

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

**Resolving the Errors-in-Variables Bias in Risk Premium
Estimation**

Kuntara Pukthuanthong, University of Missouri
Richard Roll, California Institute of Technology
Junbo Wang, University of Southern California



SOCIAL SCIENCE WORKING PAPER 1392

July 2014

Resolving the Errors-in-Variables Bias in Risk Premium Estimation

by

Kuntara Pukthuanthong, Richard Roll, and Junbo Wang

July 19, 2014

Abstract

The Fama-Macbeth (1973) rolling- β method is widely used for estimating risk premiums, but its inherent errors-in-variables bias remains an unresolved problem, particularly when using individual assets or macroeconomic factors. We propose a solution with a particular instrumental variable, β calculated from alternate observations. The resulting estimators are unbiased. In simulations, we compare this new approach with several existing methods. The new approach corrects the bias even when the sample period is limited. Moreover, our proposed standard errors are unbiased, and lead to correct rejection size in finite samples.

| | Authors' Coordinates | | |
|---------|-----------------------------------------------------------------------|----------------------------------------------------------------------------|-------------------------------------------------------------------|
| | Pukthuanthong | Roll | Wang |
| Address | Department of Finance University of Missouri Columbia, MO 65211 | California Institute of Technology Professor Emeritus, UCLA Anderson | Marshall School of Business, University of Southern California |
| Voice | (619) 807-6124 | (626) 395-3890 | (781) 258-8806 |
| E-Mail | pukthuanthongk@missouri.edu | rroll@caltech.edu | junbo@usc.edu |

1 Introduction

The methods introduced by Black, Jensen and Scholes (1972), (BJS), and refined by Fama and Macbeth (1973), (FM), are widely employed to estimate risk premiums in linear factor models. This approach involves two-pass regressions: the first pass is a time series regression of returns on the factors for each asset, which produces estimates of factor loadings, widely called “betas” in the finance literature. The second pass regresses asset returns cross-sectionally on the estimated betas. As pointed out initially by Black, Jensen and Scholes (1972), risk premiums estimates from the second pass cross-sectional regression contain an inherent errors-in-variables bias because of estimation errors in the betas from the first pass.

Given the luxury of a large number N of individual assets, one can form diversified portfolios organized by particular asset characteristics. Black, Jensen and Scholes (1972), Blume and Friend (1973), and Fama and Macbeth (1973) show that this portfolio approach reduces estimation errors in the betas because they are less affected by idiosyncratic risk; so the errors-in-variables bias is mitigated (and eliminated in the limit $N \rightarrow \infty$).

Although the finite sample properties of the portfolio grouping procedure are weak even when N is reasonably large ($N > 2000$), many papers still employ FM with portfolios to estimate risk premiums. But more troubling is portfolio diversification can mask effects that exist in individual assets. Taking a naïve example, many investors seem to believe that some assets are overpriced and others are underpriced, but any portfolio grouping by an attribute other than price itself could diversify away the mispricing, rendering it undetectable.

A more egregious defect from portfolio masking involves the cross-sectional relation between mean returns and factor exposures (“betas”.) Take the single-factor CAPM as an illustration (though the same effect is at work for any linear factor model.) The cross-sectional relation between expected returns and betas holds exactly if and only if the market index used for computing betas is on the mean/variance frontier of the individual asset universe. Errors from the beta/return line, either positive or negative, imply that the index is not on the frontier. But if the individual assets are grouped into portfolios sorted by portfolio beta and the individual errors are not related to beta, the analogous line fitted to portfolio means and betas will display much smaller errors. This could lead to a mistaken inference that the index is on the efficient frontier.

This paper contributes to the literature by proposing a method to reduce the bias in risk premium estimates without relying on portfolio grouping. We also provide simulations to gauge the magnitude of the bias in finite samples and to compare our new method with previous approaches.

To put the problem into perspective, we show using macro factors with a large number of assets ($N \approx 5000$) and sample periods ($T \approx 600$), that the risk premium bias can be 60 to 70% with BJS and up to 90% with FM. Similar biases are present when using individual assets rather than portfolios. The method introduced in this paper can virtually eliminate the bias

in either case.

Our method builds on the FM rolling- β method and the techniques introduced in Griliches and Hausman (1986) and Biorn (2000). In the first pass, β 's are estimated from two sets of non-overlapping observations. Then the β 's from one set are used as instruments for the β 's from the second set in the second pass estimation. The asymptotic distribution of the estimated risk premiums is derived using Shanken's (1992) method.

We use simulations to evaluate hypothesis tests involving estimated risk premiums, which typically boil down to a null hypothesis that the intercept in the second pass regression is zero. Simulation results suggest that with the exception of our instrumental variables approach and the covariance from Theorem 3.2 below, all other methods reject a true hypothesis too often. This high rejection rate is due to a downward bias in estimated standard errors and an upward bias in the second pass intercept.

We also apply the new approach as well as the classical approaches to estimating the risk premiums for the macroeconomic factors. With our method, the consumption growth has positive significant effect on the stock returns. However, all other methods do not lead to the same or consistent conclusion.

This paper contributes to a large literature about the errors-in-variables bias. As the length of the sample period (T) grows indefinitely, Shanken (1992) shows that the errors-in-variables bias becomes negligible because the estimated beta errors are small. Shanken also derives an asymptotic adjustment for the standard errors. Jagannathan and Wang (1998) extend this asymptotic result to the case of conditionally heterogeneous errors in the time series regression. Kan, Robbotti and Shanken (2012) and Shanken and Zhou (2007) extend the result to a misspecified model. Chen and Kan (2004) investigate the finite sample properties of the cross-sectional regression, and find that the bias can be material even if T is reasonably large ($T=600$), when using macroeconomic factors such as consumption growth.

The two papers most closely related to this paper are Kim (1995) and Gagliardini, Ossola and Scaillet (2011). Kim (1995) corrects the errors-in-variable bias using lagged β as an instrument to derive a closed-form solution for the MLE estimator of the risk premiums under the assumption that the error terms are homogeneous. The solution proposed by Kim (1995) is based on Theil's adjustment (Theil (1971), (Cf. Litzenberger and Ramaswamy (1979), and Shanken (1992)). Theil's adjustment can mitigate errors-in-variables bias when the cross-sectional residuals are weakly dependent and the number of assets is large. Its limitation depends on an estimate of the standard error of the regression residuals, which can introduce new biases. Since our method uses instrumental variables to estimate risk premiums directly via the second pass regression, it is not subject to the same difficulty.

Gagliardini, Ossola and Scaillet (2011) show that when T and N are close to each other and both converge to infinity, the errors-in-variables bias in the estimated risk premiums in the BJS method converges to zero. Following their method, this paper derives asymptotic distributions by assuming that both T and N go to infinity. However, with the macro-factor

model, even when both T and N are large, the simulations suggest that the estimator from our IV β method has a much smaller bias than the BJS method because the convergence rate for estimated risk premiums is slower in the BJS method than in the IV method.

2.The Instrumental Variable Approach

Let \mathbf{r}_t denote the $1 \times N$ row vector of excess returns on N assets in period t^1 , \mathbf{f}_t denote the $1 \times K$ vector of factor realizations, $\boldsymbol{\beta}_t$ denote the $K \times N$ matrix of factor exposures, and $\boldsymbol{\varepsilon}_t$ denote the $1 \times N$ column vector of idiosyncratic return disturbances. For convenience and without loss of generalization, it is customary to assume that the factors and disturbances have means of zero, $E(\mathbf{f}_t) = E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$. Consequently, the system in period t can be expressed as

$$\mathbf{r}_t = E(\mathbf{r}_t) + \mathbf{f}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t. \quad (2.1)$$

The no arbitrage condition of the Arbitrage Pricing Theory (APT) (Ross [1976]) stipulates

$$E(\mathbf{r}_t) = \boldsymbol{\gamma}_t \boldsymbol{\beta}_t. \quad (2.2)$$

where $\boldsymbol{\gamma}_t$ is a $1 \times K$ vector of risk premiums associated with the factors. Since 2.1 and 2.2 hold in every period, over a sample of T periods, $t=1, \dots, T$, the vectors stacked alongside each other become $\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_T]'$, $\boldsymbol{\Omega} = [\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_T]'$, and $\mathbf{FB} \equiv [\mathbf{f}_1 \boldsymbol{\beta}_1, \dots, \mathbf{f}_T \boldsymbol{\beta}_T]'$, so that the overall sample can be expressed compactly as

$$\mathbf{R} = E(\mathbf{R}) + \mathbf{FB} + \boldsymbol{\Omega}, \quad (2.1a)$$

and, similarly, defining $\boldsymbol{\Gamma B} \equiv [\boldsymbol{\gamma}_1 \boldsymbol{\beta}_1, \dots, \boldsymbol{\gamma}_T \boldsymbol{\beta}_T]'$,

$$E(\mathbf{R}) = \boldsymbol{\Gamma B}. \quad (2.2a)$$

If the factor exposures and risk premiums are time invariant, the system is simplified; i.e, for $\boldsymbol{\beta} \equiv \boldsymbol{\beta}_1 = \dots = \boldsymbol{\beta}_T$ and $\boldsymbol{\gamma} \equiv \boldsymbol{\gamma}_1 = \dots = \boldsymbol{\gamma}_T$ combining 2.1a and 2.2a, we have

$$\mathbf{R} = [(\mathbf{f}_1 + \boldsymbol{\gamma}), \dots, (\mathbf{f}_T + \boldsymbol{\gamma})]' \boldsymbol{\beta} + \boldsymbol{\Omega}, \quad (2.1b)$$

Expression 2.1b represents a set of seemingly unrelated regressions that can be used in principle to test the APT's no arbitrage condition or estimate the risk premiums. If $\boldsymbol{\beta}$ is known, 2.1b becomes a cross-sectional regression for estimating the risk premiums because $E(\mathbf{f}_t) = \mathbf{0}$. If the factors are known, (or assumed to be known), 2.1b becomes a time series regression for estimating $\boldsymbol{\beta}$.

In each regression, the intercept must be zero if there is no arbitrage. In a time series

¹Matrices and Vectors are indicated by *bold face italic*

regression, the sample means of the factors are estimates of the risk premiums if the factors are traded portfolios or indexes with market returns. See Gibbons, Ross and Shanken (1989). However, if the factors are not traded assets, (e.g., if they are macroeconomic variables) sample means of the factors are not necessarily mean returns. Also, if a joint test of the individual intercepts rejects the hypothesis that they are all zero, the accuracy of estimated risk premiums is called into question.

Another method to test Equation 2.1b is to use GMM or MLE (e.g. Gibbons (1982) and Stambaugh (1982)). However, the numbers of parameters, moment conditions and the nonlinear estimation make these methods hard to implement when N is large (See the subsection on GMM below).

A third method, based on Black, Jensen and Scholes (1972) and Fama and Macbeth (1973), uses the two-pass regression approach mentioned earlier. The first pass relates returns of each asset to pre-specified factors in a time-series regression, thereby calculating estimates of β . The second pass is a cross-sectional regression of returns on the β estimates from the first pass. This can be done repeatedly for a time series of cross-sections; then the time series means of the cross-sectional coefficients in the second pass are estimates of the risk premiums.

The second pass cross-sectional regression is inherently subject to errors-in-variables bias because its explanatory variables are estimates from the first pass. These errors introduce bias in the estimated risk premiums, the coefficients in the second pass. In addition, the error-induced noise can affect the estimated sampling variance of the risk premiums. Heretofore, applications of BJS and FM, including the original contributions, use portfolio groupings to mitigate these problems. We propose a procedure that can be implemented with individual assets; no portfolios are required.

We rely on two major assumptions:

Assumption 1: ε_t is independent and identically distributed and is independent of β_t and f_t . The covariance metrics of ε_t and f_t are Σ and Σ_F , respectively.

Assumption 2: f_t and β_t are stationary processes and are independent of each other.

Under Assumption 1, the estimated factor loadings are asymptotically consistent in the first pass regression though they are not “admissible” in the James/Stein sense (for example, Stein (1956) and James and Stein (1961)) for finite samples. Moreover, the i.i.d assumption simplifies the asymptotic standard errors for the estimated risk premium. However, as we show below, the consistency of these estimators do not require the identical distribution of errors.

Assumption 2 allows us to derive the unconditional asymptotic distribution of the estimated risk premiums.

Suppose the length of the FM rolling window is τ and F_t is the transpose of the submatrix consisting of columns $t-\tau+1$ to t of the factor observations; i.e., $F_t \equiv [f_{t-\tau+1}, \dots, f_t]'$. Similarly, let R_t and Ω_t designate the corresponding columns of R and Ω . The first pass time series regression produces the estimates

$$\hat{\beta}_t = (F_t' F_t)^{-1} F_t' R_t$$

With a total sample size of T , there are $T-\tau$ sequential overlapping rolling windows of size τ .²

The estimation error in the first pass is $\hat{\beta}_t - \beta_t = (F_t' F_t)^{-1} F_t' \Omega_t$, which depends only on the error terms from time $t-\tau+1$ to t (under assumption #1.)

The dependent variable in the second pass cross-sectional regression could be any return vector r_s for a disjoint period $s \notin [t-\tau+1, t]$, though it is often simply $t+1$. This regression can be written $r_s = \alpha I_N + \hat{\beta}_t \hat{\gamma} + \xi_s$ where α is a common intercept and I_N is a unit vector of length N . Since the true model is $r_s = \beta(f_s + \gamma) + \varepsilon_s$, the cross-sectional residuals are $\xi_s = (\beta - \hat{\beta}_t)(f_s + \gamma) + \varepsilon_s$ and hence are correlated with the explanatory variables, the $\hat{\beta}_t$'s. This induces a complicated bias in the expected values of $\hat{\gamma}$ and $\hat{\alpha}$; under the APT no arbitrage condition, the true value of α is 0.

The bias is less severe when the estimation of β is more precise. Assuming that the true value of β is time invariate, one way to improve precision is to make τ large when T is large. Asymptotically, there is ever smaller measurement error in the first stage estimates and hence little bias in the second stage.

When the number of assets is large, i.e., when $N \rightarrow \infty$ with a fixed rolling window (the time series sample period could be large), the errors-in-variables bias could still be substantial. Instrumental variables (IV) corrects the first stage bias, but as always with the IV approach, the choice of instruments is crucial. In this case, however, there are some natural candidates; viz., $\hat{\beta}$ estimated from sample observations that are non-contiguous with the sample ending at t . These could be lagged observations. For example, if the original $\hat{\beta}$'s are estimated with τ observations ending at observation 2τ , the instruments could be $\hat{\beta}$'s estimated with the τ observations from 1 to τ . Specifically, if we define $\hat{\beta}1_t \equiv [1, \hat{\beta}_t]$, the OLS estimator is:

$$\hat{\gamma}_{t+\tau}' = (\hat{\beta}1_t \hat{\beta}1_{t+\tau}')^{-1} (\hat{\beta}1_t r_{t+\tau}')$$

²Previous work typically uses sequential and equal-length rolling windows but there is no mathematical necessity for such a procedure and we shall suggest a different approach below.

Similarly, the GLS estimator is:

$$\hat{\gamma}_{t+\tau}' = (\hat{\beta}_t' \Sigma^{-1} \hat{\beta}_{t+\tau}')^{-1} (\hat{\beta}_t' \Sigma^{-1} r_{t+\tau}').$$

Of course, any other non-overlapping observations could be employed including subsequent ones.

The IV method requires that strong instruments are highly correlated with the true explanatory variables and are uncorrelated with the residuals. Since most asset returns are weakly correlated over time, residuals from factor regressions are virtually uncorrelated as well. This satisfies the second condition for strong instruments. The first condition is trivially satisfied if the true betas are time invariant provided that the estimation samples are sufficiently long. In such a circumstance, factor exposure estimates from non-contiguous samples will be strongly related. On the other hand, if there is some time variation in the true β 's, samples from non-contiguous observations widely separated in time bring the risk of weakened instruments. We have a suggestion next to counteract this possibility.

2.3 An improved IV method for risk premium estimation.

The IV method is not limited to lagged instruments. Any estimated factor loadings using non-overlapping observations can be used as the instruments. But the instruments could be weak if the factor exposures are changing over time and the non-overlapping samples are relatively far apart.

This suggests a procedure that uses non-contiguous samples constructed to be the most coincidental possible in calendar time. Here is one proposed scheme: For each asset, divide the total sample into three subsamples. The first subsample contains returns and factors for observations 1, 4, 7,..., T-3. The second subsample contains returns and factors for observations 2, 5,..., T-2, and the third subsample contains returns and factors for observations 3, 6,...T.

Given that factor model residuals are uncorrelated across time, any of the three subsamples can be used to estimate factor loadings ($\hat{\beta}$'s), while either of the other two subsamples can be used to estimate instruments for the loadings. Then, the second stage FM cross-sectional regression can be estimated for each observation in the third subsample without having the returns related in any way to the errors in the $\hat{\beta}$'s or in their instruments. The sample means of the cross-sectional coefficients then become unbiased estimates of risk premiums. Since any permutation of the three subsamples is equally suitable, all three could be used and there seems to be nothing wrong with averaging the cross-sectional coefficients over all three permutations.

To be more specific, if we define $\hat{\beta}1_{\text{sample}} \equiv [1, \hat{\beta}_{\text{sample}}]$, where sample can be one of the first, second or third subsamples, the OLS estimator is:

$$\hat{\gamma}' = (\hat{\beta}1_{\text{first}}' \hat{\beta}1_{\text{second}}')^{-1} (\hat{\beta}1_{\text{first}}' \bar{r}_{\text{third}}'),$$

Similarly, the GLS estimator is:

$$\hat{\gamma}' = (\hat{\beta}1_{\text{first}}' \Sigma^{-1} \hat{\beta}1_{\text{second}}')^{-1} (\hat{\beta}1_{\text{first}}' \Sigma^{-1} \bar{r}_{\text{third}}').$$

\bar{r}_{third} represents the sample average of the stock returns over the third subsample. In fact, since the returns from both second and third subsample are uncorrelated with $\hat{\beta}1_{\text{first}}$, we can replace \bar{r}_{third} with $\bar{r}_{\text{secondthird}}$ (obtained by taking the average over both of these two subsamples) in the estimator. This estimator is consistent when the number of stocks N is large. This will be the 3-group estimator in our simulation and empirical sections.

There is an advantage of using three-group method. Since slow variation in the true factor loadings would evolve over the entire sample, each sub-sample would roughly include the same variation, thereby strengthening the instruments relative to using, say, lagged or leading non-contiguous observations.

Finally, we note that it might not be optimal to divide up the sub-samples equally. Depending on the volatility of factors and factor model residuals, one could imagine improvements based on unequal divisions; e.g., estimating the loadings and instruments with half of the available observations and conducting the cross-sectional regressions with the other half. We reserve this refinement for later study though and stick here to a tripartite procedure.

2.4 Theil's Adjustment

To compare, we now discuss two other methods to correct the bias. The first method is Theil's adjustment, which essentially modifies the BJS method.

In BJS method, when the second pass is OLS (the method is similar when the second pass is GLS), the estimated risk premium is

$$\gamma = (\hat{\beta}1' \hat{\beta}1)^{-1} (\hat{\beta}1' \bar{r}),$$

where $\bar{r} = \frac{1}{T} \sum_{t=1}^T r_t$ is the average of the returns and $\hat{\beta}1 = [I_{1 \times N}, \hat{\beta}]$. We can show that,

$$\frac{1}{N} \hat{\beta}1' \hat{\beta}1 = \frac{1}{N} \beta1' \beta1 + T_A$$

with $T_A = \begin{bmatrix} 0 & 0_{1 \times k} \\ 0_{k \times 1} & (F'F)^{-1} (F' \begin{bmatrix} \sum \delta_{it}^2 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \sum \delta_{it}^2 \end{bmatrix} F) (F'F)^{-1} \end{bmatrix}$, where $\sum \delta_{it}^2$ is

the summation of the variances of the regression residuals $\varepsilon_t = [\varepsilon_{t1}, \dots, \varepsilon_{tN}]$. The term T_A represents the covariance of the error term in the estimated factor loadings. Since T_A is

positive semi-definite, $\frac{1}{N} \hat{\beta}' \hat{\beta} 1 > \frac{1}{N} \beta' \beta 1$. This leads to a negative bias for the estimated risk premium. In order to correct this bias, Theil (1971) and Litzenberger and Ramaswamy (1979) suggest the following method:

$$\gamma = (\hat{\beta}' \hat{\beta} 1 - N \hat{T}_A)^{-1} (\hat{\beta}' \bar{r}),$$

$$\text{with } \hat{T}_A = \begin{bmatrix} 0 & \mathbf{0}_{1 \times k} \\ \mathbf{0}_{k \times 1} & (F' F)^{-1} (F' \begin{bmatrix} \sum \hat{\delta}_{it}^2 & \dots & 0 \\ \dots & \dots & \dots \\ \dots & \dots & \sum \hat{\delta}_{iT}^2 \end{bmatrix} F) (F' F)^{-1} \end{bmatrix}, \text{ where } \sum \hat{\delta}_{it}^2 \text{ is}$$

the summation of the estimated variances of the regression residuals.

Shanken (1992) shows that this estimator is consistent when N is large under the assumption that the summation of the estimated variances of the regression residuals converges to its true value with large number of stocks. In Section 5, we also find that the finite sample bias is small with Theil's adjustment when this assumption is valid.

However, the above assumption may not always be true. For example, suppose that the summation of the variances of the regression residuals $\sum \delta_{it}^2$ is time-varying, but we estimate it by summing of the average variance of all regression residuals for all time, i.e.

$\sum \hat{\delta}_{it}^2 = \sum_{t=1, i=1}^{T, N} \hat{\delta}_{it}^2 / T$, then, $\sum (\hat{\delta}_{it}^2 - \delta_{it}^2)$ does not converge to zero since the true value $\sum \delta_{it}^2$ is time varying and the estimator $\sum \hat{\delta}_{it}^2$ is constant. In this scenario, if $\sum (\hat{\delta}_{it}^2 - \delta_{it}^2)$ is correlated with the factors, then

$$\begin{aligned} E(\hat{T}_A - T_A) &= \\ E \begin{bmatrix} 0 & \mathbf{0}_{1 \times k} \\ \mathbf{0}_{k \times 1} & (F' F)^{-1} (F' \begin{bmatrix} \sum (\hat{\delta}_{it}^2 - \delta_{it}^2) & \dots & 0 \\ \dots & \dots & \dots \\ \dots & \dots & \sum (\hat{\delta}_{iT}^2 - \delta_{iT}^2) \end{bmatrix} F) (F' F)^{-1} \end{bmatrix} \\ &\neq \mathbf{0} \end{aligned}$$

Therefore, the Theil's estimator can create a new bias when estimated variances of the residuals do not converge to the true variances. Such issue will not affect the instrumental variable method since the only assumption for this method is that the residuals are not auto correlated.

There is another issue associated with the Thiel's adjustment in estimating the standard errors of the regression residues $\sum \delta_{it}^2$, when the factor loading β is time varying. Specifically, if $\beta_1, \beta_2, \dots, \beta_T$ are not identical, the estimated variance is

$$\delta_t^2 = \frac{1}{T} \sum_{t=1}^T (R_t - f_t(F'F)^{-1}F'R)^2 = \frac{1}{T} \sum_{t=1}^T (f_t\beta_t + \varepsilon_t - f_t(F'F)^{-1}F'(FB) - f_t(F'F)^{-1}F'\varepsilon)^2$$

Here $FB = \begin{bmatrix} f_1\beta_1 \\ f_2\beta_2 \\ \dots \\ f_T\beta_T \end{bmatrix}$. The above equation can be further decomposed into

$$\frac{1}{T} \sum_{t=1}^T (\varepsilon_t - f_t(F'F)^{-1}F'\varepsilon)^2 + \frac{1}{T} \sum_{t=1}^T (f_t\beta_t - f_t(F'F)^{-1}F'(FB))^2$$

. The first part of the

$$+ \frac{1}{T} \sum_{t=1}^T (\varepsilon_t - f_t(F'F)^{-1}F'\varepsilon)(\beta_t f_t - f_t(F'F)^{-1}F'(FB))$$

decomposition is a consistent estimator of the variance of the regression residues. Moreover, when the regression residues are uncorrelated with the factors and the loadings, the expected value of the third part is zero. However, the second part is a function of β_t , and its expected value is not zero when the factor loadings are time-varying. Thus, the estimated errors contain a bias, and the bias can affect the estimated risk premiums. The time-varying factor loadings do not create such an issue for the instrumental variable approach since there is no need to estimate the standard errors of the regression residues.

2.5 GMM

Following equation (12.23) of Cochrane (2005), one can use the following moment conditions in GMM estimation:

$$E(F'(r - \alpha - FB)) = 0,$$

$$E(r - \Gamma B) = 0.$$

For N assets, there are $N(K+1)+N$ moment conditions and $NK+N+K$ parameters. Hence, the GMM is overidentified if $N > K$.

However, there are two limitations in implementing GMM. First, the method cannot be easily estimated when N is even moderately large. Suppose $N=149$ and $K=3$ (3-factor model), then there are 745 moment conditions and 599 parameters to estimate, and it becomes problematic to find the global minimal of the objective function with 599 parameters and 745 moment conditions. If $N=5000$ (as for individual stocks), then there are $N(K+1)+N=25000$ moment conditions. Hence, T must be more than 25000. Usually, we do not have data with such a large T , hence, the GMM is not implementable.

Another problem is the efficiency of the estimator. In an iterated GMM process, it is difficult to construct the efficient weighting matrix of the moment conditions that is

invertible based on the estimated parameters in previous iteration. Even when $N=25$ (the weighting matrix is 125×125), we can only use the identity matrix as a weighting matrix. However, using the identity matrix will not be efficient for the estimated parameters.

To deal with these two limitations, Shanken and Zhou (2007) make an adjustment to this method. First, they estimate β with the time series regression. Second, they use the estimated β to form the moment condition. In this case, there are only k parameters remaining in the second pass of their two-pass GMM. This adjusted method is essentially similar to the BJS method with the GLS estimation in the second pass.

3. Asymptotic Distributions

3.1 The asymptotic distribution of the β IV method.

In this section, we show the consistency of the estimator and obtain its asymptotic distribution.

Theorem 3.1 (a) The estimated risk premiums $\hat{\gamma}_t' - (\mathbf{0}, \mathbf{f}_t)'$ (rolling- β IV estimator) and $\hat{\gamma}' - (\mathbf{0}, \bar{\mathbf{f}}_{\text{secondthird}})'$ (3-group β IV estimator) are consistent.

(b) Assume that when $N \rightarrow \infty$, $\beta \mathbf{1} \Sigma^{-1} \beta' / N$ converges to an invertible matrix (denote this matrix by $\mathbf{b} \Sigma^{-1} \mathbf{b}'$). In addition, assume that $[\beta_1^1 \xi_{11}, \dots, \beta_N^1 \xi_{1N}]$ (where $[\beta_1^1, \dots, \beta_N^1] = \beta \mathbf{1} \Sigma^{-1}$) satisfies a Lindeberg condition, then the asymptotic distribution for the estimated risk premiums using the lagged IV (rolling- β IV) is:

$$\sqrt{N}(\hat{\gamma}_t' - (\mathbf{0}, \gamma + \mathbf{f}_t)') \rightarrow N(0, \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1}),$$

where $\mathbf{A} = \mathbf{b} \Sigma^{-1} \mathbf{b}'$, $\mathbf{B} = c_0 (\mathbf{b} \Sigma^{-1} \mathbf{b}' + \tilde{\mathbf{L}}_{0,t-\tau})$, where

$$\tilde{\mathbf{L}}_{0,t-\tau} = \begin{bmatrix} 0 & \mathbf{0}_{1 \times k} \\ \mathbf{0}_{k \times 1} & (\mathbf{F}_{t-\tau}' \mathbf{F}_{t-\tau})^{-1} \mathbf{F}_{t-\tau}' \mathbf{L}_\tau \mathbf{F}_{t-\tau} (\mathbf{F}_{t-\tau}' \mathbf{F}_{t-\tau})^{-1} \end{bmatrix}$$

and

$$c_0 = 1 + (\gamma + \mathbf{f}_t)(\mathbf{F}_t' \mathbf{F}_t)^{-1} \mathbf{F}_t' \mathbf{L}_\tau \mathbf{F}_t (\mathbf{F}_t' \mathbf{F}_t)^{-1} (\gamma + \mathbf{f}_t)' - 2 \frac{\tau-1}{\tau} (\gamma + \mathbf{f}_t)(\mathbf{F}_t' \mathbf{F}_t)^{-1} \mathbf{F}_t' \mathbf{I}_0$$

with $\mathbf{l}_0 = (1 - \frac{1}{\tau}, -\frac{1}{\tau}, \dots, -\frac{1}{\tau})'$ a τ -vector and $\mathbf{L}_\tau = \begin{bmatrix} \frac{\tau-1}{\tau} & -\frac{1}{\tau} & \dots & -\frac{1}{\tau} \\ -\frac{1}{\tau} & \frac{\tau-1}{\tau} & \dots & -\frac{1}{\tau} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{\tau} & \dots & -\frac{1}{\tau} & \frac{\tau-1}{\tau} \end{bmatrix}$.

The asymptotic distribution for the estimated risk premiums using the 3-group IV method is:

$$\sqrt{N}(\hat{\gamma}' - (\mathbf{0}, \gamma + \bar{f}_{\text{secondthird}})') \rightarrow N(0, \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1}).$$

Here $\mathbf{A} = \mathbf{b} \boldsymbol{\Sigma}^{-1} \mathbf{b}'$, $\mathbf{B} = c_0 (\mathbf{b} \boldsymbol{\Sigma}^{-1} \mathbf{b}' + \tilde{\mathbf{L}})$, where

$$\tilde{\mathbf{L}} = \begin{bmatrix} 0 & \mathbf{0}_{1 \times k} \\ \mathbf{0}_{k \times 1} & (\mathbf{F}_{\text{first}}' \mathbf{F}_{\text{first}})^{-1} \mathbf{F}_{\text{first}}' \mathbf{L}_{\tau^1} \mathbf{F}_{\text{first}} (\mathbf{F}_{\text{first}}' \mathbf{F}_{\text{first}})^{-1} \end{bmatrix}$$

and

$$c_0 = \frac{1}{\tau^2 \tau^3} + (\gamma + \bar{f}_{\text{secondthird}}) (\mathbf{F}_{\text{second}}' \mathbf{F}_{\text{second}})^{-1} \mathbf{F}_{\text{second}}' \mathbf{L}_{\tau^2} \mathbf{F}_{\text{second}} (\mathbf{F}_{\text{second}}' \mathbf{F}_{\text{second}})^{-1} (\gamma + \bar{f}_{\text{secondthird}})' - 2 \frac{\tau^2 - 1}{(\tau^2)^2 \tau^3} (\gamma + \bar{f}_{\text{secondthird}}) (\mathbf{F}_{\text{second}}' \mathbf{F}_{\text{second}})^{-1} \mathbf{F}_{\text{second}}' \mathbf{L}_0^2$$

with $\mathbf{L}_0^2 = (1 - \frac{1}{\tau^2}, -\frac{1}{\tau^2}, \dots, -\frac{1}{\tau^2})'$ a τ^2 -vector and $\mathbf{L}_{\tau^i} = \begin{bmatrix} \frac{\tau^i - 1}{\tau^i} & -\frac{1}{\tau^i} & \dots & -\frac{1}{\tau^i} \\ -\frac{1}{\tau^i} & \frac{\tau^i - 1}{\tau^i} & \dots & -\frac{1}{\tau^i} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{\tau^i} & \dots & -\frac{1}{\tau^i} & \frac{\tau^i - 1}{\tau^i} \end{bmatrix}$,

assuming that subsample i contains τ^i periods. Moreover, $\bar{f}_{\text{secondthird}}$ is the average of the factors over the second and the third subsample.

The proof is in the appendix.

In cross-sectional regressions, the second pass can either be OLS or GLS. The Theorem presents the asymptotic distribution when the second pass is GLS. The OLS estimation is a special case (when $\boldsymbol{\Sigma} = \delta^2 \times \mathbf{I}$).

Theorem 3.1 provides an estimation of the risk premiums conditional on the value of the

factors. The Theorem says that $\hat{\gamma}_t$ is a consistent estimator of the expected risk premium plus the unexpected factor realization at time t . If the factors are unexpected shocks or demeaned factors ($E(f) = 0$), then $\hat{\gamma}_t - (0, f_t)'$ is the consistent estimator of γ .

One important assumption in this Theorem is

$$\beta 1 \Sigma^{-1} \beta 1' / N \rightarrow b \Sigma^{-1} b'.$$

A possible issue that arises in grouping can violate this assumption. If we group the individual stocks into well diversified portfolios according to characteristics of the firms, these portfolios may have market β close to 1. Hence, $\beta 1 \Sigma^{-1} \beta 1' / N$ is not invertible, and one cannot use the rolling- β method with the instrumental variables. This situation is similar to the use less factor case in Kan and Zhang (1999) who show that the errors-in-variables bias can be amplified in finite samples. There are several approaches to control this problem. One approach is to group the stocks according to market β as well; hence, $\beta 1 \Sigma^{-1} \beta 1' / N$ is invertible. Another method is to drop the constant term in the second pass. More specifically, in the second pass, one can regress returns on the $\hat{\beta}$ on book-to-market and the $\hat{\beta}$ on size without $\hat{\beta}$ on market. This method implicitly assumes the multi-factor model is true and the intercept is 0. A third approach is to use individual stock returns to estimate the risk premiums. When N is reasonably large, we will show that using instrumental variables from non-overlapping observations is effective in correcting the bias in finite sample.

When the sample period T is large enough, the sample average of the estimated risk premiums using the lagged instrument, $\frac{1}{t-2\tau+1} \sum_{t=\tau}^{T-\tau+1} \hat{\gamma}_t$, is a consistent estimator. In the next Theorem, we provide the asymptotic distribution of the sample average of the estimated risk premiums. Notice that since the estimated time series $\{\hat{\gamma}_t\}$ is autocorrelated up to τ , the asymptotic variance of the sample average of the estimated risk premiums contains these autocorrelations.

Theorem 3.2 The sample average of the estimated risk premiums using rolling- β IV method is a consistent estimator of the risk premiums, i.e. when T is large,

$$\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T \hat{\gamma}_t \rightarrow (0, \gamma)'.$$

(a) If $\{f_t\}$ is a stationary process, the unconditional asymptotic distribution of the estimated risk premiums is:

$$\sqrt{T} \sqrt{N} \left(\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T (\hat{\gamma}_t - (0, f_t)') - (0, \gamma)' \right) \rightarrow N(0, V),$$

if $B_f = E(\sum_{t=1}^{\tau-1} (c_{t1}(\mathbf{b}\Sigma^{-1}\mathbf{b}' + \tilde{\mathbf{L}}_{t1,t-\tau})))$ in which c_{t1} 's are constants and $\tilde{\mathbf{L}}_{t1,t-\tau}$'s are the $(K+1) \times (K+1)$ matrices³ exists,

$$\mathbf{V} = \mathbf{A}_f^{-1} \mathbf{B}_f \mathbf{A}_f^{-1},$$

where $\mathbf{A}_f = \mathbf{b}\Sigma^{-1}\mathbf{b}'$.

(b) Excluding the term $(0, \mathbf{f}_t')$, we have the same asymptotic distribution of the estimated risk premiums as Theorem 2 in Shanken (1992):

$$\sqrt{T} \left(\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T \hat{\mathbf{y}}_t' - (0, \boldsymbol{\gamma})' \right) \rightarrow N(0, \tilde{\Sigma}_F), \text{ where } \tilde{\Sigma}_F = \begin{bmatrix} 0 & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 1} & \Sigma_F \end{bmatrix}.$$

The proof of this Theorem is in the appendix.

We will call \mathbf{V} the \sqrt{NT} -asymptotic covariance and $\tilde{\Sigma}_F$ the \sqrt{T} -asymptotic covariance since they represent covariances with different convergent rates. Shanken (1992) derives the asymptotic distribution of the BJS method with convergent rate of \sqrt{T} . Gagliardini, Ossola and Scaillet (2011) show that when both T and N large, the estimated risk premiums in the BJS method converge to the true value at the speed of $O(\frac{1}{T}) + O(\frac{1}{\sqrt{NT}})$ ⁴. If $N > O(T)$, then the rate of convergence is $O(\frac{1}{T})$. Theorem 3.2 shows that if one uses the β IV method and subtracts the sample average $\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T \mathbf{f}_t'$ from $\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T \hat{\mathbf{y}}_t'$, the rate of convergence for this estimator is $O(\frac{1}{\sqrt{NT}})$. This method does not depend on the relative size of T and N .

Notice that in the derivation of asymptotic distributions when the second pass is GLS, the covariance matrix of the error terms Σ is assumed to be known. In reality, this matrix is unknown. Hence, one needs a feasible version to deal with this problem. One classical approach is to assume Σ is some function of the factors, hence the estimation of the Σ becomes the estimation of the coefficients of these functions.

There are two other methods implementing GLS. The first is introduced by Shanken (1985). If $T > N+K$, one can estimate Σ by taking the sample average of the cross multiplication of the sample error terms that can be calculated through the estimation from the second pass OLS regression. Another method is that of Ferson and Harvey (1999) using weighted GLS. This paper adopts Ferson and Harvey (1999) since most of the cross-sectional correlations between the idiosyncratic risks are small.

³ The formulas of $\tilde{\mathbf{L}}_{t1,t-\tau}$ and c_{t1} will be defined in the proof. See appendix for the details.

⁴Here, for any real number X , $O(X)$ is defined as follows: there exist two positive numbers M and N , such that $MX < O(X) < NX$.

3.2 A simple way to calculate standard errors.

In Theorem 3.2, we derive the asymptotic covariance of the estimated risk premiums. If the factors are demeaned factors or the factors are the shocks from its conditional expected values, then one can use $\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T (\hat{\gamma}_t - (0, f_t))'$ as the estimator of the γ' , and V is the covariance matrix with \sqrt{NT} -convergent rate. Fama and Macbeth also provide a simpler way to calculate the asymptotic covariance by taking the sample covariance of the estimated risk premiums. This method is also applicable with the β IV method.

Define $\hat{\gamma}_t^* = \hat{\gamma}_t - (0, f_t)$ and $\bar{\gamma}^* = \frac{1}{T-2\tau+1} \sum_{t=2\tau}^T (\hat{\gamma}_t - (0, f_t))$. In this case, the Fama-Macbeth sample asymptotic covariance (with autocovariances up to the length of the rolling window) in this situation is:

$$N \sum_{t1=-\tau}^{\tau} \frac{1}{T-2\tau-t1+1} \sum_{t=2\tau+t1}^T (\hat{\gamma}_t^* - \bar{\gamma}^*)(\hat{\gamma}_{t-t1}^* - \bar{\gamma}^*).$$

Another version of the Fama-Macbeth sample covariance does not contain the autocovariances i.e., one can use

$$N \sum_{t=2\tau}^T \frac{1}{T-2\tau+1} (\hat{\gamma}_t^* - \bar{\gamma}^*)(\hat{\gamma}_t^* - \bar{\gamma}^*)$$

to estimate the covariance of the estimated risk premiums. In section 4, we will analyze different sample covariances as well as the asymptotic covariance from Theorem 3.2. The differences are small in the Fama-French three factor model, but can be large for the macro-factor model.

4 Data and Simulation Results

This section uses Fama-French portfolios and macro factors to compare empirically different methods of estimating risk premiums. We compare four methods: (1) BJS estimation without rolling betas; (2) the rolling- β method of Fama and Macbeth (1973); (3) the rolling- β method using β 's estimated with non-overlapping observations as the instrumental variable; (4) Theil's adjustment. In applying the rolling beta method, we assume the rolling window is 15 for $T=60$ and 60 for $T=600$.

Fama and French's three factors 1964 to 2009 are available on Kenneth French's data library. We use these factors along with 100 size and book-to-market portfolios plus 49 industry portfolios as well as the 25 size and book-to-market portfolios. In addition to these 25 and 149 portfolios, we obtain returns for 4,970 individual stocks from CRSP that have fewer than 20% missing observations from 2000 to 2009.

The macro variables are chosen following Chen, Roll and Ross (1986) including unexpected consumption growth, unexpected inflation and unexpected change in industrial production. The raw series are obtained from the Federal Reserve at St Louis. To measure consumption, we add the consumption of nondurables to services and divide by the US population. The Consumer Price Index for All Urban Consumers is our measured price level. Raw growth rate was estimated as the log differences of these level variables.

The macro factors are estimated shocks from conditional expected values estimated from a vector auto-regression (VAR). Specifically, define $X_t \equiv [\Delta C_t : \Delta IP_t : \Delta CPI_t]'$ where ΔC_t , ΔIP_t , and ΔCPI_t are raw consumption growth, industrial production growth and the inflation rate, respectively. Then X_t is modeled as an $AR(1)$ process that follows a vector auto-regression $X_t = A + BX_t + \zeta_t$ where ζ_t denotes the 3×1 vector of VAR innovations. The fitted value $\hat{X}_t = \hat{A} + \hat{B}\hat{X}_{t-1}$ is our the conditional expected value of X_t . The shock or unexpected value $X_t - \hat{X}_t$ is assumed to be the driving factor for assets.

4.1 Simulations

For the three Fama-French factors, we assume that the true risk premiums are the sample means of excess return factors. Shanken (1992) shows that the sample mean is a consistent estimator of the risk premiums. Using data from 1964 to 2009, we regress returns on the factors to estimate β 's for each portfolio and calculate the regression residues using returns, factors and estimated β 's. In the bootstrap simulation, the factors and the error terms are re-sampled from the pool of observed factors and error terms generated by the above regressions. Re-sampled factors, error terms and estimated β 's are used to generate the simulated (re-sampled) portfolio returns.

However, as in Kan and Zhang (1999) or Kleibergen (2009), the regression in the second pass has a multi-collinearity problem because the estimated β 's on the market are close to 1 for all N simulated portfolios. To alleviate this problem, we use the following method: we regress returns on all three factors. However, since the estimated β on the market is close to 1 for all portfolios, in the second pass, we only regress the returns on the estimated β 's on market, book-to-market and size in the second pass and omit the intercept. This essentially assumes that the Fama-French three-factor model is true and that its constant term is zero. Of course, if the model is not true, this estimation procedure is likely to produce biased risk premiums for all three factors. This is an inherent problem with portfolio grouping if market betas for the portfolios are all close to 1.0.

For the Fama-French three-factor model, we generate returns and factors 10,000 times. Each of the 10,000 samples has T of 60 or 600 and N of 25 and 149, which allows us to examine performance in various finite samples. We use the BJS method, rolling- β method and the β IV method to estimate the risk premiums. Hence, there are 10,000 estimated risk premiums for each of the methods. The reported risk premiums are the average of these estimated risk premiums. We also present the T-ratio of the difference between reported risk premiums and their true values. The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

In addition to the portfolios, we use individual stocks (with $N=4,970$) to estimate the risk premiums in the Fama-French three-factor model. With individual stocks, the β on market is not a redundant variable in the second pass; hence, we run a cross-sectional regression of the returns on 1 and the estimated β 's of all three factors in the second pass.

We also simulate macro factors and concomitant errors and investigate them with the three methods. Since the macro factors are less volatile than the traded factors, the linear factor model has larger idiosyncratic risk; therefore, the errors-in-variables bias is likely to be much larger. We examine this case in the context of monthly returns for 4,970 individual stocks from CRSP. In this examination, T is either 60 or 600 and N takes on three values, 25, 149, and 4,970. Each combination of parameters is replicated 10,000 times. To estimate the “true” risk premiums of the macro factors in the simulation, we average estimated risk premiums from the literature.⁵

4.2 Simulation Results for the Fama-French Three-Factor (Traded Factors) Model

Table 1 presents the average estimated risk premiums of three-factor model using 25 portfolios. When there are 25 portfolios, the estimated risk premiums from both the BJS and the Fama-Macbeth rolling beta methods have errors-in-variables bias. When the sample period is only 60, this bias is large. Interestingly, even when $T = 600$, and the rolling window is 60, the Fama-Macbeth rolling beta method produces risk premiums that are still about 20 percent smaller than their true values. Evidently, either the rolling window or the sample period is too short to eliminate the bias. In contrast, the β IV method produces accurate estimated risk premiums when the sample size is both 60 and 600. T-ratios of the difference are small for all estimated risk premiums, indicating that the bias is negligible in statistical sense as well.

*** Insert Table 1 here ***

For 149 portfolios, results are presented in Table 2. The errors-in-variables bias is small, though it is larger than the bias with 25 portfolios (Table 1). This is because grouping stocks into 25 well-diversified portfolios diversifies away idiosyncratic risk better than grouping the same stocks into 149 portfolios. Using T-ratios of difference, the bias for classical methods (BJS method and Fama-Macbeth method) are significant when $T=600$.

*** Insert Table 2 here ***

Estimated risk premiums using 4970 individual stocks are shown in the Table 3. For $T = 60$, estimated risk premiums from the β IV method have smallest bias. The bias is largest for the size factor, about 20% with the β IV method (Rolling IV), but this is smaller than that produced by the BJS method (about 40%), and by the Fama-Macbeth method (more than a 60%). For $T = 600$, the BJS method, β IV method and Theil's adjustment

⁵ These papers include Chen, Roll and Ross (1986), Ferson and Harvey (1991), Chan Chen and Hsieh (1985), Jagannathan and Wang (1996) and Kramer (1994).

produce consistent estimated risk premiums for the three factor model. But estimated risk premiums from the Fama-Macbeth rolling beta method can have a bias as large as a 25% . The bias for classical methods is larger with 4,970 stocks than that for 25 or 149 portfolios given larger T-ratios. Theil's adjustment has a smaller bias for the size factor than the β IV method, but has a larger bias for the book-to-market factor.

*** Insert Table 3 here ***

These simulations imply that the finite sample bias in the estimated risk premium is relatively large when the sample period is small and the idiosyncratic risk is high. The β IV method can reduce the finite sample bias, especially when the sample period is small and the number of portfolios or stocks is large. Theil's adjustment is also successful in adjusting for the finite sample bias.

4.3 Macroeconomic Factor Premiums

In addition to traded factors, we also compare the estimated risk premiums of macroeconomic factors in Tables 4, 5 and 6, respectively. The results are consistent showing β IV method can reduce the finite sample bias.

*** Insert Tables 4, 5, and 6 here ***

We also compare the estimated risk premiums for 25 and 149 portfolios in Tables 4 and 5. The estimated risk premiums for the BJS method and the Fama-Macbeth method have larger bias than with the 4,970 stocks described below suggesting the grouping might not result in a smaller finite sample bias. But the results are more different for the β IV method and Theil's adjustment. Estimated risk premiums using these two methods are far from the true risk premiums. These two methods can produce unreasonable risk premiums for some simulations due to the fact that $\hat{\beta}_{t-\tau}\hat{\beta}_t'$ or $\hat{\beta}_1\hat{\beta}_1' - N\hat{T}_A$ is not positive definite when N is not large enough and the instrumental variables are weak. One method to deal with the weak instrument is to remove the observations of the stocks with instruments and factor loadings having different signs. However, further reducing the number of stocks is not applicable when this number is small. Thus, these methods can produce consistent estimators only when the number of stocks or portfolios is reasonably large. Using simulations, we find that when number of stocks is larger than 2000, these methods can adjust the bias. (These simulation results are available from the authors upon request.)

The simulation results with 4,970 stocks are shown in Table 6. Since there is substantial idiosyncratic risk for individual stocks, the estimation error in β is large in the first pass and the bias is large for both the BJS method and the Fama-Macbeth method. The estimated risk premiums are one third or one tenth of the true value of the risk premiums for these two methods, respectively. T-ratios from the classical methods are much larger than those from the macro factors, indicating that bias is larger in statistical sense as well. However, the estimators from the BJS method with Theil's adjustment and the β IV method are much closer to the true values. T-ratios are insignificant for β IV method, though it is significant for Theil's adjustment when T=60.

4.4 Time-varying factor loadings and regression residuals

The above simulations are all based on the constant factor loadings and constant volatility of the regression residuals. We also want to examine the various methods with time-varying factor loadings and regression residuals. To do this, we estimate the factor loadings from stock returns and factors using a rolling window of 30 months. The estimated factor loadings are assumed to be the true factor loadings in the simulations. Since the estimated factor loadings using rolling windows are time-varying, the true factor loadings in the simulations are also time-varying. Similarly, we can estimate the residuals from the regression, and create a pool of residuals. To bootstrap the residuals, we first select a time point, and randomly select one of the estimated residuals in a neighborhood (either within 15 or 30 months) of this time point from the residual pool. This provides time-varying residuals in the simulations.

With these simulated time-varying factor loadings and regression residuals, we can compare various methods. In particular, we are interested in comparing the β IV method with other existing methods. The advantage of the 3-group β IV method is that it is not significantly affected by the time-varying factor loadings. The results, shown in Tables 7 and 8, indicate that the 3-group β IV method yields the best estimates and smallest T-ratios among all other alternatives.

*** Insert Tables 7 and 8 here ***

4.5 Standard Errors

In this subsection, we compare the asymptotic standard error of the estimated risk premiums from three methods: the BJS method, β IV method and Fama-Macbeth rolling- β method.⁶ First we construct the “true” standard error of the risk premiums using the bootstrap, i.e., assuming that the standard error of risk premiums across 10,000 replications is the true standard error.⁷

We compare this true standard error with the method-specific estimated standard errors, including Fama-Macbeth errors with and without the adjusting for the autocovariances, Shanken’s adjustment and estimated errors from Theorem 3.2.

There are two different asymptotic covariance from Theorem 3.2. We can obtain \sqrt{NT} - asymptotic covariance V from

$$\sqrt{T}\sqrt{N}\left(\frac{1}{T-2\tau+1}\sum_{t=2\tau}^T(\hat{\gamma}_t - (0, f_t))' - (0, \gamma)'\right) \rightarrow N(0, V) .$$

⁶ The asymptotic standard error of the Theil’s adjustment is the Shanken adjustment, which is the same as the BJS method.

⁷ The constructed “true” standard error are similar for 1,000 simulations and 10,000 simulations, indicating the convergence of the standard error when the simulation number is approximately 1,000.

We can also obtain \sqrt{T} -asymptotic covariance $\tilde{\Sigma}_F$ using $\sqrt{T}(\frac{1}{T-2\tau+1} \sum_{t=2\tau}^T \hat{y}_t' - (0, \gamma)') \rightarrow N(0, \tilde{\Sigma}_F)$. Since it is more appealing to use an estimator with faster convergent rate, we will choose the first estimator, and construct \sqrt{NT} -convergent standard errors.

Theorem

The results for standard errors are based on 6 cases: Tables 9, 10 and 11 present the standard errors for the three Fama-French factors with 25 and 149 portfolios and 4,970 stocks respectively; Tables 12, 13 and 14 present the standard errors for macro-factors with 25 and 149 portfolios and 4,970 stocks.

4.6 Traded Factors

First, we consider the standard errors for the Fama-French three factor model. In Table 9 (when $T=600$ and $N = 25$.) for the BJS method, the standard error estimated through Shanken's adjustment is closer to the true standard error than the Fama-Macbeth standard error without autocovariance. For the β IV method, the Fama-Macbeth standard errors (with or without autocovariance) and the standard error from Theorem 3.2 are all close to the true standard error. For the Fama-Macbeth rolling- β method, the Fama-Macbeth standard errors (with or without autococariances) are smaller than the true value. The results are similar for both 149 portfolios (Table 10) and 4,970 stocks (Table 11).

*** Insert Tables 9, 10, and 11 here ***

When $T = 60$, the estimated standard errors are much farther away from the true standard errors compared to the $T = 600$ case. However, Shanken's covariance and the covariance implied by Theorem 3.2 are still the closest to the true covariance for the BJS method and β IV method, respectively. This is because estimated standard errors are less accurate with smaller T . Except for $N = 4,970$ and $T = 60$, the estimated Shanken's adjusted standard errors are much smaller than the true value with the BJS method because the idiosyncratic residuals of the individual stocks are much larger than the idiosyncratic residuals of portfolios.

4.7 Macro Factors

Tables 12 and 13 present standard errors for macro-factor models. Compared with the Fama-French three-factor model, the standard errors are larger. The reason is that for 25 and 149 portfolios, the idiosyncratic errors in the macro-factor model are larger than those in the Fama-French three-factor model. We find similar results for the BJS method and rolling- β method. For rolling- β IV method, since the estimated risk premiums are affected by the weak IVs, the estimator and the standard errors are both unreliable.

*** Insert Tables 12 and 13 here ***

The more interesting results are shown in Table 14 with 4,970 individual stocks. When $T = 60$, all of the estimated standard errors are much smaller than the true standard errors, suggesting that they are inaccurate when T is small. When $T = 600$, the Shanken adjustment and the Fama-Macbeth standard errors with autocovariance are closer to the true value than those from the BJS and the Fama-Macbeth rolling- β IV methods respectively, but they are still much smaller than the true value. The results indicate that for macro-factor models, even if $T = 600$, the estimated standard errors are not accurate.⁸ The standard error from Theorem 3.2 is the closest to the true value, and the Fama-Macbeth standard error with autocovariance is also close to the standard error from Theorem 3.2 with the β IV method.

*** Insert Table 14 here ***

Ang, Liu, and Schwarz (2010) show that if one groups stocks into portfolios to estimate risk premiums, the asymptotic covariance of the estimated risk premiums is larger. The simulated standard errors from Tables 12, 13 and 14 are consistent with their theoretical findings. More specifically, the standard errors decrease with the number of portfolios.

The comparison of the standard errors leads to the following conclusions. First, one should use large enough T to make the estimated standard errors closer to the true standard errors. For the Fama-French three-factor model, $T = 600$ is large enough and for the macro-factor model, T needs to be much larger than 600 (e.g. with daily data). For smaller T (e.g., with monthly rather than daily data) but reasonably large N , using the β IV method together with standard errors from Theorem 3.2 or the Fama-Macbeth standard error with autocovariance can be more appropriate for the macro-factor model.

4.8 T Statistics

Lewellen, Nagel and Shanken (2010) show that the explanatory power can be misleading for some asset pricing models. In this paper, we examine another important issue, the size of the cross-sectional regression. Specifically, we want to examine the probability of rejecting an asset pricing model when the model is correctly specified. Following Ferson and Foster (1994) and Shanken and Zhou (2007), we compare the probability of rejecting the null hypothesis, $\alpha = 0$ (α is the constant term), when it is true. To do this, for each simulation, we calculate the t-ratios for different methods and compare them with the 95% critical value of the standard normal distribution, 1.96. Then, we calculate the number of simulations in which absolute value of the t-ratio is greater than 1.96, and divided it by the total number of simulations to obtain the probability of rejecting the true null hypothesis (rejection rate).

Since the bias of the point estimators is large for some methods, we use a "bias-corrected" T-ratio. To do this, we estimate the risk premiums in the first 1,000 trials, and take the difference between the true value and the sample average of the estimated risk premiums

⁸ We did run simulations with T ; the estimated standard errors are much closer to the true values.

as the bias. Then we adjust this bias in the simulated risk premiums to obtain the bias-corrected T-ratios for another 1,000 trials. We compare the t-ratio with different methods to estimate point and standard errors in the Fama-French three-factor model and the macro-factor model with $N = 4970$ and $T = 600$. We choose the large values of the N and T because these are the cases with the smallest bias for the BJS and the β instrumental variables methods. The results are shown in Table 15.

*** Insert Table 15 here ***

From this Table, the rejection rate is too large for most of the tests reject. When the size of the statistics is correct, the rejection rate should be close 5% (it is 2.5% for the two-sided normal distribution, but we normalize the rejection rate so that 5% represents the correct size). However, for the Fama-French three-factor model, the β IV method and the BJS method have lower rejection rate than the Fama-Macbeth rolling- β IV. For instance, the largest rejection rate for the BJS method and the β IV method is smaller than 10%, but it is more than 15% for the Fama-Macbeth rolling- β IV method. This is because the estimated standard errors are smaller than the true standard errors. Most empirical research finds that the existing models are generally misspecified with the Fama-Macbeth rolling- β IV method.⁹ This paper suggests that this finding may be due to the statistical issues. For the β IV method, the standard error from Theorem 3.2 leads to a rejection rate of 4.70% which is close to the true rejection rate of 5%. All other standard errors lead to larger rejection rates. For the BJS method, the rejection rate is also larger than 5%.

The conclusions are similar with the macro-factor model. The only exception is the BJS method with the Shanken adjustment, where the rejection rate is too small (3.3%).

To conclude, the most reliable approach for testing whether an asset pricing model can be misspecified is to use the standard error from Theorem 3.2 in the β IV method. Using the Fama-Macbeth rolling- β method can result in misleading interpretations of α .

5 Application of these methods

According to the simulation results, the instrumental variable approach can effectively remove the bias in cross-sectional regressions. Thus, it is a natural to examine whether applying this method helps researchers identify the factors that explain the expected stock returns. Instead of the traded factors, we apply this method to macro factors. The macro factors arguably affect the stock returns. However, as we can see from the simulations, the estimated risk premiums with classical methods for these macro factors are much more biased than those for the traded factors; thereby, it is more difficult for researchers to identify significant risk premiums. The instrumental variable method, which can adjust for the bias, is more likely to identify these risk factors. Moreover, the risk premiums of the

⁹ We consider the bias-corrected T-ratio here. If the bias is not correct, one can expect a higher rejection ratio with the Fama-Macbeth rolling- β method and the BJS method.

macro factors can only be estimated using the cross-sectional regression approach, while the risk premiums of the traded factors can be estimated using the sample average of the excess returns (Shanken (1992)). Therefore, the macro-factor model provides perfect specification to study various cross-sectional regression methods.

We create the same macro factors from 1964 to 2010 as before. The monthly individual stock returns for the same time horizon are available from CRSP. Since many stocks only exist for a short horizon, we exclude these short-lived stocks: i.e. stocks with less than 90 of months return data to have large enough data (for example 30 months) to estimate factor loadings in the 3-sample method. Moreover, the macro factors are less likely to significantly affect the returns of the short-lived stocks since the majority of these returns should be influenced by idiosyncratic shocks. With these returns and the factors, we estimate the risk premiums and T-statistics of different cross-sectional regression methods: BJS method, Fama-Macbeth rolling- β method, lagged Rolling IV variable method, 3-group IV method and Thiel's adjustment. For all two instrumental variable methods, the issue of weak instrumental variables can affect the estimated risk premiums. Thus, we drop the observations of the stocks for which the estimated β instruments and estimated factor loadings from two subsamples have opposite signs. The estimated risk premiums and the T-statistics are shown in Tables 16 and 17.

*** Insert Table 16 and 17 here ***

We find that the estimated risk premiums with the classical methods are generally smaller than the instrumental variables approach. For example, when the second pass is OLS, the risk premium estimates of consumption growth are 0.006 and -0.001 with BJS and Fama-Macbeth approach, respectively. Nevertheless, the estimated risk premiums for the rolling- β IV method and the 3-group IV method are 0.097 and 0.028, which are much larger. The T-statistics are also larger. For example, we find that the consumption growth can significantly affect the stock returns with two instrumental variable methods, but for the classical approaches, we cannot find such result. The findings from Thiel's adjustment are also noteworthy. When the second pass is OLS, although some estimators are even larger than those with instrumental variable approaches, the T-ratios do not lead to rejection of the null hypothesis. Moreover, with GLS as the second pass, we find that the consumption growth estimated from Thiel's adjustment negatively affects the stock returns, which is inconsistent with the findings with OLS estimators. Surprisingly, a shock in industry production negatively affects stock returns. The only significant coefficients are estimated using the lagged instrumental variable approach (for both OLS and GLS), the Fama-Macbeth (OLS), and the BJS (GLS). Thiel's adjustment using GLS is the only one that results in positive and significant risk premium from industry production shock. However, the result is not very reliable: (1) Thiel's adjustment estimator with GLS is much smaller (in absolute value) than classical methods with GLS; thus, it is not clear whether the Thiel's adjustment can control the errors-in-variables (attenuation) bias. (2) Theoretically, the OLS and GLS estimators should be close when the sample size is large enough. For the IV methods, the GLS and OLS estimators are close to each other. However, for the Thiels' adjustment, the GLS estimator are very different from the OLS estimator. (3) The OLS and

GLS estimators lead to inconsistent results for Industry production in terms of the statistical significance. Lastly, the shock in inflation is insignificant for almost all methods; although, the sign is negative.

6 Conclusion

This paper suggests an adjustment for the Fama-Macbeth β IV method. Estimated β 's from non-overlapping observations can serve as effective instruments and mitigate or entirely eliminate the errors-in-variables bias. For the cases of constant β and time-varying β , we prove consistency and derive the asymptotic distribution of the estimated risk premium when the number of portfolios N is large.

When β is a general function of conditioning information, we use simulations to compare the β IV method with the traditional BJS method and the traditional Fama-Macbeth method. For macro factors, the bias is large for the latter two methods when the number of portfolio is large and the total sample period T is small. The β IV method and Theil's adjustment method are consistent and the estimated risk premiums are much closer to the true value even when T is small (e.g. $T=60$). This is due to the \sqrt{NT} -convergent rate of these estimators. As long as N is large, the estimated risk premiums have much smaller bias even if T is small. Moreover, the traditional and still widely used Fama-Macbeth method has the largest bias among the three methods. Using the lagged estimated β as the instrumental variable corrects this bias significantly.

In addition, this paper provides the standard error estimators (Theorem 3.2) and the Fama-Macbeth standard error with autocovariance that are superior to alternative estimators when T and N are relatively large. We conduct simulations to evaluate the various standard errors and t-ratios for tests on the null hypothesis that $\alpha = 0$. The results show that most of the tests reject the null hypothesis too often in Fama-French and macro-factor models. The β IV method, combined with the standard error from Theorem 3.2, provides the best results.

Finally, the empirical applications show that the new approach can indeed reduce the error-in-variables bias and help researchers identify the factors that can affect the stock returns, while the bias drawn from other methods are either large or inconsistent.

References

- Ang Andrew, Jun Liu, and Krista Schwarz 2010, Using Stocks or Portfolio in Tests of Factor Models, working paper.
- Bansal Ravi, 2004, Long Run Risks and Risk Compensation in Equity Markets, working paper, Duke University.
- Bansal Ravi and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481-1509.
- Biorn Eric, 2000, Panel data with measurement errors: instrumental variables and GMM procedures combining levels and differences, *Econometrics Review*, 19, 391-424.
- Black Fisher, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, Michael C. Jensen, ed: *Studies in the Theory of Capital Markets*, 79-121.
- Blume Marshall, and Irwin Friend, 1973, A New Look at the Capital Asset Pricing Model, *Journal of Finance*, 28, 19-34.
- Braun Phillip, Daniel Nelson and Alain Sunier, 1995, Good news, bad news volaitiltiy and betas, *Journal of Finance*, 50, 1575-1604.
- Chan K. C., Nai-Fu Chen, and David A. Hsieh, 1985, An Exploratory Investigation of the Firm Size Effect, *Journal of Financial Economics*, 14, 451-471.
- Chen Robert, and Raymond Kan, 2004, Finite Sample Analysis of Two-Pass Cross-Sectional Regressions.
- Chen Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic Forces and the Stock Market, *Journal of Business*, 59, 383-403.
- Cochrane John H, 2001, *Asset Pricing*, (Princeton University Press).
- Fama Eugene F, and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.
- Fama Eugene F, and James D. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.
- Ferson Wayne E, and Stephen R. Foster, 1994, Finite sample properties of the generalized method of moments in tests of conditional asset pricing models, *Journal of Financial Economics*, 36, 29-55.
- Ferson Wayne E, and Campbell Harvey, 1991, The variation of economic risk

premiums, *Journal of Political Economy*, 99, 385–415.

Ferson Wayne E, and Robert A. Korajczyk, 1995, Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business*, 68, 309-349.

Ferson Wayne E, and Campbell R. Harvey, 1999, Conditioning Variables and Cross-section of Stock Returns, *Journal of Finance* 54, 1325-1360.

Foster Dean P, and Dan B. Nelson, 1996, Continuous Record Asymptotics for Rolling Sample Variance Estimators, *Econometrica*, 64, 139-174.

Griliches Zvi, and Jerry A. Hausman, 1986, Error in variables in panel data, *Journal of Econometrics*, 31, 93-118.

Gibbons Michael R, 1982, Multivariate tests of Financial models: A new approach, *Journal of Financial Economics*, 10, 3-27.

Gibbons Michael R, Stephen A. Ross, and Jay. Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica*, 57, 1121-1152.

Ghysels Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance*, 53, 549-573.

Hansen Lars P, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029-1054.

Hansen Lars P, John Heaton, and Amir Yaron, 1996, Finite-sample properties of some alternative gmm estimators, *Journal of Business and Economic Statistics*, 14, 262-280.

Hansen Lars P, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy*, 99, 225-262.

Hansen Lars P, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance*, 62, 557-590.

Hansen Lars P, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica*, 55, 587-613.

Hansen Lars P, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica*, 50, 1269-1286.

Jagannathan Ravi, and Zhenyu Wang, 1996, The conditional capm and the cross-section of expected returns, *Journal of Finance*, 51, 3-53.

Jagannathan Ravi, and Zhenyu Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance*, 53, 1285-1309.

Jagannathan Ravi, and Zhenyu Wang, 2002, Empirical evaluation of asset-pricing models: A comparison of the sdf and beta methods, *Journal of Finance*, 57, 2337-2367.

James Willard, and Charles. Stein, 1961, Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* 1, 361-379.

Kan Raymond, Cesare Robotti and Jay Shanken, 2010, Pricing model performance and the two-pass cross-sectional regression methodology.

Kan Raymond and Chu Zhang, 1999, Two-pass tests of asset pricing models with useless factors, *Journal of Finance* 54, 203-235.

Kan Raymond, and Guofu Zhou, 1999, A critique of the stochastic discount factor methodology, *Journal of Finance*, 54, 1021-1048.

Kamer Charles, 1994, Macroeconomic Seasonality and the January Effect, *Journal of Finance*, 49, 1883-1891.

Kleibergen Frank, 2009, Tests of risk premia in linear factor models, *Journal of Econometrics*, 149, 149-173.

Lewellen Jonathan, Stefan Nagel and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics*, 96, 175-194.

Litzenberger Robert H, and Krishna Ramaswamy, 1979, The effect of personal taxes and dividends of capital asset prices: The theory and evidence, *Journal of Financial Economics* 7, 163-196.

Merton Robert C, 1973, An intertemporal asset pricing model, *Econometrica* 41, 867-888.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Rosenberg, Barr and Vinay Marathe, 1975, Test of asset pricing hypothesis, *Research in Finance* 1, 115-223.

Shanken Jay, 1985, Multivariate tests of the zero-beta capm, *Journal of Financial Economics*, 14, 327-348.

Shanken Jay, 1987, Multivariate proxies and asset pricing relations : Living with

the roll critique, *Journal of Financial Economics*, 18, 91-110.

Shanken Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies*, 5, 1-33.

Shanken Jay, and Guofu Zhou, 2007, Estimating and testing beta pricing models: alternative methods and their performance in simulations, *Journal of Financial Economics*, 84, 40-86.

Stein Charles, 1956, Inadmissibility of the usual estimator for the mean of a multivariate distribution, *Proc. Third Berkeley Symp. Math. Statist. Prob.* 1, 197-206.

Welch Ivo, 2000, Views of financial economists on the equity premium and on professional controversies, *Journal of Business* 73, 501-537.

Welch Ivo, 2008, The consensus estimate for the equity premium by academic financial economists in december 2007, Working Paper, Brown University.

Appendix: Proof of the Theorems

Proof of Theorem 3.1.

We prove the consistency and asymptotic theorem for the case using the lagged estimated factor loadings as instrumental variable. The case using the general instrumental variable can be proved in the exactly same way.

To illustrate the prove, let the second step be OLS, i.e. $\Sigma = \mathbf{I}$ one has:

$$\hat{\gamma}_t' - (0, \gamma + f_t)' = (\hat{\beta}1_{t-\tau} \hat{\beta}1_t')^{-1} (\hat{\beta}1_{t-\tau} \xi_t').$$

The consistency is established since $E(\xi_t' \hat{\beta}1_{t-\tau}') = \mathbf{0}$ and the Lindeberg condition.

Since $\beta\beta'/N$ converges to bb' when $N \rightarrow \infty$ and $[\beta_1 \xi_{1t}, \dots, \beta_N \xi_{Nt}]$ satisfies the Lindeberg condition, one can apply the Lindeberg Central Limit Theorem.

Define $u_t = (F_t' F_t)^{-1} F_t' \Omega_t$. To get the asymptotic covariance, notice that as $N \rightarrow \infty$,

$$1/N(\hat{\beta}1_{t-\tau} \hat{\beta}1_t') \rightarrow bb',$$

and on the other hand,

$$\begin{aligned} & E(1/N(\hat{\beta}1_{t-\tau} \xi_t' \xi_t \hat{\beta}1_{t-\tau}') | F) \\ &= E\left(\frac{1}{N}(\beta 1_{t-\tau}(-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t)\beta 1_t') | F\right) \\ &+ E\left(\frac{1}{N}(u_{t-\tau}(-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t)u_{t-\tau}') | F\right) \\ &\rightarrow \frac{1}{N}\beta 1 E((-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t) | F)\beta 1' \\ &+ \frac{1}{N}E(\tilde{u}_{t-\tau} E_{t-\tau}((-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t) | F)\tilde{u}_{t-\tau}' | F). \end{aligned}$$

$E_{t-\tau}(\cdot | F)$ takes the expected value of a random variable at time $t - \tau$ conditioning on all the information F and $\tilde{u}_{t-\tau}' = [\mathbf{0}'_{1 \times N}, u_{t-\tau}']$. One can show that¹⁰

$$\begin{aligned} & E_{t-\tau}((-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t) | F) \\ &= E((-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t) | F) = c_0 \mathbf{I}, \end{aligned}$$

hence,

$$E(1/N(\hat{\beta}1_{t-\tau} \xi_t' \xi_t \hat{\beta}1_{t-\tau}') | F)$$

¹⁰We will show that in the next page.

$$\rightarrow c_0 \mathbf{b}\mathbf{b}' + c_0 \tilde{\mathbf{L}}_{0,t-\tau}.$$

Then the asymptotic covariance matrix can be written as

$$\text{Acov}(\hat{\gamma}_t - (\gamma + (0, f_t)) | F) = c_0 (\mathbf{b}\mathbf{b}')^{-1} (\mathbf{b}\mathbf{b}' + \tilde{\mathbf{L}}_{0,t-\tau}) (\mathbf{b}\mathbf{b}')^{-1}.$$

In the second step is GLS, one has:

$$\hat{\gamma}_t' - (0, \gamma + f_t)' = (\hat{\beta} 1_{t-\tau} \Sigma^{-1} \hat{\beta} 1_t')^{-1} (\hat{\beta} 1_{t-\tau} \Sigma^{-1} \xi_t').$$

The consistency can be proved in same way as before. To derive the asymptotic covariance, one can show that

$$\begin{aligned} & E\left(\frac{1}{N} \hat{\beta} 1_{t-\tau} \Sigma^{-1} \xi_t' \xi_t \Sigma^{-1} \hat{\beta} 1_{t-\tau}' | F\right) \\ &= E\left(\frac{1}{N} \beta 1 \Sigma^{-1} \beta 1' + \frac{1}{N} \tilde{u}_{t-\tau} \Sigma^{-1} \tilde{u}_{t-\tau}' | F\right) \rightarrow c_0 (\mathbf{b} \Sigma^{-1} \mathbf{b}' + \tilde{\mathbf{L}}_{0,t-\tau}), \end{aligned}$$

so the conditional asymptotic distribution can be derived similarly as before.

Note that the key step in this proof is that

$$E(\beta(-(\gamma + f_t)u_t + \varepsilon_t)'(-(\gamma + f_t)u_t + \varepsilon_t)\beta' | F) = c_0 \beta \beta'.$$

This result follows the proof in Shanken (1992). The details are shown below:

Since $u_t = (F_t' F_t)^{-1} F_t' \Omega_t$,

$$u_t'(\gamma + f_t)' = (I \otimes (\gamma + f_t)(F_t' F_t)^{-1} F_t') \text{Vec}(\tilde{E}_t),$$

where $\text{Vec}(\tilde{E}_t)$ reshapes the $\tau \times N$ matrix \tilde{E}_t into the a $\tau \times N$ column vector, i.e

$$\text{Vec}(\tilde{E}_t) = (\varepsilon_{t-\tau,1} - \bar{\varepsilon}_1^t, \dots, \varepsilon_{t,1} - \bar{\varepsilon}_1^t, \varepsilon_{t-\tau,2} - \bar{\varepsilon}_2^t, \dots, \varepsilon_{t,2} - \bar{\varepsilon}_2^t, \dots, \varepsilon_{t-\tau,N} - \bar{\varepsilon}_N^t, \dots, \varepsilon_{t,N} - \bar{\varepsilon}_N^t)',$$

with $\bar{\varepsilon}_i^t$ the average of $\varepsilon_{t-\tau,i}, \dots, \varepsilon_{t,i}$.

Applying this formula, one has

$$\begin{aligned} & E(u_t'(\gamma + f_t)'(\gamma + f_t)u_t | F) \\ &= (I \otimes (\gamma + f_t)(F_t' F_t)^{-1} F_t')(I \otimes L_\tau)(I \otimes (\gamma + f_t)(F_t' F_t)^{-1} F_t')' \\ &= (\gamma + f_t)(F_t' F_t)^{-1} F_t' L_\tau ((\gamma + f_t)(F_t' F_t)^{-1} F_t')' I. \end{aligned}$$

Using this method, one has

$$E\left(\frac{1}{N}\boldsymbol{\beta}1'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)\boldsymbol{\beta}1' \mid \boldsymbol{F}\right) = c_0\boldsymbol{b}\boldsymbol{b}',$$

similarly, one can show that

$$\begin{aligned} & E\left(\frac{1}{N}\boldsymbol{u}_{t-\tau}'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)\boldsymbol{u}_{t-\tau}' \mid \boldsymbol{F}\right) \\ &= c_0(\boldsymbol{F}_{t-\tau}'\boldsymbol{F}_{t-\tau})^{-1}\boldsymbol{F}_{t-\tau}'\boldsymbol{L}_\tau\boldsymbol{F}_{t-\tau}(\boldsymbol{F}_{t-\tau}'\boldsymbol{F}_{t-\tau})^{-1} \end{aligned}$$

Proof of Theorem 3.2.

From Theorem 3.1,

$$\frac{1}{T-2\tau+1}\sum_{t=2\tau}^T\hat{\boldsymbol{y}}_t' = \boldsymbol{\gamma}' + \frac{1}{T-2\tau+1}\sum_{t=2\tau}^T(0, -\boldsymbol{f}_t)',$$

the consistency is established because $\frac{1}{T-2\tau+1}\sum_{t=2\tau}^T(0, -\boldsymbol{f}_t)' \rightarrow 0$ when T is large.

When the second step is OLS, to derive the unconditional asymptotic distribution, notice that

$$\begin{aligned} & \frac{1}{T-2\tau+1}\left(\sum_{t=2\tau}^T\hat{\boldsymbol{y}}_t' - \boldsymbol{\gamma}'\right) \\ &= \frac{1}{T-2\tau+1}\sum_{t=2\tau}^T((\hat{\boldsymbol{\beta}}_{t-\tau}'\hat{\boldsymbol{\beta}}_{t-\tau})^{-1}(\hat{\boldsymbol{\beta}}_{t-\tau}'\boldsymbol{\xi}_t') + (0, \boldsymbol{f}_t)'). \end{aligned}$$

where $\boldsymbol{\xi}_t' = -\boldsymbol{u}_t'(\boldsymbol{\gamma} + \boldsymbol{f}_t)' + \boldsymbol{\varepsilon}_t'$ and $\hat{\boldsymbol{\beta}}_{t-\tau}' = \boldsymbol{\beta}' + \tilde{\boldsymbol{u}}_{t-\tau}'$.

By the assumption that \boldsymbol{f}_t is a stationary process and $\boldsymbol{\beta}$ satisfies the Lindeberg condition, it is clear that one can apply the Central Limit Theorem to derive the asymptotic distribution of $\sum_{t=2\tau}^T(\hat{\boldsymbol{y}}_t - (0, \boldsymbol{f}_t))'$, i.e.

$$\sqrt{N}\sqrt{T}\frac{1}{T-2\tau+1}\left(\sum_{t=2\tau}^T(\hat{\boldsymbol{y}}_t - (0, \boldsymbol{f}_t))' - (0, \boldsymbol{\gamma})'\right) \rightarrow N(0, \boldsymbol{V}).$$

The asymptotic variance can be written as the summation of variance and autocovariance of the error term.

To be more specific, first, notice that conditional on \boldsymbol{F} :

$$E\left(\frac{1}{N}\hat{\boldsymbol{\beta}}_{t-\tau}'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)'(-(\boldsymbol{\gamma} + \boldsymbol{f}_t)\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t)\hat{\boldsymbol{\beta}}_{t-\tau}' \mid \boldsymbol{F}\right)$$

$$\rightarrow c_0 \mathbf{b} \mathbf{b}' + c_0 \tilde{\mathbf{L}}_{0,t-\tau}.$$

In addition, for any integer t_1 between -1 and $\tau-1$,

$$\begin{aligned} & E\left(\frac{1}{N} \hat{\boldsymbol{\beta}}_{t-\tau} (-(\boldsymbol{\gamma} + \mathbf{f}_t) \mathbf{u}_t + \boldsymbol{\varepsilon}_t)' (-(\boldsymbol{\gamma} + \mathbf{f}_{t-t_1}) \mathbf{u}_{t-t_1} + \boldsymbol{\varepsilon}_{t-t_1}) \hat{\boldsymbol{\beta}}_{t-t_1} \mid \mathbf{F}\right) \\ & \rightarrow c_{t_1} \mathbf{b} \mathbf{b}' + c_{t_1} \tilde{\mathbf{L}}_{t_1,t-\tau}^1, \end{aligned}$$

where $c_{t_1} = ((\boldsymbol{\gamma} + \mathbf{f}_t) \tilde{\mathbf{L}}_{t_1,t} (\boldsymbol{\gamma} + \mathbf{f}_{t-t_1})' - (\boldsymbol{\gamma} + \mathbf{f}_t) (\mathbf{F}_t' \mathbf{F}_t)^{-1} \mathbf{F}_t' \mathbf{I}_{t_1+1})$. Here,

$$\tilde{\mathbf{L}}_{t_1,t-\tau}^1 = \begin{bmatrix} 0 & \mathbf{0}_{1 \times k} \\ \mathbf{0}_{k \times 1} & \mathbf{L}_{t_1,t-\tau}^1 \end{bmatrix},$$

$$\mathbf{L}_{t_1,t-\tau}^1 = (\mathbf{F}_{t-\tau}' \mathbf{F}_{t-\tau})^{-1} \mathbf{F}_{t-\tau}' \mathbf{M}_x \mathbf{F}_{t-\tau-t_1} (\mathbf{F}_{t-\tau-t_1}' \mathbf{F}_{t-\tau-t_1})^{-1},$$

where $\mathbf{M}_x = \mathbf{I}(-t_1) \mathbf{L}_{\tau-|t_1|} + \mathbf{I}(t_1) \mathbf{L}_{\tau-|t_1|}'$ with

$$\mathbf{L}_{\tau-|t_1|} = \begin{bmatrix} -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & \frac{\tau-t_1}{\tau^2} & \dots & \frac{\tau-t_1}{\tau^2} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & \dots & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \frac{\tau-t_1}{\tau^2} \\ 1 - \frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} \\ -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & 1 - \frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & 1 - \frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} & \dots & -\frac{1}{\tau} + \frac{\tau-t_1}{\tau^2} \end{bmatrix}_{\tau \times \tau},$$

the term $1 - \frac{1}{\tau} + \frac{\tau-t_1}{\tau^2}$ in this matrix are the $(t_1+1,1), (t_1+2,2), \dots, (\tau, \tau-t_1)$ 'th elements, $\frac{\tau-t_1}{\tau^2}$'s are in the upper triangular portion of the $\tau \times \tau$ matrix. (The triangular matrix is from $(1, \tau-|t_1|+1)$ to (t_1, τ) to $(1, \tau)$ 'th element.), and

$$\mathbf{I}_{|t_1|+1} = \left(-\frac{\tau-t_1}{\tau^2}, \dots, 1 - \frac{\tau-t_1}{\tau^2}, \dots, -\frac{\tau-t_1}{\tau^2}\right)',$$

where one is the $|t_1|+1$ 'th element. In addition, \mathbf{I} is an indicator function: $\mathbf{I}(x) = 1$ if $x > 0$, $\mathbf{I}(x) = 0$ if $x < 0$ and $\mathbf{I}(x) = \frac{1}{2}$ if $x = 0$.

The final asymptotic variance is as follows:

$$\mathbf{V} = (\mathbf{b}\mathbf{b}')^{-1} E\left(\sum_{t_l=-\tau+1}^{\tau-1} (c_{t_l}(\mathbf{b}\mathbf{b}' + \tilde{\mathbf{L}}_{t_l, t-\tau}))\right)(\mathbf{b}\mathbf{b}')^{-1}.$$

When the second pass estimation is GLS,

$$\begin{aligned} & E\left(\frac{1}{N}(\hat{\boldsymbol{\beta}}\mathbf{1}_{t-\tau}'\boldsymbol{\Sigma}^{-1}(-(\boldsymbol{\gamma} + \mathbf{f}_t)\mathbf{u}_t + \boldsymbol{\varepsilon}_t)'(-(\boldsymbol{\gamma} + \mathbf{f}_{t-t_l})\mathbf{u}_{t-t_l} + \boldsymbol{\varepsilon}_{t-t_l})\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\beta}}\mathbf{1}_{t-\tau-t_l}') \mid \mathbf{F}\right) \\ & \rightarrow c_{t_l}(\mathbf{b}\boldsymbol{\Sigma}^{-1}\mathbf{b}' + E\left(\frac{1}{N}(\mathbf{u}_{t-\tau}'\boldsymbol{\Sigma}^{-1}\mathbf{u}_{t-\tau-t_l}') \mid \mathbf{F}\right), \end{aligned}$$

This implies that

$$\mathbf{V} = (\mathbf{b}\boldsymbol{\Sigma}^{-1}\mathbf{b}')^{-1} E\left(\sum_{t_l=-\tau+1}^{\tau-1} (c_{t_l}(\mathbf{b}\boldsymbol{\Sigma}^{-1}\mathbf{b}' + \tilde{\mathbf{L}}_{t_l, t-\tau}))\right)(\mathbf{b}\boldsymbol{\Sigma}^{-1}\mathbf{b}')^{-1}.$$

In addition, since

$$\sqrt{T} \frac{1}{T - 2\tau + 1} \sum_{t=2\tau}^T ((\hat{\boldsymbol{\beta}}_{t-\tau}\hat{\boldsymbol{\beta}}\mathbf{1}_t')^{-1}(\hat{\boldsymbol{\beta}}\mathbf{1}_{t-\tau}'\boldsymbol{\xi}_t')) = O\left(\frac{1}{\sqrt{N}}\right),$$

it is easy to show that

$$\sqrt{T}(\hat{\boldsymbol{\gamma}} - (0, \boldsymbol{\gamma})) \rightarrow N(0, \tilde{\boldsymbol{\Sigma}}_F).$$

Table 1
Three methods to estimate Fama-French three-factor risk premiums using the bootstrap
(25 portfolios, monthly data)

This Table uses the FF three-factor model to generate stock returns. The first pass is a time series regression of returns on market, size, and book-to-market factors for each asset, which produces beta estimates. The second pass regresses asset returns cross-sectionally on the market β , size β and book-to market β . This Table presents the estimated risk premiums with 25 portfolios using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data. 25 portfolios include 25 size and book-to-market portfolios for 1964 to 2009. In applying the rolling beta method, we assume the rolling window is 15 for T=60 and 60 for T=600. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values. The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | |
|--------------------------------|---------------|-------------|-----------|
| Factor | Market | Size | BM |
| True | 0.41 | 0.26 | 0.42 |
| BJS | 0.43 | 0.25 | 0.40 |
| T-Diff | 0.23 | -0.16 | -0.14 |
| Rolling | 0.45 | 0.23 | 0.35 |
| T-Diff | 0.44 | -0.32 | -0.49 |
| Rolling IV | 0.41 | 0.26 | 0.43 |
| T-Diff | -0.05 | -0.03 | 0.08 |
| T=600 Rolling window 60 | | | |
| True | 0.41 | 0.26 | 0.42 |
| BJS | 0.42 | 0.26 | 0.41 |
| T-Diff | 0.25 | -0.16 | -0.16 |
| Rolling | 0.43 | 0.25 | 0.40 |
| T-Diff | 0.69 | -0.49 | -0.43 |
| Rolling IV | 0.42 | 0.26 | 0.42 |
| T-Diff | -0.04 | 0.01 | 0.02 |

Table 2
Three methods to estimate Fama-French three-factor risk premiums using the bootstrap
(149 portfolios, monthly data)

This Table uses the FF three-factor model to generate stock returns. The first pass is a time series regression of returns on market, size, and book-to-market factors for each asset, which produces beta estimates. The second pass regresses asset returns cross-sectionally on the market β , size β and book-to market β . This Table presents the estimated risk premiums with 149 portfolios using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data. 149 portfolios include 100 size and book-to-market portfolios combined with 49 industry portfolios. In applying the rolling beta method, we assume the rolling window is 15 for T=60 and 60 for T=600. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values. The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | |
|--------------------------------|---------------|-------------|-----------|
| Factor | Market | Size | BM |
| True | 0.41 | 0.26 | 0.42 |
| BJS | 0.44 | 0.25 | 0.36 |
| T-Diff | 0.53 | -0.22 | -0.63 |
| Rolling | 0.51 | 0.18 | 0.22 |
| T-Diff | 1.19 | -0.70 | -1.21 |
| Rolling IV | 0.41 | 0.27 | 0.43 |
| T-Diff | -0.06 | 0.05 | 0.09 |
| T=600 Rolling window 60 | | | |
| True | 0.41 | 0.26 | 0.42 |
| BJS | 0.42 | 0.25 | 0.41 |
| T-Diff | 1.77 | -0.73 | -2.23 |
| Rolling | 0.44 | 0.24 | 0.35 |
| T-Diff | 1.62 | -1.10 | -2.31 |
| Rolling IV | 0.41 | 0.26 | 0.42 |
| T-Diff | -0.04 | -0.04 | -0.02 |

Table 3

Three methods to estimate Fama-French three-factor risk premiums using the bootstrap
(4,970 stocks, monthly data)

This Table uses the FF three-factor model to generate stock returns. The first pass is a time series regression of returns on market, size, and book-to-market factors for each asset, which produces beta estimates. The second pass regresses asset returns cross-sectionally on the market β , size β and book-to market β . This Table presents the estimated risk premiums with 4,970 individual stocks using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), lagged beta IV method (Rolling IV), and Theil adjustment (Theil). The estimation is based on monthly data from 1964 to 2009. In applying the rolling beta method, we assume the rolling window is 15 for T=60 and 60 for T=600. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values. The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | | |
|---------------------------------|-----------------|---------------|-------------|-----------|
| Factor | constant | Market | Size | BM |
| True | 0.00 | 0.41 | 0.26 | 0.42 |
| BJS | 0.07 | 0.32 | 0.15 | 0.29 |
| T-Diff | 1.20 | -1.16 | -0.97 | -1.06 |
| Rolling | 0.17 | 0.17 | 0.09 | 0.12 |
| T-Diff | 0.99 | -1.04 | -0.68 | -1.13 |
| Rolling IV | 0.02 | 0.36 | 0.19 | 0.47 |
| T-Diff | 0.47 | -1.15 | -1.08 | 0.77 |
| Theil | -0.00 | 0.36 | 0.22 | 0.53 |
| T-Diff | -0.05 | -0.67 | -0.38 | 0.96 |
| T=600 Rolling window =60 | | | | |
| True | 0.00 | 0.41 | 0.26 | 0.42 |
| BJS | 0.01 | 0.41 | 0.25 | 0.40 |
| T-Diff | 1.02 | -0.60 | -1.60 | -1.64 |
| Rolling | 0.06 | 0.37 | 0.19 | 0.27 |
| T-Diff | 2.90 | -1.72 | -1.89 | -3.68 |
| Rolling IV | 0.00 | 0.41 | 0.26 | 0.42 |
| T-Diff | 0.02 | -0.04 | 0.03 | -0.05 |
| Theil | -0.00 | 0.42 | 0.26 | 0.40 |
| T-Diff | -0.02 | 1.07 | 0.27 | -1.74 |

Table 4

Three methods to estimate macro three-factor risk premiums using the bootstrap
(25 portfolios, monthly data)

This Table presents the estimated risk premiums with 25 portfolios using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), lagged beta IV method (Rolling IV), and Theil's adjustment (Theil). The estimation is based on monthly data. 25 portfolios include 25 size and book-to-market portfolios for 1964 to 2009. In applying rolling beta method, we assume the rolling window is 15 for total period $T=60$ and 60 for $T=600$. The three factors are ΔC : consumption growth, ΔCPI : change in inflation, and ΔIP : change in industrial production. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values (True). The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| True | 0.00 | 0.20 | -0.10 | 1.20 |
| BJS | -0.03 | 0.02 | -0.03 | 0.09 |
| T-Diff | -0.18 | -3.39 | 1.72 | -5.39 |
| Rolling | -0.01 | 0.00 | -0.01 | 0.02 |
| T-Diff | -0.15 | -4.95 | 2.76 | -8.28 |
| Rolling IV | -0.17 | 0.31 | 0.93 | -1.43 |
| T-Diff | -0.00 | 0.01 | 0.01 | -0.01 |
| Theil | -0.26 | 0.66 | -1.09 | 0.97 |
| T-Diff | -0.54 | 1.01 | -1.78 | 3.11 |
| T=600 Rolling window 60 | | | | |
| True | 0.00 | 0.20 | -0.10 | 1.20 |
| BJS | -0.04 | 0.10 | -0.01 | 0.57 |
| T-Diff | -1.50 | -3.79 | 2.33 | -6.49 |
| Rolling | -0.07 | 0.02 | -0.03 | 0.09 |
| T-Diff | -4.08 | -16.09 | 9.94 | -24.63 |
| Rolling IV | 17.62 | -7.35 | -3.70 | 9.18 |
| T-Diff | 0.00 | -0.00 | -0.01 | 0.01 |
| Theil | 0.30 | 0.15 | -0.84 | 1.56 |
| T-Diff | 0.32 | 0.76 | -0.39 | 0.79 |

Table 5
Three methods to estimate macro three-factor risk premiums using the bootstrap
(149 portfolios, monthly data)

This Table presents the estimated risk premiums with 149 portfolios using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), the rolling beta IV method (Rolling IV), and Theil's adjustment (Theil). The estimation is based on monthly data from 1964 to 2009. 149 portfolios include 100 size and book-to-market portfolios combined with 49 industry portfolios. In applying rolling beta method, we assume the rolling window is 15 for total period T=60 and 60 for T=600. The three factors are ΔC : consumption growth, ΔCPI : change in inflation, ΔIP : change in industrial production. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values (True). The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.08 | 0.05 | -0.04 | 0.20 |
| T-Diff | -0.93 | -5.16 | 3.15 | -7.88 |
| Rolling | -0.08 | 0.01 | -0.01 | 0.05 |
| T-Diff | -0.86 | -6.15 | 3.72 | -10.46 |
| Rolling IV | 1.44 | -0.55 | 0.12 | -0.76 |
| T-Diff | 0.01 | -0.01 | 0.00 | -0.01 |
| Theil | -0.36 | 0.17 | 0.20 | 0.26 |
| T-Diff | -3.96 | -0.73 | 7.46 | -7.01 |
| T=600 Rolling window 60 | | | | |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.06 | 0.15 | -0.14 | 0.79 |
| T-Diff | -1.50 | -4.00 | 2.52 | -6.41 |
| Rolling | -0.12 | 0.04 | -0.04 | 0.19 |
| T-Diff | -3.93 | -16.09 | 9.94 | -24.63 |
| Rolling IV | 0.01 | 0.18 | 0.49 | 2.40 |
| T-Diff | 0.00 | -0.01 | 0.01 | 0.00 |
| Theil | 0.01 | 0.21 | -0.21 | 1.25 |
| T-Diff | 0.30 | 0.45 | -0.31 | 0.80 |

Table 6
Three methods to estimate macro three-factor risk premiums using the bootstrap
(4,970 stocks, monthly data)

This Table presents the estimated risk premiums with 4,970 individual stocks using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), the rolling beta IV method (Rolling IV), and Theil's adjustment (Theil). The estimation is based on monthly data from 1964 to 2009. In applying rolling beta method, we assume the rolling window is 15 for total period $T=60$ and 60 for $T=600$. The three factors are ΔC : consumption growth, ΔCPI : change in inflation, and ΔIP : change in industrial production. The true risk premiums (True) are the sample means of excess return factors. We also report the T-ratio (T-Diff) of the difference between reported risk premiums and their true values (True). The standard errors used to construct the T-ratio are based on the sample covariance of the estimated risk premiums from 10,000 trials.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.05 | 0.07 | -0.06 | 0.43 |
| T-Diff | -1.51 | -5.88 | 3.21 | -6.40 |
| Rolling | -0.04 | 0.02 | -0.02 | 0.14 |
| T-Diff | -1.25 | -6.64 | 3.90 | -9.84 |
| Rolling IV | -0.03 | 0.21 | -0.21 | 1.31 |
| T-Diff | -0.49 | 0.27 | -0.09 | 0.45 |
| Theil | -0.03 | 0.23 | -0.25 | 1.42 |
| T-Diff | -0.97 | 1.39 | -1.16 | 1.83 |
| T=600 Rolling window 60 | | | | |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.01 | 0.17 | -0.16 | 1.02 |
| T-Diff | -0.69 | -9.55 | 5.96 | -9.35 |
| Rolling | 0.35 | 0.07 | -0.06 | 0.45 |
| T-Diff | 36.89 | 17.00 | 10.22 | -19.26 |
| Rolling IV | -0.00 | 0.20 | -0.20 | 1.20 |
| T-Diff | -0.05 | -0.02 | -0.04 | 0.04 |
| Theil | -0.00 | 0.20 | -0.20 | 1.21 |
| T-Diff | -0.13 | 0.16 | -0.04 | 0.09 |

Table 7

Various methods to estimate Fama-French three-factor risk premiums with time-varying factor loadings and regression residuals (4,970 stocks monthly data)

This Table uses the FF three-factor model to generate stock returns with time-varying factor loadings and regression residuals, and presents the estimated risk premiums with 4,970 individual stocks using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), the rolling beta IV method (Rolling IV) and 3-group IV. We assume that the true risk premiums (True) are the sample mean of each excess return factor. The total periods are T=60 and 600. In applying the rolling beta method, we assume the rolling window is 15 for T=60 and 60 for T=600.

| T=60 Rolling window 15 | | | | |
|---------------------------------|-----------------|---------------|-------------|-----------|
| Factor | constant | Market | Size | BM |
| True | 0.00 | 0.41 | 0.26 | 0.42 |
| BJS | 0.11 | 0.28 | 0.12 | 0.25 |
| T-Diff | 1.71 | -1.20 | -0.95 | -1.09 |
| Rolling | 0.20 | 0.13 | 0.08 | 0.10 |
| T-Diff | 1.09 | -1.14 | -0.66 | -1.30 |
| Rolling IV | 0.02 | 0.33 | 0.17 | 0.49 |
| T-Diff | 0.49 | -1.25 | -1.05 | 0.87 |
| 3-group IV | 0.02 | 0.38 | 0.24 | 0.40 |
| T-Diff | 0.39 | -0.69 | -0.11 | -0.22 |
| Thiel | -0.00 | 0.33 | 0.18 | 0.53 |
| T-Diff | -0.06 | -0.75 | -0.35 | 1.03 |
| T=600 Rolling window =60 | | | | |
| True | 0.00 | 0.41 | 0.26 | 0.42 |
| BJS | 0.02 | 0.39 | 0.23 | 0.38 |
| T-Diff | 1.19 | -1.58 | -2.31 | -2.29 |
| Rolling | 0.10 | 0.35 | 0.18 | 0.23 |
| T-Diff | 3.88 | -2.82 | -1.91 | -5.38 |
| Rolling IV | 0.00 | 0.39 | 0.23 | 0.39 |
| T-Diff | 0.05 | -0.90 | -1.21 | -1.25 |
| 3-group IV | 0.00 | 0.41 | 0.25 | 0.42 |
| T-Diff | 0.03 | -0.36 | -0.21 | -0.30 |
| Thiel | -0.00 | 0.38 | 0.23 | 0.40 |
| T-Diff | -0.04 | -1.67 | -0.97 | -1.84 |

Table 8

Three methods to estimate macro three-factor risk premiums using the bootstrap with time-varying factor loadings and regression residuals (4,970 stocks, monthly data)

This Table uses the three macro-factor model to generate stock returns with time-varying factor loadings and regression residuals, and presents the estimated risk premiums with 4,970 individual stocks using the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), the rolling beta IV method (Rolling IV) and 3-group IV method. We assume that the true risk premiums (True) are the sample mean of each excess return factor. The total periods are $T=60$ and 600. In applying rolling beta method, we assume the rolling window is 15 for total period $T=60$ and 60 for $T=600$. The three factors are ΔC : consumption growth, ΔCPI : change in inflation, and ΔIP : change in industrial production.

| T=60 Rolling window 15 | | | | |
|---------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.07 | 0.06 | -0.05 | 0.33 |
| T-Diff | -1.91 | -6.18 | 3.89 | -8.03 |
| Rolling | -0.07 | 0.02 | -0.02 | 0.12 |
| T-Diff | -1.93 | -7.01 | 3.98 | -10.12 |
| Rolling IV | -0.04 | 0.22 | -0.24 | 1.38 |
| T-Diff | -0.54 | 0.40 | -0.42 | 0.63 |
| 3-group IV | -0.03 | 0.21 | -0.22 | 1.28 |
| T-Diff | -0.31 | 0.25 | -0.13 | 0.23 |
| Thiel | -0.04 | 0.25 | -0.27 | 1.48 |
| T-Diff | -1.08 | 1.55 | -1.45 | 2.02 |
| T=600 Rolling window =60 | | | | |
| True | 0.00 | 0.20 | -0.20 | 1.20 |
| BJS | -0.01 | 0.15 | -0.13 | 0.92 |
| T-Diff | -0.78 | -10.18 | 7.03 | -11.35 |
| Rolling | 0.55 | 0.07 | -0.05 | 0.40 |
| T-Diff | 48.29 | -17.39 | 10.90 | -19.89 |
| Rolling IV | -0.00 | 0.21 | -0.19 | 1.18 |
| T-Diff | -0.25 | -0.60 | -0.62 | -0.74 |
| 3-group IV | -0.00 | 0.20 | -0.19 | 1.20 |
| T-Diff | -0.08 | -0.11 | -0.40 | -0.21 |
| Thiel | -0.00 | 0.21 | -0.18 | 1.11 |
| T-Diff | -0.13 | 1.26 | -1.30 | -2.99 |

Table 9
The standard error for Fama-French three-factor risk premiums
(25 portfolios monthly data)

This Table presents the standard error of the estimated risk premiums for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data. 25 portfolios include 25 size and book-to-market portfolios for 1964 to 2009. The true standard error (True) is calculated using the bootstrapped covariance of the estimated risk premiums. The estimated standard error is the average of the bootstrapped Fama-Macbeth standard errors. The Fama-Macbeth standard errors contain the autocovariance (Auto) of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period T=60 and 60 for T=600. We also present Fama-Macbeth standard errors ignoring the autocovariance. For the BJS method, we also present the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | |
|--------------------------------|---------------|-------------|-----------|
| Factor | Market | Size | BM |
| BJS Method | | | |
| True | 0.07 | 0.09 | 0.11 |
| Estimated (No Auto) | 0.06 | 0.08 | 0.10 |
| Estimated (Shanken) | 0.06 | 0.08 | 0.10 |
| Rolling Method | | | |
| True | 0.08 | 0.11 | 0.13 |
| Estimated (No Auto) | 0.06 | 0.08 | 0.10 |
| Estimated (With Auto) | 0.06 | 0.08 | 0.09 |
| Rolling IV method | | | |
| True | 0.11 | 0.15 | 0.18 |
| Estimated (No Auto) | 0.09 | 0.12 | 0.15 |
| Estimated (With Auto) | 0.07 | 0.09 | 0.12 |
| Estimated (Theory 3.2) | 0.13 | 0.17 | 0.21 |
| T=600 Rolling window 60 | | | |
| BJS Method | | | |
| True | 0.02 | 0.03 | 0.03 |
| Estimated (No Auto) | 0.02 | 0.03 | 0.03 |
| Estimated (Shanken) | 0.02 | 0.03 | 0.03 |

| Rolling Method | | | |
|---------------------------|------|------|------|
| True | 0.02 | 0.03 | 0.04 |
| Estimated (No Auto) | 0.02 | 0.03 | 0.03 |
| Estimated (With Auto) | 0.02 | 0.03 | 0.03 |
| Rolling IV method | | | |
| True | 0.02 | 0.03 | 0.04 |
| Estimated (No Auto) | 0.02 | 0.03 | 0.04 |
| Estimated (With Auto) | 0.02 | 0.03 | 0.04 |
| Estimated (Theory 3.2) | 0.02 | 0.03 | 0.04 |

Table 10
The standard error for Fama-French three-factor risk premiums
(149 portfolios monthly data)

This Table presents the standard error of the estimated risk premiums for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data from 1964 to 2009. 149 portfolios include 100 size and book-to-market portfolios combined with 49 industry portfolios. The true standard error (True) is calculated through taking the covariance of the estimated risk premiums in each simulation. The estimated standard error is the average of the Fama-Macbeth standard errors in each simulation. The Fama-Macbeth standard errors contain the autocovariance of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period T=60 and 60 for T=600. We also compare the Fama-Macbeth standard errors without the autocovariance. For BJS method, we also compare the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | |
|--------------------------------|---------------|-------------|-----------|
| Factor | Market | Size | BM |
| BJS Method | | | |
| True | 0.05 | 0.07 | 0.09 |
| Estimated (No Auto) | 0.04 | 0.06 | 0.07 |
| Estimated (Shanken) | 0.04 | 0.06 | 0.07 |
| Rolling Method | | | |
| True | 0.08 | 0.12 | 0.16 |
| Estimated (No Auto) | 0.05 | 0.07 | 0.09 |
| Estimated (With Auto) | 0.04 | 0.06 | 0.07 |
| Rolling IV method | | | |
| Estimated (No Auto) | 0.07 | 0.10 | 0.12 |
| Estimated (With Auto) | 0.05 | 0.08 | 0.10 |
| Estimated (Theory 3.2) | 0.09 | 0.13 | 0.17 |
| T=600 Rolling window 60 | | | |
| BJS Method | | | |
| True | 0.01 | 0.02 | 0.03 |
| Estimated (No Auto) | 0.01 | 0.02 | 0.02 |
| Estimated (Shanken) | 0.01 | 0.02 | 0.03 |

| Rolling Method | | | |
|---------------------------|------|------|------|
| True | 0.02 | 0.02 | 0.03 |
| Estimated (No Auto) | 0.01 | 0.02 | 0.02 |
| Estimated (With Auto) | 0.01 | 0.02 | 0.03 |
| Rolling IV method | | | |
| True | 0.02 | 0.02 | 0.03 |
| Estimated (No Auto) | 0.02 | 0.02 | 0.03 |
| Estimated (With Auto) | 0.02 | 0.02 | 0.03 |
| Estimated (Theory 3.2) | 0.02 | 0.02 | 0.03 |

Table 11
The standard error for Fama-French three-factor risk premiums
(4,970 portfolios monthly data)

This Table presents the standard error of the estimated risk premiums for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The true standard error (True) is calculated through taking the covariance of the estimated risk premiums in each simulation. The estimated standard error is the average of the Fama-Macbeth standard errors in each simulation. The Fama-Macbeth standard errors contain the autocovariance of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period $T=60$ and 60 for $T=600$. We also compare the Fama-Macbeth standard errors without the autocovariance. For BJS method, we also compare the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|---------------|-------------|-----------|
| Factor | constant | Market | Size | BM |
| BJS Method | | | | |
| True | 0.06 | 0.08 | 0.12 | 0.12 |
| Estimated (No Auto) | 0.02 | 0.02 | 0.02 | 0.02 |
| Estimated (Shanken) | 0.02 | 0.02 | 0.02 | 0.02 |
| Rolling Method | | | | |
| True | 0.17 | 0.24 | 0.25 | 0.26 |
| Estimated (No Auto) | 0.04 | 0.06 | 0.06 | 0.06 |
| Estimated (With Auto) | 0.08 | 0.11 | 0.11 | 0.12 |
| Rolling IV method | | | | |
| True | 0.04 | 0.05 | 0.07 | 0.07 |
| Estimated (No Auto) | 0.03 | 0.04 | 0.06 | 0.06 |
| Estimated (With Auto) | 0.02 | 0.03 | 0.05 | 0.05 |
| Estimated (Theory 3.2) | 0.04 | 0.05 | 0.06 | 0.06 |
| T=600 Rolling window 60 | | | | |
| BJS Method | | | | |
| True | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (No Auto) | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (Shanken) | 0.01 | 0.01 | 0.01 | 0.01 |

| Rolling Method | | | | |
|---------------------------|------|------|------|------|
| True | 0.02 | 0.03 | 0.04 | 0.04 |
| Estimated (No Auto) | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (With Auto) | 0.01 | 0.02 | 0.03 | 0.03 |
| Rolling IV method | | | | |
| True | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (No Auto) | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (With Auto) | 0.01 | 0.01 | 0.01 | 0.01 |
| Estimated (Theory 3.2) | 0.01 | 0.01 | 0.01 | 0.01 |

Table12

The standard error for macro-factor risk premiums (25 portfolios monthly data)

This Table presents the standard error of the estimated risk premiums for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data. 25 portfolios include 25 size and book-to-market portfolios for 1964 to 2009. The true standard error (True) is calculated through taking the covariance of the estimated risk premiums in each simulation. The estimated standard error is the average of the Fama-Macbeth standard errors in each simulation. The Fama-Macbeth standard errors contain the autocovariance of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period T=60 and 60 for T=600. We also compare the Fama-Macbeth standard errors without the autocovariance. For BJS method, we also compare the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-------------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| BJS Method | | | | |
| True | 0.20 | 0.06 | 0.10 | 0.21 |
| Estimated (No Auto) | 0.17 | 0.04 | 0.08 | 0.16 |
| Estimated (Shanken) | 0.18 | 0.05 | 0.09 | 0.21 |
| Rolling Method | | | | |
| True | 0.20 | 0.04 | 0.07 | 0.14 |
| Estimated (No Auto) | 0.17 | 0.03 | 0.04 | 0.09 |
| Estimated (With Auto) | 0.15 | 0.03 | 0.04 | 0.09 |
| Rolling IV method | | | | |
| True | 162.46 | 41.32 | 62.53 | 118.45 |
| Estimated (No Auto) | 162.49 | 41.32 | 62.52 | 118.52 |
| Estimated (With Auto) | 107.01 | 27.83 | 45.40 | 86.04 |
| Estimated (Theory 3.2) | DNE ¹¹ | DNE | 1.56 | DNE |
| T=600 Rolling window 60 | | | | |
| BJS Method | | | | |
| True | 0.10 | 0.04 | 0.09 | 0.18 |
| Estimated (No Auto) | 0.07 | 0.03 | 0.06 | 0.12 |
| Estimated | 0.10 | 0.05 | 0.08 | 0.24 |

¹¹ DNE represents does not exist—the covariance is negative.

| | | | | |
|--------------------------|---------|--------|--------|---------|
| (Shanken) | | | | |
| Rolling Method | | | | |
| True | 0.06 | 0.02 | 0.03 | 0.07 |
| Estimated | 0.06 | 0.02 | 0.03 | 0.05 |
| (No Auto) | | | | |
| Estimated | 0.06 | 0.02 | 0.03 | 0.06 |
| (With Auto) | | | | |
| Rolling IV method | | | | |
| True | 1168.50 | 444.60 | 490.40 | 1458.00 |
| Estimated | 1168.50 | 444.60 | 490.40 | 1458.00 |
| (No Auto) | | | | |
| Estimated | 1119.50 | 425.20 | 469.70 | 1394.70 |
| (With Auto) | | | | |
| Estimated | DNE | DNE | DNE | DNE |
| (Theory 3.2) | | | | |

Table 13

The standard error for macro-factor risk premiums (149 portfolios monthly data)

This Table presents the standard error of the estimated risk premiums for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data from 1964 to 2009. 149 portfolios include 100 size and book-to-market portfolios combined with 49 industry portfolios. The true standard error (True) is calculated through taking the covariance of the estimated risk premiums in each simulation. The estimated standard error is the average of the Fama-Macbeth standard errors in each simulation. The Fama-Macbeth standard errors contain the autocovariance of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period $T=60$ and 60 for $T=600$. We also compare the Fama-Macbeth standard errors without the autocovariance. For BJS method, we also compare the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|------------|--------------|-------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| BJS Method | | | | |
| True | 0.09 | 0.03 | 0.05 | 0.14 |
| Estimated (No Auto) | 0.07 | 0.01 | 0.03 | 0.05 |
| Estimated (Shanken) | 0.08 | 0.02 | 0.03 | 0.07 |
| Rolling Method | | | | |
| True | 0.09 | 0.03 | 0.05 | 0.11 |
| Estimated (No Auto) | 0.07 | 0.01 | 0.02 | 0.04 |
| Estimated (With Auto) | 0.06 | 0.02 | 0.03 | 0.06 |
| Rolling IV method | | | | |
| True | 186.91 | 63.90 | 74.50 | 189.58 |
| Estimated (No Auto) | 187.02 | 63.83 | 75.00 | 188.81 |
| Estimated (With Auto) | 150.99 | 50.72 | 55.51 | 145.60 |
| Estimated (Theory 3.2) | 6.44 | 2.09 | 2.36 | 7.52 |
| T=600 Rolling window 60 | | | | |
| BJS Method | | | | |
| True | 0.04 | 0.01 | 0.03 | 0.06 |
| Estimated (No Auto) | 0.03 | 0.01 | 0.02 | 0.04 |
| Estimated (Shanken) | 0.04 | 0.01 | 0.03 | 0.06 |

| Rolling Method | | | | |
|---------------------------|-------|-------|-------|--------|
| True | 0.03 | 0.01 | 0.02 | 0.04 |
| Estimated (No Auto) | 0.02 | 0.01 | 0.09 | 0.02 |
| Estimated (With Auto) | 0.02 | 0.01 | 0.01 | 0.03 |
| Rolling IV method | | | | |
| True | 36.08 | 13.17 | 49.07 | 222.01 |
| Estimated (No Auto) | 36.08 | 13.16 | 49.04 | 222.02 |
| Estimated (With Auto) | 33.68 | 12.49 | 45.73 | 206.93 |
| Estimated (Theory 3.2) | DNE | DNE | DNE | DNE |

Table14

The standard error for macro-factor risk premiums (4,970 stocks, monthly data)

This Table presents the standard error of the estimated risk premiums of individual 4,970 stocks for the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The true standard error (True) is calculated through taking the covariance of the estimated risk premiums in each simulation. The estimated standard error is the average of the Fama-Macbeth standard errors in each simulation. The Fama-Macbeth standard errors contain the autocovariance of estimated risk premiums in each period up to the length of rolling windows. This length of the rolling window is 15 for total period $T=60$ and 60 for $T=600$. We also compare the Fama-Macbeth standard errors without the autocovariance. For BJS method, we also compare the Shanken adjustment in estimated standard errors.

| T=60 Rolling window 15 | | | | |
|--------------------------------|-----------------|------------|--------------|-------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| BJS Method | | | | |
| True | 0.03 | 0.02 | 0.04 | 0.12 |
| Estimated (No Auto) | 0.01 | 0.00 | 0.00 | 0.01 |
| Estimated (Shanken) | 0.01 | 0.00 | 0.01 | 0.01 |
| Rolling Method | | | | |
| True | 0.03 | 0.03 | 0.05 | 0.11 |
| Estimated (No Auto) | 0.02 | 0.01 | 0.01 | 0.02 |
| Estimated (With Auto) | 0.02 | 0.01 | 0.02 | 0.05 |
| Rolling IV method | | | | |
| True | 0.07 | 0.04 | 0.11 | 0.25 |
| Estimated (No Auto) | 0.04 | 0.03 | 0.09 | 0.19 |
| Estimated (With Auto) | 0.06 | 0.04 | 0.12 | 0.20 |
| Estimated (Theory 3.2) | 0.03 | 0.01 | 0.04 | 0.09 |
| T=600 Rolling window 60 | | | | |
| BJS Method | | | | |
| True | 0.01 | 0.00 | 0.01 | 0.02 |
| Estimated (No Auto) | 0.01 | 0.00 | 0.00 | 0.00 |
| Estimated (Shanken) | 0.01 | 0.00 | 0.00 | 0.0081 |
| Rolling Method | | | | |
| True | 0.01 | 0.01 | 0.01 | 0.04 |
| Estimated | 0.01 | 0.00 | 0.00 | 0.01 |

| | | | | |
|--------------------------|------|------|------|------|
| (No Auto) | | | | |
| Estimated | 0.01 | 0.01 | 0.01 | 0.04 |
| (With Auto) | | | | |
| Rolling IV method | | | | |
| True | 0.01 | 0.00 | 0.01 | 0.02 |
| Estimated | 0.01 | 0.00 | 0.00 | 0.01 |
| (No Auto) | | | | |
| Estimated | 0.01 | 0.00 | 0.01 | 0.01 |
| (With Auto) | | | | |
| Estimated | 0.01 | 0.00 | 0.01 | 0.02 |
| (Theory 3.2) | | | | |

Table 15
The rejection ratio of t-statistics with non-hypothesis $\alpha = 0$
(4,970 stocks, monthly data)

This Table presents the t-ratio for $\alpha = 0$ of 4,970 individual stocks with the BJS estimation without rolling beta (BJS), the Fama-Macbeth rolling beta method (Rolling), and the lagged beta IV method (Rolling IV). The estimation is based on monthly data from 1964 to 2009. For each simulation, we calculate the t-ratios for different methods and compare them with the 95% critical value. Then, we calculate the number of the simulations such that the absolute value of t-ratio is greater than the critical value and divide this number by the number of simulations to get the probability of rejecting the true non-hypothesis (rejection ratio). For different methods, there are different standard errors and t-ratios. For the BJS method, we use the Fama-Macbeth standard error without autocovariance and Shanken's standard error. For the Fama-Macbeth rolling-beta method, we use the Fama-Macbeth standard error(with and without autocovariance). For the beta IV method, we use the Fama-Macbeth standard error and the standard error that derived from Theorem 3.2. There are two cases: the Fama-French three-factor model and the macro-factor model with $N = 4970$ and $T = 600$ since these are the cases with smallest bias for the BJS and the lagged beta IV method methods. FF3 represents the Fama-French three-factor model and the MF represents macro-factor model.

| T=600 Rolling window 60 | | | | |
|--------------------------------|--------------|---------------------|--------------|---------------------|
| Std method | Model | Reject Ratio | Model | Reject Ratio |
| BJS Method | | | | |
| Estimated (No Auto) | FF3 | 0.08 | MF | 0.30 |
| Estimated (Shanken) | FF3 | 0.07 | MF | 0.03 |
| Rolling Method | | | | |
| Estimated (No Auto) | FF3 | 0.54 | MF | 0.35 |
| Estimated (With Auto) | FF3 | 0.16 | MF | 0.16 |
| Rolling IV method | | | | |
| Estimated (Theory 3.2) | FF3 | 0.05 | MF | 0.05 |
| Estimated (No Auto) | FF3 | 0.17 | MF | 0.32 |
| Estimated (With Auto) | FF3 | 0.07 | MF | 0.16 |

Table16
Estimated risk premiums for three macro factors (OLS)

This Table presents the estimated risk premiums and T-ratios for three macro factor model using individual stock returns with the BJS estimation (BJS), the Fama-Macbeth rolling beta method (Rolling), the lagged beta IV method (Rolling IV), Thiel's adjustment and the 3-group method. We use the standard errors from Fama-Macbeth (FM), Theorem 3.1 and Theorem 3.2 to calculate T-ratios. The length of the rolling window is 60 for total period T=552.

| T=552 Rolling window 60 | | | | |
|--------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| BJS Method | | | | |
| Risk Premium | 1.09 | 0.01 | 0.00 | -0.01 |
| T-ratio (FM) | 5.77 | 0.54 | 0.11 | -0.54 |
| Rolling Method | | | | |
| Risk Premium | 0.82 | -0.00 | -0.00 | -0.07 |
| T-ratio (FM) | 4.95 | -0.02 | -0.04 | -1.74 |
| Rolling IV method | | | | |
| Risk Premium | 1.06 | 0.10 | -0.04 | -0.13 |
| T-ratio (FM) | 11.72 | 4.41 | -1.87 | -2.52 |
| T-ratio (Theorem 3.2) | 4.77 | 2.78 | -1.42 | -2.09 |
| Thiel's adjustment | | | | |
| Risk Premium | 0.79 | 0.12 | -0.22 | 0.01 |
| T-ratio (FM) | 3.28 | 0.53 | -0.57 | 0.60 |
| 3-group method | | | | |
| Risk Premium | 0.31 | 0.03 | -0.00 | -0.04 |
| T-ratio (FM) | 4.46 | 1.99 | -0.25 | -2.59 |
| T-ratio (Theorem 3.1) | 4.35 | 1.70 | -0.08 | -1.29 |

Table17
Estimated risk premiums for three macro factors (GLS)

This Table presents the estimated risk premiums and T-ratios for three macro factor model using individual stock returns with the BJS estimation (BJS), the Fama-Macbeth rolling beta method (Rolling), the lagged beta IV method (Rolling IV), Thiel's adjustment and the 3-group method. We use the standard errors from Fama-Macbeth (FM), Theorem 3.1 and Theorem 3.2 to calculate T-ratios. The length of the rolling window is 60 for total period T=552.

| T=552 Rolling window 60 | | | | |
|--------------------------------|-----------------|------------------------------|--------------------------------|-------------------------------|
| Factor | constant | ΔC | ΔCPI | ΔIP |
| BJS Method | | | | |
| Risk Premium | 0.26 | 0.02 | -0.00 | -0.03 |
| T-ratio (FM) | 7.54 | 1.42 | -0.38 | -2.35 |
| Rolling Method | | | | |
| Risk Premium | 0.29 | 0.04 | -0.02 | -0.07 |
| T-ratio (FM) | 5.87 | 1.13 | -0.87 | -1.78 |
| Rolling IV method | | | | |
| Risk Premium | 0.91 | 0.07 | -0.04 | -0.19 |
| T-ratio (FM) | 11.34 | 2.60 | -1.63 | -2.67 |
| T-ratio (Theorem 3.2) | 6.19 | 2.13 | -1.67 | -3.26 |
| Thiel's adjustment | | | | |
| Risk Premium | 0.38 | -0.00 | -0.00 | 0.00 |
| T-ratio (FM) | 7.07 | -1.62 | -0.16 | 2.39 |
| 3-group method | | | | |
| Risk Premium | 0.18 | 0.04 | -0.01 | -0.09 |
| T-ratio (FM) | 5.92 | 2.76 | -1.59 | -4.69 |
| T-ratio (Theorem 3.1) | 4.48 | 2.60 | -0.57 | -3.67 |