stock-specific characteristic.

2.5 The Approach of Fama and French (1992)

An approach that uses all individual stocks but computes betas using test portfolios is Fama and French (1992). Their approach seems to have the advantage of more precisely estimated factor loadings, which come from portfolios, with the greater efficiency of using all stocks as observations. Fama and French run cross-sectional regressions using all stocks, but they use portfolios to estimate factor loadings. First, they create P portfolios and estimate betas, $\hat{\beta}_p$, for each portfolio p. Fama and French assign the estimated beta of an individual stock to be the fitted beta of the portfolio to which that stock is assigned. That is,

$$\hat{\beta}_i = \hat{\beta}_p \qquad \forall\, i \in p. \tag{17}$$

The Fama-MacBeth (1973) cross-sectional regression is then run over all stocks $i = 1, \ldots, N$ but using the portfolio betas instead of the individual stock betas. In Appendix D we show that in the context of estimating only factor risk premia, this procedure results in exactly the same risk premium coefficients as running a cross-sectional regression using the portfolios $p = 1, \ldots, P$ as test assets. Thus, estimating a pure factor premium using the approach of Fama and French (1992) on all stocks is no different from estimating a factor model using portfolios as test assets. Consequently, our treatment of portfolios nests the Fama and French (1992) approach.

2.6 Intuition Behind Efficiency Losses Using Portfolios

Since the maximum likelihood estimates achieve the Cramér-Rao lower bound, creating subsets of this information can only do the same at best and usually worse.[7] In this section, we present the intuition for why creating portfolios leads to higher standard errors than using all individual stocks. To illustrate the reasoning most directly, assume that $\sigma_i = \sigma$ is the same across stocks

[7] Berk (2000) also makes the point that the most effective way to maximize the cross-sectional differences in expected returns is to not sort stocks into groups. However, Berk focuses on first forming stocks into groups and then running cross-sectional tests within each group. In this case the cross-sectional variance of expected returns within groups is lower than the cross-sectional variation of expected returns using all stocks. Our results are different because we consider the efficiency losses of using portfolios created from all stocks, rather than just using stocks or portfolios within a group.

and the idiosyncratic shocks are uncorrelated across stocks. In this case the standard errors of $\hat{\alpha}$, $\hat{\lambda}$, and $\hat{\beta}_i$ in equations (8)-(10) simplify to

$$\begin{aligned}
\text{var}(\hat{\alpha}) &= \frac{\sigma^2}{NT}\frac{\sigma_m^2+\lambda^2}{\sigma_m^2}\frac{\mathrm{E}_c(\beta^2)}{\text{var}_c(\beta)} \\
\text{var}(\hat{\lambda}) &= \frac{\sigma^2}{NT}\frac{\sigma_m^2+\lambda^2}{\sigma_m^2}\frac{1}{\text{var}_c(\beta)} \\
\text{var}(\hat{\beta}_i) &= \frac{1}{T}\frac{\sigma^2}{(\sigma_F^2+\lambda^2)}\left(1+\frac{\lambda^2}{N\sigma^2\sigma_F^2}\frac{\mathrm{E}_c(\beta^2)-2\beta_i\mathrm{E}_c(\beta)+\beta_i^2}{\text{var}_c(\beta)}\right).
\end{aligned} \tag{18}$$

Assume that beta is normally distributed. We create portfolios by partitioning the beta space into P sets, each containing an equal proportion of stocks. We assign all portfolios to have $1/P$ of the total mass. Appendix E derives the appropriate moments for equation (18) when using P portfolios. We refer to the variance of $\hat{\alpha}$ and $\hat{\lambda}$ computed using P portfolios as $\text{var}_p(\hat{\alpha})$ and $\text{var}_p(\hat{\lambda})$, respectively, and the variance of the portfolio beta, β_p, as $\text{var}(\hat{\beta}_p)$.

The literature's principle motivation for grouping stocks into portfolios is that "estimates of market betas are more precise for portfolios" (Fama and French, 1993, p. 430). This is true and is due to the diversification of idiosyncratic risk in portfolios. In our setup, equation (14) shows that the variance for $\hat{\beta}_i$ is directly proportional to idiosyncratic variance, ignoring the small second term if the cross section is large. This efficiency gain in estimating the factor loadings is tremendous.

Figure 1 considers a sample size of $T = 60$ with $N = 1000$ stocks under a single factor model where the factor shocks are $F_t \sim N(0, (0.15)^2/12)$ and the factor risk premium $\lambda = 0.06/12$. We graph various percentiles of the true beta distribution with black circles. For individual stocks, the standard error of $\hat{\beta}_i$ is 0.38 assuming that betas are normally distributed with mean 1.1 and standard deviation 0.7 with $\sigma = 0.5/\sqrt{12}$. We graph two-standard error bands of individual stock betas in black through each circle. When we create portfolios, $\text{var}(\hat{\beta}_p)$ shrinks by approximately the number of stocks in each portfolio, which is N/P. The top plot of Figure 1 shows the position of the $P = 25$ portfolio betas, which are plotted with small crosses linked by the red solid line. The two-standard error bands for the portfolio betas go through the red crosses and are much tighter than the two-standard error bands for the portfolios. In the bottom plot, we show $P = 5$ portfolios with even tighter two-standard error bands where the standard error of $\hat{\beta}_p$ is 0.04.

However, this substantial reduction in the standard errors of portfolio betas does not mean that the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ are lower using portfolios. In fact, aggregating information into portfolios increases the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$. Grouping stocks into portfolios has two effects on var$(\hat{\alpha})$ and var$(\hat{\lambda})$. First, the idiosyncratic volatilities of the portfolios change. This does not lead any efficiency gain for estimating the risk premium. Note that the term σ^2/N using all individual stocks in equation (18) remains the same using P portfolios since each portfolio contains equal mass $1/P$ of the stocks:

$$\frac{\sigma_p^2}{P} = \frac{(\sigma^2 P/N)}{P} = \frac{\sigma^2}{N}. \tag{19}$$

Thus, when idiosyncratic risk is constant, forming portfolios shrinks the standard errors of factor loadings, but this has no effect on the efficiency of the risk premium estimate. In fact, the formulas (18) involve the total amount of idiosyncratic volatility diversified by all stocks and forming portfolios does not change the total composition.[8] Equation (19) also shows that it is not simply a denominator effect of using a larger number of assets for individual stocks compared to using portfolios that makes using individual stocks more efficient.

The second effect in forming portfolios is that the cross-sectional variance of the portfolio betas, var$_c(\beta_p)$, changes compared to the cross-sectional variance of the individual stock betas, var$_c(\beta)$. Forming portfolios destroys some of the information in the cross-sectional dispersion of beta making the portfolios less efficient. When idiosyncratic risk is constant across stocks, the only effect that creating portfolios has on var$(\hat{\lambda})$ is to reduce the cross-sectional variance of beta compared to using all stocks, that is var$_c(\beta_p) <$ var$_c(\beta)$. Figure 1 shows this effect. The cross-sectional dispersion of the $P = 25$ betas is similar to, but smaller than, the individual beta

[8] Kandel and Stambaugh (1995) and Grauer and Janmaat (2008) show that repackaging the tests assets by linear transformations of N assets into N portfolios does not change the position of the mean-variance frontier. In our case, we form $P < N$ portfolios, which leads to inefficiency.

dispersion. In the bottom plot, the $P = 5$ portfolio case clearly shows that the cross-sectional variance of betas has shrunk tremendously. It is this shrinking of the cross-sectional dispersion of betas that causes var($\hat{\alpha}$) and var($\hat{\lambda}$) to increase when portfolios are used.

Our analysis so far forms portfolios on factor loadings. Often in practice, and as we investigate in our empirical work, coefficients on firm-level characteristics are estimated as well as coefficients on factor betas.[9] We show in Appendix B that the same results hold for estimating the coefficient on a firm-level characteristic using portfolios versus individual stocks. Grouping stocks into portfolios destroys cross-sectional information and inflates the standard error of the cross-sectional coefficients.

What drives the identification of α and λ is the cross-sectional distribution of betas. Intuitively, if the individual distribution of betas is extremely diverse, there is a lot of information in the betas of individual stocks and aggregating stocks into portfolios causes the information contained in individual stocks to become more opaque. Thus, we expect the efficiency losses of creating portfolios to be largest when the distribution of betas is very disperse.

3 Data and Efficiency Losses

In our empirical work, we use first-pass OLS estimates of betas and estimate risk premia coefficients in a second-pass cross-sectional regression. We work in non-overlapping five-year periods, which is a trade-off between a long enough sample period for estimation but over which an average true (not estimated) stock beta is unlikely to change drastically (see comments by Lewellen and Nagel, 2006; Ang and Chen, 2007). Our first five-year period is from January 1971 to December 1975 and our last five-year period is from January 2011 to December 2015. We consider each stock to be a different draw from equation (1). Our data are sampled monthly and we take all non-financial stocks listed on NYSE, AMEX, and NASDAQ with share type

[9] We do not focus on the question of the most powerful specification test of the factor structure in equation (1) (see, for example, Daniel and Titman, 1997; Jagannathan and Wang, 1998; Lewellen, Nagel and Shanken, 2010) or whether the factor lies on the efficient frontier (see, for example, Roll and Ross, 1994; Kandel and Stambaugh, 1995). Our focus is on testing whether the model intercept term is zero, $H_0^{\alpha=0}$, whether the factor is priced given the model structure, $H_0^{\lambda=0}$, and whether the factor cross-sectional mean is equal to its time-series average, $H_0^{\lambda=\mu}$.

codes of 10 or 11. In order to include a stock in our universe it must have data for at least three of the years in each five-year period, have a price that is above \$0.5 and market capitalization of at least \$0.75 million. Our stock returns are in excess of the Ibbotson one-month T-bill rate. In our empirical work we use regular OLS estimates of betas over each five-year period. Our simulations also follow this research design and specify the sample length to be 60 months.

We estimate a one-factor market model using the CRSP universe of individual stocks or using portfolios. Our empirical strategy mirrors the data generating process (1) and looks at the relation between realized factor loadings and realized average returns. We take the CRSP value-weighted excess market return to be the single factor. We do not claim that the unconditional CAPM is appropriate or truly holds, rather our purpose is to illustrate the differences on parameter estimates and the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ when the entire sample of stocks is used compared to creating test portfolios.

3.1 Distribution of Betas and Idiosyncratic Volatility

Table 1 reports summary statistics of the betas and idiosyncratic volatilities across firms. The full sample contains 30,833 firm observations. As expected, betas are centered approximately at one, but are slightly biased upwards due to smaller firms tending to have higher betas. The cross-sectional beta distribution has a mean of 1.14 and a cross-sectional standard deviation of 0.76. The average annualized idiosyncratic volatility is 0.50 with a cross-sectional standard deviation of 0.31. Average idiosyncratic volatility has generally increased over the sample period from 0.43 over 1971-1975 to 0.65 over 1995-2000, as Campbell et al. (2001) find, but it declines later consistent with Bekaert, Hodrick and Zhang (2010). Stocks with high idiosyncratic volatilities tend to be stocks with high betas, with the correlation between beta and σ equal to 0.26.

In Figure 2, we plot empirical histograms of beta (top panel) and $\ln \sigma$ (bottom panel) over all firm observations. The distribution of beta is positively skewed, with a skewness of 0.70, and fat-tailed with an excess kurtosis of 4.44. This implies there is valuable cross-sectional dispersion information in the tails of betas which forming portfolios may destroy. The distribution of $\ln \sigma$ is fairly normal, with almost zero skew at 0.17 and excess kurtosis of 0.04.

3.2 Efficiency Losses Using Portfolios

We compute efficiency losses using P portfolios compared to individual stocks using the variance ratios

$$\frac{\text{var}_p(\hat{\alpha})}{\text{var}(\hat{\alpha})} \quad \text{and} \quad \frac{\text{var}_p(\hat{\lambda})}{\text{var}(\hat{\lambda})}, \tag{20}$$

where we denote the variances of $\hat{\alpha}$ and $\hat{\lambda}$ computed using portfolios as $\text{var}_p(\hat{\alpha})$ and $\text{var}_p(\hat{\lambda})$, respectively. We compute these variances using Monte Carlo simulations allowing for progressively richer stochastic environments. First, we form portfolios based on true betas, which are allowed to be cross-sectionally correlated with idiosyncratic volatility. Second, we form portfolios based on estimated betas. Third, we specify that firms with high betas tend to have high idiosyncratic volatility, as is observed in data. Finally, we allow entry and exit of firms in the cross section. We show that each of these variations further contributes to efficiency losses when using portfolios compared to individual stocks.

3.2.1 Cross-Sectionally Correlated Betas and Idiosyncratic Volatility

Consider the following one-factor model at the monthly frequency:

$$R_{it} = \beta_i \lambda + \beta_i F_t + \varepsilon_{it}, \tag{21}$$

where $\varepsilon_{it} \sim N(0, \sigma_i^2)$. We specify the factor returns $F_t \sim N(0, (0.15)^2/12)$, $\lambda = 0.06/12$ and specify a joint normal distribution for $(\beta_i, \ln \sigma_i)$ (not annualized):

$$\begin{pmatrix} \beta_i \\ \ln \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 1.14 \\ -2.09 \end{pmatrix}, \begin{pmatrix} 0.58 & 0.13 \\ 0.13 & 0.28 \end{pmatrix} \right), \tag{22}$$

which implies that the cross-sectional correlation between betas and $\ln \sigma_i$ is 0.31. These parameters come from the one-factor betas and residual risk volatilities reported in Table 1. From this generated data, we compute the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ in the estimated process (1), which are given in equations (12) and (13).

We simulate small samples of size $T = 60$ months with $N = 5000$ stocks. We use OLS beta estimates to form portfolios using the ex-post betas estimated over the sample. Note that these portfolios are formed ex post at the end of the period and are not tradable portfolios. In each simulation, we compute the variance ratios in equation (20). We simulate 10,000 small samples and report the mean and standard deviation of variance ratio statistics across the generated small samples. Table 2 reports the results. In all cases the mean and medians are very similar.

Panel A of Table 2 forms P portfolios ranking on true betas and shows that forming as few as $P = 10$ portfolios leads to variances of the estimators about 3 times larger for $\hat{\alpha}$ and $\hat{\lambda}$. Even when 250 portfolios are used, the variance ratios are still around 2.5 for both $\hat{\alpha}$ and $\hat{\lambda}$. The large variance ratios are due to the positive correlation between idiosyncratic volatility and betas in the cross section. Creating portfolios shrinks the absolute value of the $-\text{cov}_c(\beta^2/\sigma^2, 1/\sigma^2)$ term in equations (12) and (13). This causes the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ to significantly increase using portfolios relative to the case of using all stocks. When the correlation of beta and $\ln \sigma$ is set higher than our calibrated value of 0.31, there are further efficiency losses from using portfolios.

Forming portfolios based on true betas yields the lowest efficiency losses; the remaining panels in Table 2 form portfolios based on estimated betas.[10] In Panel B, where we form portfolios on estimated betas with the same data-generating process as Panel A, the efficiency losses increase. For $P = 25$ portfolios the mean variance ratio $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda})$ is 4.9 in Panel B compared to 2.8 in Panel A when portfolios are formed on the true betas. For $P = 250$ portfolios formed on estimated betas, the mean variance ratio for $\hat{\lambda}$ is still 4.2. Thus, the efficiency

[10] We confirm the findings of Shanken and Zhou (2007) that the maximum likelihood estimates are very close to the two-pass cross-sectional estimates and portfolios formed on maximum likelihood estimates give very similar results to portfolios formed on the OLS betas.

losses increase considerably once portfolios are formed on estimated betas. More sophisticated approaches to estimating betas, such as Avramov and Chordia (2006) and Meng, Hu and Bai (2007), do not make the performance of using portfolios any better because these methods can be applied at both the stock and the portfolio level.

3.2.2 Cross-Sectionally Correlated Residuals

We now extend the simulations to account for cross-sectional correlation in the residuals. We extend the data generating process in equation (21) by assuming

$$\varepsilon_{it} = \xi_i u_t + \sigma_{vi} v_{it}, \tag{23}$$

where $u_t \sim N(0, \sigma_u^2)$ is a common, zero-mean, residual factor that is not priced and v_{it} is a stock-specific shock. This formulation introduces cross-sectional correlation across stocks by specifying each stock i to have a loading, ξ_i, on the common residual shock, u_t.

To simulate the model we draw $(\beta_i \, \xi_i \, \ln \sigma_{vi})$ from

$$\begin{pmatrix} \beta_i \\ \xi_i \\ \ln \sigma_{vi} \end{pmatrix} \sim N \left(\begin{pmatrix} 1.14 \\ 1.01 \\ -2.09 \end{pmatrix}, \begin{pmatrix} 0.58 & 0.22 & 0.13 \\ 0.22 & 1.50 & 0.36 \\ 0.13 & 0.36 & 0.28 \end{pmatrix} \right), \tag{24}$$

and set $\sigma_u = 0.09/\sqrt{12}$. In this formulation, stocks with higher betas tend to have residuals that are more correlated with the common shock (the correlation between β and ξ is 0.24) and higher idiosyncratic volatility (the correlation of β with $\ln \sigma_{vi}$ is 0.33).

We report the efficiency loss ratios of $\hat{\alpha}$ and $\hat{\lambda}$ in Panel C of Table 2. The loss ratios are much larger, on average, than Panels A and B and are 30 for $\text{var}_p(\hat{\alpha})/\text{var}(\hat{\alpha})$ and 17 for $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda})$ for $P = 25$ portfolios. Thus, introducing cross-sectional correlation makes the efficiency losses in using portfolios worse compared to the case with no cross-sectional correlation. The intuition is that cross-sectionally correlated residuals induces further noise in the estimated beta loadings. The increased range of estimated betas further reduces the dispersion of true portfolio betas.

3.2.3 Entry and Exit of Individual Firms

One reason that portfolios may be favored is that they permit analysis of a fixed cross section of assets with potentially much longer time series than individual firms. However, this particular argument is specious because assigning a stock to a portfolio must be made on some criteria; ranking on factor loadings requires an initial, "pre-ranking" beta to be estimated on individual stocks. If a firm meets this criteria, then analysis can be done at the individual stock level. Nevertheless, it is still an interesting and valid exercise to compute the efficiency losses using stocks or portfolios with a stochastic number of firms in the cross section.

We consider a log-logistic survivor function for a firm surviving to month T after listing given by

$$P(T > t) = \left[1 + ((0.0323)T)^{1.2658}\right]^{-1}, \tag{25}$$

which is estimated on all CRSP stocks taking into account right-censoring. The implied median firm duration is 31 months. We simulate firms over time and at the end of each $T = 60$ month period, we select stocks with at least $T = 36$ months of history. In order to have a cross section of 5,000 stocks, on average, with at least 36 observations, the average total number of firms is 6,607. We start with 6,607 firms and as firms delist, they are replaced by new firms. Firm returns follow the data-generating process in equation (21) and as a firm is born, its beta, common residual loading, and idiosyncratic volatility are drawn from equation (24).

Panel D of Table 2 reports the results. The efficiency losses are a bit larger than Panel C with a fixed cross section. For example, with 25 portfolios, $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda}) = 19$ compared to 17 for Panel C. Thus, with firm entry and exit, forming portfolios results in greater efficiency losses. Although the number of stocks is, on average, the same as in Panel C, the cross section now contains stocks with fewer than 60 observations (but at least 36). This increases the estimation error of the betas, which accentuates the same effect as Panel B. There is now larger error in assigning stocks with very high betas to portfolios and creating the portfolios masks the true cross-sectional dispersion of the betas. In using individual stocks, the information in the beta cross section is preserved and there is no efficiency loss.

3.2.4 Summary

Potential efficiency losses are large for using portfolios instead of individual stocks. The efficiency losses become larger when residual shocks are cross-sectionally correlated across stocks and when the number of firms in the cross section changes over time.

4 Empirical Analysis

We now investigate the differences in using portfolios versus individual stocks in the data with actual historical stock returns from 1971 to 2015. First, we estimate factor risk premia using all of the stocks in our sample as the test assets. Then, we compare the efficiency of our factor risk premia estimates from using all stocks to estimates from using portfolios as test assets.

We form portfolios based on two types of sorting procedures, ex-post and ex-ante. To create ex-post portfolios, we rank stocks into portfolios based on same-sample factor loadings. To create ex-ante portfolios, we rank stocks into portfolios based on factor loadings formed just prior to rebalancing. Once the stocks are sorted into ex-post and ex-ante portfolios, we compute the same-sample realized betas for each portfolio type. We then relate these realized betas to same-sample returns in order to form factor risk premia estimates for all of our test assets.

In estimating factor risk premia, we find that the efficiency losses predicted by our analytical framework are borne out in the data. When stocks are grouped into portfolios, the estimated factor loadings show less variance, which translates into higher variance of the risk premia estimates. The more cross-sectional dispersion that stocks lose when grouped into portfolios, either due to the sorting method or to the number of portfolios formed, the more extreme the effect.

We compare estimates of a one-factor market model on the CRSP universe in Section 4.1 and the Fama-French (1993) three-factor model in Section 4.2, for all stocks and for the two types of portfolio sorts. We compute standard errors for the factor risk premia estimates using maximum likelihood, which assumes normally distributed residuals, and also using GMM, which is distribution free. The standard errors account for cross-correlated residuals, which are modeled by a common factor and also using industry factors. These models are described in Appendix F. In order to present a concise discussion in this section, we refer to the results for the common factor residual model alone. The results using the industry classification are similar, and we present both models in the tables for completeness and as an additional robustness check. The coefficient estimates are all annualized by multiplying the monthly estimates by 12.

4.1 One-Factor Model

4.1.1 Using All Stocks

The factor model in equation (1) implies a relation between realized firm excess returns and realized firm betas. Thus, we stack all stocks' excess returns from each five-year period into one panel and run a regression using average realized firm excess returns over each five-year period as the regressand, with a constant and the estimated betas for each stock as the regressors. Panel A of Table 3 reports the estimates and standard errors of α and λ in equation (1), using all 30,833 firm observations.

Using all stocks produces risk premia estimates of $\hat{\alpha} = 8.54\%$ and $\hat{\lambda} = 4.79\%$. The GMM standard errors are 1.40 and 1.05, respectively, with t-statistics of 6.1 and 4.6, respectively. The maximum likelihood t-statistics, which assume normally distributed residuals, are larger, at 53.9 and 29.8, respectively. With either specification, the CAPM is firmly rejected since $H_0^{\alpha=0}$ is overwhelmingly rejected. We also clearly reject $H_0^{\lambda=0}$, and so we find that the market factor is priced. The market excess return is $\mu = 6.43\%$, which is close to the cross-sectional estimate $\hat{\lambda} = 4.79\%$, over our 1971-2015 sample period. We formally test $H_0^{\lambda=\mu}$ below.

Using individual stocks as test assets to estimate the relationship between realized returns and realized factor loadings gives t-statistics that are comparable in magnitude to other studies with the same the experimental design like Ang, Chen and Xing (2006). The set-up of many factor model studies in the literature differ in two important ways. First, portfolios are often used as test assets instead of stocks, and second, the portfolios are often sorted on predicted rather than realized betas. In this section, we investigate empirically the potential impact of these two specification differences on the size of the $\hat{\alpha}$ and $\hat{\lambda}$ t-statistics.

Our theoretical results in section 2 show that there could be a large loss of efficiency in the estimation of factor risk premia using portfolios as test assets instead of individual stocks. Thus, our empirical focus is on the increase in standard errors, or the decrease in absolute values of the t-statistics, resulting from the choice of test asset (stocks versus portfolios, and the type and size of portfolio). The various types of standard errors (maximum likelihood versus GMM) also differ, but our focus is on the relative differences for the various test assets within each type of standard error. We now investigate these effects.

4.1.2 "Ex-Post" Portfolios

We first form "ex-post" portfolios. For each five-year period we group stocks into P portfolios, based on realized OLS estimated betas over those five years. Within each portfolio, all stocks are equally weighted at the end of the five-year period. While these portfolios are formed ex-post and are not tradeable, they represent valid test assets to estimate the cross-sectional model (1). Once the portfolios are formed, we regress the average realized portfolio excess returns onto the realized portfolio betas, following the same estimation procedure as in the all stocks case above.

In the last four columns of Table 3, we report statistics of the cross-sectional dispersion of the factor loadings for each of the various test assets. Specifically, we show the mean asset beta value, $E_c(\hat{\beta})$, the cross-sectional standard deviation, $\sigma_c(\hat{\beta})$, and the beta values corresponding to the 5%- and 95%-tiles of the distribution. These statistics allow us to compare the cross-sectional information available to estimate risk premia for all stocks as compared to different sizes of the ex-post and ex-ante portfolios.

For $P = 5$ ex-post portfolios, the cross-sectional standard deviation of beta is $\sigma_c(\hat{\beta}_p)$ is 0.69, whereas over all stocks, the cross-sectional standard deviation of beta $\sigma_c(\hat{\beta})$ is 0.76, showing the shrinkage in the cross-sectional distribution of beta when stocks are grouped into portfolios. The lower variance of factor loading estimates indicates the loss of cross-sectional information in the beta estimates from grouping stocks into portfolios, which ultimately produces larger standard errors in the second-stage risk premia estimation. For $P = 50$, $\sigma_c(\hat{\beta}_p)$ is 0.74, closer to the variance using all stocks. The decline in standard deviation with an increasing number of portfolios is indicative of the ultimate convergence of portfolios to individual stocks. As the number of portfolios increases, it will eventually equal the number of stocks, thus producing identical results.

In Panel B of Table 3, we show the implications for factor risk premia that come from the trade-off in precision versus cross-sectional dispersion of beta estimates discussed above. Panel B shows that the ex-post portfolio $\hat{\alpha}$ and $\hat{\lambda}$ estimates are quite close to those computed using all stocks. However, as implied in our analytical framework, the standard errors using portfolios are larger than those computed using all stocks. As P increases, the standard errors in Panel B of Table 3 decrease (and the t-statistics increase) to approach the values using individual stocks. For example, for $P = 5$ portfolios, the maximum likelihood standard error on $\hat{\lambda}$ is 0.72 compared with 0.16 using all stocks. For $P = 50$ ex-post portfolios, the maximum likelihood standard error for $\hat{\lambda}$ falls to 0.41. The loss of efficiency from forming ex-post portfolios is also apparent for the GMM standard errors, reported in Table 3, but is smaller than for the maximum likelihood standard errors.

The data show empirical support for our analytical model's proposition that the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ depend importantly on the cross-sectional distribution of the factor loading estimates. The distribution is increasingly truncated as a greater number of stocks are grouped into a smaller number of portfolios. The more stocks that are grouped into each portfolio, the smaller the variance in factor loadings, and the larger the standard errors for the risk premia estimates.

4.1.3 "Ex-Ante" Portfolios

Next, we form "ex-ante" tradeable portfolios. We group stocks into portfolios at the beginning of each calendar year, ranking on the market beta estimated over the previous five years. Once the portfolios are formed based on the pre-formation betas, they are held for twelve months to produce ex-ante portfolio returns. We rebalance the ex-ante portfolios annually, weighting stocks equally within each portfolio. Then, we compute the first-pass realized OLS market betas of each portfolio, in each non-overlapping five-year period, in parallel to the procedure used for the ex-post portfolios. These realized portfolio betas are the factor loadings for the ex-ante portfolios. Finally, to estimate the ex-ante portfolio α and λ in Panel C of Table 3 we run a second-pass cross-sectional regression of excess returns onto the ex-ante factor loadings. Thus, we examine the same realized beta–realized return relation as in the case of all stocks and ex-post portfolios, in Panels A and B, respectively, over the same sample.

The difference between the ex-post and ex-ante portfolios is in the sorting procedure used to form the portfolios. This has an important implication for our estimation since portfolios sorted on realized betas achieve the maximum cross-sectional dispersion in realized portfolio betas. However, even sorting on realized betas, as in the ex-post portfolios, leaves open the possibility of error in assigning stocks to a portfolio since the betas are estimated. Sorting on any variable other than the realized beta, such as the pre-formation beta that we use for ex-ante portfolios or a characteristic such as book-to-market or size, used by Fama French (1993) and others, will produce a smaller realized portfolio beta distribution.

The last four columns of Panel C inTable 3, illustrate the reduced dispersion in realized betas using ex-ante portfolios. For $P = 5$, the cross-sectional standard deviation of beta is only $\sigma_c(\hat{\beta}_p) = 0.35$, compared to $\sigma_c(\hat{\beta}) = 0.76$ using all stocks and $\sigma_c(\hat{\beta}) = 0.69$ for $P = 5$ ex-post portfolios. This means that the loss in the cross-sectional beta information available to form risk premia estimates is greater than the loss when forming ex-post portfolios. The severe shrinkage in the beta distribution means that the ex-ante portfolios miss substantial information in the tails; the 5% and 95%-tiles for $P = 5$ ex-ante portfolios are 0.62 and 1.64, respectively, compared to 0.22 and 2.26 for the same number of ex-post portfolios (Panel B), and 0.12 and 2.44 for all stocks (Panel A).

The cross-section of the realized betas relate to returns in the second-stage estimation of risk premia in tests of factor models. The truncated distribution of the ex-ante portfolio factor loadings produces much larger standard errors in the cross-sectional estimation of λ than using either ex-post portfolios or the full stock universe. For all portfolio sizes, the ex-ante portfolio standard errors exceed those of the ex-post standard errors, for both GMM and MLE. For example, the GMM standard error of $\hat{\lambda}$ for $P = 5$ ex-ante portfolios is 2.81, compared to 1.37 for $P = 5$ ex-post portfolios and 1.05 for all stocks. The ex-ante portfolios fail to reject $H_0^{\lambda=0}$, except for MLE standard errors for $P = 50$, in contrast to the overwhelming rejection when using any of the ex-post portfolios or all stocks. This underscores the importance of the information in the realized-beta distribution, which is largely preserved in ex-post portfolios and entirely preserved using all stocks.

Panel C also shows that the estimates of α and λ from ex-ante portfolios are quite dissimilar to the estimates in Panels A and B. Using ex-ante portfolios as test assets produces an estimate of α around 14% and an estimate of λ around 1-2%. In contrast, both all stocks (Panel A) and ex-post portfolios (Panel B) produce alpha estimates around 8% and estimates of λ around 4-5%. This marked difference in $\hat{\alpha}$ and $\hat{\lambda}$ is driven by the dramatic distributional shrinkage in realized betas that stems from forming portfolios on pre-formation betas rather realized betas.

The ex-ante portfolio $\hat{\lambda}$ eventually converges to the estimate using all stocks, but the convergence rate is slow. Figure 3 plots the evolution of $\hat{\lambda}$ as the number of ex-ante portfolios grows larger (and the number of stocks in each portfolio decreases). For $P = 2000$ ex-ante portfolios, each of which contains only one or two stocks, $\hat{\lambda} = 3.79\%$, a full percentage point lower than for using the full distribution of stocks. In the data, stocks have a finite life. When the history of a firm's return is short, there is larger error in assigning the stock to a portfolio, which potentially exacerbates the beta distribution's shrinkage. Indeed, the steep and concave curve shows that the point estimates can be severely altered when many stocks are grouped into each portfolio, markedly shrinking the beta distribution. Even for $P = 100$ ex-ante portfolios, $\hat{\lambda} = 1.91\%$, which is quite different compared to $\hat{\lambda} = 4.79\%$ for the full stock universe.

4.1.4 Tests of Cross-Sectional and Time-Series Estimates

We end our analysis of the one-factor model by testing $H_0^{\lambda=\mu}$, which tests equality of the cross-sectional risk premium and the time-series mean of the market factor portfolio. Table 4 presents the results. Using all stocks, $\hat{\lambda} = 4.79\%$ is fairly close to the time-series estimate, $\hat{\mu} = 6.43\%$, but the small standard errors of maximum likelihood cause $H_0^{\lambda=\mu}$ to be rejected with a t-statistic of 10.16. With GMM standard errors, we fail to reject $H_0^{\lambda=\mu}$ with a t-statistic of 1.56. Similarly, the ex-post portfolios reject $H_0^{\lambda=\mu}$ with maximum-likelihood standard errors at the 5% level (expect for $P = 5$), but fail to reject with GMM standard errors (except for $P = 50$). In contrast, the ex-ante portfolio estimates all reject $H_0^{\lambda=\mu}$, at least at the 10% level, with either maximum likelihood or GMM standard errors.

4.1.5 Summary

We overwhelmingly reject $H_0^{\alpha=0}$ and hence the one-factor model using all stocks or portfolios. For all stocks and for ex-post portfolios, we also reject $H_0^{\lambda=0}$, thus finding the market factor priced. Using all stocks we estimate $\hat{\lambda} = 4.79\%$. Ex-post portfolios preserve most of the cross-sectional spread in betas and produce similar risk premium point estimates to the all stocks case, although with larger standard errors. The fewer the portfolios, the smaller the factor loading dispersion.

Using all stocks or portfolios can produce quite different point estimates of cross-sectional risk premia. In particular, the formation of ex-ante portfolios on past estimated betas severely pares the tails of the realized betas. The $\hat{\lambda}$ produced by between 5 and 50 ex-ante portfolios range from only 1.14% to 1.73%. Further, this loss of information in the cross section of ex-ante portfolio factor loadings leads us to fail to reject $H_0^{\lambda=0}$, for all except the largest number of portfolios, $P = 50$ with maximum likelihood standard errors. For the test of $H_0^{\lambda=\mu}$, all specifications of ex-ante portfolios reject the hypothesis at the 10 percent level, while results are mixed using ex-post portfolios and stocks. The loss of dispersion in portfolio factor loadings can come from the number of portfolios formed or from the portfolio formation method. In either case, individual stocks retain the entire distribution and thus give the most precise risk premia estimates.

4.2 Fama-French (1993) Model

This section estimates the Fama and French (1993) model:

$$R_{it} = \alpha + \beta_{MKT,i}\lambda_{MKT} + \beta_{SMB,i}\lambda_{SMB} + \beta_{HML,i}\lambda_{HML} + \sigma_i\varepsilon_{it}, \tag{26}$$

where MKT is the excess market return, SMB is a size factor, and HML is a value/growth factor. We follow the same estimation procedure as Section 4.1 in that we stack all observations into one panel of non-overlapping five-year periods to estimate the cross-sectional coefficients α, λ_{MKT}, λ_{SMB}, and λ_{HML}.

4.2.1 Factor Loadings

We now compare the Fama French model factor loadings of all stocks to those of ex-post and ex-ante portfolios. We form the portfolios by using the same procedures described in subsections 4.1.2 and 4.1.3, for ex-post and ex-ante portfolios, respectively. We sort stocks into $n \times n \times n$ portfolios sequentially, ranking first on $\hat{\beta}_{MKT}$, then on $\hat{\beta}_{SMB}$, and lastly on $\hat{\beta}_{HML}$, which gives us the same number of stocks in each portfolio.

Table 5 reports summary statistics for the distribution of the factor loadings $\hat{\beta}_{MKT}$, $\hat{\beta}_{SMB}$, and $\hat{\beta}_{HML}$ for all specifications of test assets The mean of each factor loading type is almost the same for all stocks and for portfolios, both ex-post and ex-ante. The market betas are centered around one after controlling for SMB and HML, and the SMB and HML factor loadings are between 0 and 1. SMB and HML are zero-cost portfolios, but the beta estimates are not centered around zero since the break points used by Fama and French (1993) to construct SMB and HML are based on NYSE stocks alone rather than on all stocks. Small stocks tend to skew the SMB and HML loadings to be positive, especially for the SMB loadings which have a mean of 0.94 for all stocks.

The notable difference for portfolios as compared to stocks is in the distribution of the factor loading estimates. Table 5 shows three important effects on the distribution of factor loadings that result from portfolio formation, similar to those found for the one-factor model in Section 4.1.

First, forming portfolios reduces the cross-sectional variance in the factor loadings; the effect is modest for the ex-post portfolios, but it is severe for the ex-ante portfolios. For example, the $\hat{\beta}_{SMB}$ and $\hat{\beta}_{HML}$ cross-sectional standard deviation is 1.21 for all stocks, and it is 0.85 for the $2 \times 2 \times 2$ ex-post portfolios, but it is cut by more than one-half to 0.37 and 0.29, respectively, for the ex-ante $2 \times 2 \times 2$ portfolios. As described for the one-factor model in Section 4.1, the mechanism amplifying the cross-sectional shrinkage in factor loading dispersion is that ex-ante portfolios are formed on pre-formation betas, and thus the realized betas have less dispersion than if the portfolios were formed directly on the realized betas.

Second, forming portfolios truncates the tails of the factor loading distribution. This information loss is most pronounced for the ex-ante portfolios; the 5%-tile to 95%-tile range for $\hat{\beta}_{MKT}$ shifts from -0.01 to 2.24 for all stocks to 0.60 to 1.37 for the $3 \times 3 \times 3$ ex-ante portfolio. Such a difference in the distribution of factor loadings for ex-ante portfolios could produce quite different cross-sectional factor risk premia estimates.

Finally, for ex-post and ex-ante portfolios, the fewer stocks that are grouped into each portfolio, the less shrinkage there is in the dispersion of factor loading estimates and the less tail information that is lost. This follows the intuition that the effect of forming portfolios on risk premia estimation diminishes as portfolios converge to individual stocks (once there are enough portfolios to put each stock into its own portfolio). We now estimate Fama-French (1993) factor risk premia for ex-ante and ex-post portfolios of different sizes.

4.2.2 Cross-Sectional Factor Risk Premia

Table 6 reports estimates of the Fama-French (1993) factor risk premia. Using all stocks in Panel A, we find a positive and significant estimate of the market risk premium, $\hat{\lambda}_{MKT} = 5.05\%$ (very close to the one-factor model estimate in Table 3), a positive and significant size factor premium estimate, $\hat{\lambda}_{SMB} = 6.79\%$, and $\hat{\lambda}_{HML} = 0.01$, not significantly different from 0 at the 5% level. The ex-post portfolios in Panel B also have positive $\hat{\lambda}_{SMB}$ and $\hat{\lambda}_{MKT}$, with similar magnitudes to all stocks, but the ex-post portfolio $\hat{\lambda}_{HML}$ is negative. The ex-ante portfolios in Panel C yield very different estimates of factor risk premia in comparison to all stocks and the ex-post portfolios. Notably, the ex-ante portfolio $\hat{\lambda}_{MKT}$ are negative. Thus, the sign of the $\hat{\lambda}_{MKT}$ and the $\hat{\lambda}_{HML}$ risk premia depend on the particular choice of test asset used in the Fama-French (1993) model.

As in the one-factor model estimation in Section 4.1, the size of the standard errors on the risk premia estimates shrink and the t-statistics increase, both for maximum likelihood and GMM, as the number of test assets grows. This supports the main prediction of our analytical model, that the loss of information from grouping stocks produces less efficient risk premia estimates. It also follows the intuition of the model; efficiency loss in the cross-sectional estimation of factor risk premia is directly related to the drop in cross-sectional dispersion of the factor loadings that comes from grouping individual assets into portfolios. The cross-sectional information loss outweighs the efficiency gain from estimating the factor loadings with portfolios.

4.2.3 Tests of Cross-Sectional and Time-Series Estimates

We report the results of the tests of the null $H_0^{\lambda=\mu}$ for the Fama-French (1993) model in Table 7. For all three types of test assets we firmly reject the hypothesis that the cross-sectional risk premia are equal to the mean factor portfolio returns, for the market risk premium and SMB, using either maximum likelihood or GMM standard errors. For all stocks and the ex-post portfolios we also reject $H_0^{\lambda=\mu}$ for HML. Using ex-ante portfolios, the hypothesis $H_0^{\lambda=\mu}$ for HML is rejected using maximum likelihood standard errors, but not with GMM standard errors. All in all, while the market and size factors are cross-sectionally priced, there is little evidence that the cross-sectional risk premia are consistent with the time-series of factor returns.

4.2.4 Summary

Like the CAPM, the Fama-French (1993) model is strongly rejected in testing $H_0^{\alpha=0}$ using both individual stocks and portfolios. We find that the MKT and SMB Fama-French factors do help in pricing the cross section of stocks with large rejections of $H_0^{\lambda=0}$ for stocks and ex-post portfolios. However, tests of $H_0^{\lambda=\mu}$ reject the hypothesis that the cross-sectional risk premium estimates are equal to the mean factor returns.

Using individual stocks versus portfolios makes a difference in the precision with which factor risk premia are estimated. With individual stocks, the MKT and the HML factor premium are positive, though the latter is not significantly different from zero. In contrast, the sign of the MKT and the HML factor premia flip, depending on whether stocks are sorted into portfolios ex-ante or ex-post. Even though the sorting procedure and thus the risk premia test results differ for these types of portfolios, both portfolio types eventually must converge to the all-stock case as the number of portfolios converges to the number of individual stocks.

5 Conclusion

The finance literature takes two approaches to specifying base assets in tests of cross-sectional factor models. One approach is to aggregate stocks into portfolios. Another approach is to use individual stocks. The motivation for creating portfolios is originally stated by Blume (1970): betas are estimated with error and this estimation error is diversified away by aggregating stocks into portfolios. Numerous authors, including Black, Jensen and Scholes (1972), Fama and MacBeth (1973), and Fama and French (1993), use this motivation to choose portfolios as base assets in factor model tests. The literature suggests that more precise estimates of factor loadings should translate into more precise estimates and lower standard errors of factor risk premia.

We show analytically and confirm empirically that this motivation is wrong. The sampling uncertainty of factor loadings is markedly reduced by grouping stocks into portfolios, but this does not translate into lower standard errors for factor risk premia estimates. An important determinant of the standard error of risk premia is the cross-sectional distribution of risk factor loadings. Intuitively, the more dispersed the cross section of betas, the more information the cross section contains to estimate risk premia. Aggregating stocks into portfolios loses information by reducing the cross-sectional dispersion of the betas. While creating portfolios does reduce the sampling variability of the estimates of factor loadings, the standard errors of factor risk premia actually increase. It is the decreasing dispersion of the cross section of beta when stocks are grouped into portfolios that leads to potentially large efficiency losses in using portfolios versus individual stocks.

In data, the point estimates of the cross-sectional market risk premium using individual stocks are positive and highly significant. This is true in both a one-factor market model specification and the three-factor Fama and French (1993) model. For the one-factor model using all stocks, the cross-sectional market risk premium estimate of 4.79% per annum is close to the time-series average of the market excess return, at 6.43% per annum. In contrast, the market risk premium is insignificant and sometimes has a negative sign when portfolios are constructed on factor loadings that are estimated ex ante. Thus, using stocks or portfolios as base test assets can result in very different conclusions regarding whether a particular factor carries a significant price of risk. Test results from using portfolios converge to those with all stocks as the number of portfolios becomes large enough to equal the number of individual stocks.

The most important message of our results is that using individual stocks permits more efficient tests of whether factors are priced. When just two-pass cross-sectional regression coefficients are estimated there should be no reason to create portfolios and the asset pricing tests should be run on individual stocks instead. Thus, the use of portfolios in cross-sectional regressions should be carefully motivated.

Appendix

A Derivation of Maximum Likelihood Asymptotic Variances

The maximum likelihood estimators for α, λ, and β_i are given by:[11]

$$\hat{\alpha} = \frac{1}{T}\frac{\sum_t 1'\Omega_\varepsilon^{-1}(R_t - \hat{\beta}(F_t + \hat{\lambda}))}{1'\Omega_\varepsilon^{-1}1} \quad \text{(A-1)}$$

$$\hat{\lambda} = \frac{1}{T}\frac{\sum_t \hat{\beta}'\Omega_\varepsilon^{-1}(R_t - \hat{\alpha} - \hat{\beta}F_t)}{\beta'\Omega_\varepsilon^{-1}\beta} \quad \text{(A-2)}$$

$$\hat{\beta}_i = \frac{\sum_t (R_{it} - \hat{\alpha})(\hat{\lambda} + F_t)}{\sum_t (\hat{\lambda} + F_t)^2}. \quad \text{(A-3)}$$

The information matrix is given by

$$\left(\mathrm{E}\left[-\frac{\partial^2 L}{\partial\Theta\partial\Theta'}\right]\right)^{-1} = \frac{1}{T}\begin{pmatrix} 1'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}\beta & 1'\Omega_\varepsilon^{-1}\lambda \\ \beta'\Omega_\varepsilon^{-1}1 & \beta'\Omega_\varepsilon^{-1}\beta & \beta'\Omega_\varepsilon^{-1}\lambda \\ \lambda'\Omega_\varepsilon^{-1}1 & \lambda'\Omega_\varepsilon^{-1}\beta & (\lambda^2 + \sigma_F^2)\Omega_\varepsilon^{-1} \end{pmatrix}^{-1}, \quad \text{(A-4)}$$

where under the null $\frac{1}{T}\sum_t R_t \to \alpha + \beta\lambda$.

To invert this we partition the matrix as:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} Q^{-1} & -Q^{-1}BD^{-1} \\ -D^{-1}CQ^{-1} & D^{-1}(I + CQ^{-1}BD^{-1}) \end{pmatrix},$$

where $Q = A - BD^{-1}C$, and

$$A = \begin{pmatrix} 1'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}\beta \\ \beta'\Omega_\varepsilon^{-1}1 & \beta'\Omega_\varepsilon^{-1}\beta \end{pmatrix}, \quad B = \begin{pmatrix} 1'\Omega_\varepsilon^{-1}\lambda \\ \beta'\Omega_\varepsilon^{-1}\lambda \end{pmatrix}, \quad C = B', \quad D = (\lambda^2 + \sigma_F^2)\Omega_\varepsilon^{-1}.$$

We can write $Q = A - BD^{-1}B'$ as

$$\left(1 - \frac{\lambda^2}{\lambda^2 + \sigma_F^2}\right)\begin{pmatrix} 1'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}\beta \\ \beta'\Omega_\varepsilon^{-1}1 & \beta'\Omega_\varepsilon^{-1}\beta \end{pmatrix}.$$

The inverse of Q is

$$Q^{-1} = \frac{\sigma_F^2 + \lambda^2}{\sigma_F^2}\frac{1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}\begin{pmatrix} \beta'\Omega_\varepsilon^{-1}\beta & -1'\Omega_\varepsilon^{-1}\beta \\ -\beta'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}1 \end{pmatrix}. \quad \text{(A-5)}$$

This gives the variance of $\hat{\alpha}$ and $\hat{\lambda}$ in equations (8) and (9).

To compute the term $D^{-1}(I + CQ^{-1}BD^{-1})$ we evaluate

$$\begin{aligned} D^{-1}B'Q^{-1}BD^{-1} &= \frac{\lambda^2}{\sigma_F^2(\lambda^2 + \sigma_F^2)}\frac{1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2} \\ &\quad \times\Omega_\varepsilon\begin{pmatrix} \beta'\Omega_\varepsilon^{-1}\beta & -1'\Omega_\varepsilon^{-1}\beta \\ -\beta'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}1 \end{pmatrix}\begin{pmatrix} 1'\Omega_\varepsilon^{-1}\lambda \\ \beta'\Omega_\varepsilon^{-1}\lambda \end{pmatrix}\Omega_\varepsilon \\ &= \frac{\lambda^2}{\sigma_F^2(\lambda^2 + \sigma_F^2)}\frac{(\beta'\Omega_\varepsilon^{-1}\beta)11' - (1'\Omega_\varepsilon^{-1}\beta)\beta 1' - (1'\Omega_\varepsilon^{-1}\beta)1\beta' + (1'\Omega_\varepsilon^{-1}1)\beta\beta'}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}. \end{aligned}$$

[11] In our empirical work we use consistent OLS estimates. Any consistent estimator can be used to evaluate the Cramér-Rao lower bound.

Thus,

$$D^{-1} + D^{-1}CQ^{-1}BD^{-1} = \frac{1}{\lambda^2+\sigma_F^2}\left[\Omega + \frac{\lambda^2}{\sigma_F^2}\frac{(\beta'\Omega_\varepsilon^{-1}\beta)11' - (1'\Omega_\varepsilon^{-1}\beta)\beta 1' - (1'\Omega_\varepsilon^{-1}\beta)1\beta' + (1'\Omega_\varepsilon^{-1}1)\beta\beta'}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}\right]. \quad \text{(A-6)}$$

This gives the variance of $\hat{\beta}_i$ in equation (10).

To compute the covariances between $(\hat{\alpha}, \hat{\lambda})$ and $\hat{\beta}_i$, we compute

$$-Q^{-1}BD^{-1} = \frac{\lambda}{\sigma_F^2}\frac{1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}\begin{pmatrix} (1'\Omega_\varepsilon^{-1}\beta)\beta' - (\beta'\Omega_\varepsilon^{-1}\beta)1' \\ (\beta'\Omega_\varepsilon^{-1}1)1' - (1'\Omega_\varepsilon^{-1}1)\beta' \end{pmatrix}. \quad \text{(A-7)}$$

This yields the following asymptotic covariances:

$$\begin{aligned}
\text{cov}(\hat{\alpha}, \hat{\lambda}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{-\text{E}_c(\beta/\sigma^2)}{\text{var}_c(\beta/\sigma^2) - \text{cov}_c(\beta^2/\sigma^2, 1/\sigma^2)} \\
\text{cov}(\hat{\alpha}, \hat{\beta}_i) &= \frac{1}{NT}\frac{\lambda}{\sigma_F^2}\frac{\beta_i\text{E}_c(\beta/\sigma^2) - \text{E}_c(\beta^2/\sigma^2)}{\text{var}_c(\beta/\sigma^2) - \text{cov}_c(\beta^2/\sigma^2, 1/\sigma^2)} \\
\text{cov}(\hat{\lambda}, \hat{\beta}_i) &= \frac{1}{NT}\frac{\lambda}{\sigma_F^2}\frac{\text{E}_c(\beta/\sigma^2) - \beta_i\text{E}_c(1/\sigma^2)}{\text{var}_c(\beta/\sigma^2) - \text{cov}_c(\beta^2/\sigma^2, 1/\sigma^2)}. \quad \text{(A-8)}
\end{aligned}$$

B Factor Risk Premia and Characteristics

Consider the following cross-sectional regression:

$$R_{it} = \alpha + \beta_i\lambda + z_i\gamma + \beta_i F_t + \sigma_i\varepsilon_{it}, \quad \text{(B-1)}$$

where z_i is a firm-specific characteristic, the variance of F_t is σ_F^2, and ε_{it} is IID $N(0,1)$ with ε_{it} uncorrelated across stocks i for simplicity. Assume that α, σ_i, and σ_i are known and the parameters of interest are $\Theta = (\lambda \ \gamma \ \beta_i)$. We assume the intercept term α is known to make the computations easier. The information matrix is given by

$$\left(\text{E}\left[-\frac{\partial^2 L}{\partial\Theta\partial\Theta'}\right]\right)^{-1} = \frac{1}{T}\begin{pmatrix} \sum_i \frac{\beta_i^2}{\sigma_i^2} & \sum_i \frac{\beta_i z_i}{\sigma_i^2} & \frac{\beta_i\lambda}{\sigma_i^2} \\ \sum_i \frac{\beta_i z_i}{\sigma_i^2} & \sum_i \frac{z_i^2}{\sigma_i^2} & \frac{z_i\lambda}{\sigma_i^2} \\ \frac{\beta_i\lambda}{\sigma_i^2} & \frac{z_i\lambda}{\sigma_i^2} & \frac{\lambda^2+\sigma_F^2}{\sigma_i^2} \end{pmatrix}^{-1}. \quad \text{(B-2)}$$

Using methods similar to Appendix A, we can derive var$(\hat{\lambda})$ and var$(\hat{\gamma})$ to be

$$\begin{aligned}
\text{var}(\hat{\lambda}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\text{E}_c(z^2/\sigma^2)}{\text{var}_c(z\beta/\sigma^2) - \text{cov}_c(\beta^2/\sigma^2, z^2/\sigma^2)} \\
\text{var}(\hat{\gamma}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\text{E}_c(\beta^2/\sigma^2)}{\text{var}_c(z\beta/\sigma^2) - \text{cov}_c(\beta^2/\sigma^2, z^2/\sigma^2)}, \quad \text{(B-3)}
\end{aligned}$$

where we define the cross-sectional moments

$$\begin{aligned}
\mathrm{E}_c(z^2/\sigma^2) &= \frac{1}{N}\sum_j \frac{z_j^2}{\sigma_j^2} \\
\mathrm{E}_c(\beta^2/\sigma^2) &= \frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2} \\
\mathrm{var}_c(z\beta/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{z_j^2\beta_j^2}{\sigma_j^4}\right) - \left(\frac{1}{N}\sum_j \frac{z_j\beta_j}{\sigma_j^2}\right)^2 \\
\mathrm{cov}_c(z^2/\sigma^2, \beta^2/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{z_j^2\beta_j^2}{\sigma_j^4}\right) - \left(\frac{1}{N}\sum_j \frac{z_j^2}{\sigma_j^2}\right)\left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2}\right). \qquad \text{(B-4)}
\end{aligned}$$

C Testing Time-Series Means

In this section we derive a test for $H_0^{\lambda=\mu} : \tilde{\lambda} \equiv (\lambda - \mu) = 0$. In Section C.1, we work in the context of the same model of Appendix A using maximum likelihood and show it to have the same standard error as the test for $H_0^{\lambda=0} : \lambda = 0$. In Section C.2, we contrast our test with the approach of Shanken (1992), which involves directly estimating both λ and μ, whereas we only need to directly estimate λ. Our test is consequently much more efficient. Finally, in Section C.3 we couch our new test in GMM and contrast it with the moment conditions for the traditional Shanken (1992) approach. This is also the easiest method computationally for dealing with multiple factors.

C.1 Likelihood Function

Consider the model of $N \times 1$ returns in vector notation

$$R_t = \alpha + \beta\lambda + \beta(\tilde{F}_t - \mu) + \Omega_\varepsilon^{1/2}\varepsilon_t. \qquad \text{(C-1)}$$

The difference with equation (3) in the main text is that now the cross-sectional risk premium, λ, is potentially different from the time-series mean of the factor, μ. The factor shocks $F_t \equiv (\tilde{F}_t - \mu)$ are mean zero.

Let $\tilde{\lambda} = \lambda - \mu$. Then, we can write equation (C-1) as

$$R_t = \alpha + \beta\tilde{\lambda} + \beta\tilde{F}_t + \Omega_\varepsilon^{1/2}\varepsilon_t. \qquad \text{(C-2)}$$

This has exactly the same likelihood as equation (7) except replacing $\tilde{\lambda}$ and $\tilde{F}_t$ for λ and F_t, respectively. Hence, the standard errors for the estimators $\hat{\alpha}$ and $\hat{\tilde{\lambda}}$ are identical to equations (8) and (9), respectively, except we replace λ with $\tilde{\lambda}$ in the latter case. Thus, the test for $H_0^{\lambda=\mu}$ involves standard errors for $\hat{\tilde{\lambda}}$ that are identical to the standard errors for the estimator $\hat{\lambda}$.

The intuition behind this result is that the cross section only identifies the combination $(\lambda - \mu)$. In the case of an APT, the implied econometric assumption is that μ is effectively known as the factor shocks, F_t, are mean zero. The hypothesis $H_0^{\lambda=\mu}$ does not require λ to be separately estimated; only the combination $\tilde{\lambda} - \mu$ needs to be tested. Economically speaking, the cross section is identifying variation of stock returns relative to the base level of the factor – it cannot identify the pure component of the factor itself. If we need to identify the actual level of λ itself together with μ, we could impose that $\lambda = \mu$, which would be the case from the CAPM for a one-factor market model. Another way is to use the time-series mean of a traded set of factors to identify μ. This is the approach of Shanken (1992), to which we now compare our test.

C.2 Shanken (1992)

We work with the following log likelihood (ignoring the constant) of a one-factor model in vector notation for N stocks and the factor F_t:

$$L = -\sum_t (R_t - \alpha - \beta(F_t + \lambda))'\Omega_\varepsilon^{-1}(R_t - \alpha - \beta(F_t + \lambda)) + \sum_t \frac{1}{2\sigma_F^2}(F_t - \mu)^2, \tag{C-3}$$

There are two differences between equation (C-3) and the factor model in equation (7). First, λ and μ are now treated as separate parameters because we have not specified the shocks to be zero mean by construction as in an APT. Second, we identify μ by including F_t as another asset where $\alpha = 0$ and $\beta = 1$, or μ is estimated by the time-series mean of F_t.

In constructing the Hessian matrix for $\theta = (\alpha\,\lambda\,\mu\,\beta)$, it can be shown that the standard errors for $\hat{\alpha}$ and $\hat{\lambda}$ are given by

$$\begin{aligned} \text{var}(\hat{\alpha}) &= \frac{1}{T}\frac{\sigma_F^2 + \lambda^2}{\sigma_F^2}\frac{\beta'\Omega_\varepsilon^{-1}\beta}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2} \\ \text{var}(\hat{\lambda}) &= \frac{\sigma_F^2}{T} + \frac{1}{T}\frac{\sigma_F^2 + \lambda^2}{\sigma_F^2}\frac{1'\Omega_\varepsilon^{-1}1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}. \end{aligned} \tag{C-4}$$

These are the maximum likelihood standard errors derived by Shanken (1992) when including both a cross-sectional risk premium, λ, and a time-series mean of the factors, μ. We observe that var$(\hat{\alpha})$ is identical to equation (8), but var$(\hat{\lambda})$ differs from equation (9) by an additive term, $\frac{1}{T}\sigma_F^2$. The intuition that var$(\hat{\alpha})$ is unaffected by separating λ and σ is that when μ is estimated, a mean-zero change to the residual of one individual stock,

$$R_{it} - \alpha - \beta_i(F_t - \lambda + \mu),$$

changes only the estimate of λ. This result is exactly the same as saying that only the combination $(\lambda - \mu)$ is identified by the cross section of stock returns.

To understand why var$(\hat{\lambda})$ carries an additional term compared to the case where μ is not estimated, note that the maximum likelihood estimator for μ and the standard error for $\hat{\mu}$ are given by:

$$\begin{aligned} \hat{\mu} &= \frac{1}{T}\sum_t F_t \\ \text{var}(\hat{\mu}) &= \frac{\sigma_F^2}{T}. \end{aligned} \tag{C-5}$$

The likelihood function in equation (C-3) has two independent estimates, $\hat{\lambda} - \hat{\mu}$ and $\hat{\mu}$. The independence arises from the independence of ε_{it} and F_t. Thus,

$$\text{var}(\hat{\lambda}) = \text{var}((\hat{\lambda} - \hat{\mu}) + \hat{\mu}) = \text{var}(\hat{\lambda} - \hat{\mu}) + \text{var}(\hat{\mu}).$$

Note that $\hat{\tilde{\lambda}} = \hat{\lambda} - \hat{\mu}$ is exactly the same as the variance when μ is not estimated from Section C.1. This makes clear that the greater efficiency of the test in Section C.1 is that it tests $H_0^{\lambda=\mu} : \lambda - \mu$ without having to directly estimate μ. Testing the hypothesis $H_0^{\lambda=\mu}$ by estimating μ incurs the additional variance of μ, which is a nuisance parameter.

Finally, consider the likelihood function, without the constant, of the system with $\tilde{\lambda}$ augmented with the non-zero mean $\tilde{F}_t$:

$$L = -\sum_t (R_t - \alpha - \beta(\tilde{F}_t + \tilde{\lambda}))'\Omega_\varepsilon^{-1}(R_t - \alpha - \beta(\tilde{F}_t + \tilde{\lambda})) + \sum_t \frac{1}{2\sigma_F^2}(\tilde{F}_t - \mu)^2. \tag{C-6}$$

For the parameter vector $\theta = (\alpha\,\tilde{\lambda}\,\mu\,\beta)$, the information matrix is given by:

$$\left(\mathrm{E}\left[-\frac{\partial^2 L}{\partial\Theta\partial\Theta'}\right]\right)^{-1} = \frac{1}{T}\begin{pmatrix} 1'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}\beta & 0 & 1'\Omega_\varepsilon^{-1}(\tilde{\lambda}+\mu) \\ \beta'\Omega_\varepsilon^{-1}1 & \beta'\Omega_\varepsilon^{-1}\beta & 0 & \beta'\Omega_\varepsilon^{-1}(\tilde{\lambda}+\mu) \\ 0 & 0 & \sigma_F^2 & 0 \\ (\tilde{\lambda}+\mu)'\Omega_\varepsilon^{-1}1 & (\tilde{\lambda}+\mu)'\Omega_\varepsilon^{-1}\beta & 0 & ((\tilde{\lambda}+\mu)^2+\sigma_F^2)\Omega_\varepsilon^{-1} \end{pmatrix}^{-1}. \tag{C-7}$$

This explicitly shows that the estimate $\hat{\mu}$ is uncorrelated with $\hat{\tilde{\lambda}}$ and since $\tilde{\lambda}+\mu=\lambda$, the standard errors for the system with λ and this system with $\tilde{\lambda}$ are identical. Whatever the mean of $\tilde{F}_t$, $\tilde{\lambda}\neq 0$ implies that the factor is priced.

C.3 GMM

We work with the data-generating process for

$$R_t = \alpha + B\tilde{\lambda} + B\tilde{F}_t + \varepsilon_t, \tag{C-8}$$

with the distribution-free assumption that $\mathrm{E}[\varepsilon_t]=0$ for K factors in $\tilde{F}_t$ with mean μ and N stocks in R_t. We write this as

$$\tilde{R}_t \equiv R_t - B\tilde{F}_t = X\gamma + \varepsilon_t, \tag{C-9}$$

for $\gamma = [\alpha\,\tilde{\lambda}]$ which is $K+1$ and $X=[1\ B]$ which is $N\times(K+1)$. We test $H_0^{\lambda=\mu}$ by testing $\tilde{\lambda}=0$.

The Fama-MacBeth (1973) estimator is given by running cross-sectional regressions at time t:

$$\hat{\gamma}_t = (\hat{X}'W\hat{X})^{-1}\hat{X}'W\tilde{R}_t,$$

for weighting matrix W, $\hat{X}=[1\ \hat{B}]$, and then averaging across all $\hat{\gamma}_t$:

$$\hat{\gamma} = \frac{1}{T}\sum\gamma_t = (\hat{X}'W\hat{X})^{-1}\hat{X}'W\bar{\tilde{R}}, \tag{C-10}$$

where $\bar{\tilde{R}} = \frac{1}{T}\sum\tilde{R}_t$. The beta estimates are given by time-series regressions:

$$\hat{B} = \left[\frac{1}{T}\sum(\tilde{R}_t-\bar{\tilde{R}})(\tilde{F}_t-\bar{\tilde{F}})'\right]\hat{\Sigma}_F^{-1}, \tag{C-11}$$

where $\bar{\tilde{F}} \equiv \hat{\mu} = \frac{1}{T}\sum\tilde{F}_t$ and $\hat{\Sigma}_F = \frac{1}{T}\sum(\tilde{F}_t-\bar{\tilde{F}})(\tilde{F}_t-\bar{\tilde{F}})'$.

Assume the moment conditions

$$\begin{aligned} \mathrm{E}[h_{1t}] &= \mathrm{E}[\tilde{R}_t - \mathrm{E}\tilde{R}_t] = 0 \quad (N\times 1) \\ \mathrm{E}[h_{2t}] &= \mathrm{E}\left[[(\tilde{F}_t-\mathrm{E}\tilde{F}_t)'\Sigma_F^{-1}\lambda]\varepsilon_t\right] = 0 \quad (N\times 1), \end{aligned} \tag{C-12}$$

with $h_t = (h_{1t}\,h_{2t})$ satisfying the Central Limit Theorem

$$\frac{1}{\sqrt{T}}\sum h_t \xrightarrow{d} N(0,\Sigma_h),$$

where

$$\Sigma_h = \begin{bmatrix}\Sigma_\varepsilon & 0 \\ 0 & (\lambda'\Sigma_F^{-1}\lambda)\Sigma_\varepsilon\end{bmatrix}.$$

The Fama-MacBeth estimator is consistent, as shown by Cochrane (1991) and Jagannathan, Skoulakis and Wang (2002), among others. To derive the limiting distribution of $\hat{\gamma}$, define $D=(X'WX)^{-1}X'W$ with its sample counterpart $\hat{D}$ and write

$$\begin{aligned}
\hat{\gamma}_t &= \hat{D}\tilde{R}_t \\
&= \hat{D}[\hat{X}\gamma + (B-\hat{B})\lambda + \tilde{R}_t - X\gamma] \\
\hat{\gamma}_t - \gamma &= \hat{D}[(B-\hat{B})\lambda + (\tilde{R}_t - \mathrm{E}\tilde{R}_t)].
\end{aligned}$$

Thus, the asymptotic distribution is given by

$$\begin{aligned}
\sqrt{T}\left(\frac{1}{T}\sum \hat{\gamma}_t - \gamma\right) &= \hat{D}\left[-\frac{1}{\sqrt{T}}\sum \varepsilon_t(\tilde{F}_t - \bar{\tilde{F}})'\hat{\Sigma}_F^{-1}\lambda + \frac{1}{\sqrt{T}}\sum(\tilde{R}_t - \mathrm{E}\tilde{R}_t)\right] \\
&\stackrel{d}{\to} D\begin{bmatrix} I_N & 0 \\ 0 & -I_N \end{bmatrix}\frac{1}{\sqrt{T}}\sum h_t \\
&\stackrel{d}{\to} N(0, \Sigma_\gamma), \qquad \text{(C-13)}
\end{aligned}$$

where

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda)D\Omega_\varepsilon D'. \qquad \text{(C-14)}$$

Note the $\mathrm{E}[h_{2t}]$ set of moment conditions define the factor betas. We refer to the case where $W = I$ as "GMM" standard errors, which are given by

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda)(X'X)^{-1}X'\Omega_\varepsilon X(X'X)^{-1}. \qquad \text{(C-15)}$$

For choice of $W = \Omega_\varepsilon^{-1}$ we have

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda)(X'\Omega_\varepsilon^{-1}X)^{-1}, \qquad \text{(C-16)}$$

which is the same as maximum likelihood. Equation (C-16) is the matrix counterpart of equations (8) and (9) in the main text for a single factor model. We use equation (C-16) to compute standard errors for multiple factors.

It is instructive to note the difference with Shanken (1982). Consider the model

$$R_t = \alpha + B\lambda + B(\tilde{F}_t - \mu) + \varepsilon_t.$$

To derive the Shanken (1982) standard errors for the Fama-MacBeth estimates $\hat{\gamma} = [\hat{\alpha}\,\hat{\lambda}]$, set up the moment conditions

$$\begin{aligned}
\mathrm{E}[h_{1t}^*] &= \mathrm{E}[R_t - \mathrm{E}R_t] = 0 \\
\mathrm{E}[h_{2t}^*] &= \mathrm{E}\left[[(\tilde{F}_t - \mathrm{E}\tilde{F}_t)'\Sigma_F^{-1}\lambda]\varepsilon_t\right] = 0.
\end{aligned}$$

The difference between the Shanken test and our test is that we use the moment conditions $\mathrm{E}[h_{1t}]$ which utilize $\tilde{R}_t$ in equation (C-12) rather than the moment conditions $\mathrm{E}[h_{1t}^*]$. Both cases use the same Fama-MacBeth estimator in equation (C-10). With the following Central Limit Theorem for $h_t = (h_{1t}^*\, h_{2t}^*)$:

$$\frac{1}{\sqrt{T}}\sum h_t^* \stackrel{d}{\to} N(0, \Sigma_h^*),$$

where

$$\Sigma_h^* = \begin{bmatrix} B\Sigma_F B' + \Sigma_\varepsilon & 0 \\ 0 & (\lambda'\Sigma_F^{-1}\lambda)\Sigma_\varepsilon \end{bmatrix},$$

we can derive the Shanken (1982) standard errors (see also Jagannathan, Skoulakis and Wang, 2002). For the case of $K = 1$, the standard errors of $\hat{\gamma}$ reduce to those in equation (C-4).

D The Approach of Fama and French (1992)

In the second-stage of the Fama and MacBeth (1973) procedure, excess returns, R_i, are regressed onto estimated betas, $\hat{\beta}_i$ yielding a factor coefficient of

$$\hat{\lambda} = \frac{\text{cov}(R_i, \hat{\beta}_i)}{\text{var}(\hat{\beta}_i)}.$$

In the approach of Fama and French (1992), P portfolios are first created and then the individual stock betas are assigned to be the portfolio beta to which that stock belongs, as in equation (17). The numerator of the Fama-MacBeth coefficient can be written as:

$$\begin{aligned}
\text{cov}(R_i, \hat{\beta}_i) &= \frac{1}{N}\sum_i (R_i - \bar{R})(\hat{\beta}_i - \bar{\beta}) \\
&= \frac{1}{P}\sum_p \left(\frac{1}{(N/P)} \sum_{i \in p} (R_i - \bar{R}) \right) (\hat{\beta}_p - \bar{\beta}) \\
&= \frac{1}{P}\sum_{p=1}^{P} (\hat{R}_p - \bar{R})(\hat{\beta}_p - \bar{\beta}) \\
&= \text{cov}(\hat{R}_p, \hat{\beta}_p), \qquad \text{(D-1)}
\end{aligned}$$

where the first to the second line follows because of equation (17). The denominator of the estimated risk premium is

$$\begin{aligned}
\text{var}(\hat{\beta}_i) &= \frac{1}{N}\sum_i (\hat{\beta}_i - \bar{\beta})^2 \\
&= \frac{1}{P}\sum_p \frac{1}{(N/P)} \sum_{i \in p} (\hat{\beta}_i - \bar{\beta})^2 \\
&= \frac{1}{P}\sum_{p=1}^{P} (\hat{\beta}_p - \bar{\beta})^2 \\
&= \text{var}(\hat{\beta}_p), \qquad \text{(D-2)}
\end{aligned}$$

where the equality in the third line comes from $\hat{\beta}_p = \hat{\beta}_i$ for all $i \in p$, with N/P stocks in portfolio p having the same value of β_p for their fitted betas. Thus, the Fama and French (1992) procedure will produce the same Fama-MacBeth (1973) coefficient as using only the information from $p = 1, \ldots, P$ portfolios.

E Cross-Sectional Moments For Normally Distributed Betas

We assume that stocks have identical idiosyncratic volatility, σ, and so idiosyncratic volatility does not enter into any cross-sectional moments with beta. If beta is normally distributed with mean μ_β and standard deviation σ_β, the relevant cross-sectional moments are:

$$\begin{aligned}
\text{E}_c(\beta^2) &= \sigma_\beta^2 + \mu_\beta^2 \\
\text{var}_c(\beta^2) &= \sigma_\beta^2. \qquad \text{(E-1)}
\end{aligned}$$

We form P portfolios each containing equal mass of ordered betas. Denoting $N(\cdot)$ as the cumulative distribution function of the standard normal, the critical points δ_p corresponding to the standard normal are

$$N(\delta_p) = \frac{p}{P}, \quad p = 1, ..., P-1, \qquad \text{(E-2)}$$

and we define $\delta_0 = -\infty$ and $\delta_P = +\infty$. The points ζ_p, $p = 1, \ldots, P-1$ that divide the stocks into different portfolios are given by

$$\zeta_p = \mu_\beta + \sigma_\beta \delta_p. \tag{E-3}$$

The beta of portfolio p, β_p, is given by:

$$\beta_p = \frac{\int_{\delta_{p-1}}^{\delta_p} (\mu_\beta + \sigma_\beta \delta) e^{-\frac{\delta^2}{2}} \frac{d\delta}{\sqrt{2\pi}}}{\int_{\delta_{p-1}}^{\delta_p} e^{-\frac{\delta^2}{2}} \frac{d\delta}{\sqrt{2\pi}}} = \mu_\beta + \frac{P\sigma_\beta}{\sqrt{2\pi}} \left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}} \right). \tag{E-4}$$

Therefore, the cross-sectional moments for the P portfolio betas are:

$$\begin{aligned}
\mathrm{E}_c[\beta_p] &= \mu_\beta \\
\mathrm{E}_c[\beta_p^2] &= \frac{1}{P} \sum_{p=1}^{P} \left(\mu_\beta + \frac{P\sigma_\beta}{\sqrt{2\pi}} \left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}} \right) \right)^2 \\
&= \mu_\beta^2 + P \frac{\sigma_\beta^2}{2\pi} \sum_{p=1}^{P} \left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}} \right)^2 \\
\mathrm{var}_c[\beta_p] &= P \frac{\sigma_\beta^2}{2\pi} \sum_{p=1}^{P} \left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}} \right)^2.
\end{aligned} \tag{E-5}$$

F Standard Errors with Cross-Correlated Residuals

We compute standard errors taking into account cross-correlation in the residuals using two methods: specifying a one-factor model of residual comovements and using industry factors.

F.1 Residual One-Factor Model

For the one-factor model, we assume that the errors for stock or portfolio i in month t have the structure

$$\varepsilon_{it} = \xi_i u_t + v_{it} \tag{F-1}$$

where $u_t \sim N(0, \sigma_u^2)$ and $v_{it} \sim N(0, \sigma_{vi}^2)$ is IID across stocks $i = 1, ..., N$. We write this in matrix notation for N stocks:

$$\varepsilon_t = \Xi u_t + \Sigma_v v_t, \tag{F-2}$$

where Ξ is a $N \times 1$ vector of residual factor loadings, Σ_v is a diagonal matrix containing $\{\sigma_{vi}^2\}$, and $v_t = (v_{1t}, ..., v_{Nt})$ is a $N \times 1$ vector of shocks. The residual covariance matrix, Ω_ϵ, is then given by

$$\Omega_\varepsilon = \Xi \sigma_u^2 \Xi' + \Sigma_v. \tag{F-3}$$

We estimate u_t by the following procedure. We denote e_{it} as the fitted residual for asset i at time t in the first-pass regression

$$e_{it} = R_{it} - \hat{a}_i - \hat{\beta}_i F_t. \tag{F-4}$$

We take an equally weighted average of residuals, $\tilde{u}_t$,

$$\tilde{u}_t = \frac{1}{N} \sum_i e_{it}, \tag{F-5}$$

and construct u_t to be the component of $\tilde{u}_t$ orthogonal to the factors, F_t, in the regression

$$\tilde{u}_t = c_0 + c_1 F_t + u_t. \tag{F-6}$$

We set $\hat{\sigma}_u^2$ to be the sample variance of u_t. To estimate the error factor loadings, ξ_i, we regress e_{it} onto u_t for each asset i. The fitted residuals are used to obtain estimates of σ_{vi}^2. This procedure obtains estimates $\hat{\Xi}$ and $\hat{\Sigma}_v$.

F.2 Industry Residual Model

In the industry residual model, we specify ten industry portfolios: durables, nondurables, manufacturing, energy, high technology, telecommunications, shops, healthcare, utilities, and other. The SIC definitions of these industries follow those constructed by Kenneth French at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html. We assume that the errors for stock or portfolio i have the structure

$$\varepsilon_{it} = \xi_i' u_t + v_{it}, \tag{F-7}$$

where ξ_i is a 10×1 vector of industry proportions, the jth element of which is the fraction of stocks in portfolio i that belong to industry j. If i is simply a stock, then one element of ξ_i is equal to one corresponding to the industry of the stock and all the other elements are equal to zero. The industry factors are contained in u_t, which is a 10×1 vector of industry-specific returns. We assume $u_t \sim N(0, \Sigma_u)$. We can stack all N stocks to write in matrix notation:

$$\Omega_\varepsilon = \Xi \Sigma_u \Xi' + \Sigma_v, \tag{F-8}$$

where Ξ is $N \times 10$ and Σ_v is a diagonal matrix containing $\{\sigma_{vi}^2\}$.

The industry residuals are specified to be uncorrelated with the factors F_t. To estimate Σ_u, we regress each of the ten industry portfolios onto F_t in time-series regressions, giving industry residual factors u_{jt} for industry j. We estimate Σ_u as the sample covariance matrix of $\{u_{jt}\}$.

To estimate Σ_v, we take the residuals e_{it} for asset i in equation (F-4) and define

$$\hat{v}_{it} = e_{it} - \hat{\xi}_i' u_t. \tag{F-9}$$

We estimate Σ_v to be the sample covariance matrix of $\{\hat{v}_{it}\}$.

References

[1] Ang, A., and J. Chen, 2007, "CAPM Over the Long-Run: 1926-2001," *Journal of Empirical Finance*, 14, 1-40.

[2] Ang, A., J. Chen, and Y. Xing, 2006, "Downside Risk," *Review of Financial Studies*, 19, 1191-1239.

[3] Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The Cross Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259-299.

[4] Avramov, D., and T. Chordia, 2006, "Asset Pricing Models and Financial Market Anomalies," *Review of Financial Studies*, 19, 1001-1040.

[5] Bekaert, G., R. J. Hodrick, and X. Zhang, 2010, "Aggregate Idiosyncratic Volatility?" working paper, Columbia University.

[6] Berk, J. B., 2000, "Sorting Out Sorts," *Journal of Finance*, 55, 407-427.

[7] Black, F., M. Jensen, and M. Scholes, 1972, "The Capital Asset Pricing Model: Some Empirical Tests," in Jensen, M., ed., *Studies in the Theory of Capital Markets*, Praeger, New York, 79-121.

[8] Blume, M. E., 1973, "Portfolio Theory: A Step Toward Its Practical Application," *Journal of Business*, 43, 152-73.

[9] Brennan, M. J., T. Chordia, and A. Subrahmanyam, 1998, "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*, 49, 345-373.

[10] Burnside, A. C., 2007, "Empirical Asset Pricing and Statistical Power in the Presence of Weak Risk Factors," NBER working paper 13357.

[11] Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 56, 1-43.

[12] Cochrane, J. H., 2001, *Asset Pricing*, Princeton University Press, NJ.

[13] Daniel, K., and S. Titman, 1997, "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns," *Journal of Finance*, 52, 1-34.

[14] Fama, E. F., and K. R. French, 1992, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427-465.

[15] Fama, E. F., and K. R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3-56.

[16] Fama, E. F., and K. R. French, 2008, "Dissecting Anomalies," *Journal of Finance*, 63, 1653-1678.

[17] Fama, E. F., and J. D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 71, 607-636.

[18] Gibbons, M. R., S. Ross, and J. Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121-1152.

[19] Grauer, R. R., and J. A. Janmaat, 2004, "The Unintended Consequences of Grouping in Tests of Asset Pricing Models," *Journal of Banking and Finance*, 28, 2889-2914.

[20] Grauer, R. R., and J. A. Janmaat, 2008, "Cross-Sectional Tests of the CAPM and Fama-French Three-Factor Model," working paper, UBC.

[21] Hou, K., and R. L. Kimmel, 2006, "On Estimation of Risk Premia in Linear Factor Models," working paper, Ohio State University.

[22] Huberman, G., and S. Kandel, 1987, "Mean-Variance Spanning," *Journal of Finance*, 42, 873-888.

[23] Jagannathan, R., G. Skoulakis, and Z. Wang, 2002, "The Analysis of the Cross Section of Security Returns," forthcoming Aït-Sahalia, Y., and Hansen, L. P., eds., *Handbook of Financial Econometrics*, Elsevier.

[24] Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3-53.

[25] Jagannathan, R., and Z. Wang, 1998, "An Asymptotic Theory for Estimating Beta-Pricing Models using Cross-Sectional Regression," *Journal of Finance*, 53, 1285-1309.

[26] Jobson, D., and R. Korkie, 1982, "Potential Performance and Tests of Portfolio Efficiency," *Journal of Financial Economics*, 10, 433-436.

[27] Kan, R., 1998, "On the Explanatory power of Asset Pricing Models Across and Within Portfolios," working paper, University of Toronto.

[28] Kan, R., and C. Zhang, 1999, "Two-Pass Tests of Asset Pricing Models with Useless Factors," *Journal of Finance*, 54, 203-35.

[29] Kan, R., and C. Robotti, 2006, "Specification Tests of Asset Pricing Models Using Excess Returns," Federal Reserve Bank of Atlanta Working Paper 2006-10.

[30] Kandel, S., and R. Stambaugh, 1995, "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Finance*, 50, 157-184.

[31] Lettau, M., and S. Ludvigson, 2001a, "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815-849.

[32] Lettau, M., and S. Ludvigson, 2001b, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time Varying," *Journal of Political Economy*, 109, 1238-1287.

[33] Lewellen, J., S. Nagel, and J. Shanken, 2010, "A Skeptical Appraisal of Asset-Pricing Tests," *Journal of Financial Economics*, 96, 175-194.

[34] Lewellen, J., and S. Nagel, 2006, "The Conditional CAPM Does Not Explain Asset-Pricing Anomalies," *Journal of Financial Economics*, 82, 289-314.

[35] Litzenberger, R. H., and K. Ramaswamy, 1979, "The Effects of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence," *Journal of Financial Economics*, 7, 163-195.

[36] Mehra, R., and E. Prescot, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.

[37] Meng, J. G., G. Hu, and J. Bai, 2007, "A Simple Method for Estimating Betas when Factors are Measured with Error," working paper, NYU.

[38] Merton, R. C., 1980, "On Estimating the Expected Return on the Market," *Journal of Financial Economics*, 8, 323-361.

[39] Pástor, L., and R. F. Stambaugh, 2003, "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111, 642-685.

[40] Ross, S., 1976, "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341-360.

[41] Roll, R., and S. A. Ross, 1994, "On the Cross-Sectional Relation between Expected Returns and Betas," *Journal of Finance*, 49, 101-121.

[42] Santos, T., and P. Veronesi, 2006, "Labor Income and Predictable Stock Returns," *Review of Financial Studies*, 19, 1-44.

[43] Shanken, J., 1985, "Multivariate Tests of the Zero-Beta CAPM," *Journal of Financial Economics*, 14, 327-348.

[44] Shanken, J., 1992, "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 1-33.

[45] Shanken, J., and G. Zhou, 2007, "Estimating and Testing Beta Pricing Models: Alternative Methods and Their Performance in Simulations," *Journal of Financial Economics*, 84, 40-86.

[46] Velu, R., and G. Zhou, 1999, "Testing Multi-Beta Asset Pricing Models," *Journal of Empirical Finance*, 6, 219-241.

[47] Zhou, G., 1991, "Small Sample Tests of Portfolio Efficiency," *Journal of Financial Economics*, 30, 165-191.

Table 1: Summary Statistics of Betas and Idiosyncratic Volatilities

	Means			Stdev			Correlations		
	$\hat{\beta}$	$\hat{\sigma}$	$\ln\hat{\sigma}$	$\hat{\beta}$	$\hat{\sigma}$	$\ln\hat{\sigma}$	$(\hat{\beta},\hat{\sigma})$	$(\hat{\beta},\ln\hat{\sigma})$	No Obs
1970-1975	1.24	0.43	-0.93	0.55	0.19	0.42	0.48	0.49	2934
1975-1980	1.24	0.39	-1.04	0.58	0.20	0.47	0.38	0.45	3218
1980-1985	1.08	0.45	-0.92	0.63	0.27	0.49	0.25	0.30	3500
1985-1990	1.03	0.49	-0.85	0.52	0.29	0.52	0.05	0.13	3764
1990-1995	0.94	0.52	-0.82	0.95	0.37	0.56	-0.05	0.06	4000
1995-2000	1.01	0.65	-0.57	0.78	0.38	0.50	0.51	0.55	4363
2000-2005	1.35	0.54	-0.75	1.04	0.30	0.53	0.52	0.55	3493
2005-2010	1.33	0.51	-0.82	0.71	0.31	0.51	0.45	0.47	3009
2010-2015	1.20	0.40	-1.08	0.70	0.26	0.55	0.21	0.28	2552
Overall	1.14	0.50	-0.85	0.76	0.31	0.53	0.26	0.31	30833

The table reports the summary statistics of estimated betas ($\hat{\beta}$) and idiosyncratic volatility ($\hat{\sigma}$) over each five year sample and over the entire sample. We estimate betas and idiosyncratic volatility in each five-year non-overlapping period using time-series regressions of monthly excess stock returns onto a constant and monthly excess market returns. The idiosyncratic stock volatilities are annualized by multiplying by $\sqrt{12}$. The last column reports the number of stock observations.

Table 2: Variance Ratio Efficiency Losses in Monte Carlo simulations

	α Efficiency Loss					λ Efficiency Loss				
Number of Portfolios P	10	25	50	100	250	10	25	50	100	250
Panel A: Sorting on True Betas, Correlated Betas and Idiosyncratic Volatility										
Mean	2.93	2.79	2.73	2.66	2.54	2.97	2.80	2.73	2.67	2.54
StDev	0.14	0.13	0.13	0.12	0.11	0.13	0.12	0.12	0.11	0.10
Panel B: Correlated Betas and Idiosyncratic Volatility										
Mean	5.17	5.07	4.96	4.78	4.37	4.97	4.86	4.75	4.60	4.23
StDev	0.44	0.42	0.40	0.38	0.33	0.40	0.39	0.37	0.35	0.31
Panel C: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals										
Mean	38.9	30.2	23.1	16.3	9.4	21.2	16.7	13.0	9.7	6.5
StDev	20.9	15.7	11.6	7.5	3.5	16.6	11.8	8.0	4.8	2.0
Panel D: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals Entry and Exit of Firms										
Mean	43.0	34.2	26.9	19.6	11.8	24.1	19.4	15.7	12.0	8.3
StDev	22.0	16.8	12.7	8.5	4.1	18.5	13.2	9.2	5.6	2.4

The table reports the efficiency loss variance ratios $\text{var}_p(\hat{\theta})/\text{var}(\hat{\theta})$ for $\theta = \alpha$ or λ where $\text{var}_p(\hat{\theta})$ is computed using P portfolios and $\text{var}(\theta)$ is computed using all stocks. We simulate 10,000 small samples of $T = 60$ months with $N = 5,000$ stocks using the model in equation (21). Panel A sorts stocks by true betas in each small sample and the remaining panels sort stocks by estimated betas. All the portfolios are formed equally weighting stocks at the end of the period. Panels B-D estimate betas in each small sample by regular OLS and the standard error variances are computed using the true cross-sectional betas and idiosyncratic volatilities. Panels A and B assume correlated betas and idiosyncratic volatility following the process in equation (22). Panel C introduces cross-sectionally correlated residuals across stocks following equation (24). In Panel D, firms enter and exit stochastically and upon entry have a log-logistic model for duration given by equation (25). To take a cross section of 5,000 firms that have more than 36 months of returns, on average, requires a steady-state firm universe of 6,607 stocks.

Table 3: Estimates of a One-Factor Model

			Residual Factor Model				Industry Residual Model							
			Max Lik		GMM		Max Lik		GMM		$\hat{\beta}$ Cross Section			
Num Ports P		Estimate (%)	SE	|t-stat|	SE	|t-stat|	SE	|t-stat|	SE	|t-stat|	$E_c(\hat{\beta})$	$\sigma_c(\hat{\beta})$	5%	95%
Panel A: All Stocks														
	$\hat{\alpha}$	8.54	0.16	53.86	1.40	6.12	0.34	24.85	0.73	11.71	1.14	0.76	0.12	2.44
	$\hat{\lambda}_{MKT}$	4.79	0.16	29.76	1.05	4.56	0.18	27.22	0.55	8.73				
Panel B: "Ex-Post" Portfolios														
5	$\hat{\alpha}$	8.21	0.79	10.46	1.48	5.56	1.32	6.20	1.68	4.89	1.16	0.69	0.22	2.26
	$\hat{\lambda}_{MKT}$	5.07	0.72	7.03	1.37	3.70	1.17	4.33	1.60	3.18				
10	$\hat{\alpha}$	8.26	0.63	13.13	1.42	5.84	1.04	7.96	1.36	6.06	1.16	0.72	0.12	2.57
	$\hat{\lambda}_{MKT}$	5.03	0.60	8.36	1.30	3.86	0.89	5.67	1.27	3.95				
25	$\hat{\alpha}$	8.43	0.49	17.07	1.35	6.25	0.81	10.39	1.07	7.85	1.16	0.74	0.13	2.49
	$\hat{\lambda}_{MKT}$	4.88	0.49	10.05	1.20	4.05	0.65	7.47	0.95	5.11				
50	$\hat{\alpha}$	8.44	0.41	20.40	1.33	6.37	0.70	12.06	0.93	9.11	1.16	0.74	0.13	2.42
	$\hat{\lambda}_{MKT}$	4.87	0.41	11.75	1.16	4.21	0.53	9.13	0.79	6.17				
Panel C: "Ex-Ante" Portfolios														
5	$\hat{\alpha}$	14.72	1.09	13.50	2.81	5.23	2.33	6.31	3.04	4.85	1.12	0.35	0.62	1.64
	$\hat{\lambda}_{MKT}$	1.14	1.50	0.76	2.81	0.41	2.19	0.52	2.88	0.40				
10	$\hat{\alpha}$	14.24	0.91	15.61	2.63	5.42	1.82	7.82	2.38	5.99	1.12	0.36	0.60	1.67
	$\hat{\lambda}_{MKT}$	1.58	1.30	1.22	2.65	0.60	1.68	0.94	2.23	0.71				
25	$\hat{\alpha}$	14.13	0.73	19.42	2.45	5.76	1.40	10.07	1.80	7.87	1.12	0.36	0.58	1.70
	$\hat{\lambda}_{MKT}$	1.69	1.05	1.61	2.50	0.68	1.27	1.33	1.65	1.02				
50	$\hat{\alpha}$	14.08	0.62	22.63	2.37	5.94	1.20	11.77	1.52	9.24	1.12	0.36	0.59	1.70
	$\hat{\lambda}_{MKT}$	1.73	0.85	2.03	2.42	0.72	1.06	1.64	1.38	1.26				

Note to Table 3

The point estimates of α and λ for the single factor, MKT, in equation (1) are reported over all stocks (Panel A) and various portfolio sortings (Panels B and C). The betas are estimated by running a first-pass OLS regression of monthly excess stock returns onto monthly excess market returns over non-overlapping five-year samples beginning in January 1971 and ending in December 2015. All stock returns in each five-year period are stacked and treated as one panel. We use a second-pass cross-sectional regression to compute $\hat{\alpha}$ and $\hat{\lambda}$. Using these point estimates we compute the various standard errors (SE) and absolute values of t-statistics (|t-stat|). We compute the maximum likelihood standard errors (equations (12) and (13)) in the columns labeled "Max Lik" and GMM standard errors, detailed in Appendix C, in the columns labeled "GMM". We allow for cross-correlated residuals computed using a one-factor model or industry classifications, which are described in Appendix F. The three last columns labeled "$\hat{\beta}$ Cross Section" list various statistics of the cross-sectional beta distribution: the cross-sectional standard deviation, $\sigma_c(\hat{\beta})$, and the beta values corresponding to the 5%- and 95%-tiles of the cross-sectional distribution of beta. In Panel B we form "ex-post" portfolios, which are formed in each five-year period by grouping stocks into equally-weighted P portfolios based on realized estimated betas over those five years. In Panel C we form "ex-ante" portfolios by grouping stocks into portfolios at the beginning of each calendar year, ranking on the estimated market beta over the previous five years. Equally-weighted portfolios are created and the portfolios are held for twelve months to produce monthly portfolio returns. The portfolios are rebalanced annually at the beginning of each calendar year. The first estimation period is January 1966 to December 1970 to produce monthly returns for the calendar year 1971 and the last estimation period is January 2010 to December 2014 to produce monthly returns for 2015. Thus, the sample period is exactly the same as using all stocks and the ex-post portfolios. After the ex-ante portfolios are created, we follow the same procedure as Panels A and B to compute realized OLS market betas in each non-overlapping five-year period and then estimate a second-pass cross-sectional regression. In both Panels B and C, the second-pass cross-sectional regression is run only on the P portfolio test assets. All estimates $\hat{\alpha}$ and $\hat{\lambda}$ are annualized by multiplying the monthly estimates by 12.

Table 4: Tests for $H_0^{\lambda=\mu}$ (|T-statistics|) for the One-Factor Model

		Residual Factor		Industry Residuals	
Num Ports P	$\hat{\lambda}$ (%)	Max Lik	GMM	Max Lik	GMM
$\hat{\mu}_{MKT} = 6.43\%$					
All Stocks	4.79	10.16	1.56	9.32	2.98
"Ex-Post" Portfolios					
5	5.07	1.95	1.02	1.20	0.88
10	5.03	2.34	1.08	1.58	1.10
25	4.88	3.20	1.29	2.38	1.63
50	4.87	3.77	1.35	2.93	1.98
"Ex-Ante" Portfolios					
5	1.14	3.23	1.72	2.22	1.68
10	1.58	3.73	1.83	2.88	2.17
25	1.69	4.52	1.90	3.74	2.87
50	1.73	5.50	1.94	4.44	3.41

The table reports absolute values of t-statistics for testing if the cross-sectional risk premium, λ, is equal to the time-series mean of the factor portfolio, μ, which is the hypothesis test $H_0^{\lambda=\mu}$ for the one-factor model. The maximum likelihood test and the GMM test, in the columns labeled "Max Lik" and "GMM", respectively, are detailed in Appendix C. We allow for cross-correlated residuals computed using a one-factor model or industry classifications, which are described in Appendix F. The column labeled "$\hat{\lambda}$" reports the annualized estimate of the cross-sectional market risk premium, obtained by multiplying the monthly estimate by 12. The data sample is January 1971 to December 2015.

Table 5: Cross-Sectional Distribution of Fama-French (1993) Factor Loadings

	Factor Loadings	$E_c(\hat{\beta})$	$\sigma_c(\hat{\beta})$	5%	95%
All Stocks	$\hat{\beta}_{MKT}$	1.02	0.73	-0.01	2.24
	$\hat{\beta}_{SMB}$	0.94	1.21	-0.52	2.91
	$\hat{\beta}_{HML}$	0.18	1.21	-1.71	1.93
"Ex-Post" Portfolios					
$2 \times 2 \times 2$	$\hat{\beta}_{MKT}$	1.03	0.54	0.25	1.75
	$\hat{\beta}_{SMB}$	0.94	0.85	-0.07	2.17
	$\hat{\beta}_{HML}$	0.17	0.85	-1.05	1.51
$3 \times 3 \times 3$	$\hat{\beta}_{MKT}$	1.03	0.61	0.14	2.02
	$\hat{\beta}_{SMB}$	0.94	0.96	-0.27	2.53
	$\hat{\beta}_{HML}$	0.17	0.97	-1.30	1.74
"Ex-Ante" Portfolios					
$2 \times 2 \times 2$	$\hat{\beta}_{MKT}$	1.01	0.20	0.69	1.34
	$\hat{\beta}_{SMB}$	0.88	0.37	0.35	1.54
	$\hat{\beta}_{HML}$	0.22	0.29	-0.27	0.67
$3 \times 3 \times 3$	$\hat{\beta}_{MKT}$	1.01	0.23	0.60	1.37
	$\hat{\beta}_{SMB}$	0.88	0.43	0.24	1.65
	$\hat{\beta}_{HML}$	0.22	0.34	-0.30	0.74

The table reports cross-sectional summary statistics of estimated Fama-French (1993) factor loadings, $\hat{\beta}_{MKT}$, $\hat{\beta}_{SMB}$, and $\hat{\beta}_{HML}$. We report cross-sectional means, standard deviations ($\sigma_c(\hat{\beta})$), and and the estimated factor loadings corresponding to the 5%- and 95%-tiles of the cross-sectional distribution. The factor loadings are estimated by running a multivariate OLS regression of monthly excess stock returns onto the monthly Fama-French (1993) factors (MKT, SMB, and HML) over non-overlapping five-year samples beginning in January 1971 and ending in December 2015. All of the factor loadings in each five-year period are stacked and treated as one panel. The "ex-post" portfolios are formed in each five-year period by grouping stocks into P equally-weighted portfolios based on realized estimated factor loadings over those five years. We form $n \times n \times n$ portfolios using sequential sorts of n portfolios ranked on each of the Fama-French factor loadings at the end of each five-year period. We sort first on $\hat{\beta}_{MKT}$, then on $\hat{\beta}_{SMB}$, and then finally on $\hat{\beta}_{HML}$. The "ex-ante" portfolios are formed by grouping stocks into portfolios at the beginning of each calendar year ranking on the estimated factor loadings over the previous five years. Equally-weighted, sequentially sorted portfolios are created and the portfolios are held for twelve months to produce monthly portfolio returns. The portfolios are rebalanced annually at the beginning of each calendar year. The first estimation period is January 1966 to December 1970 to produce monthly returns for the calendar year 1971 and the last estimation period is January 2010 to December 2014 to produce monthly returns for 2015.

Table 6: Estimates of the Fama-French (1993) Model

			Residual Factor Model				Industry Residual Model			
			Max Lik		GMM		Max Lik		GMM	
Num Ports P		Estimate (%)	SE	|t-stat|	SE	|t-stat|	SE	|t-stat|	SE	|t-stat|
Panel A: All Stocks										
	$\hat{\alpha}$	2.43	0.16	14.76	0.91	2.67	0.33	7.41	0.80	3.03
	$\hat{\lambda}_{MKT}$	5.05	0.16	31.18	0.63	8.05	0.18	28.58	0.47	10.72
	$\hat{\lambda}_{SMB}$	6.79	0.10	67.17	0.84	8.10	0.11	61.81	0.64	10.59
	$\hat{\lambda}_{HML}$	0.01	0.11	0.11	0.56	0.02	0.12	0.10	0.44	0.03
Panel B: "Ex-Post" Portfolios										
$2 \times 2 \times 2$	$\hat{\alpha}$	6.88	0.54	12.64	0.96	7.17	0.86	7.99	1.18	5.84
	$\hat{\lambda}_{MKT}$	4.06	0.51	7.95	0.90	4.50	0.70	5.77	1.01	4.00
	$\hat{\lambda}_{SMB}$	3.73	0.36	10.24	0.65	5.76	0.45	8.31	0.58	6.43
	$\hat{\lambda}_{HML}$	-2.99	0.32	-9.35	0.62	-4.80	0.47	-6.31	0.66	-4.51
$3 \times 3 \times 3$	$\hat{\alpha}$	6.47	0.38	16.82	0.82	7.92	0.65	10.01	0.87	7.44
	$\hat{\lambda}_{MKT}$	4.05	0.37	11.05	0.79	5.11	0.48	8.51	0.71	5.69
	$\hat{\lambda}_{SMB}$	4.12	0.25	16.54	0.61	6.78	0.30	13.69	0.43	9.58
	$\hat{\lambda}_{HML}$	-2.77	0.24	-11.76	0.56	-4.97	0.33	-8.46	0.49	-5.64
Panel C: "Ex-Ante" Portfolios										
$2 \times 2 \times 2$	$\hat{\alpha}$	11.01	4.50	2.45	2.49	4.41	1.86	5.93	2.45	4.49
	$\hat{\lambda}_{MKT}$	-5.54	5.19	-1.07	3.51	-1.58	2.16	-2.57	2.87	-1.93
	$\hat{\lambda}_{SMB}$	11.50	3.74	3.08	1.97	5.85	1.22	9.44	1.40	8.22
	$\hat{\lambda}_{HML}$	1.64	3.91	0.42	1.93	0.85	1.40	1.17	1.85	0.89
$3 \times 3 \times 3$	$\hat{\alpha}$	10.51	2.52	4.17	2.09	5.03	1.30	8.06	1.74	6.03
	$\hat{\lambda}_{MKT}$	-4.87	2.80	-1.74	2.89	-1.68	1.37	-3.56	1.98	-2.46
	$\hat{\lambda}_{SMB}$	11.50	2.01	5.72	1.66	6.93	0.76	15.08	0.98	11.70
	$\hat{\lambda}_{HML}$	0.86	1.95	0.44	1.59	0.54	0.91	0.94	1.23	0.70

Note to Table 6

The point estimates $\hat{\alpha}$, $\hat{\lambda}_{MKT}$, $\hat{\lambda}_{SMB}$, and $\hat{\lambda}_{HML}$ in equation (26) are reported over all stocks (Panel A) and various portfolio sortings (Panels B and C). The betas are estimated by running a first-pass multivariate OLS regression of monthly excess stock returns onto the monthly Fama-French (1993) factors (MKT, SMB, and HML) over non-overlapping five-year samples beginning in January 1971 and ending in December 2015. All of the stock returns in each five-year period are stacked and treated as one panel. We use a second-pass cross-sectional regression to compute the cross-sectional coefficients. Using these point estimates we compute the various standard errors (SE) and absolute values of t-statistics (|t-stat|). We compute the maximum likelihood standard errors (equations (12) and (13)) in the columns labeled "Max Lik" and GMM standard errors, detailed in Appendix C, in the columns labeled "GMM". We allow for cross-correlated residuals computed using a one-factor model or industry classifications, which are described in Appendix F. In Panel B we form "ex-post" portfolios, which are formed in each five-year period by grouping stocks into P equally-weighted portfolios based on realized estimated factor loadings over those five years. We form $n \times n \times n$ portfolios using sequential sorts of n portfolios ranked on each of the Fama-French factor loadings at the end of each five-year period. We sort first on $\hat{\beta}_{MKT}$, then on $\hat{\beta}_{SMB}$, and then finally on $\hat{\beta}_{HML}$. In Panel C we form "ex-ante" portfolios by grouping stocks into portfolios at the beginning of each calendar year, ranking on the estimated factor loadings over the previous five years. Equally-weighted, sequentially sorted portfolios are created and the portfolios are held for twelve months to produce monthly portfolio returns. The portfolios are rebalanced annually at the beginning of each calendar year. The first estimation period is January 1966 to December 1970 to produce monthly returns for the calendar year 1971 and the last estimation period is January 2010 to December 2014 to produce monthly returns for 2015. Thus, the sample period is exactly the same as using all stocks and the ex-post portfolios. After the ex-ante portfolios are created, we follow the same procedure as Panels A and B to compute realized OLS factor loadings in each non-overlapping five-year period and then estimate a second-pass cross-sectional regression. In both Panels B and C, the second-pass cross-sectional regression is run only on the P portfolio test assets. All estimates are annualized by multiplying the monthly estimates by 12.

Table 7: Tests for $H_0^{\lambda=\mu}$ ($|$T-statistics$|$) for the Fama-French (1993) Model

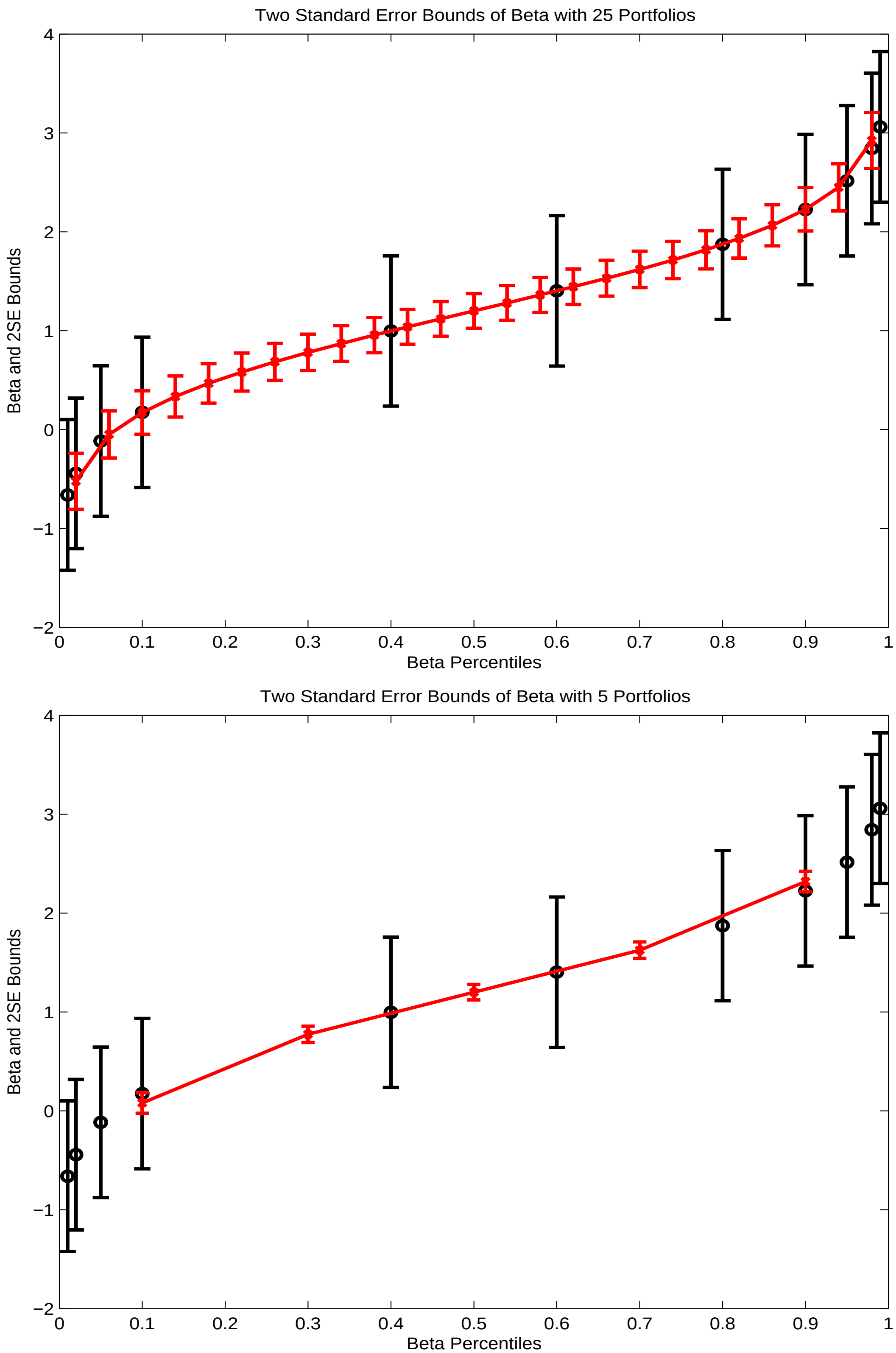
			Residual Factor		Industry Residuals	
Num Ports P		Estimate (%)	Max Lik	GMM	Max Lik	GMM
$\hat{\mu}_{MKT} = 6.43\%$, $\hat{\mu}_{SMB} = 2.16\%$, $\hat{\mu}_{HML} = 3.90\%$						
All Stocks	$\hat{\lambda}_{MKT}$	5.05	8.51	2.20	7.80	2.92
	$\hat{\lambda}_{SMB}$	6.79	45.83	5.53	42.17	7.23
	$\hat{\lambda}_{HML}$	0.01	35.91	6.90	31.94	8.90
"Ex-Post" Portfolios						
$2 \times 2 \times 2$	$\hat{\lambda}_{MKT}$	4.06	4.64	2.63	3.37	2.33
	$\hat{\lambda}_{SMB}$	3.73	4.32	2.43	3.50	2.71
	$\hat{\lambda}_{HML}$	-2.99	21.52	11.04	14.52	10.39
$3 \times 3 \times 3$	$\hat{\lambda}_{MKT}$	4.05	6.49	3.00	5.00	3.34
	$\hat{\lambda}_{SMB}$	4.12	7.87	3.22	6.51	4.56
	$\hat{\lambda}_{HML}$	-2.77	28.31	11.96	20.37	13.58
"Ex-Ante" Portfolios						
$2 \times 2 \times 2$	$\hat{\lambda}_{MKT}$	-5.54	7.22	3.41	5.55	4.18
	$\hat{\lambda}_{SMB}$	11.50	8.88	4.75	7.67	6.68
	$\hat{\lambda}_{HML}$	1.64	2.37	1.17	1.61	1.22
$3 \times 3 \times 3$	$\hat{\lambda}_{MKT}$	-4.87	10.38	3.91	8.25	5.71
	$\hat{\lambda}_{SMB}$	11.50	13.75	5.63	12.25	9.50
	$\hat{\lambda}_{HML}$	0.86	4.75	1.91	3.33	2.48

The table reports absolute values of t-statistics for testing if the cross-sectional risk premium, λ, is equal to the time-series mean of the factor portfolio, μ, which is the hypothesis test $H_0^{\lambda=\mu}$ for the Fama and French (1993) three-factor model. The maximum likelihood test and the GMM test, in the columns labeled "Max Lik" and "GMM", respectively, are detailed in Appendix C. We allow for cross-correlated residuals computed using a one-factor model or industry classifications, which are described in Appendix F. Estimates of the cross-sectional factor risk premia are annualized by multiplying the monthly estimate by 12. The data sample is January 1971 to December 2015.

Note to Table 7

We estimate the Fama-French (1993) model (equation (26)) using all stocks (Panel A), 5×5 ex-ante portfolios sorted on market beta and book-to-market ratios (upper part of Panel B), and 5×5 ex-ante portfolios sorted on size and book-to-market ratios (lower part of Panel B). The betas are estimated by running a first-pass multivariate OLS regression of monthly excess stock returns onto the monthly Fama-French (1993) factors (MKT, SMB, and HML) over non-overlapping five-year samples beginning in January 1971 and ending in December 2015. The stock returns in each five-year period are stacked and treated as one panel. We use a second-pass cross-sectional regression to compute the cross-sectional coefficients. Using these point estimates we compute the various standard errors (SE) and absolute values of t-statistics (|t-stat|). We compute the maximum likelihood standard errors (equations (12) and (13)) in the columns labeled "Max Lik" and GMM standard errors, detailed in Appendix C, in the columns labeled "GMM". We allow for cross-correlated residuals computed using a one-factor model or industry classifications, which are described in Appendix F. The stock universe in this table differs from Tables 6 and 7 as we require all stocks to have observable market capitalization and book-to-market ratios. The stock universe in Panels A and B is the same. Panel A considers a cross-sectional regression with a constant and only factor loadings and also a specification which includes the book-to-market ratio (B/M). In Panel B, we form "ex-ante" portfolios by grouping stocks into portfolios at the beginning of each calendar year, ranking on market betas and book-to-market ratios or market capitalization and book-to-market ratios. The book-to-market ratios are constructed from COMPUSTAT as the ratio of book equity divided by market value. Book equity is defined as total assets (COMPUSTAT Data 6) minus total liabilities (COMPUSTAT Data 181). Market value is constructed from CRSP and defined as price times shares outstanding. We match fiscal year-end data for book equity from the previous year, $t - 12$, with time t market data. Equally-weighted portfolios are created and the portfolios are held for twelve months to produce monthly portfolio returns. After the ex-ante portfolios are created, we follow the same procedure as Panel A to compute realized OLS factor loadings in each non-overlapping five-year period and then estimate a second-pass cross-sectional regression. In Panel B, the second-pass cross-sectional regression is run only on the P portfolio test assets. The coefficients on α, β_{MKT}, β_{SMB}, and β_{HML} are annualized by multiplying the monthly estimates by 12.

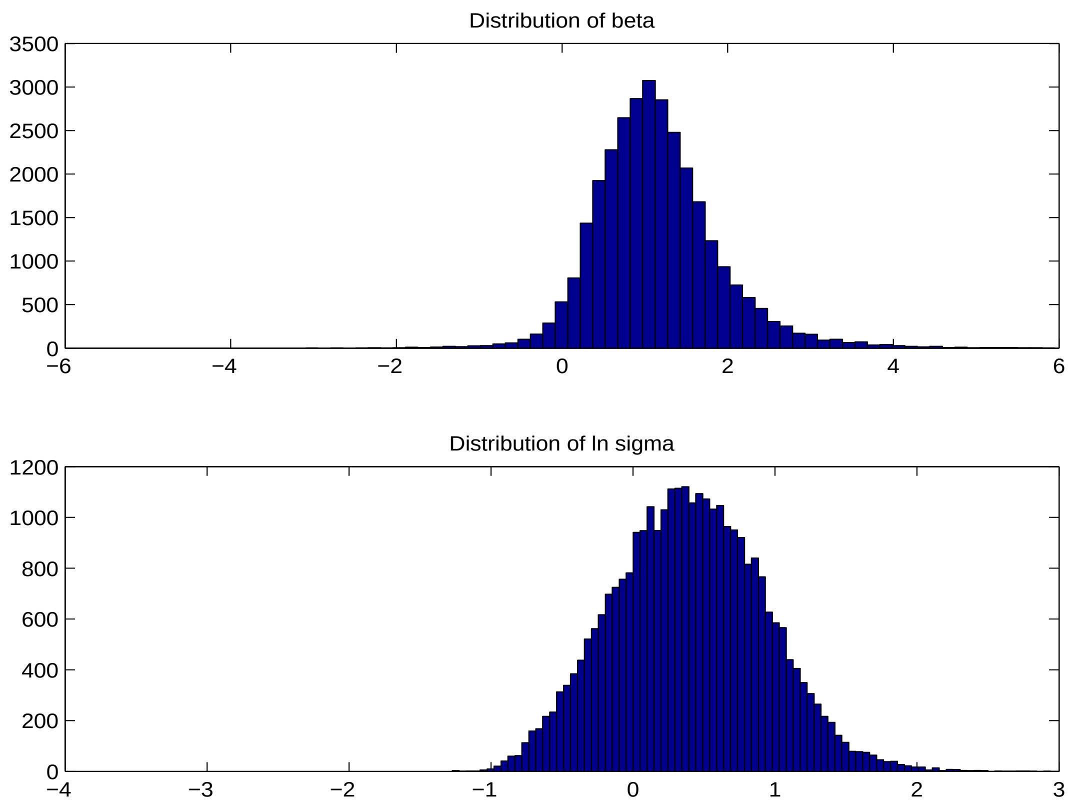
Figure 1: Standard Errors for $\hat{\beta}$ Using All Stocks or Portfolios



Note to Figure 1

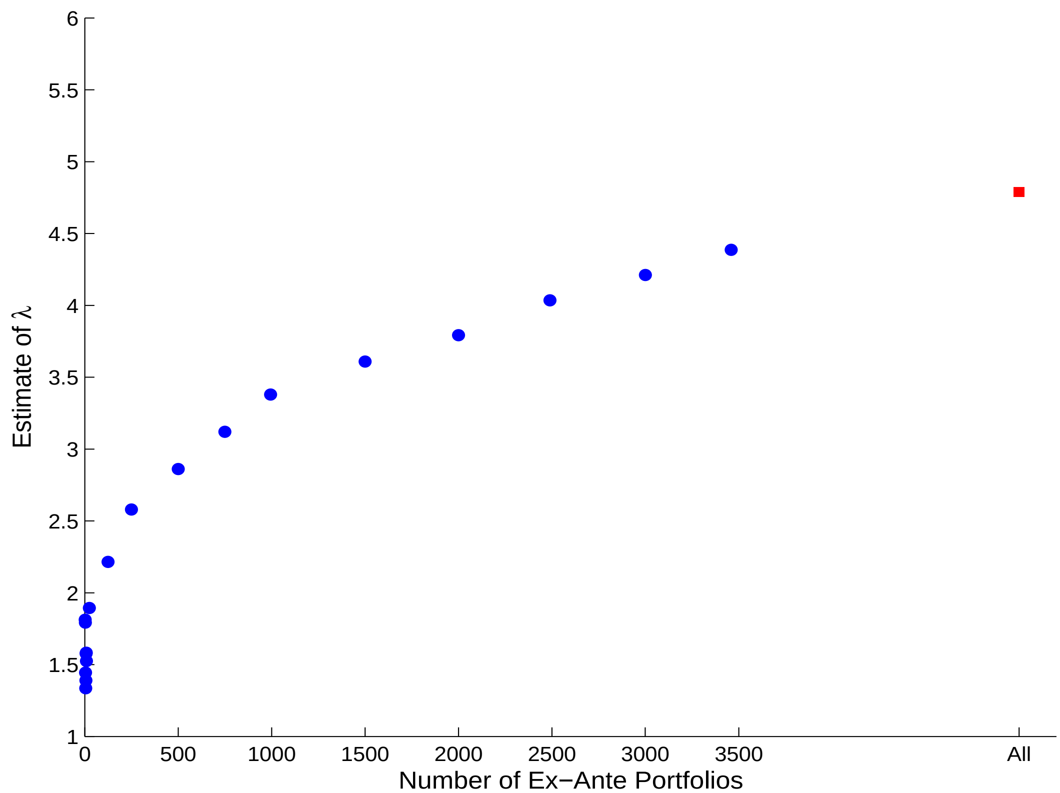
We assume a single factor model where $F_t \sim N(0, (0.15)^2/12)$ and the factor risk premium $\lambda = 0.06/12$. Betas are drawn from a normal distribution with mean $\mu_\beta = 1.1$ and standard deviation $\sigma_\beta = 0.7$ and idiosyncratic volatility across stocks is constant at $\sigma_i = \sigma = 0.5/\sqrt{12}$. We assume a sample of size $T = 60$ months with $N = 1000$ stocks. We graph two standard error bars of $\hat{\beta}$ for the various percentiles of the true distribution marked in circles for percentiles 0.01, 0.02, 0.05, 0.1, 0.4, 0.6, 0.8, 0.9, 0.95, 0.98, and 0.99. These are two-standard error bands for individual stock betas. The standard error bands for the portfolio betas for $P = 25$ portfolios (top panel) and $P = 5$ portfolios (bottom panel) are marked with small crosses and connected by the red line. These are graphed at the percentiles which correspond to the mid-point mass of each portfolio. The formula for var$(\hat{\beta})$ is given in equation (18) and the computation for the portfolio moments are given in Appendix E.

Figure 2: Empirical Distributions of Betas and Idiosyncratic Volatilities



The figure plots an empirical histogram over the 15,256 firms in non-overlapping five year samples from 1971-2015, computed by OLS estimates. Panel A plots the histogram of market betas while Panel B plots the histogram of annualized log idiosyncratic volatility.

Figure 3: One-Factor Risk Premium Estimates with Ex-Ante Portfolios



The figure plots $\hat{\lambda}$ in a one-factor model using P "ex-ante" portfolios in blue circles. The ex-ante portfolios are formed by grouping stocks into portfolios at the beginning of each calendar year ranking on the estimated market beta over the previous five years. Equally-weighted portfolios are created and the portfolios are held for twelve months to produce monthly portfolio returns. The estimate obtained using all individual stocks is labeled "All" on the x-axis and is graphed in the red square. The first-pass beta estimates are obtained using non-overlapping five-year samples from 1971-2015 with OLS.