# Using Stocks or Portfolios in Tests of Factor Models[1]

Andrew Ang[2]
Columbia University and Blackrock and NBER

Jun Liu[3]
UCSD

Krista Schwarz[4]
University of Pennsylvania

This Version: August 7, 2018

JEL Classification: G12, C13

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

[1]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 and an anonymous referee for comments. The usual disclaimer applies.

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

[3]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/.

[4]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 (EIV) problem of estimated betas as regressors when using individual stocks:

> *...If an investor's assessments of $\alpha_i$ and $\beta_i$ 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 $\alpha_i$'s and $\beta_i$'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 $\alpha_i$ and $\beta_i$ 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 diversified portfolios allows for more precise 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 (1972) 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.

Forming portfolios dramatically reduces the standard errors of factor loadings due to decreasing idiosyncratic risk. Estimating risk premia with more precise factor loadings will produce less biased risk premia estimates, ceteris paribus. However, there is a tradeoff. In this paper, we show that the more precise estimates of *portfolio 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 the number of portfolios used for a given number of stocks, the smaller the standard errors of the portfolio factor loadings 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 portfo-

[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), Gibbons, Ross and Shanken (1985), Zhou (1991), Shanken (1992) and 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.

lios 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 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 dispersed 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. Forming portfolios leaves the researcher with fewer factor loading estimates than would the case if the test assets had not been combined into portfolios. 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.

An early motivator for portfolio formation was the aim of reducing bias in the risk premium point estimate by grouping stocks and thus reducing beta measurement error. With a risk premium estimator that does not adjust for bias, the tradeoff in bias versus efficiency changes with the number of portfolios. For a given number of stocks, when fewer portfolios are formed, bias tends to diminish but the estimator becomes less efficient. We use Monte Carlo simulations to compare the relative efficiency and bias in several variations of our simple analytical setting, over a range of portfolio sizes and individual stocks. We find that relaxing standard assumptions can have a large effect on both bias and efficiency. Importantly, higher cross-sectional variation in factor loadings motivates grouping fewer stocks into portfolios both on bias and efficiency grounds. Depending on the data environment, our results argue for the use of either a large number of portfolios (with few stocks in each portfolio) or individual stocks when estimating risk premia. The case tilts in favor of individual stocks when estimation errors in factor loadings are large or the variance in true factor loadings is low.

Our analysis on estimation efficiency complements Kim's (1995) study of the problem of bias in the point estimate. He proposes a maximum likelihood-based correction for bias. Kim uses results on classical measurement error to adjust two-pass Fama MacBeth estimators for bias from EIV that is otherwise larger when using stocks as test assets as compared to portfolios. When full-blown maximum likelihood estimation is used, there is no clear motivation for forming portfolios. Thus, risk premium estimation methods that account for bias present the clearest case for using individual stocks over portfolios due to the efficiency gains.

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.

Our paper focuses on the efficiency losses from forming portfolios. Still this is by no means the only argument for using individual stocks. Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999, 2003) and Berk (2000) emphasize that the particular way in which stocks are grouped into portfolios is subject to potential data-snooping biases. For instance, sorting firms by characteristics that are known to be correlated with returns in sample can affect risk premiums. Portfolios formed on anomaly factors can embed spurious risk premiums. Grauer and Janmaat (2004) show that portfolio grouping under the alternative when a factor model is false may cause the model to appear correct. Likewise, Lewellen, Nagel and Shanken (2010) show that the Fama French portfolios have a strong factor structure which biases the researcher in favor of factor models. These problems are all avoided by working with individual stocks which takes away the potential for data mining from the construction of portfolios. As emphasized by Campbell, Lo and MacKinlay (1997), data-mining difficulties are tough to circumvent completely. Still, avoiding the formation of portfolios as test assets removes a key element of the problem.

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 assumes that we are concerned with efficiency in a correctly-specified model. A large literature discusses misspecification tests 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, but that is not the setting considered in this paper. 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]

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. We also show that using portfolios versus individual stocks matters in actual data.

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 the effects on efficiency and bias when using portfolios as opposed to individual stocks in Section 3. Section 4 compares the performance of portfolios versus stocks in the CRSP

[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.

database. Finally, Section 5 concludes.

# 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 $\sigma_F^2$. We specify the shocks $\varepsilon_{it}$ 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, $\alpha$ is a scalar, $1$ is a $N \times 1$ vector of ones, $\beta = (\beta_1 \ \dots \ \beta_N)'$ is an $N \times 1$ vector of betas, $\Omega_\varepsilon$ is an $N \times N$ invertible covariance matrix, and $\varepsilon_t$ 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 $\alpha$ and $\lambda$ in equations (1)-(3). If the

[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.

zero-beta expected return equals the risk free rate,

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

A rejection of $H_0^{\alpha=0}$ means that the factor cannot explain the average level of stock returns in excess of the riskfree rate. This is often the case for factors based on consumption-based asset pricing models because of the Mehra and 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. Well-known 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, 2001), 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 Appendix A, we derive the statistical properties of the estimators of $\alpha$, $\lambda$, and $\beta_i$ 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 Appendix B. 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 log-likelihood of equation (3) is given by

$$L = -\sum_t (R_t - \alpha - \beta(F_t + \lambda))' \Omega_\varepsilon^{-1} (R_t - \alpha - \beta(F_t + \lambda)) \tag{6}$$

ignoring the constant and the determinant of the covariance terms. For notational simplicity, we assume that $\sigma_F$ and $\Omega_\varepsilon$ 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, $\beta$, must be estimated and not taking the estimation error into account results in incorrect standard errors of the estimates of $\alpha$ and $\lambda$. 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 Appendix A, we state the maximum likelihood estimators, $\hat{\Theta}$.

[5] Consistent estimators are given by the sample formulas

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

## 2.3 Standard Errors

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

$$
\begin{aligned}
\mathrm{var}(\hat{\alpha}) &= \frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\beta'\Omega_\varepsilon^{-1}\beta}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2} && (7)\\
\mathrm{var}(\hat{\lambda}) &= \frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{1'\Omega_\varepsilon^{-1}1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2} && (8)\\
\mathrm{var}(\hat{\beta}) &= \frac{1}{T}\frac{1}{\lambda^2+\sigma_F^2}\\
&\times \left[\Omega_\varepsilon+\frac{\lambda^2}{\sigma_F^2}\frac{(\beta'\Omega_\varepsilon^{-1}\beta)11'-(1'\Omega_\varepsilon^{-1}\beta)\beta 1'-(1'\Omega_\varepsilon^{-1}\beta)1\beta'+(1'\Omega_\varepsilon^{-1}1)\beta\beta'}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2}\right]. && (9)
\end{aligned}
$$

We provide a full derivation in Appendix A.

To obtain some intuition, consider the case where idiosyncratic risk is uncorrelated across stocks so $\Omega_\varepsilon$ 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 and the summations are across $N$ stocks:

$$
\begin{aligned}
\mathrm{E}_c(\beta/\sigma^2) &= \frac{1}{N}\sum_j \frac{\beta_j}{\sigma_j^2}\\
\mathrm{E}_c(\beta^2/\sigma^2) &= \frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2}\\
\mathrm{E}_c(1/\sigma^2) &= \frac{1}{N}\sum_j \frac{1}{\sigma_j^2}\\
\mathrm{var}_c(\beta/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^4}\right)-\left(\frac{1}{N}\sum_j \frac{\beta_j}{\sigma_j^2}\right)^2\\
\mathrm{cov}_c(\beta^2/\sigma^2, 1/\sigma^2) &= \left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^4}\right)-\left(\frac{1}{N}\sum_j \frac{\beta_j^2}{\sigma_j^2}\right)\left(\frac{1}{N}\sum_j \frac{1}{\sigma_j^2}\right). && (10)
\end{aligned}
$$

In the case of uncorrelated idiosyncratic risk across stocks, the standard errors of $\hat{\alpha}$, $\hat{\lambda}$, and

$\hat{\beta}_i$ in equations (7)-(9) simplify to

$$\begin{aligned}
\mathrm{var}(\hat{\alpha}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\mathrm{E}_c(\beta^2/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)} && (11)\\
\mathrm{var}(\hat{\lambda}) &= \frac{1}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\mathrm{E}_c(1/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)} && (12)\\
\mathrm{var}(\hat{\beta}_i) &= \frac{1}{T}\frac{\sigma_i^2}{(\sigma_F^2+\lambda^2)}\left(1+\frac{\lambda^2}{N\sigma_i^2\sigma_F^2}\frac{\mathrm{E}_c(\beta^2/\sigma^2)-2\beta_i\mathrm{E}_c(\beta/\sigma^2)+\beta_i^2\mathrm{E}_c(1/\sigma^2)}{\mathrm{var}_c(\beta/\sigma^2)-\mathrm{cov}_c(\beta^2/\sigma^2,1/\sigma^2)}\right). && (13)
\end{aligned}$$

**Comment 2.1** *The standard errors of $\hat{\alpha}$ and $\hat{\lambda}$ depend on the cross-sectional distributions of betas and idiosyncratic volatility (cross-sectional means, variances and covariances as in equations (11)-(13).*

In equations (11) and (12), 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 $\beta_i$ and $\lambda$ 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 $\alpha$ and $\lambda$ but primarily only time-series data is useful for estimating $\beta_i$.*

In equations (11) and (12), 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 $\alpha$ and $\lambda$. Thus, the standard errors follow the same convergence properties as a pooled regression with IID time-series observations, as noted by Cochrane (2005). In contrast, the variance of $\hat{\beta}_i$ in equation (13) 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, $\sigma_i^2$, 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 (13) 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 $\alpha$ and $\lambda$ 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 $\hat{\lambda}$.*

Appendix A shows that the term $(\sigma_F^2 + \lambda^2)/\sigma_F^2$ in equations (11) and (12) 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.

## 2.4 Unknown factor mean

In this paper, we consider the case where the factor is known to have zero mean. But more generally, we could consider the model of $N \times 1$ returns in vector notation

$$R_t = \alpha 1 + \beta\lambda + \beta(\tilde{F}_t - \mu) + \Omega_\varepsilon^{1/2}\varepsilon_t \tag{14}$$

[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 $\hat{\lambda} \to \lambda$.

where $\mu = \mathrm{E}(\tilde{F}_t)$ and the factor shocks $F_t \equiv (\tilde{F}_t - \mu)$ are mean zero. Let $\tilde{\lambda} = \lambda - \mu$. Then, we can write equation (14) as

$$R_t = \alpha 1 + \beta\tilde{\lambda} + \beta\tilde{F}_t + \Omega_\varepsilon^{1/2}\varepsilon_t. \tag{15}$$

This has exactly the same likelihood as equation (3) except replacing $\tilde{\lambda}$ and $\tilde{F}_t$ for $\lambda$ and $F_t$, respectively. Hence, the standard errors for the estimators $\hat{\alpha}$ and $\hat{\tilde{\lambda}}$ are identical to equations (7) and (8), respectively, except that we replace $\lambda$ with $\tilde{\lambda}$ in the latter case. If the factors are not traded, we cannot identify the expected risk premiums without separately estimating $\mu$.

In models where the factor is tradeable, we can test if the mean of the factor is equal to the expected risk premium from the cross section:

$$H_0^{\tilde{\lambda}=0}: \quad \tilde{\lambda} = 0. \tag{16}$$

The efficient 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 hypothesis $H_0^{\tilde{\lambda}=0}$ does not require $\mu$ to be separately estimated.

If a traded 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 $\lambda$. Put another way, the estimated cross-sectional risk premium, $\lambda$, on a traded factor is not the same as the mean return, $\mu$, on the factor portfolio.

If the factor is not traded, it is necessary to use the time-series mean of a traded set of factors to identify $\mu$. This is the approach of Shanken (1992). In this approach, we work with the following log likelihood (neglecting a constant):

$$L = -\sum_t (R_t - \alpha - \beta(\tilde{F}_t - \mu + \lambda))'\Omega_\varepsilon^{-1}(R_t - \alpha - \beta(\tilde{F}_t - \mu + \lambda)) + \sum_t \frac{1}{2\sigma_F^2}(\tilde{F}_t - \mu)^2, \tag{17}$$

There are two differences between equation (17) and the factor model in equation (6). First, $\lambda$ and $\mu$ are now treated as separate parameters. Second, we identify $\mu$ by including $F_t$ as another asset where $\alpha = 0$ and $\beta = 1$, and $\mu$ is estimated by the time-series mean of $F_t$.

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

$$\begin{aligned} \mathrm{var}(\hat{\alpha}) &= \frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\beta'\Omega_\varepsilon^{-1}\beta}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2} \\ \mathrm{var}(\hat{\lambda}) &= \frac{\sigma_F^2}{T}+\frac{1}{T}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{1'\Omega_\varepsilon^{-1}1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta)-(1'\Omega_\varepsilon^{-1}\beta)^2}. \end{aligned} \tag{18}$$

These are the maximum likelihood standard errors derived by Shanken (1992) when including both a cross-sectional risk premium, $\lambda$, and a time-series mean of the factors, $\mu$. We observe that var$(\hat{\alpha})$ is identical to equation (7), but var$(\hat{\lambda})$ differs from equation (8) by an additive term, $\frac{1}{T}\sigma_F^2$. This term shows how the variance of $\hat{\lambda}$ is affected by the estimation of $\mu$, as pointed out by Jagannathan and Wang (2002).

Finally, consider the likelihood function of the system with $\tilde{\lambda}$ augmented by the non-zero mean $\tilde{F}_t$ in equation (17). For the parameter vector $\theta$, the information matrix is given by

$$\left(\mathrm{E}\left[-\frac{\partial^2 L}{\partial\Theta\partial\Theta'}\right]\right)^{-1} = \frac{1}{T}\begin{pmatrix} 1'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}\beta & 1'\Omega_\varepsilon^{-1}(\tilde{\lambda}+\mu) & 0 \\ \beta'\Omega_\varepsilon^{-1}1 & \beta'\Omega_\varepsilon^{-1}\beta & \beta'\Omega_\varepsilon^{-1}(\tilde{\lambda}+\mu) & 0 \\ (\tilde{\lambda}+\mu)'\Omega_\varepsilon^{-1}1 & (\tilde{\lambda}+\mu)'\Omega_\varepsilon^{-1}\beta & ((\tilde{\lambda}+\mu)^2+\sigma_F^2)\Omega_\varepsilon^{-1} & 0 \\ 0 & 0 & 0 & \sigma_F^2 \end{pmatrix}^{-1}. \tag{19}$$

This explicitly shows that $\hat{\mu}$ is uncorrelated with $\hat{\tilde{\lambda}}$ and since $\tilde{\lambda}+\mu=\lambda$, the standard errors for the system with $\lambda$ and this system with $\tilde{\lambda}$ are identical.

## 2.5 Portfolios and Factor Loadings

From the properties of maximum likelihood, the estimators using all stocks are most efficient with standard errors given by equations (11)-(13). If we use only $P$ portfolios as test assets, what is the efficiency loss? Let the portfolio weights be $\phi_{pi}$, where $p = 1, \ldots, P$ and $i = 1, \ldots, N$.

The returns for portfolio $p$ are given by

$$R_{pt} = \alpha + \beta_p \lambda + \beta_p F_t + \sigma_p \varepsilon_{pt}, \tag{20}$$

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

$$\begin{aligned} \beta_p &= \sum_i \phi_{pi} \beta_i \\ \sigma_p &= \left( \sum_i \phi_{pi}^2 \sigma_i^2 \right)^{1/2} \end{aligned} \tag{21}$$

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.

## 2.6 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,

$$\hat{\beta}_i = \hat{\beta}_p \qquad \forall\, i \in p. \tag{22}$$

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 C 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.

## 2.7 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. 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 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 (7)-(9) simplify to

$$\begin{aligned}
\mathrm{var}(\hat{\alpha}) &= \frac{\sigma^2}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{\mathrm{E}_c(\beta^2)}{\mathrm{var}_c(\beta)} \\
\mathrm{var}(\hat{\lambda}) &= \frac{\sigma^2}{NT}\frac{\sigma_F^2+\lambda^2}{\sigma_F^2}\frac{1}{\mathrm{var}_c(\beta)} \\
\mathrm{var}(\hat{\beta}_i) &= \frac{1}{T}\frac{\sigma^2}{(\sigma_F^2+\lambda^2)}\left(1+\frac{\lambda^2}{N\sigma^2\sigma_F^2}\frac{\mathrm{E}_c(\beta^2)-2\beta_i\mathrm{E}_c(\beta)+\beta_i^2}{\mathrm{var}_c(\beta)}\right).
\end{aligned} \tag{23}$$

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. The cross-sectional moments for equation (23) when using $P$ portfolios are:

$$
\begin{aligned}
\mathrm{E}_c[\beta_p] &= \mu_\beta \\
\mathrm{E}_c[\beta_p^2] &= \frac{1}{P}\sum_{p=1}^{P}\left(\mu_\beta + \frac{P\sigma_\beta}{\sqrt{2\pi}}\left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}}\right)\right)^2 \\
&= \mu_\beta^2 + P\frac{\sigma_\beta^2}{2\pi}\sum_{p=1}^{P}\left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}}\right)^2 \\
\mathrm{var}_c[\beta_p] &= P\frac{\sigma_\beta^2}{2\pi}\sum_{p=1}^{P}\left(e^{-\frac{\delta_{p-1}^2}{2}} - e^{-\frac{\delta_p^2}{2}}\right)^2. \qquad (24)
\end{aligned}
$$

as derived in Appendix D. We refer to the variance of $\hat{\alpha}$ and $\hat{\lambda}$ computed using $P$ portfolios as $\mathrm{var}_p(\hat{\alpha})$ and $\mathrm{var}_p(\hat{\lambda})$, respectively, and the variance of the portfolio beta, $\beta_p$, as $\mathrm{var}(\hat{\beta}_p)$.

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 (13) 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, $\mathrm{var}(\hat{\beta}_p)$ 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 individual stocks. 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 to any efficiency gain for estimating the risk premium. Note that the term $\sigma^2/N$ using all individual stocks in equation (23) remains the same using $P$ portfolios since each portfolio contains equal mass $1/P$ of the stocks:

$$\frac{\sigma_p^2}{P} = \frac{(\sigma^2 P/N)}{P} = \frac{\sigma^2}{N}. \tag{25}$$

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 (23) involve the total amount of idiosyncratic volatility diversified by all stocks and forming portfolios does not change the total composition. Equation (25) 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 dispersion. In the bottom plot, the $P = 5$ portfolio case clearly shows that the cross-sectional variance of betas has increased tremendously. It is this increased 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.[7] We show in Appendix E 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 $\alpha$ and $\lambda$ 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 be lost. Thus, we expect the efficiency losses of creating portfolios to be largest when the distribution of betas is very dispersed.

# 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 tradeoff between a long enough sample period for estimation but over which an average true (not estimated) stock beta is unlikely to change drastically (and is a standard practice going back to Blume (1970)). 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 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

[7] 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). 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, $H_0^{\lambda=\mu}$.

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 estimated betas and 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 $\sigma$ 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

$$\frac{\text{var}_p(\hat{\alpha})}{\text{var}(\hat{\alpha})} \quad \text{and} \quad \frac{\text{var}_p(\hat{\lambda})}{\text{var}(\hat{\lambda})}, \tag{26}$$

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 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 allow for cross-sectionally correlated residuals. Fourthly, we allow entry and exit of firms in the cross section. Finally, we consider two different cases in which returns are affected by a characteristic in addition to the factor and the portfolios are sorted on this characteristic. In one case, we sort stocks into portfolios on a characteristic that is correlated with beta, and in the other case we sort on a characteristic that is uncorrelated with beta. 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:

$$R_{it} = \beta_i \lambda + \beta_i F_t + \varepsilon_{it}, \tag{27}$$

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):

$$\begin{pmatrix} \beta_i \\ \ln \sigma_i \end{pmatrix} \sim N\left( \begin{pmatrix} 1.14 \\ -2.09 \end{pmatrix}, \begin{pmatrix} 0.41 & 0.13 \\ 0.13 & 0.28 \end{pmatrix} \right), \tag{28}$$

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, except that the cross-sectional variance of betas is adjusted by subtracting off the estimated noise variance from the time series regressions. 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 (11) and (12).

We simulate small samples of size $T = 60$ months with $N = 5000$ stocks. In each simulation, we compute the variance ratios in equation (26) using portfolios relative to using all stocks. 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 (11) and (12). 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 next 3 panels in Table 2 form portfolios based on estimated betas. In Panel B, we form portfolios on estimated betas with the same data-generating process as Panel A. However, to ensure that the portfolios are actual tradeable portfolios, we generate a pre-sample of 60 months of data for each stock that we use solely to estimate ex-ante betas for sorting into portfolios. Forming portfolios on estimated betas, the efficiency losses increase. For $P = 25$ portfolios the mean variance ratio $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda})$ is 5.4 in Panel B compared to 2.7 in Panel A when portfolios are formed on the true betas. For $P = 500$ portfolios formed on estimated betas, the mean variance ratio for

$\hat{\lambda}$ is still 4.1. Thus, the efficiency losses increase considerably once portfolios are formed on estimated betas.

### 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 equations (27)-(28) by assuming

$$\varepsilon_{it} = \xi_i u_t + \sigma_{vi} v_{it}, \tag{29}$$

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, $\xi_i$, on the common residual shock, $u_t$.

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

$$\begin{pmatrix} \beta_i \\ \xi_i \\ \ln \sigma_{vi} \end{pmatrix} \sim N \left( \begin{pmatrix} 1.14 \\ 1.01 \\ -2.09 \end{pmatrix}, \begin{pmatrix} 0.41 & 0.22 & 0.13 \\ 0.22 & 1.50 & 0.36 \\ 0.13 & 0.36 & 0.28 \end{pmatrix} \right), \tag{30}$$

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 $\beta$ and $\xi$ is 0.24) and higher idiosyncratic volatility (the correlation of $\beta$ with $\ln \sigma_{vi}$ is 0.33). As in panel B, a pre-sample of 60 months data were used for estimating the betas for the purpose of sorting stocks into tradeable portfolios.

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 18 for $\text{var}_p(\hat{\alpha})/\text{var}(\hat{\alpha})$ and 20 for $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda})$ for $P = 25$ portfolios. Thus, cross-sectional correlation worsens the efficiency losses from using portfolios. Greater estimation error of the betas leads to less precise portfolio assignment of the individual assets. To the extent that the residuals are correlated with one another, the within-portfolio benefit of offsetting errors is reduced. Intuitively, one can think about the extreme case of perfectly correlated residuals; there will be no reduction in noise from grouping assets and yet there will be a compressed distribution of observations from which to draw inference. With correlated residuals, the efficiency gains in estimating beta loadings are more quickly offset by the loss of information about the true distribution of betas.

### 3.2.3 Entry and Exit of Individual Firms

We also compute efficiency losses using stocks or portfolios with entry and exit of individual firms, giving a stochastic number of firms in the cross-section. We consider a log-logistic survivor function for a firm surviving to month $T$ at time $t$ after listing given by

$$Pr(T > t) = \left[1 + (0.0323t)^{1.2658}\right]^{-1}, \tag{31}$$

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 (27) and as a firm is born, its beta, common residual loading, and idiosyncratic volatility are drawn from equation (30). In this simulation, the firms are sorted into portfolios based on their estimated betas over the sample, rather than an additional pre-sample, because of the structure of entry and exit. This means that in panel D, the portfolios are ex-post portfolios, unlike the tradeable ex-ante portfolios in panels B and C.

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, $\text{var}_p(\hat{\lambda})/\text{var}(\hat{\lambda}) = 25$ compared to 20 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 Sorting on a characteristic

Consider the model in equation (27) and let $c_i$ be a characteristic of each stock such that $c_i \sim N(0, \sigma_c^2)$ where $c_i$ and $\beta_i$ have correlation $\rho_{c\beta}$. We assume that the econometrician can observe $c_i$ and think of this as representing a characteristic of the individual stock, such as size. We then form portfolios sorting on the characteristic $c_i$. As the characteristic is observable and constant over time for each stock, these are clearly tradeable portfolios. So as not to overstate the cross-sectional variation of the estimated betas, we subtract the variance of the time-series betas estimation from the cross-sectional beta variance.

Panel E of Table 2 reports results when $\rho_{c\beta} = 0.5$, indicating moderate correlation between the characteristic and beta. Panel F of Table 2 forms $P$ portfolios ranking on $c_i$ when $\rho_{c\beta} = 0$. In both cases, there is a large efficiency loss from forming portfolios that can be avoided by using individual stocks. When $\rho_{c\beta} = 0$ the efficiency losses from forming portfolios are enormous. This is because there is a relationship between betas and returns for individual stocks, but the portfolios are sorted on the characteristic, and the characteristic is uncorrelated with betas. This destroys almost all the variation in betas across portfolios and renders the parameters virtually unidentified.

#### 3.2.5 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, when the number of firms in the cross section changes over time and when portfolios are formed based on characteristics.

### 3.3 Bias-Efficiency Tradeoff

Studies may decide on using stocks versus portfolios based on efficiency, the focus in this paper, but the potential for bias in the point estimate of factor risk premia should also be considered. For the progressively richer environments considered in the simulations in the previous subsection, we also compute bias and report the results in Table 3.

Grouping stocks into portfolios comes at the expense of having fewer test assets in the cross section to estimate risk premia, and thus less efficient estimates. However, when stocks are not aggregated into portfolios, the beta estimates are more subject to measurement error, which could bias the risk premium point estimate. The problems of bias can be large. Several researchers have examined the nature of bias effects and proposed methods to mitigate them (Litzenberger and Ramaswamay (1979), Gibbons (1982), Shanken (1992), Kim (1995), Shanken and Zhou (2007), Jegadeesh et al. (2017)). Kim (1995) constructs a bias correction with a maximum likelihood interpretation, and shows that once the measurement error in the beta estimates is accounted for, factor risk premium estimation is $N$-consistent. Our analytical framework shows that using all test assets available in the cross section gives the most efficient estimates of risk premia. Thus, estimation methods that account for bias present the clearest case for using individual stocks rather than grouping stocks into a smaller number of test assets.

In practice, however, many researchers use risk premium estimation methods where the choice of test asset has a large influence on bias in the point estimate. To provide some guidance on the potential bias-efficiency tradeoff in the presence of EIV effects, we consider the standard two-step Fama MacBeth estimation. In the first stage of the Fama MacBeth procedure, betas can be estimated for each stock separately. Then, the second-stage regressors are formed from the first-stage beta estimates. It is the measurement accuracy of the beta estimate and the cross-sectional dispersion in the true factor loading that matter for bias due to classical measurement error.

### 3.3.1 Bias in the Point Estimate

The potential reduction in measurement error of the beta estimates has motivated the practice of grouping stocks into portfolios as test assets. The more stocks that are grouped to form a portfolio, the more idiosyncratic risk can be diversified away in the first stage beta estimates. Consistent with this, the bias in estimates of $\alpha$ and $\lambda$ across the various empirical settings in Table 3 generally becomes less severe as the number of portfolios declines (as the number of stocks in each portfolio becomes larger).

The attenuation bias in classical measurement error scales the true slope coefficient by a factor:

$$\frac{\sigma_*^2}{\sigma_*^2 + \sigma_v^2} \tag{32}$$

where $\sigma_*^2$ denotes the variance of the true regressor and $\sigma_v^2$ denotes the variance of the measurement error. This bias worsens as the variance of the measurement error in the first-stage betas increases and also as the variance of the true second-stage regressor shrinks, two potential countervailing effects to the diversification benefit of forming portfolios.

In comparing Panels A and B of Table 3, the only difference is the variable on which stocks are sorted into portfolios. By sorting stocks on estimated rather than true betas, there is a less diverse set of portfolios, giving less variation across true factor loadings. Panel B shows that bias is still less severe when estimating risk premia with portfolios as compared to individual stocks, but less so than in Panel A.

Panel C shows bias in the case of cross-correlated residuals, which directly dampens the benefit of offsetting errors with portfolio formation. Panel D allows for entry and exit of firms, which results in less well-measured first stage beta estimates, giving higher variance in the measurement error and inducing more error in sorting stocks into portfolios. Similar to the case in Panel B when sorting stocks on estimated betas, variance in true factor loadings declines as portfolios become less dissimilar. The bias reduction benefit to forming portfolios still outweighs the cost in Panels C and D, but to a lesser extent than in the baseline empirical setting shown in

Panels A and B.

If stocks are sorted into portfolios in a way that sufficiently shrinks variation across true portfolio factor loadings, it can fully offset the diversification benefit from forming portfolios. An extreme example is to sort stocks into portfolios at random. The true portfolio factor loadings will not be meaningfully differentiable, greatly worsening attenuation bias. Panel E shows the case of stocks sorted into portfolios based on a characteristic that is correlated with the factor loading. The bias in the $\alpha$ and $\lambda$ estimates is least pronounced with the smallest number of portfolios (portfolios that contain the greatest number of stocks), but forming a large number of portfolios (with few stocks in each) shows more bias than using stocks individually. The case shown in Panel F sorts on a characteristic that is uncorrelated with betas, generating no diversification benefit from forming portfolios. The bias in the $\alpha$ and $\lambda$ estimates is worse when using any number of portfolios as compared to using individual stocks.

### 3.3.2 Comparing Bias and Eficiency Effects

One way to measure the relative importance of efficiency as compared to bias on risk premia estimates is to consider the mean squared error (MSE), which incorporates both effects. Table 4 reports the MSE for each simulation and number of portfolios. We see that the minimum MSE is achieved in some of our empirical environments by using a large number of portfolios (250 to 500) and in other environments by using individual stocks as test assets.

The effect of bias is very small when just a few portfolios are formed (when each portfolio has many stocks), contributing little to the MSE. Yet, the efficiency loss from using very few portfolios as test assets and thus reducing cross-sectional variation for the second-stage estimation means that forming fewer than around 250 portfolios, each containing roughly 20 stocks, is not optimal in a MSE sense in any of our empirical settings. We emphasize that we do not aim to propose an optimal portfolio size or an optimal number of portfolios. Rather, we encourage the researcher to consider the effect of realistic data features on the bias-efficiency tradeoff when faced with choices in estimation and interpretation of results.

Sorting stocks into portfolios based on true betas, as in Panel A of Tables 2 through 4, preserves the most cross-sectional variation possible in portfolio estimates. In this case, the efficiency loss generated by the positive cross-sectional correlation between idiosyncratic volatility and the betas is gradual as the number of portfolios declines. The results for portfolios sorted on estimated betas show a steeper decline in efficiency in the transition from using individual stocks to forming portfolios containing very few stocks, but the outcome is similar to Panel A. Estimation in the presence of cross-correlated residuals (Panel C) or firm entry and exit over time (Panel D) further reduces the cross-sectional variation in portfolio betas. The severity of the cross-sectional information loss increases sharply as fewer portfolios are formed (the number of stocks grouped into each portfolio grows). There is relatively little efficiency loss when using a very large number of portfolios. Meanwhile, bias with a large number of portfolios is close to the size of bias with individual stocks. The combined effect achieves the lowest MSE with a very large number of portfolios (500) in Panels C and D.

Panels E and F show results for portfolios that are sorted on characteristics rather than the first-stage beta estimates. In Panel E, portfolios are formed on a characteristic that is correlated with the first-stage estimates, and so the net effect on the MSE is similar to a case of sorting on very noisy beta estimates. Panel F shows the more extreme case where the characteristic is entirely uncorrelated with the beta. The complete lack of bias reduction from forming portfolios means that MSE is minimized by using individual stocks as test assets.

#### 3.3.3 Summary

The accuracy with which the individual assets are assigned to portfolios matters to the bias-variance tradeoff. Mis-grouping individual stock beta estimates can shrink the cross-sectional dispersion in test assets when using portfolios, without an offsetting reduction in measurement error of the estimates. The particular choice of variable on which portfolios are sorted can adversely affect risk premia variance and bias. Since using individual stocks does not require sorting, this source of negative influence on bias and variance of risk premia estimates only affects estimation with portfolios.

In each empirical setting we consider (Table 2, Panels A through F), efficiency gains in risk premium estimates increase as the number of portfolios approaches the number of individual stocks. Diversification across estimation errors in the cross section can lessen measurement error in beta estimates and thus reduce bias, and the effect is larger when portfolios contain a greater number of stocks. But, grouping does not necessarily mean less bias. The importance of various data features (e.g. cross-sectional correlation among errors) can lessen the diversification benefit from grouping stocks into portfolios. Poorly constructed portfolios lessen variation across true portfolio factor loadings, and this could potentially worsen bias in the risk premium point estimate relative to the case of using individual stocks. Variance and bias of estimates may be adversely affected by the particular choice of stock grouping for portfolios.

# 4 Empirical Analysis

We now investigate the differences in using portfolios versus individual stocks in the data with actual historical stock returns. First, we use the past three to five years of monthly returns to estimate ex-ante betas for each stock in each year, (e.g. the 1970 ex-ante beta is formed with returns from January 1966 to December 1970). We rank the stocks into portfolios based on these ex-ante betas in December of each year. Using these portfolio groupings, we calculate the rolling 12-month portfolio return (e.g. portfolio returns from January 1971 to December 1971 are calculated for portfolios formed in December 1970).

We use the rolling portfolio returns to estimate the contemporaneous portfolio betas, in 5-year non-overlapping windows. The first 5-year beta is formed with returns from 1971 to 1975. We then relate these contemporaneously estimated portfolio betas to same-sample returns (1971 to 1975 in the first 5-year period) in order to form portfolio factor risk premia estimates. Using betas estimated from 1971 to 2015, we compare the portfolio factor risk premia to those estimated with individual stocks as test assets. In a balanced panel, the portfolio betas would be equal to the weighted-average of the individual stock beta estimates, but the sample period returns are not a balanced panel.

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 betas 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, 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 portfolios. 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 firm excess returns and estimated firm betas. Thus, we stack all stocks' excess returns from each five-year period into one panel and run a regression using average 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 5 reports the estimates and standard errors of $\alpha$ and $\lambda$ 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 returns and estimated betas 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 that portfolios are often used as test assets instead of stocks. In this section, we investigate empirically the potential impact of this specification difference on the size of the $\hat{\alpha}$ and $\hat{\lambda}$ t-statistics.

#### 4.1.2 Using Portfolios

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 portfolio size). 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.

To form 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 portfolio returns. We rebalance the portfolios annually, weighting stocks equally within each portfolio. Then, we compute the first-pass OLS market betas of each portfolio, in each non-overlapping five-year period. These portfolio betas are the factor loadings for the portfolios. Finally, to estimate the portfolio $\alpha$ and $\lambda$ in Panel B of Table 5 we run a second-pass cross-sectional regression of excess returns onto the betas. Thus, we examine the same beta–realized return relation in the case of all stocks and portfolios, in Panels A and B, respectively, over the same sample.

In the last four columns of Table 5, we report statistics of the cross-sectional dispersion of the betas 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 different test assets. With $P = 5$ portfolios, 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. The severe shrinkage in the beta distribution means that the portfolios miss substantial information in the tails; the 5%- and 95%-tiles for $P = 5$ portfolios are 0.62 and 1.64, respectively, compared to 0.12 and 2.44 for all stocks.

The cross-section of the estimated betas relate to returns in the second-stage estimation of risk premia in tests of factor models. The truncated distribution of the portfolio betas produces much larger standard errors in the cross-sectional estimation of $\lambda$ than using the full stock universe. For all portfolio sizes, the portfolio standard errors exceed those of the individual stocks standard errors, for both GMM and MLE. The portfolios fail to reject $H_0^{\lambda=0}$, except for MLE standard errors for $P = 50$, which is significant at the 10% level, in contrast to the overwhelming rejection when using all stocks. This underscores the importance of the information in the beta distribution, which is entirely preserved using all stocks.

Panel B also shows that the estimates of $\alpha$ and $\lambda$ from portfolios are quite dissimilar to the estimates in Panel A. Using portfolios as test assets produces an estimate of $\alpha$ around 14% and an estimate of $\lambda$ around 1-2%. In contrast, all stocks (Panel A) produce alpha estimates around 8% and estimates of $\lambda$ around 4-5%. This marked difference in $\hat{\lambda}$ is not driven by individual stocks having greater attenuation bias due to classical measurement error in the beta estimates because that would go the other way. Rather, it indicates that portfolios may provide noisier point estimates over this sample.

Figure 3 plots the evolution of $\hat{\lambda}$ as the number of portfolios grows larger (and the number of stocks in each portfolio decreases) with two-standard error bounds around the point estimate. There is a huge variance in the distribution from which portfolio estimates are drawn, especially when only a small number of portfolios are formed. The portfolio $\lambda$ point estimate is within error bounds of the $\lambda$ point estimate for individual stocks only when $P > 500$ portfolios. However, the $\lambda$ point estimate from individual stocks is within the error bounds of the portfolio $\lambda$ point estimate for all portfolio sizes. The value of $\lambda$ is estimated with much greater certainty when individual stocks are used as test assets.

With the smallest number of portfolios, $\hat{\lambda}$ is most divergent from $\hat{\lambda}$ with individual stocks, but the divergence drops as the number of portfolios is increased. In addition to truncating the tails of the beta distribution, grouping many stocks into a small number of portfolios has large implications for the higher moments of the distribution. Skewness for the beta distribution when using $P = 10$ is 0.85, greater than 0.70 for individual stocks. Excess kurtosis for $P = 10$ is 2.27, about half of that for individual stocks. 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. We require stocks to have at least 3 years of returns to be included in the analysis, but the firm/year panel is still far from balanced.

#### 4.1.3 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 6 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. In contrast, the portfolio estimates all reject $H_0^{\lambda=\mu}$, at least at the 10% level, with either maximum likelihood or GMM standard errors.

### 4.1.4 Summary

We overwhelmingly reject $H_0^{\alpha=0}$ and hence the one-factor model using all stocks or portfolios. For all stocks, we also reject $H_0^{\lambda=0}$, thus finding the market factor priced. Using all stocks we estimate $\hat{\lambda} = 4.79\%$. But using portfolios can produce quite different point estimates of cross-sectional risk premia. The $\hat{\lambda}$ produced by between 5 and 50 portfolios range from only $1.14\%$ to $1.73\%$. Further, the loss of information in the cross section of 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 $H_0^{\lambda=\mu}$, portfolios reject the hypothesis at the 10 percent level, while results are mixed using individual stocks.

## 4.2 Fama-French (1993) Model

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

$$R_{it} = \alpha + \beta_{MKT,i}\lambda_{MKT} + \beta_{SMB,i}\lambda_{SMB} + \beta_{HML,i}\lambda_{HML} + \sigma_i\varepsilon_{it}, \tag{33}$$

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 $\alpha$, $\lambda_{MKT}$, $\lambda_{SMB}$, and $\lambda_{HML}$.

### 4.2.1 Factor Loadings

We now compare the Fama French model factor loadings of all stocks to those of portfolios. We form the portfolios by using the same procedures described in subsection 4.1.2. 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 7 reports summary statistics for the distribution of the betas $\hat{\beta}_{MKT}$, $\hat{\beta}_{SMB}$, 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. The market betas are centered around one after controlling for $SMB$ and $HML$, and the $SMB$ and $HML$ betas 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 beta estimates. Table 7 shows three important effects on the distribution of betas that result from portfolio formation, similar to those found for the one-factor model in Section 4.1.

First, forming portfolios severely reduces the cross-sectional variance in the betas. For example, the $\hat{\beta}_{SMB}$ and $\hat{\beta}_{HML}$ cross-sectional standard deviation is 1.21 for all stocks, but it is cut by more than one-half to 0.37 and 0.29, respectively, for the $2 \times 2 \times 2$ portfolios.

Second, forming portfolios truncates the tails of the beta distribution. 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 $3 \times 3 \times 3$ portfolio. Such a difference in the distribution of betas for portfolios could produce quite different cross-sectional factor risk premia estimates.

Finally, 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 portfolios of different sizes.

#### 4.2.2 Cross-Sectional Factor Risk Premia

Table 8 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 5), 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 portfolios in Panel B yield very different estimates of factor risk premia in comparison to all stocks. Notably, the 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 9. For both individual stocks and portfolios 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 we also reject $H_0^{\lambda=\mu}$ for HML. Using 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 individual stocks. 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.

# 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. This is because grouping stocks into portfolios also diversifies away information contained in individual stock factor loadings. 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 when using portfolios. 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. Moreover, using portfolios creates an additional layer of potential for data-mining biases. There is still a bias-variance tradeoff to consider in deciding between using portfolios and individual stocks if the standard two-pass methodology is employed, but minimizing mean square error calls for using a very large number of portfolios, or all stocks. Furthermore, the bias motivation for portfolios is unclear once full-blown maximum likelihood estimation, or some other bias-adjusted approach, is used. Thus, the use of portfolios in cross-sectional regressions ought to be carefully motivated.

# Appendix

## A Derivation of Maximum Likelihood Asymptotic Variances

We consider the case where the mean of the factors is known. The maximum likelihood estimators for $\alpha$, $\lambda$, and $\beta_i$ are given by:

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

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

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

The information matrix is given by

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

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

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

This gives the variance of $\hat{\alpha}$ and $\hat{\lambda}$ in equations (7) and (8).

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

$$\begin{aligned} D^{-1}B'Q^{-1}BD^{-1} &= \frac{\lambda^2}{\sigma_F^2(\lambda^2 + \sigma_F^2)}\frac{1}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2} \\ &\quad \times\Omega_\varepsilon\begin{pmatrix} \beta'\Omega_\varepsilon^{-1}\beta & -1'\Omega_\varepsilon^{-1}\beta \\ -\beta'\Omega_\varepsilon^{-1}1 & 1'\Omega_\varepsilon^{-1}1 \end{pmatrix}\begin{pmatrix} 1'\Omega_\varepsilon^{-1}\lambda \\ \beta'\Omega_\varepsilon^{-1}\lambda \end{pmatrix}\Omega_\varepsilon \\ &= \frac{\lambda^2}{\sigma_F^2(\lambda^2 + \sigma_F^2)}\frac{(\beta'\Omega_\varepsilon^{-1}\beta)11' - (1'\Omega_\varepsilon^{-1}\beta)\beta1' - (1'\Omega_\varepsilon^{-1}\beta)1\beta' + (1'\Omega_\varepsilon^{-1}1)\beta\beta'}{(1'\Omega_\varepsilon^{-1}1)(\beta'\Omega_\varepsilon^{-1}\beta) - (1'\Omega_\varepsilon^{-1}\beta)^2}. \end{aligned}$$

Thus,

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

This gives the variance of $\hat{\beta}_i$ in equation (9).

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:

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

# B Multiple Factors and GMM

We work with the data-generating process with potentially multiple factors:

$$R_t = \alpha + B\tilde{\lambda} + B\tilde{F}_t + \varepsilon_t, \quad \text{(B-1)}$$

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

$$\tilde{R}_t \equiv R_t - B\tilde{F}_t = X\gamma + \varepsilon_t, \quad \text{(B-2)}$$

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 $\hat{\gamma}_t$:

$$\hat{\gamma} = \frac{1}{T}\sum \gamma_t = (\hat{X}'W\hat{X})^{-1}\hat{X}'W\bar{\tilde{R}}, \quad \text{(B-3)}$$

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

$$\hat{B} = \left[\frac{1}{T}\sum(\tilde{R}_t - \bar{\tilde{R}})(\tilde{F}_t - \bar{\tilde{F}})'\right]\hat{\Sigma}_F^{-1}, \quad \text{(B-4)}$$

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

$$\begin{aligned}
\mathrm{E}[h_{1t}] &= \mathrm{E}[\tilde{R}_t - \mathrm{E}\tilde{R}_t] = 0 \quad (N\times 1) \\
\mathrm{E}[h_{2t}] &= \mathrm{E}\left[[(\tilde{F}_t - \mathrm{E}\tilde{F}_t)'\Sigma_F^{-1}\lambda]\varepsilon_t\right] = 0 \quad (N\times 1), \quad \text{(B-5)}
\end{aligned}$$

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 (2005) 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

$$\begin{aligned} \sqrt{T}\left(\frac{1}{T}\sum \hat{\gamma}_t - \gamma\right) &= \hat{D}\left[-\frac{1}{\sqrt{T}}\sum \varepsilon_t(\tilde{F}_t - \bar{\tilde{F}})'\hat{\Sigma}_F^{-1}\lambda + \frac{1}{\sqrt{T}}\sum(\tilde{R}_t - \mathrm{E}\tilde{R}_t)\right] \\ &\xrightarrow{d} D\begin{bmatrix} I_N & 0 \\ 0 & -I_N \end{bmatrix}\frac{1}{\sqrt{T}}\sum h_t \\ &\xrightarrow{d} N(0, \Sigma_\gamma), \end{aligned} \tag{B-6}$$

where

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda) D\Omega_\varepsilon D'. \tag{B-7}$$

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

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda)(X'X)^{-1}X'\Omega_\varepsilon X(X'X)^{-1}. \tag{B-8}$$

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

$$\Sigma_\gamma = (1 + \lambda'\Sigma_F^{-1}\lambda)(X'\Omega_\varepsilon^{-1}X)^{-1}, \tag{B-9}$$

which is the same as maximum likelihood. Equation (B-9) is the matrix counterpart of equations (7) and (8) in the main text for a single factor model. We use equation (B-9) to compute maximum likelihood standard errors for multiple factors.

It is instructive to note the difference with Shanken (1992). Consider the model

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

To derive the Shanken (1992) 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 (B-5) rather than the moment conditions $\mathrm{E}[h_{1t}^*]$. Both cases use the same Fama-MacBeth estimator in equation (B-3). With the following Central Limit Theorem for $h_t = (h_{1t}^*\, h_{2t}^*)$:

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

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 (1992) 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 (18).

# C 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 (22). The numerator of the Fama-MacBeth coefficient can be written as:

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

where the first to the second line follows because of equation (22). The denominator of the estimated risk premium is

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

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 $\beta_p$ 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.

# D Cross-Sectional Moments For Normally Distributed Betas

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

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

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 $\delta_p$ corresponding to the standard normal are

$$N(\delta_p) = \frac{p}{P}, \quad p = 1, ..., P-1, \tag{D-2}$$

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

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

The beta of portfolio $p$, $\beta_p$, is given by:

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

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

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

# E Factor Risk Premia and Characteristics

Consider the following cross-sectional regression:

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

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

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

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

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

where we define the cross-sectional moments

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

# 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 $\Xi$ is a $N \times 1$ vector of residual factor loadings, $\Sigma_v$ 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, $\Omega_\epsilon$, 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, $\xi_i$, we regress $e_{it}$ onto $u_t$ for each asset $i$. The fitted residuals are used to obtain estimates of $\sigma_{vi}^2$. 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 $\xi_i$ is a $10 \times 1$ vector of industry proportions, the $j$th 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 $\xi_i$ 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 $10 \times 1$ 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 $\Xi$ is $N \times 10$ and $\Sigma_v$ is a diagonal matrix containing $\{\sigma_{vi}^2\}$.

The industry residuals are specified to be uncorrelated with the factors $F_t$. To estimate $\Sigma_u$, 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 $\Sigma_u$ as the sample covariance matrix of $\{u_{jt}\}$.

To estimate $\Sigma_v$, 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 $\Sigma_v$ to be the sample covariance matrix of $\{\hat{v}_{it}\}$.

# References


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

[2] 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.

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

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

[5] 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.

[6] Blume, M. E., 1970, "Portfolio Theory: A Step Toward Its Practical Application," *Journal of Business*, 43, 152-73.

[7] 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.

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

[9] 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.

[10] Campbell, J. Y., A.W. Lo and A.C. MacKinlay, 1996, *The Econometrics of Financial Markets*, Princeton University Press, NJ.

[11] Cochrane, J. H., 2005, *Asset Pricing*, Princeton University Press, NJ.

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

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

[14] 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.

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

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

[17] Ferson, W.E. and T. Simin and S.Sarkissian, 2003, "The Alpha Factor Asset Pricing Model: A Parable," emph Journal of Financial Markets, 2, 49-6

[18] Ferson, W.E. and T. Simin and S.Sarkissian, 2003, "Spurious regressions in Financial Economics?," emph Journal of Finance, 58, 1393-1414.

[19] Gibbons, M.R., 1982, "Multivariate tests of financial models: A new approach", *Journal of Financial Economics*, 10, 3-27.

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

[21] 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.

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

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

[24] 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.

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

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

[27] Jegadeesh, N., J. Noh, K. Pukthuanthong, R. Roll, and J. Wang, 2017, "Empirical Tests of Asset Pricing Models with Individual Assets: Resolving the Errors-in-Variables Bias in Risk Premium Estimation," working paper, Emory University.

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

[29] Kan, R., 2004, "On the Explanatory power of Asset Pricing Models Across and Within Portfolios," working paper, University of Toronto.

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

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

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

[33] Kim, D., 1995, "The Errors in the variables problem in the cross-section of expected stock returns." *Journal of Finance*, 50, 1605-1634.

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

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

[36] 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.

[37] Lo, A., W., and A. C. MacKinlay, 1990, "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431-468.

[38] MacKinlay, A. C., 1987, "On Multivariate Tests of the CAPM," *Journal of Financial Economics*, 18, 341-372.

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

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

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

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

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

[44] 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.

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

[46] 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

| | $\alpha$ Efficiency Loss | | | | | | $\lambda$ Efficiency Loss | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Portfolios $P$ | 10 | 25 | 50 | 250 | 500 | 2,500 | 10 | 25 | 50 | 250 | 500 | 2,500 |
| **Panel A: Sorting on True Betas, Correlated Betas and Idiosyncratic Volatility** | | | | | | | | | | | | |
| Mean | 2.85 | 2.68 | 2.61 | 2.43 | 2.28 | 1.51 | 2.88 | 2.69 | 2.62 | 2.43 | 2.28 | 1.51 |
| StDev | 0.14 | 0.13 | 0.13 | 0.11 | 0.10 | 0.05 | 0.12 | 0.11 | 0.11 | 0.10 | 0.08 | 0.04 |
| **Panel B: Sorting on Estimated Betas, Correlated Betas and Idiosyncratic Volatility** | | | | | | | | | | | | |
| Mean | 6.16 | 6.00 | 5.84 | 5.08 | 4.40 | 1.94 | 5.54 | 5.39 | 5.25 | 4.62 | 4.05 | 1.89 |
| StDev | 0.60 | 0.56 | 0.53 | 0.44 | 0.37 | 0.10 | 0.51 | 0.48 | 0.45 | 0.38 | 0.32 | 0.09 |
| **Panel C: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals** | | | | | | | | | | | | |
| Mean | 22.4 | 18.3 | 15.0 | 7.7 | 5.5 | 2.0 | 25.8 | 20.3 | 15.9 | 7.6 | 5.4 | 2.0 |
| StDev | 15.8 | 12.2 | 9.3 | 3.1 | 1.5 | 0.2 | 21.2 | 15.1 | 10.2 | 2.5 | 1.1 | 0.9 |
| **Panel D: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals Entry and Exit of Firms** | | | | | | | | | | | | |
| Mean | 27.5 | 22.6 | 18.4 | 9.5 | 6.8 | 2.4 | 31.7 | 25.0 | 19.6 | 9.3 | 6.6 | 2.4 |
| StDev | 19.3 | 14.9 | 11.4 | 3.8 | 1.9 | 0.2 | 26.1 | 18.6 | 12.7 | 3.1 | 1.4 | 0.1 |
| **Panel E: Sorting on characteristics correlated with betas** | | | | | | | | | | | | |
| Mean | 13.2 | 12.5 | 12.1 | 10.0 | 8.2 | 2.6 | 12.4 | 11.7 | 11.4 | 9.4 | 7.7 | 2.5 |
| StDev | 1.0 | 0.9 | 0.9 | 0.7 | 0.5 | 0.1 | 0.8 | 0.8 | 0.8 | 0.6 | 0.5 | 0.1 |
| **Panel F: Sorting on characteristics uncorrelated with betas** | | | | | | | | | | | | |
| Mean | 2380.8 | 756.0 | 350.7 | 61.9 | 28.2 | 3.3 | 2169.0 | 689.2 | 319.9 | 56.7 | 26.1 | 3.2 |
| StDev | 1621.2 | 243.5 | 77.5 | 6.5 | 2.3 | 0.2 | 1427.8 | 221.3 | 70.2 | 5.8 | 2.1 | 0.2 |

The table reports the efficiency loss variance ratios var$_p(\hat{\theta})/\text{var}(\hat{\theta})$ for $\theta = \alpha$ or $\lambda$ where var$_p(\hat{\theta})$ is computed using $P$ portfolios and 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 (27). Panel A sorts stocks by true betas in each small sample and the panels B-D 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 (28). Panel C introduces cross-sectionally correlated residuals across stocks following equation (30). In Panel D, firms enter and exit stochastically and upon entry have a log-logistic model for duration given by equation (31). 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. In Panels E and F, stocks are sorted by characteristics; these characteristics have a correlation of 0.5 with true betas and panel E, but are uncorrelated with true betas in panel F.

Table 3: Biases in Monte Carlo simulations

| Number of Portfolios $P$ | 10 | 25 | 50 | 250 | 500 | 2,500 | All Stocks |
|---|---|---|---|---|---|---|---|
| **Panel A: Sorting on True Betas, Correlated Betas and Idiosyncratic Volatility** | | | | | | | |
| $\alpha$ | 0.013 | 0.021 | 0.034 | 0.131 | 0.236 | 0.698 | 0.885 |
| $\lambda$ | 0.026 | 0.017 | 0.002 | -0.110 | -0.229 | -0.756 | -0.971 |
| **Panel B: Sorting on Estimated Betas Correlated Betas and Idiosyncratic Volatility** | | | | | | | |
| $\alpha$ | 0.029 | 0.050 | 0.081 | 0.286 | 0.469 | 0.900 | 0.885 |
| $\lambda$ | 0.011 | -0.009 | -0.038 | -0.238 | -0.421 | -0.920 | -0.971 |
| **Panel C: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals** | | | | | | | |
| $\alpha$ | 0.035 | 0.059 | 0.094 | 0.314 | 0.505 | 0.953 | 0.950 |
| $\lambda$ | -0.070 | -0.094 | -0.127 | -0.343 | -0.536 | -1.064 | -1.138 |
| **Panel D: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals Entry and Exit of Firms** | | | | | | | |
| $\alpha$ | -0.036 | -0.009 | 0.039 | 0.338 | 0.588 | 1.169 | 1.200 |
| $\lambda$ | -0.055 | -0.081 | -0.127 | -0.417 | -0.667 | -1.330 | -1.431 |
| **Panel E: Sorting on characteristics correlated with betas** | | | | | | | |
| $\alpha$ | 0.019 | 0.063 | 0.135 | 0.539 | 0.839 | 1.187 | 0.885 |
| $\lambda$ | 0.021 | -0.021 | -0.088 | -0.469 | -0.758 | -1.179 | -0.971 |
| **Panel F: Sorting on characteristics uncorrelated with betas** | | | | | | | |
| $\alpha$ | 2.508 | 2.714 | 2.609 | 2.447 | 2.281 | 1.486 | 0.885 |
| $\lambda$ | -2.162 | -2.345 | -2.256 | -2.134 | -2.010 | -1.435 | -0.971 |

The table reports the simulated biases of the maximum likelihood estimators of $\alpha$ and $\lambda$ using $P$ portfolios and all stocks. We simulate 10,000 small samples of $T = 60$ months with $N = 5,000$ stocks using the model in equation (27). The different sorting methods are defined in Table 2. The biases are expressed as annualized percentage rates.

Table 4: MSE in Monte Carlo simulations

| Number of Portfolios $P$ | 10 | 25 | 50 | 250 | 500 | 2,500 | All Stocks |
|---|---|---|---|---|---|---|---|
| **Panel A: Sorting on True Betas, Correlated Betas and Idiosyncratic Volatility** | | | | | | | |
| $\alpha$ | 0.351 | 0.330 | 0.323 | 0.316 | 0.336 | 0.674 | 0.906 |
| $\lambda$ | 0.299 | 0.280 | 0.272 | 0.265 | 0.289 | 0.729 | 1.047 |
| **Panel B: Sorting on Estimated Betas Correlated Betas and Idiosyncratic Volatility** | | | | | | | |
| $\alpha$ | 0.758 | 0.741 | 0.724 | 0.706 | 0.761 | 1.048 | 0.906 |
| $\lambda$ | 0.575 | 0.560 | 0.547 | 0.537 | 0.598 | 1.042 | 1.047 |
| **Panel C: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals** | | | | | | | |
| $\alpha$ | 2.777 | 2.278 | 1.869 | 1.054 | 0.940 | 1.151 | 1.027 |
| $\lambda$ | 2.914 | 2.303 | 1.808 | 0.969 | 0.896 | 1.356 | 1.409 |
| **Panel D: Correlated Betas, Idiosyncratic Volatility, Cross-Correlated Residuals Entry and Exit of Firms** | | | | | | | |
| $\alpha$ | 3.408 | 2.798 | 2.289 | 1.289 | 1.189 | 1.665 | 1.563 |
| $\lambda$ | 3.581 | 2.825 | 2.225 | 1.221 | 1.194 | 2.044 | 2.162 |
| **Panel E: Sorting on characteristics correlated with betas** | | | | | | | |
| $\alpha$ | 1.618 | 1.539 | 1.500 | 1.518 | 1.708 | 1.728 | 0.906 |
| $\lambda$ | 1.285 | 1.221 | 1.187 | 1.200 | 1.380 | 1.654 | 1.047 |
| **Panel F: Sorting on characteristics uncorrelated with betas** | | | | | | | |
| $\alpha$ | 298.9 | 100.3 | 49.9 | 13.6 | 8.679 | 2.621 | 0.906 |
| $\lambda$ | 229.8 | 77.0 | 38.3 | 10.4 | 6.752 | 2.396 | 1.047 |

The table reports the simulated mean square error of the maximum likelihood estimators of $\alpha$ and $\lambda$ using $P$ portfolios and all stocks. We simulate 10,000 small samples of $T = 60$ months with $N = 5,000$ stocks using the model in equation (27). The different sorting methods are defined in Table 2.

Table 5: 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: 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 5**

The point estimates of $\alpha$ and $\lambda$ for the single factor, $MKT$, in equation (1) are reported over all stocks (Panel A) and various portfolio sortings (Panel B). 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 (11) and (12)) in the columns labeled "Max Lik" and GMM standard errors, detailed in Appendix B, 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 four last columns labeled "$\hat{\beta}$ Cross Section" list various statistics of the cross-sectional beta distribution: the cross-sectional mean, $E_c(\hat{\beta})$, the cross-sectional standard deviation, $\sigma_c(\hat{\beta})$, and the beta values corresponding to the 5%- and 95%-tiles of the cross-sectional distribution of beta. In Panel B 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. After the ex-ante portfolios are created, we follow the same procedure as Panel A to compute realized OLS market betas 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. All estimates $\hat{\alpha}$ and $\hat{\lambda}$ are annualized by multiplying the monthly estimates by 12.

Table 6: 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 |
| 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, $\lambda$, is equal to the time-series mean of the factor portfolio, $\mu$, 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 the text and Appendix B. 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 7: 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 |
| 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 ($E_c(\hat{\beta})$), 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 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 8: 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: 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 8**

The point estimates $\hat{\alpha}$, $\hat{\lambda}_{MKT}$, $\hat{\lambda}_{SMB}$, and $\hat{\lambda}_{HML}$ in equation (33) are reported over all stocks (Panel A) and various portfolio sortings (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. 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 (11) and (12)) in the columns labeled "Max Lik" and GMM standard errors, detailed in the text and in Appendix B, 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-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. 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 Panel B, 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 9: Tests for $H_0^{\lambda=\mu}$ ($|$T-statistics$|$) for the Fama-French (1993) Model

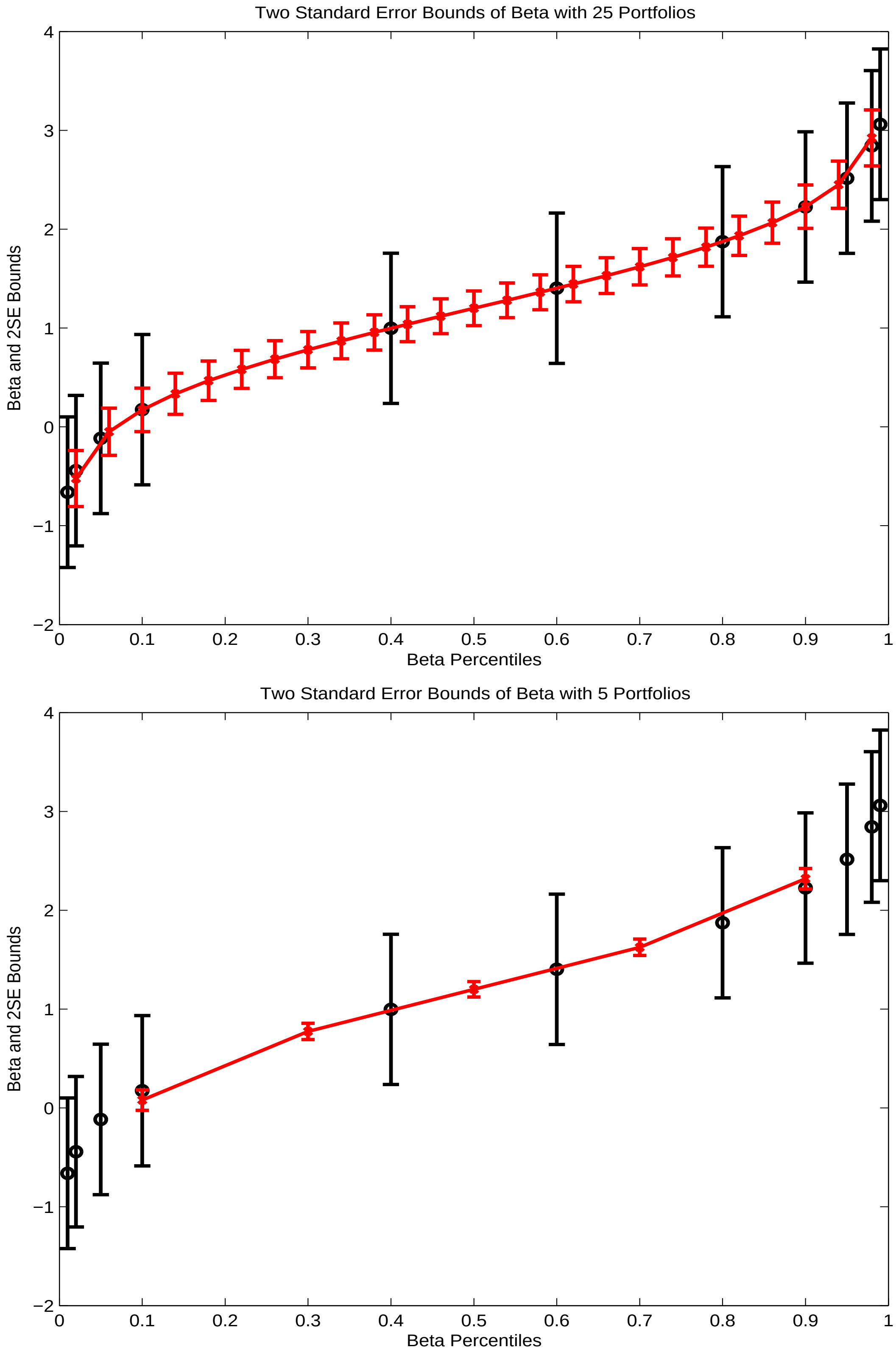
| | | | Residual Factor | | Industry Residuals | |
|---|---|---|---|---|---|---|
| Num Ports $P$ | Estimate (%) | | Max Lik | GMM | Max Lik | GMM |
| $\hat{\mu}_{MKT} = 6.43\%$, $\hat{\mu}_{SMB} = 2.16\%$, $\hat{\mu}_{HML} = 3.90\%$ | | | | | | |
| All Stocks | $\hat{\lambda}_{MKT}$ | 5.05 | 8.51 | 2.20 | 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 |
| 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, $\lambda$, is equal to the time-series mean of the factor portfolio, $\mu$, 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 the text and Appendix B. 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 9**

We estimate the Fama-French (1993) model (equation (33)) using all stocks (Panel A), $5 \times 5$ ex-ante portfolios sorted on market beta and book-to-market ratios (upper part of Panel B), and $5 \times 5$ 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 (11) and (12)) in the columns labeled "Max Lik" and GMM standard errors, detailed in the text and Appendix B, 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 8 and 9 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 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 $\alpha$, $\beta_{MKT}$, $\beta_{SMB}$, and $\beta_{HML}$ are annualized by multiplying the monthly estimates by 12.

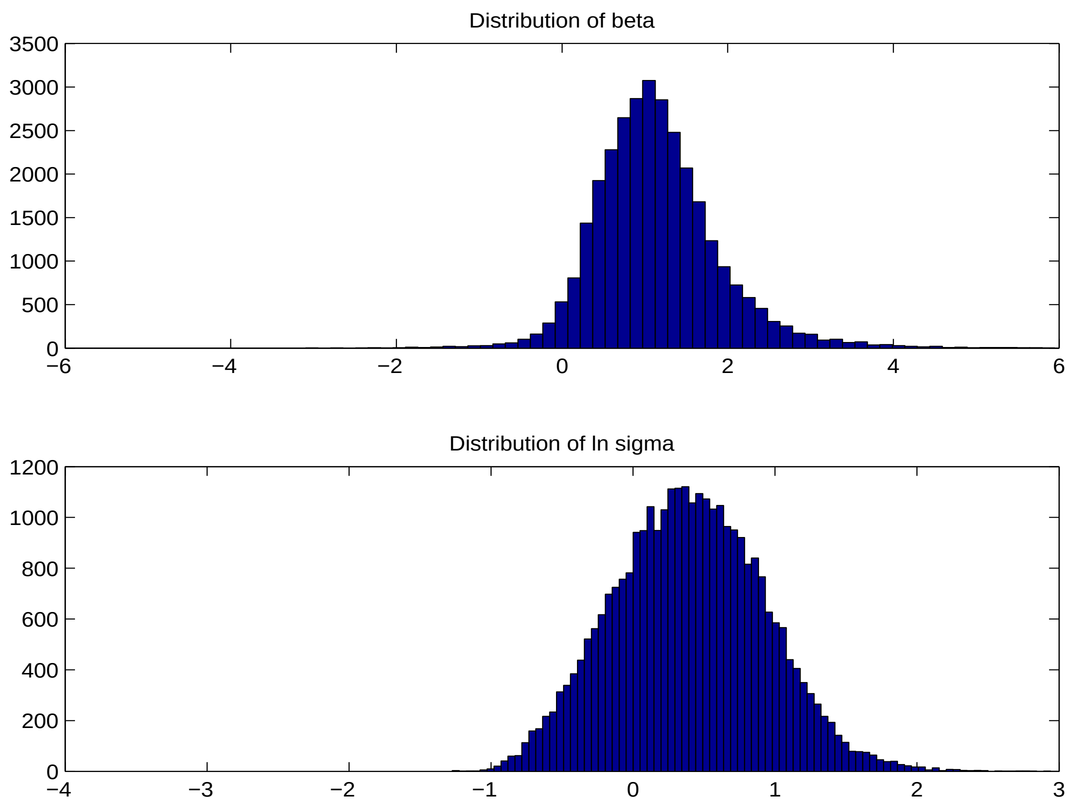
Figure 1: Standard Errors for $\hat{\beta}$ Using All Stocks or Portfolios



**Note to Figure 1**

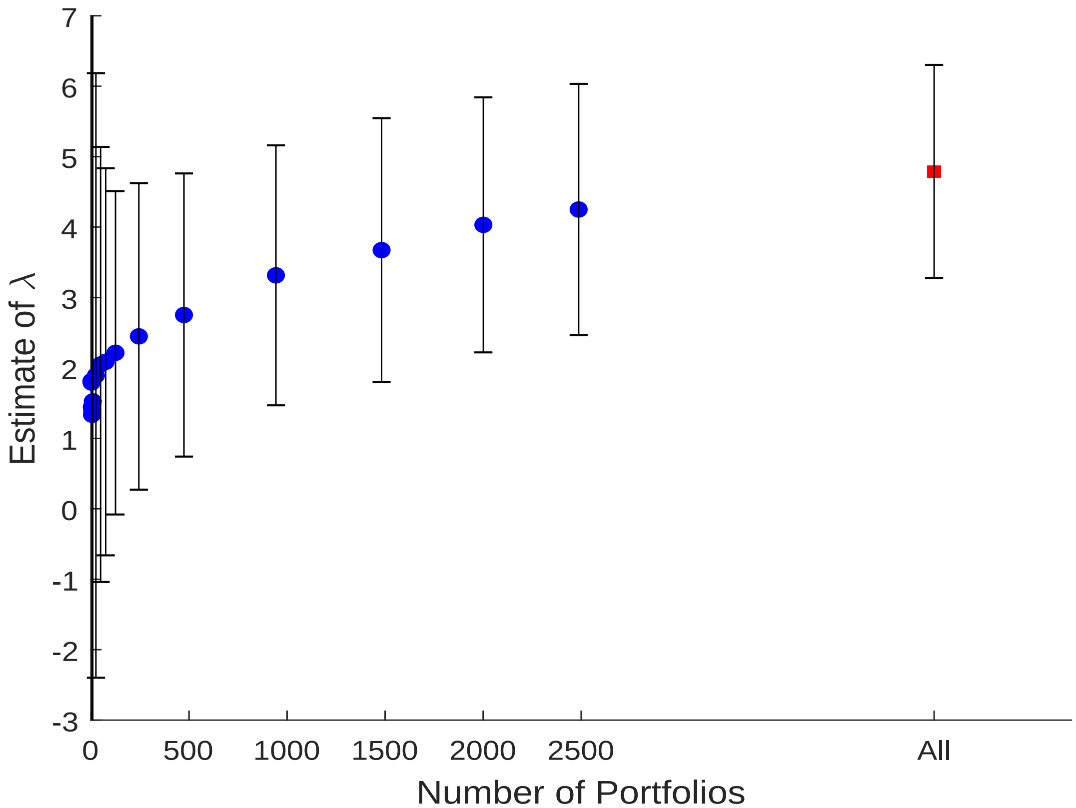
We assume a single factor model where $F_t \sim N(0, (0.15)^2/12)$ and the factor risk premium $\lambda = 0.06/12$. Betas are drawn from a normal distribution with mean $\mu_\beta = 1.1$ and standard deviation $\sigma_\beta = 0.7$ and idiosyncratic volatility across stocks is constant at $\sigma_i = \sigma = 0.5/\sqrt{12}$. 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 (23) and the computation for the portfolio moments are given in Appendix D.

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 Portfolios




The figure plots $\hat{\lambda}$ in a one-factor model using $P$ portfolios in blue circles. Two-standard error bands are marked as black lines intersecting the point estimates. The 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.