Olive: a simple method for estimating betas when factors are measured with error.

Meng, Ginger and Hu, Gang and Bai, Jushan

Stonehill College, Babson College, Columbia University

March 2007

OLIVE: A SIMPLE METHOD FOR ESTIMATING BETAS WHEN FACTORS ARE MEASURED WITH ERROR [†]

J. Ginger Meng
Stonehill College

Gang Hu
Babson College

Jushan Bai
Columbia University

[†] We thank the referee, Tim Simin, and the editors, Gerald Gay and Jayant Kale, for their detailed constructive comments and suggestions. For helpful comments and discussions, we thank Tom Chemmanur, Wayne Ferson, Qiang Kang, Arthur Lewbel, Zhijie Xiao, and seminar participants at Boston College, New York University, and the 2007 Financial Management Association meetings in Orlando. We thank Martin Lettau and Sydney Ludvigson for making their data available and answering our questions. We thank Michelle Graham for editorial assistance. All remaining errors and omissions are our own.

Abstract

We propose a simple and intuitive method for estimating betas when factors are measured with error: ordinary least squares instrumental variable estimator (OLIVE). OLIVE performs well when the number of instruments becomes large, while the performance of conventional instrumental variable methods becomes poor or even infeasible. In an empirical application, OLIVE beta estimates improve R-squared significantly. More importantly, our results help resolve two puzzling findings in the prior literature: first, the sign of average risk premium on the beta for market return changes from negative to positive; second, the estimated value of average zero-beta rate is no longer too high.

JEL Classifications: C30, G12.

I. Introduction

In financial economics, we often need to estimate asset return betas (factor loadings). Ordinary least squares (OLS) is the simplest and most widely used method by both academic researchers and practitioners. However, factors, especially those constructed using macroeconomic data, are known to contain large measurement error (e.g., Chen, Roll, and Ross 1986; Connor and Korajczyk 1986, 1991; Ferson and Harvey 1999). In addition, even when a factor is measured accurately, it may still be different from the true underlying factor. For example, the return on the stock market index is perhaps measured reasonably accurately, but it may still contain large "measurement error" in the sense that it may be an imperfect proxy for the return on the true market portfolio (Roll 1977). Under these circumstances, the OLS beta estimator will be inconsistent. Furthermore, in the Fama and MacBeth (1973) two-pass framework, if the first-pass beta estimates are inconsistent because of measurement error in factors, the second-pass risk premia and zero-beta rate estimates will be inconsistent as well.

Instrumental variable (IV) estimation is the usual solution to the measurement error problem. Intuitively, because all asset returns vary together with a common set of factors, one can use information contained in other asset returns to improve the beta estimate for a given asset. This is often a large N and small T setting, because there are typically more assets or stocks than periods. Ideally, we would want to use all available information, that is, all valid instruments (the other (N-1) asset returns), but conventional instrumental variable estimators such as two-stage least squares (2SLS) perform poorly when the number of instruments is large. This is similar to the "weak instruments" problem (Hahn and Hausman 2002). Furthermore, these methods cannot accommodate more instruments than the sample size.

In this article, we propose a simple method for estimating betas when factors are measured with error: ordinary least squares instrumental variable estimator (OLIVE). OLIVE easily allows for large numbers of instruments (can be larger than the sample size). It is intuitive, easy to implement, and achieves better performance in simulations than other instrumental variable estimators such as 2SLS, bias-corrected two-stage least squares (B2SLS), limited information maximum likelihood (LIML), and the Fuller (1977) estimator (FULLER), especially when the number of potential instruments (N-1) is large and the sample size (T) is small.

We show that OLIVE is a consistent estimator under the assumption that idiosyncratic errors are cross-sectionally independent (Proposition 1). Consistency is obtained when the number of assets (N) is fixed or goes to infinity. When idiosyncratic errors are cross-sectionally correlated, returns of other assets as instruments are invalid in the conventional sense because they are correlated with the regression errors. We show that even in this case, OLIVE beta estimates remain consistent, provided that N is large (Proposition 2). In a sense, we exploit the large N of panel data to arrive at a consistent estimator. Because conventional generalized method of moments (GMM) breaks down for $N > T$, and consistency in the absence of valid instruments requires large N, OLIVE's ability to handle large N is appealing.

OLIVE can be viewed as a one-step GMM estimator using the identity weighting matrix. When N is larger than T, the optimal weighting matrix in the GMM estimation cannot be consistently estimated in the usual unconstrained way. However, in our setting we are able to derive the two-step equation-by-equation GMM estimator, as well as the joint GMM estimator, based on the restrictions implied by the model. Even though the two-step GMM estimator is asymptotically optimal, it performs worse than OLIVE in simulations. This is because the two-step GMM estimator has poor finite sample properties caused by imprecise estimation of the optimal weighting matrix.

Previous studies also show that the two-step GMM estimator that is optimal in the asymptotic sense can be severely biased in finite samples of reasonable size (e.g., Ferson and Foerster 1994; Hansen, Heaton, and Yaron 1996; Newey and Smith 2004; Doran and Schmidt 2006). One-step GMM estimators use weighting matrices that are independent of estimated parameters, whereas the efficient two-step GMM estimator weighs the moment conditions by a consistent estimate of their covariance matrix.[1] Given the difficulty in estimating the optimal weighting matrix, especially when *N* is large, using the identity weighing matrix becomes an intuitive option. OLIVE can be viewed as a GMM estimator using the identity weighting matrix.

Fama and MacBeth's (1973) two-pass method can be modified by using OLIVE instead of OLS to estimate betas in the first pass. As an empirical application, we reexamine Lettau and Ludvigson's (2001b) test of the conditional/consumption capital asset pricing model ((C)CAPM) using this modified Fama-MacBeth method. Lettau and Ludvigson's factor *cay* is found to have strong forecasting power for excess returns on aggregate stock market indices. The factor *cay* is the cointegrating residual between log consumption c, log asset wealth a, and log labor income y. Macroeconomic variables usually contain large measurement error. We find that in regressions where macroeconomic factors are included, using OLIVE instead of OLS improves the R^2 significantly (e.g., from 31% to 80%).

More important, our results based on OLIVE beta estimates help resolve two puzzling findings in the prior literature. If we use OLS when factors are measured with error, both the

[1] Wyhowski (1998) performs simulations and shows the GMM estimator performs well if the true optimal weighting matrix is used. Methods to correct the bias problem include, for example, using a subset of the moment conditions and normal quasi-MLE. Other solutions to this problem use higher order expansions to construct weighting matrix estimators, or use generalized empirical likelihood (GEL) estimators as in Newey and Smith (2004). Doran and Schmidt (2006) suggest using principal components of the weighting matrix.

first-pass beta estimates and the second-pass risk premia and zero-beta rate estimates will be inconsistent. Conversely, because OLIVE beta estimates are consistent even when factors contain measurement error, the risk premia and zero-beta rate can be consistently estimated in the second pass if OLIVE is used in the first pass to estimate betas. First, Lettau and Ludvigson (2001b) find that their estimated average risk price on the beta for the value-weighted return is negative. Jagannathan and Wang (1996) report a similar finding. Using OLIVE instead of OLS estimation in the first pass changes the sign of the average risk premium on the beta for the value-weighted market index from negative to positive, which is in accordance with the theory. Second, in Lettau and Ludvigson, the estimated value of the average zero-beta rate is too high. As the authors observe, this finding is not uncommon in studies that use macroeconomic factors. We find that when OLIVE beta estimates from the first pass are used, the estimated value of the average zero-beta rate in the second pass is no longer too high (e.g., from 5.19% to 1.91% per quarter). Our results suggest that measurement error in factors is the cause of this problem.

In contrast, it makes almost no difference whether we use OLIVE or OLS to estimate betas for the Fama-French three-factor model, where the factors may contain little measurement error as they are constructed from stock returns. Overall, our results from this empirical application validate the use of OLIVE to help improve beta estimation when factors are measured with error.

Many existing empirical asset pricing models implicitly assume that macroeconomic variables are measured without error, for example, Chen, Roll, and Ross (1986). Previous studies have noted the measurement error problem in this context (e.g., Ferson and Harvey 1999). Connor and Korajczyk (1991) develop and apply a procedure similar to 2SLS. However, since the fitted values are linear combinations of statistical factors, they do not contain any more information beyond statistical factors, which lack clear economic interpretations. Wei, Lee, and Chen (1991) also note the presence of errors-in-variables problem in factors. They use the standard econometric treatment: instrumental variables approach (IV or 2SLS). Both their factors and instruments are size-based portfolios. Even if there is measurement error in size-based portfolio returns, the problem would not be solved by using other size-based portfolio returns as instruments. As one would expect, they find extremely high first-stage R^2. This means their IV results will be very similar to OLS results, and indeed that is what they find.

II. Estimation Framework

Model Setup

To describe the model, we begin by assuming that asset returns are generated by a linear multi-factor model:

$$y_{it} = x_t *' B_i + e_{it}, \tag{1}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$, y_{it} is asset i's return at time t, x_t* is an $M \times 1$ vector of true factors at time t, and β_i is an $M \times 1$ vector of factor loadings for asset i. However, the true factors x_t* *are* observed with error:

$$x_t = x_t * + v_t, \tag{2}$$

where v_t is an $M \times 1$ vector of measurement error. This is similar to the setup in Connor and Korajczyk (1991) and Wansbeek and Meijer (2000). Using (2), we can rewrite (1) as:

$$y_{it} = x_t ' B_i + \varepsilon_{it}, \tag{3}$$

where $\varepsilon_{it} = e_{it} - v_t ' B_i$.

We cannot use OLS to estimate β_i equation-by-equation, even though x_t is observable, because the error term ε_{it} is correlated with the observable factors x_t due to the measurement error v_t.

For a fixed asset i, rewrite (3) as

$$Y_i = XB_i + \varepsilon_i, \tag{4}$$

where Y_i is a T vector of asset returns, $X \equiv [\iota, (x_1, \ldots x_T)']$ is a $T \times (M+1)$ matrix of observable factors (ι is a T vector of 1's), and B_i is an $(M+1)$ vector of factor loadings. As noted before, OLS produces inconsistent estimates of factor loadings:

$$\hat{B}_i^{OLS} = (X'X)^{-1} X'Y_i. \tag{5}$$

Let $Y_{-i} \equiv [Y_1, \ldots Y_{i-1}, Y_{i+1}, \ldots, Y_N]$ be a $T \times (N-1)$ matrix of all asset returns excluding the ith asset. Then Y_{-i} can serve as instrumental variables. Let $Z_i \equiv [\iota, Y_{-i}]$, multiply both sides of equation (4) by Z_i to obtain:

$$Z_i'Y_i = Z_i'XB_i + Z_i'\varepsilon_i. \tag{6}$$

It can be shown that the usual IV or 2SLS is equivalent to running Feasible GLS on (6); that is,

$$\hat{B}_i^{2SLS} = (X'Z_i(Z_i'Z_i)^{-1}Z_i'X)^{-1} X'Z_i(Z_i'Z_i)^{-1}Z_i'Y_i. \tag{7}$$

The idea of 2SLS is first to project the regressors (X) onto the space of instruments (Z_i), and then to regress the dependent variables (Y_i) on fitted values of regressors instead of regressors themselves. It is well known that two-staged least squares (2SLS) estimators may perform poorly when the instruments are weak or when number of instruments is large. In this case 2SLS tends to suffer from substantial small sample biases.

OLIVE

The motivation behind our approach begins with the fact that 2SLS only works when *N* (number of instruments) is much smaller than *T* (sample size), which is not the case for most finance applications. To illustrate the problem, imagine the case where *N* = *T*. Then the fitted values are the same as original regressors, and 2SLS becomes the same as OLS. This problem of 2SLS is related to the "weak instruments" literature in econometrics, which has grown rapidly in recent years; see for example Hahn and Hausman (2002).

We propose to estimate factor loadings B_i by simply running OLS on equation (6). We call it Ordinary Least-squares Instrumental Variable Estimator (OLIVE):

$$\widehat{B}_i^{OLIVE} = (X'Z_iZ_i'X)^{-1}X'Z_iZ_i'Y_i. \tag{8}$$

Proposition 1. *Under the assumption that idiosyncratic errors* e_{it} *are cross-sectionally independent, then for either fixed N or N going to infinity, the OLIVE estimator is* $\sqrt{T}$ *consistent and asymptotically normal.*

See Appendix A for a proof of Proposition 1. Proposition 1 relies on the assumption of valid instruments. That is, e_{jt} is uncorrelated with e_{it} ($j \neq i$). However, if the idiosyncratic errors are also cross-sectionally correlated, none of the instruments will be valid in the conventional sense. For example, if the objective is to estimate B_1, by equation (3), $\varepsilon_{1t} = e_{1t} - B_1'v_t$. When e_{1t} is correlated with e_{jt}, y_{jt} will be correlated with ε_{1t}. Thus y_{jt} will not be a valid instrument. However, we can still establish the consistency of the OLIVE, provided that the cross-sectional correlation is not too strong and *N* is large. To this end, let $\gamma_{ij} = E\left(e_{it}e_{jt}\right)$. We assume

$$\sum_{j=1}^{N} \left| \gamma_{ij} \right| \leq C < \infty \tag{9}$$

for each i. This condition is analogous to the sum of autocovariances being bounded in the time series context, a requirement for a time series being weakly correlated. Bai (2003) shows that the condition implies (3) being an approximate factor model of Chamberlain and Rothschild (1983).

Proposition 2. *Under the assumption of weak cross-sectional correlation for the idiosyncratic errors as stated in (9), if $\sqrt{T}/N \to 0$, then the OLIVE estimator is $\sqrt{T}$ consistent and asymptotically normal.*

A proof of Proposition 2 is provided in Appendix B. Mere consistency would only require $1/N \to 0$. It is the $\sqrt{T}$ consistency and asymptotic normality that require $\sqrt{T}/N \to 0$. Note that under fixed N, all IV estimators discussed in the next section including OLIVE (using y_{jt} as instruments) will be inconsistent due to the lack of valid instruments. In a sense, we exploit the large N of panel data to arrive at a consistent estimator. Far from being a nuisance, large N is clearly beneficial. In view that conventional GMM breaks down for $N > T$ and consistency in the absence of valid instruments requires large N, OLIVE's ability to handle large N is appealing.

Let $\hat{\varepsilon}_i = Y_i - X\hat{B}_i^{OLIVE}$ and $\hat{\sigma}_i^2 = \dfrac{1}{T-M-1}\hat{\varepsilon}_i{}'\hat{\varepsilon}_i$, the variance-covariance matrix of $\hat{B}_i^{OLIVE}$ is a $(M+1)\times(M+1)$ matrix:

$$\hat{\Omega}_i = \hat{\sigma}_i^2 (X'Z_iZ_i'X)^{-1}(X'Z_iZ_i'Z_iZ_i'X)(X'Z_iZ_i'X)^{-1}. \tag{10}$$

The above estimation is done for each i = 1, …, N. With the B_i obtained for each i, we can estimate x_t* using a cross-section regression based on equation (1). This is done for each t =

1, …, T. Given estimated x_t*, the estimated risk premia can then be recovered as in Black, Jensen, and Scholes (1972) and Campbell, Lo, and MacKinlay (1997, Chapter 6).

The above setup also allows us to test the validity of the multi-factor models. When the instrumental variables are $Z_i \equiv [\iota, Y_{-i}]$, the constant regressor ι itself is an instrumental variable. The test for the constant coefficient's being zero is $t = \frac{\hat{a}_i}{\sqrt{\widehat{\Omega}^{11}}}$, where $\widehat{\Omega}^{11}$ is the first diagonal element of the inverse matrix $\widehat{\Omega}_i^{-1}$.

There is an alternative method for estimating the true factors, i.e., the method of Connor and Korajczyk (1991). They first regress the observed factors on APC estimated statistical factors and use the fitted values as estimates of the true factors (rotate observed factors onto statistical factors). They find the R-squared to be quite small, and they interpret this as evidence for much measurement error in the observed factors. APC should have good performance theoretically and empirically. However, the statistical factors using the principle-components method lack clear economic interpretations. In contrast, note that estimated factors x_t* using OLIVE has the same interpretations as x_t, the observable factors. Thus the estimated risk premia also have economic interpretations.

Other IV Estimators

We compare the performance of OLIVE with OLS and several well known IV estimators: 2SLS, LIML, B2SLS, as well as FULLER. OLS is to be considered as a benchmark. 2SLS is the most widely used IV estimator. It has finite sample bias that depends on the number of instruments used (K) and inversely on the R^2 of the first-stage regression (Hahn and Hausman 2002). The higher-order mean bias of 2SLS is proportional to the number of instruments K. However 2SLS can have smaller higher-order mean squared error (MSE) than LIML using the second-order approximations when the number of instruments is not too large. LIML is known not to have finite sample moments of any order. LIML is also known to be median unbiased to second order and to be admissible for median unbiased estimators (Rothenberg 1983). The higher-order mean bias for LIML does not depend on K. B2SLS denotes a bias adjusted version of 2SLS.

The formulae for these estimators are as follows:

Let $P = Z(Z'Z)^{-1}Z'$ be the idempotent projection matrix, $M = I - P$, $W \equiv [Y, X]$, $Z = (y_1, y_2, \cdots, y_{i-1}, y_{i+1}, \cdots, y_N)$, then:

$$\begin{aligned}
\hat{\beta}_{OLS} &= (X'X)^{-1}X'Y \\
\hat{\beta}_{2SLS} &= (X'PX)^{-1}X'PY \\
\hat{\beta}_{LIML} &= (X'(P-\hat{\lambda}M)X)^{-1}X'(P-\hat{\lambda}M)Y \\
\hat{\beta}_{B2SLS} &= (X'(P-\bar{\lambda}M)X)^{-1}X'(P-\bar{\lambda}M)Y \\
\hat{\beta}_{FULLER} &= (X'(P-\tilde{\lambda}M)X)^{-1}X'(P-\tilde{\lambda}M)Y \\
\hat{\beta}_{OLIVE} &= (X'ZZ'X)^{-1}X'ZZ'Y
\end{aligned} \quad . \tag{11}$$

For the above equations, 2SLS, LIML, B2SLS, and FULLER can all be regarded as κ-class estimators given by $\dfrac{X'PY - \kappa X'MY}{X'PX - \kappa X'MX}$. For $\kappa = 0$, we get 2SLS. For $\kappa = \hat{\lambda}$, which is the smallest eigenvalue of the matrix $W'PW(W'MW)^{-1}$, we obtain LIML. For $\kappa = \bar{\lambda}$, which equals $\dfrac{K-2}{T}$, we obtain B2SLS. For $\kappa = \tilde{\lambda}$, which equals $\hat{\lambda} - \dfrac{\alpha}{T-K}$, we obtain FULLER. Following Hahn, Hausman, and Kuersteiner (2004), we consider the choice of α to be either 1 or 4 in our simulation studies later (Section IV). The choice of $\alpha = 1$ is advocated by Davidson and McKinnon (1993), which has the smallest bias, while $\alpha = 4$ has a nonzero higher mean bias, but a smaller MSE according to calculation based on Rothenberg's (1983) analysis.

There are other solutions to the errors-in-variables problem, for example, Coën and Racicot (2007) propose a higher moment estimator, and find that estimators based on moments of order higher than two performed better than ordinary least squares estimators in terms of root mean squared errors and also in terms of size of type I errors of standard tests. The estimator may be interpreted in its simplest version as a linear matrix combination of the generalized version Durbin's estimator (1954) and Pal's estimator (1980). Kim (1995) proposes a correction for the EIV problem in the estimation of the price of beta risk within the two-pass estimation framework. The intuition is to incorporate the extracted information about the relationship between the measurement error variance and the idiosyncratic error variance into the maximum likelihood estimation under either homoscedasticity or heteroscedasticity of the disturbance term of the market model. Chao and Swanson (2005) show that the use of many weak instruments may improve the performance of certain point estimators since the consistent estimation depends importantly on the strength and the number of instruments. Hussman (1993) demonstrates that using monthly returns data, the cross-sectional regression approach will accept the null hypothesis of no relation between β and stock returns even when the underlying model is true, because the average excess market return is typically small relative to its standard error.

Using portfolios of asset returns as instruments to reduce errors-in-variables is another interesting and feasible alternative. Starting with Fama and MacBeth (1973), studies in the two-pass tradition try to solve the EIV problem by grouping the firms into portfolios. In a recent paper, Barnes and Hughes (2002) propose a quantile regression and show the quantile estimator is inconsistent under EIV. When the ordering of the instruments is given, Donald and Newey (2001) propose an information criterion approach to choose the number of instruments. Since we do not assume the ordering of the instruments to be known, there are too many potential models, hence exhaustive search of optimal instruments are not possible. Our method makes use of all available instruments without the need to assume any ordering of the instruments or determine the optimal number of instruments. The simplicity of our method is therefore appealing.

III. Efficient Two-Step GMM

What makes OLIVE appealing is its ease of use. Since OLIVE is a GMM estimator when setting the weighting matrix to an identity matrix, it is natural to try to improve the efficiency of the estimator by using the optimal weighting matrix. Traditional unconstrained GMM will break down when *N*>*T* (the estimated weighting matrix is not invertible). We will derive the theoretical weighting matrix, which depends on far fewer number of parameters. Replacing the unknown parameters by their estimated counterparts will result in an estimated theoretical weighting matrix, which is invertible even for *N*>*T*.

Equation-by-Equation GMM

Consider estimating B_i for equation *i*. By definition, $\varepsilon_{it} = y_{it} - x_t{}'B_i = e_{it} - v_t{}'B_i$. For every *j* $(j \neq i)$, y_{jt} can serve as an instrument. Let $u_{ijt} = y_{jt}\varepsilon_{it} = (x_t{}^*{}'B_j + e_{jt})(e_{it} - v_t{}'B_i)$. Under the assumption that e_{it}, e_{jt}, v_t, and $x_t{}^*$ are mutually independent, the moment conditions, or orthogonality conditions, will be satisfied at the true value of B_i:

$$E(u_{ijt}) = E\left[y_{jt} \left(y_{it} - x_t{}'B_i \right) \right] = 0 \,. \tag{12}$$

Each of the (*N*-1) moment equations corresponds to a sample moment, and we write these (*N*-1) sample moments as:

$$\bar{u}_{ij}(B_i) = \frac{1}{T}\sum_{t=1}^{T} u_{ijt}(B_i). \tag{13}$$

Let $\bar{u}_i(B_i)$ be defined by stacking $\bar{u}_{ij}(B_i)$ over j. For a given weighting matrix W_i, the equation-by-equation GMM is estimated by minimizing: $\min_{B_i} \bar{u}_i(B_i)' W_i^{-1} \bar{u}_i(B_i)$. For each i, t, let u_{it} be the (N-1) vector by stacking u_{ijt} over j. The optimal weighting matrix is $W_i = E(u_{it}u_{it}')$. Given the above functional form for u_{ijt}, W_i can be parameterized in terms of $\text{var}(e_{it})$ and B_i for each i, $\text{var}(v_t)$, and $\text{var}(x_t^*) = \text{var}(x_t) - \text{var}(v_t)$.

We now derive the expression of W_i. The (j, k)th element of W_i ($j \neq i, k \neq i$) is given by:

$$\begin{aligned}
&E(u_{ijt}u_{ikt}) = E\left[y_{jt}(e_{it} - v_t' B_i)(e_{it} - B_i' v_t) y_{kt} \right] \\
&= E\left[y_{jt}\left(\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i\right) y_{kt} \right] \\
&= E\left[\left(B_j' x_t^* + e_{jt}\right)\left(x_t^{*\prime} B_k + e_{kt}\right)\left(\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i\right) \right] \\
&= E\left[\left(B_j'\text{var}(x_t^*)B_k + \delta_{jk}\,\text{var}(e_{jt})\right)\left(\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i\right) \right] \\
&= \left(B_j'\text{var}(x_t^*)B_k + \delta_{jk}\,\text{var}(e_{jt})\right)\left(\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i\right)
\end{aligned} \quad , \tag{14}$$

where $\delta_{jk} = 1$ if $j = k$, and zero otherwise. In the last equality, B_js are assumed non random coefficients.

For example, suppose $i = 1$, then the above covariance matrix is simply the following. Let

$$\Lambda_{-1} = \begin{bmatrix} B_2 \\ B_3 \\ \vdots \\ B_N \end{bmatrix}, \tag{15}$$

then the (N-1) by (N-1) covariance matrix W_1 is given by:

$$W_1 = \left(\text{var}(e_{1t}) + B_1'\text{var}(v_t)B_1\right)\left(\Lambda_{-1}\,\text{var}(x_t^*)\Lambda_{-1}' + \Omega_{-1}\right), \tag{16}$$

where Ω_{-1} is a diagonal matrix of dimension (N-1), that is

$$\Omega_{-1} = diag(\text{var}(e_{2t}), \cdots, \text{var}(e_{Nt})). \tag{17}$$

Note that $\left(\text{var}(e_{1t}) + B_1'\text{var}(v_t)B_1\right)$ is a scalar, which is the variance of the OLS residual ε_{1t}, thus can be estimated by $\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2$.

For a general i, the formula for W_i becomes:

$$W_i = \left(\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i\right)\left(\Lambda_{-i}\,\text{var}(x_t^*)\Lambda_{-i}' + \Omega_{-i}\right). \tag{18}$$

The analytical expression for the inverse of W_i is:

$$\begin{aligned} W_i^{-1} &= \frac{\left(\Lambda_{-i}\,\text{var}(x_t^*)\Lambda_{-i}^{-1} + \Omega_{-i}\right)^{-1}}{\text{var}(e_{it}) + B_i'\text{var}(v_t)B_i} \\ &= \frac{\Omega_{-i}^{-1} - \Omega_{-i}^{-1}\Lambda_{-i}\left(\left(\text{var}(x_t^*)^{-1}\right) + \Lambda_{-i}'\Omega_{-i}^{-1}\Lambda_{-i}\right)^{-1}\Lambda_{-i}'\Omega_{-i}^{-1}}{\text{var}(\varepsilon_{it})} \end{aligned} \tag{19}$$

The estimation procedure is then as follows.

First we use OLIVE to obtain, for each asset i, $\widetilde{B}_i$ and $\hat{\varepsilon}_{it} = y_{it} - x_t'\widetilde{B}_i$, which equals an estimate of $e_{it} - v_t'B_i$. The denominator of W_i^{-1} is computed by the sample variance of $\hat{\varepsilon}_{it}$.

Second, given $\widetilde{B}_i$, we run cross-sectional regression to obtain x_t^* for each t, and then estimate $\text{var}(x_t^*)$. Also, given x_t^*, we can estimate $\hat{e}_{it} = y_{it} - x_t^*\widetilde{B}_i$, so that $\text{var}(e_{it})$ are computed for each i.

Third, we use the above estimates to construct a consistent estimate of $E(u_t u_t')$, and use that to do two-step GMM. For each asset i, there is an (N-1)×(N-1) weighing matrix W_i.

The estimate of beta is:

$$\widehat{B}_i = (X'Z_i\widehat{W}_i^{-1}Z_i'X)^{-1}X'Z_i\widehat{W}_i^{-1}Z_i'Y_i. \tag{20}$$

The choice of W_i is optimal in the sense that it leads to the smallest asymptotic variance matrix for the GMM estimate. However, a number of papers have found that GMM estimators using all of the available moment conditions may have poor finite sample properties in highly identified models. With many moment conditions, the optimal weighting matrix is poorly estimated. The problem becomes more severe when many of the moment conditions (implicit instruments) are "weak." The poor finite sample performance of the estimates has two aspects, as noted by Doran and Schmidt (2006). First, the estimates may be seriously biased. This is generally believed to be a result of correlation between the estimated weighting matrix $\widehat{W}_i$ and the sample moment conditions in equation (13). Second, the asymptotic variance expression may seriously understate the finite sample variance of the estimates, so that the estimates are spuriously precise.

Joint GMM

In this subsection, we discuss joint GMM estimation of $B = \left(B_1', B_2', \cdots, B_N'\right)'$. Let u_t be the vector with elements u_{ijt} for all *i, j* pairs $\left(j \neq i\right)$. The optimal GMM weighting matrix, $E(u_t u_t')$, is difficult to estimate in the usual unconstrained way because the number of moment conditions, $N(N\text{-}1)$, can be much larger than T. Under our model specification, however, $E(u_t u_t')$ can also be parameterized in terms of $\operatorname{var}(e_{it})$, B_i, $\operatorname{var}(v_t)$, and $\operatorname{var}(x_t{}^*)$.

The $N(N\text{-}1)$ by $N(N\text{-}1)$ weighting matrix W can be partitioned into N^2 block matrices, each being (N-1) by (N-1). We denote these block matrices $W_{ih} = E(u_{it} u_{ht}')$, for all *i, h* = 1, …, N. The block diagonal matrix W_{ii} corresponds to the equation-by-equation weighting matrix W_i, as derived in equation (14) in the previous subsection. In short, the (j,k)th element of the block diagonal matrix W_{ii} (denoted as w_{jk}^{ii}) is:

$$\begin{aligned} w_{jk}^{ii} &= E(u_{ijt}u_{ikt}) = E\left[y_{jt}(e_{it} - v_t{}'B_i)(e_{it} - B_i{}'v_t)y_{kt} \right] \\ &= \left(B_j{}'\operatorname{var}(x_t^*)B_k + \delta_{jk}\operatorname{var}(e_{jt}) \right)\left(\operatorname{var}(e_{it}) + B_i{}'\operatorname{var}(v_t)B_i \right) \end{aligned}. \tag{21}$$

The block off-diagonal matrix W_{ih} $(i \neq h)$ represents the variance-covariance matrix between the orthogonality conditions for assets *i* and *h*. This matrix is nonzero because an instrument used for asset *i* may also be used for asset *h*. In addition, asset *i* is also an instrument for asset *h* and vice versa. Thus the orthogonality conditions associated with different equations are correlated. The (j,k)th element of this matrix, w_{jk}^{ih}, equals $E(u_{ijt}u_{hkt}) = E\left[y_{jt}(e_{it} - v_t{}'B_i)(e_{ht} - B_h{}'v_t)y_{kt} \right]$, where $j \neq i$ and $k \neq h$ by definition of IV. We derive the formulae for w_{jk}^{ih}, the (j,k)th element of the block off-diagonal matrix W_{ih}, in each of the four possible cases in Appendix C.

We now have the whole weighting matrix *W*. GMM is estimated by minimizing $\min_B \overline{u}(B)'W^{-1}\overline{u}(B)$. The estimate of beta is:

$$\widehat{B} = ((I \otimes X)'Z\widehat{W}^{-1}Z'(I \otimes X))^{-1}(I \otimes X)'Z\widehat{W}^{-1}Z'Y, \tag{22}$$

where $Y = \left(Y_1{}', \cdots, Y_N{}' \right)'$. Z is a block diagonal matrix, with $Z = diag(Z_1, Z_2, ..., Z_N)$, where $Z_i \equiv [\iota, Y_{-i}]$. Joint GMM will not be used later in this paper because the number of moment conditions, *N*(*N*-1), is too large. But if *N* is small, joint GMM will be useful.

IV. Simulation Study

Simulation Design

We conduct a Monte Carlo simulation study to compare the performance of our simple OLIVE estimator with other estimators. The data generating process (DGP) for our simulation study is as follows. We assume no intercept, i.e., arbitrage pricing theory (APT) or capital asset pricing model (CAPM) holds, as in Connor and Korajczyk (1993) and Jones (2001). Although the estimation framework is general for any factor model, we implement our simulation with a stock market application in mind. The DGP below is very similar to the one in Connor and Korajczyk (1993).

We first generate a security y_{0t} (as in the following equation), with a true beta of one, which is to be estimated.

$$\begin{aligned} & y_{0t} = \beta_0' x_t^* + e_{0t}, i = 1, \ldots, N \\ & x_t^* \sim MVN(\pi, \sigma_x^2 I^J) \\ & \beta_0 = \iota^J \\ & e_{0t} \sim N(0, \sigma_e^2) \end{aligned} \tag{23}$$

Then we generate $K = N$-1 instruments using the following:

$$\begin{aligned} & x_t = x_t^* + v_t, t = 1, \ldots, T \\ & y_{it} = \beta_i' x_t^* + e_{it} = \beta_i' x_t + (-\beta_i' v_t + e_{it}) = \beta_i' x_t + \varepsilon_{it}, i = 1, \ldots, N \\ & x_t^* \sim MVN(\pi, \sigma_x^2 I^J) \\ & v_t \sim MVN(0^J, \sigma_v^2 I^J) \\ & \beta_i \sim MVN(\iota^J, \sigma_\beta^2 I^J) \\ & e_t \sim MVN(0^N, \sigma_e^2 I^N) \\ & y_{-i,t} = (y_{1t}, y_{2t}, \ldots, y_{i-1,t}, y_{i+1,t}, \ldots, y_{Nt}) \end{aligned} \tag{24}$$

We use K = (2, 10, 45, 150, 600), T = 60, π = 0.1, σ_x = 0.1, σ_β = 1 and 1000 replications. Without loss of generality, we assume J, the number of explanatory variables to be 1, which makes the model specification equivalent to the CAPM for the excess return. We allow x and β

to be normally generated. One advantage of OLIVE is that when K is larger than T, it still works while most other IV estimators do not.

Two important parameters for the performance of the estimators are the standard deviation of the error in returns, σ_e, and the standard deviation of the measurement error, σ_v. We allow these two parameters to change from low (0.01), medium (0.1), to high (1), i.e., $\sigma_e \in$ (0.01, 0.1, 1) and $\sigma_v \in$ (0.01, 0.1, 1). When σ_e increases from 0.01 to 1, the instruments becomes weaker. When σ_v increases from 0.01 to 1, the magnitude of measurement error increases. Panel A of Table 1 presents simulation results when both σ_v and σ_e are set equal to 0.1, which is the medium measurement error and medium instruments case. Panel B of Table 1 presents simulation results when both σ_v and σ_e are set equal to 1, which is the large measurement error and weak instruments case.[2]

[INSERT TABLE 1 HERE]

We further conduct simulation study allowing for weak cross-sectional correlation among securities. The setup is similar except that the cross-sectional error term e_{it} is generated as an AR(1) process, i.e., $e_{it} = a_t e_{i-1,t} + \eta_{it}$, where $a_t \sim U(-0.5, 0.5)$ and $\eta_{it} \sim MVN(0,1)$. These simulation results are reported in Table 2.

Simulation Results

In Tables 1 and 2, a variety of summary statistics is computed for each estimator. When K is set from 1 to 45 (*K<T)*, all estimators are computed. When *K>T,* only OLS, OLIVE, and the two-step equation-by-equation GMM estimator (2GMM) are computed because other IV estimators

[2] We also run simulations using different levels of these two parameters and show that our findings are robust. The results are not reported due to space constraints, but are available upon request.

become infeasible. Following Donald and Newey (2001), we compute the mean bias and the mean absolute deviation (AD), for each estimator from the true value of β generated. We examine dispersion of each estimator using both the inter-quartile range (IQR) and the difference between the 1st and 9th deciles (Dec. Rge) in the distribution of each estimator. Throughout, OLS offers the smallest dispersion in terms of both IQR and Dec. Rge. This finding is consistent with Hahn, Hausman, and Kuersteiner (2004). We also report the coverage rate of a nominal 95% confidence interval (Cov. Rate). Panel A of Table 1 presents simulation results when both σ_v and σ_e are set equal to 0.1, which is the medium measurement error and medium instruments case. Panel B of Table 1 presents simulation results when both σ_v and σ_e are set equal to 1, which is the large measurement error and weak instruments case. Table 2 reports results when we allow for weak cross-sectional correlation among securities.

We first focus our discussion on simulation results in Table 1 Panel A, the medium measurement error and medium instruments case. When there is only one instrument, 2SLS, LIML, OLIVE, and 2GMM are all equivalent. Throughout, both OLS and 2GMM seem to be biased downwards. As Newey and Smith (2004) point out, the asymptotic bias of GMM often grows with the number of moment restrictions. Our simulation results show that the performance of the two-step GMM estimator becomes worse as the number of instruments grows. As the number of instruments becomes very large (e.g., when K = 150 and 600), 2GMM has even worse performance than OLS.

As expected, LIML performs well in terms of median Bias when it is feasible (when K = 2, 10, and 45). In terms of mean Bias, FULLER1 usually performs well (when K = 2, 10, and 45). In general, OLIVE does quite well in terms of bias. It is comparable to these “unbiased” estimators and sometimes the bias of OLIVE is even smaller (for example, when K = 2, 10, and 45 for mean bias).

As the number of instruments increase, the advantage of OLIVE in terms of absolute deviation becomes more significant. When K equals 10 and larger, OLIVE has the smallest median and mean absolute deviations. Moreover, when K is larger than 10, OLIVE also has the smallest mean squared error.

When the number of instruments is larger than the number of time periods ($K>T$), instrumental variable estimators such as 2SLS, LIML, B2SLS, and FULLER all become infeasible. Among the three estimators that are still feasible, OLIVE performs significantly better than both OLS and 2GMM in terms of median and mean bias, median and mean absolute deviation, and mean squared error.

Overall, when the number of instruments increases, the advantage of OLIVE becomes more and more significant (this is also true in the supplemental tables). The performance of OLIVE improves almost monotonically as the number of instruments increases (levels off when K becomes very large). On the other hand, other IV estimators usually peak at a certain number of instruments then deteriorate as the number of instruments further increase. This demonstrates another advantage of OLIVE: one can simply use all valid instruments at hand without having to select instruments or determine the optimal number of instruments.

Table 1 Panel B presents simulation results for the large measurement error and weak instruments case. It is not surprising that when measurement error is large and instruments are weak, none of the instrumental variable estimators perform well. In fact, they do not perform better than the OLS estimator. In this case, OLIVE, like other instrumental variable estimators, does not perform well either. Table 2 presents results when we allow for weak cross-sectional correlation among securities. These results are qualitatively similar to those in Table 1 Panel A. OLIVE performs well compared to other instrumental variable estimators, especially when the number of instruments (K) is large. These simulation results confirm our theoretical prediction in Proposition 2.

[INSERT TABLE 2 HERE]

V. Empirical Application

Background

One of the most successful multifactor models for explaining the cross-section of stock returns is the Fama-French three-factor model. Fama and French (1993) argue that the new factors they identify, "small-minus-big" (SMB) and "high-minus-low" (HML), proxy for unobserved common risk factors. However, both SMB and HML are based on returns on stock portfolios sorted by firm characteristics, and it is not clear what underlying economic risk factors they proxy for. On the other hand, even though macroeconomic factors are theoretically easy to motivate and intuitively appealing, they have had little success in explaining the cross-section of stock returns.

Lettau and Ludvigson (2001b) specify a macroeconomic model that does almost as well as the Fama-French three-factor model in explaining the 25 Fama-French portfolio returns. They explore the ability of conditional versions of the CAPM and the Consumption CAPM (CCAPM) to explain the cross-section of average stock returns. They express a conditional linear factor model as an unconditional multifactor model in which additional factors are constructed by scaling the original factors. This methodology builds on the work in Cochrane (1996), Campbell and Cochrane (1999), and Ferson and Harvey (1999). The choice of the conditioning (scaling) variable in Lettau and Ludvigson (2001b) is unique: *cay* - a cointegrating residual between log consumption *c*, log asset wealth *a*, and log labor income *y*. Lettau and Ludvigson (2001a) finds that *cay* has strong forecasting power for excess returns on aggregate stock market indices. Lettau and Ludvigson (2001b) argue that *cay* may have important advantages as a scaling variable in cross-sectional asset pricing tests because it summarizes investor expectations about the entire market portfolio.

We conjecture that, as with most factors constructed using macroeconomic data, *cay* may contain measurement error. If so, our OLIVE method should improve the findings in Lettau and Ludvigson (2001b). Indeed, our empirical results suggest the presence of large measurement error in *cay* and other macroeconomic factors, but not in return-based factors, such as the Fama-French factors.

Data and Methodology

Our sample is formed using data from the third quarter of 1963 to the third quarter of 1998. We choose the same time period as Lettau and Ludvigson (2001b), so that our results are directly comparable. As in Lettau and Ludvigson (2001b), the returns data are for the 25 Fama-French (1992, 1993) portfolios. These data are value-weighted returns for the intersections of five size portfolios and five book-to-market equity (BE/ME) portfolios on NYSE, AMEX and NASDAQ stocks in CRSP and Compustat. We convert the monthly portfolio returns to quarterly data. The Fama-French factors, SMB and HML, are constructed the same way as in Fama and French (1993). R_{vw} is the value-weighted CRSP index return. The conditioning variable, *cay*, is constructed as in Lettau and Ludvigson (2001a, b). We use the measure of labor income growth, Δy, advocated by Jagannathan and Wang (1996). Labor income growth is measured as the growth in total personal, per capita income less dividend payments from the National Income and Product Accounts published by the Bureau of Economic Analysis. Labor income is lagged one month to capture lags in the official reports of aggregate income.

Our methodology can be viewed as a modified version of Fama and MacBeth's (1973) two-pass method. Lettau and Ludvigson (2001b) discuss different methods available, and argue that the Fama-MacBeth procedure has important advantages for their application. In the first pass, the time-series betas are computed in one multiple regression of the portfolio returns on the factors. In addition to estimating betas by running time-series OLS regressions like in Lettau and Ludvigson (2001b), we also use OLIVE to estimate betas. For a given portfolio (R_i), returns on the other portfolios serve as "instruments" (R_{-i}). As shown by our simulation results, if factors contain measurement error, betas estimated using OLIVE are much more precise than betas estimated using OLS (and more precise than other IV methods).

In the second pass, cross-sectional OLS regressions using 25 Fama-French portfolio returns are run on betas estimated using either OLS or OLIVE in the first pass to draw comparisons:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_i{}'\lambda\,. \tag{25}$$

Empirical Results

Tables 3 and 4 report the Fama-MacBeth cross-sectional regression (second pass) coefficients, $\boldsymbol{\lambda}$, with two *t*-statistics in parentheses for each coefficient estimate. The top *t*-statistic uses uncorrected Fama-MacBeth standard errors, and the bottom *t*-statistic uses the Shanken (1992) correction. The cross-sectional R^2 is also reported. Table 3 (Table 4) corresponds to Table 1 (Table 3) in Lettau and Ludvigson (2001b), with the same row numbers representing the same models. For each row, the OLS results are replications of Lettau and Ludvigson (2001b). After numerous correspondences with the authors (we are grateful for their timely responses), we are able to obtain very similar results, though not completely identical. The OLIVE results are based on our OLIVE beta estimates in the first pass.

[INSERT TABLE 3 HERE]

Unconditional Models. Following Lettau and Ludvigson (2001b), we begin by presenting results from three unconditional models.

Row 1 of Table 3 presents results from the static CAPM, with the CRSP value-weighted return, R_{vw}, used as a proxy for the unobservable market return. This model implies the following cross-sectional specification:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{vwi}\lambda_{vw}. \tag{26}$$

The OLS results in Row 1 highlight the failure of the static CAPM, as documented by previous studies (e.g., Fama and French 1992). Only 1% of the cross-sectional variation in average returns can be explained by the beta for the market return. The estimated value of λ_{vw} is statistically insignificant and has the wrong sign (negative instead of positive) according the CAPM theory. The constant term, which is an estimate of the zero-beta rate, is too high (4.18% per quarter). Estimating betas using OLIVE instead of OLS provides little improvement in terms of cross-sectional explanatory power: the R^2 is still 1%. However, the sign of the estimated value of λ_{vw} changes from negative to positive, though still statistically insignificant, and the estimated zero-beta rate decreases from 4.18% to 3.48% per quarter. We expect the advantage of OLIVE estimation to be small here, since R_{vw} is a return-based factor likely with little measurement error.

Row 2 of Table 3 presents results for the human capital CAPM, which adds the beta for labor income growth, Δy, into the static CAPM (Jagannathan and Wang 1996):

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{vwi}\lambda_{vw} + \beta_{\Delta yi}\lambda_{\Delta y}\,. \tag{27}$$

The human capital CAPM performs much better than the static CAPM, explaining 58% of the cross-sectional variation in returns. Labor income growth is a macroeconomic factor, which probably contains measurement error. When OLIVE is used to estimate betas, the R^2 jumps from 58% to 78%. However, for both OLS and OLIVE results, the estimated value of λ_{vw} has the wrong sign and the estimated zero-beta rate is too high.

Row 3 of Table 3 presents results for the Fama-French three-factor model:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{vwi}\lambda_{vw} + \beta_{SMBi}\lambda_{SMB} + \beta_{HMLi}\lambda_{HML}\,. \tag{28}$$

This specification performs extremely well with OLS estimated betas: the R^2 becomes 81%; the estimated value of λ_{vw} has the correct positive sign; and the estimated zero-beta rate is reasonable (1.76% per quarter). The Fama-French factors should contain little measurement error, since they are constructed from stock returns. As one would expect, using OLIVE estimated betas yields almost identical coefficient estimates. The R^2 only marginally improves to 83%.

Conditional/Scaled Factor Models. Row 4 of Table 3 reports results from the scaled, conditional CAPM with one fundamental factor, the market return, and a single scaling variable, $\widehat{cay}$:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{cayi}\lambda_{cay} + \beta_{vwi}\lambda_{vw} + \beta_{vwcayi}\lambda_{vwcay}\,. \tag{29}$$

Under this specification, using OLIVE instead of OLS to estimate betas dramatically improves the cross-sectional explanatory power from 31% to 80%, which is similar to the performance of the Fama-French three-factor model. This is consistent with our conjecture that since $\widehat{cay}$ is constructed using macroeconomic data, it contains large measurement error. Using OLIVE also changes the sign of the estimated value of λ_{vw} from negative to positive, though the estimated coefficients are close to zero for both OLS and OLIVE. Using OLIVE also reduces the estimated zero-beta rate from 3.69% to 3.09% per quarter, though they are still too high.

Rows 5 and 5' are variations of Row 4. Given the finding that the estimated value of λ_{cay} is not statistically different from zero in Row 4, Row 5 omits β_{cayi} as an explanatory variable in the second-pass cross-sectional regressions, but still includes $\widehat{cay}$ in the first-pass time-series regressions. Row 5' further excludes $\widehat{cay}$ in the first-pass time-series regressions. Results in Rows 5 and 5' are very similar to those in Row 4, suggesting that the time-varying component of the intercept is not an important determinant of cross-sectional returns. The impact of using OLIVE to estimate betas is also very similar: the cross-sectional R^2 jumps from about 30% to about 80%.

Row 6 of Table 3 reports results from the scaled, conditional version of the human capital CAPM:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{cayi}\lambda_{cay} + \beta_{vwi}\lambda_{vw} + \beta_{\Delta yi}\lambda_{\Delta y} + \beta_{vwcayi}\lambda_{vwcay} + \beta_{\Delta ycayi}\lambda_{\Delta ycay}. \quad (30)$$

We focus our discussions on this "complete" specification. Using OLIVE instead of OLS in the first pass to estimate betas improves the second-pass cross-sectional R^2 from 77% to 83% (similar to the performance of the Fama-French three-factor model).

More importantly, our results here help to resolve two puzzling findings by Lettau and Ludvigson (2001b) and Jagannathan and Wang (1996). First, Lettau and Ludvigson (2001b) note that "a problem with this model, however, is that there is a negative average risk price on the beta for the value-weighted return." Jagannathan and Wang (1996) report a similar finding for the signs of the risk prices on the market and human capital betas. Indeed, in our OLS results in Row 6 of Table 3, the estimated value of λ_{vw} (coefficient on the market return beta) is -2.00, and the estimated value of $\lambda_{\Delta ycay}$ (coefficient on the scaled human capital beta) is -0.17, both negative which is inconsistent with the theory. However, when we use OLIVE to estimate betas in the first pass, the estimated value of λ_{vw} becomes positive (1.33), and the estimated value of $\lambda_{\Delta ycay}$ becomes close to zero (-0.0005), more consistent with the theory.

Second, Lettau and Ludvigson (2001b) state that "the average zero-beta rate should be between the average 'riskless' borrowing and lending rates, and the estimated value is implausibly high for the average investor." Jagannathan and Wang (1996) report similar findings. The authors note that "it is possible that the greater sampling error we find in the estimated betas of the scaled models with macro factors is contributing to an upward bias in the zero-beta estimates of those models relative to the estimates for models with only financial factors." They also note that "such arguments for large zero-beta estimates have a long tradition in the cross-sectional asset pricing literature (e.g., Black et al. 1972; Miller and Scholes 1972)." However, the authors conclude that "procedures for discriminating the sampling error explanation for these large estimates of the zero-beta rate from others are not obvious, and its development is left to future research." Our results suggest that measurement error in factors is the cause of this problem. Sampling error is a second-order issue; it becomes negligible as the sample size T becomes large. Unlike sampling error, the measurement error problem does not diminish as the sample size T becomes large. When macroeconomic factors with measurement error are included in the model, OLIVE can provide more precise beta estimates in the first pass, which lead to more precise estimates of the zero-beta rate in the second pass. In Row 6 of Table 2, the estimated zero-beta rate based on OLS estimated betas is too high at 5.19% per quarter.

However, when we use OLIVE to estimate betas, the estimated zero-beta rate drops dramatically to a reasonable 1.91% per quarter.

Rows 7 and 7' are variations of Row 6. Row 7 omits β_{cayi} as an explanatory variable in the second-pass cross-sectional regressions, but still includes $\widehat{cay}$ in the first-pass time-series regressions. Row 7' further excludes $\widehat{cay}$ in the first-pass time-series regressions. Results in Rows 7 and 7' are very similar to those in Row 6. The impact of using OLIVE instead of OLS to estimate betas is also very similar: the cross-sectional R^2 increases; the sign of the estimated value of λ_{vw} changes from negative to positive; and the estimated zero-beta rate drops significantly to a reasonable magnitude.

To summarize, our results in Table 3 confirm the existence of large measurement error in macroeconomic factors, such as $\widehat{cay}$ and labor income growth, and validate the use of OLIVE to help improve beta estimation under these circumstances.

Consumption CAPM. Table 4 presents, for the consumption CAPM, the same results presented in Table 3 for the static CAPM and the human capital CAPM. The scaled multifactor consumption CAPM, with $\widehat{cay}$ as the single conditioning variable takes the form:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{cayi}\lambda_{cay} + \beta_{\Delta ci}\lambda_{\Delta c} + \beta_{\Delta ccayi}\lambda_{\Delta ccay}\,, \tag{31}$$

where Δc denotes consumption growth (log difference in consumption), as measured in Lettau and Ludvigson (2001a).

[INSERT TABLE 4 HERE]

As a comparison, Row 1 of Table 4 reports results of the unconditional consumption CAPM. The performance of this specification is poor, explaining only 16% of the cross-sectional variation in portfolio returns. Using OLIVE beta estimates seems to have made the performance even worse.

Row 2 of Table 4 presents the results of estimating the scaled specification in equation (31). The R^2 jumps to 70%, in sharp contrast to the unconditional results in Row 1. When OLIVE is used to estimate betas, the R^2 further increases to 82%. For both OLS and OLIVE results, the estimated value of $\lambda_{\Delta ccay}$ (scaled consumption growth) is positive and statistically significant.

Row 3 excludes β_{cayi} as an explanatory variable in the second-pass cross-sectional regressions, but still includes $\widehat{cay}$ in the first-pass time-series regressions. This seems to have made very little difference, as the results in Row 3 are very similar to those in Row 2. Again, when OLIVE estimated betas are used, the R^2 increases from 69% to 81%.

Row 3' further excludes $\widehat{cay}$ in the first-pass time-series regressions. As noted by Lettau and Ludvigson (2001b), the results here are somewhat sensitive to this exclusion (see their footnote 25). The R^2 drops to 27% for OLS results and 34% for OLIVE results. These results suggest that including the scaling variable $\widehat{cay}$ as a factor in the pricing kernel can be important even when the beta for this factor is not priced in the cross-section.

Our results in Table 4 suggest that using OLIVE instead of OLS to estimate betas in the conditional consumption CAPM generally increases the cross-sectional variation of portfolio returns explained by the model, as measured by the R^2. However, unlike in Table 3, the estimated zero-beta rates remain high.

VI. Conclusion

In this paper, we put forth a simple method for estimating betas (factor loadings) when factors are measured with error, which we call OLIVE. OLIVE uses all available instruments at hand, and is intuitive and easy to implement. OLIVE achieves better performance in simulations than OLS and other instrumental variable estimators such as 2SLS, B2SLS, LIML, and FULLER, when the number of instruments is large. OLIVE can be interpreted as a GMM estimator when setting the weighting matrix equal to the identity matrix and it has better finite sample properties than the efficient two-step GMM estimator. OLIVE also has an important advantage over the Asymptotic Principle Components (APC) because the statistical factors of the principle components method lack clear economic interpretations, while OLIVE directly makes use of the observed economic factors.

OLIVE has many potential empirical applications and is especially suitable for estimating asset return betas when factors are measured with error, since this is often a large N and small T setting. Intuitively, since all asset returns vary together with a common set of factors, one can use information contained in other asset returns to improve the beta estimate for a given asset.

As an empirical application, we reexamine Lettau and Ludvigson's (2001b) test of the (C)CAPM using OLIVE in addition to OLS to estimate betas. Lettau and Ludvigson's factor *cay* has been found to have strong forecasting power for excess returns on aggregate stock market indices, but may contain measurement error. We find that in regressions where macroeconomic factors are included, using OLIVE instead of OLS improves the R^2 significantly. Perhaps more importantly, our results from OLIVE estimation help to resolve two puzzling findings by Lettau and Ludvigson (2001b) and Jagannathan and Wang (1996): first, the sign of the average risk premium on the beta for the market return changes from negative to positive, which is in accordance with the theory; second, the estimated value of average zero-beta rate is no longer too high. These results suggest that when macroeconomic factors with measurement error are included in the model, OLIVE can provide more precise beta estimates in the first pass, which lead to more precise estimates of the risk premia and zero-beta rate in the second pass. Our results from this empirical application validate the use of OLIVE to help improve beta estimation when factors are measured with error. Our findings are also consistent with the theme in Ferson, Sarkissian, and Simin (2008) that the (C)CAPMs might work better than previously recognized in the literature.

Appendix A. Proof of Proposition 1

To simplify notation, we consider a more abstract setting. Let

$$y_t = \beta' x_t + \varepsilon_t = x_t' \beta + \varepsilon_t , \tag{A1}$$

where x_t and β are $M \times 1$ vectors, $E(x_t \varepsilon_t) \neq 0$, and $x_t = x_t^* + v_t$. Let $z_{it} = \beta_i' x_t^* + e_{it}$ be instruments (i = 1, …, N; t = 1, …, T). Here we assume there are N instruments (i.e., N+1 assets). For example, to estimate B_1 in the notation of Section II, we let $\beta = B_1$, and $y_t = y_{1t}$, $\varepsilon_t = e_{1t}$, and $z_{it} = y_{i+1,t}$ for $i \geq 1$. Then

$$\frac{1}{T}\sum_{t=1}^{T} z_{it} y_t = \frac{1}{T}\sum_{t=1}^{T} z_{it} x_t' \beta + \frac{1}{T}\sum_{t=1}^{T} z_{it}\varepsilon_t , \tag{A2}$$

or it can be simplified as

$$\bar{y}_i = \bar{x}_i' \beta + \bar{\varepsilon}_i , \tag{A3}$$

where $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_t z_{it}'$, $\bar{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} z_{it}\varepsilon_t$, and $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} z_{it} y_t$. The estimator OLIVE is $\hat{\beta}^{OLIVE} = \left(\bar{x}'\bar{x}\right)^{-1}\bar{x}'\bar{y}$. Now

$$\begin{aligned} E(\bar{x}_i \bar{\varepsilon}_i) &= E\left(\frac{1}{T^2}\sum_{t=1}^{T} z_{it} x_t' \sum_{s=1}^{T} z_{is}\varepsilon_s\right) \\ &= \frac{1}{T} E\left(\frac{1}{T}\sum_{t=1}^{T} z_{it} x_t' z_{it}\varepsilon_t\right) + \frac{1}{T}\left[\frac{1}{T}\right]\sum_{\substack{t=1 \\ t\neq s}}^{T}\sum_{s=1}^{T} E\left(z_{it} x_t' z_{is}\varepsilon_s\right). \\ &= O\left(\frac{1}{T}\right) \to 0 \quad as \quad t \to \infty \end{aligned} \tag{A4}$$

Therefore,

$$\begin{aligned} \hat{\beta}^{OLIVE} - \beta &= \left(\sum_{i=1}^{N} \bar{x}_i \bar{x}_i'\right)^{-1} \sum_{i=1}^{N} \bar{x}_i \bar{\varepsilon}_i \\ &= \left(\frac{1}{N}\sum_{i=1}^{N} \bar{x}_i \bar{x}_i'\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} \bar{x}_i \bar{\varepsilon}_i\right) \end{aligned} \tag{A5}$$

$$\sqrt{T}\left(\widehat{\beta}^{OLIVE}-\beta\right)=\left(\frac{1}{N}\sum_{i=1}^{N}\bar{x}_i\bar{x}_i'\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{T}\,\bar{x}_i\bar{\varepsilon}_i\right)$$
$$=\left(\frac{1}{N}\sum_{i=1}^{N}\bar{x}_i\bar{x}_i'\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{1}{T}\sum_{t=1}^{T}x_t z_{it}'\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}z_{it}\varepsilon_t\right)\right],\qquad \text{(A6)}$$
$$=A_N^{-1}\frac{1}{N}\sum_{i=1}^{N}D_{iT}\xi_{iT}$$

where $A_N=\frac{1}{N}\sum_{i=1}^{N}\bar{x}_i\bar{x}_i'$, $D_{iT}=\frac{1}{T}\sum_{t=1}^{T}x_t z_{it}'$, and $\xi_{iT}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}z_{it}\varepsilon_t$. Note that ξ_{iT} and ξ_{jT} are dependent through the common term $\sum_{t=1}^{T}x_t*\varepsilon_t$, see (A8) below. The instruments z_{it} are determined by true factor x_t*:

$$z_{it}=\beta_i'x_t*+e_{it}, \qquad \text{(A7)}$$

therefore,

$$\xi_{iT}=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\beta_i'x_t*\varepsilon_t+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\varepsilon_t, \qquad \text{(A8)}$$

and

$$D_{iT}\xi_{iT}=D_{iT}\beta_i'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_t*\varepsilon_t+D_{iT}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\varepsilon_t. \qquad \text{(A9)}$$

$$\frac{1}{N}\sum_{i=1}^{N}D_{iT}\xi_{iT}=\left[\frac{1}{N}\sum_{i=1}^{N}D_{iT}\beta_i'\right]\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_t*\varepsilon_t\right]+\frac{1}{N}\sum_{i=1}^{N}\left[D_{iT}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\varepsilon_t\right]. \qquad \text{(A10)}$$

We have

$$\left[\frac{1}{N}\sum_{i=1}^{N}D_{iT}\beta_i'\right]\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_t*\varepsilon_t\right]\to\Gamma N(0,\Omega), \qquad \text{(A11)}$$

where $\left[\frac{1}{N}\sum_{i=1}^{N}D_{iT}\beta_i'\right]\to\Gamma$, and $\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_t*\varepsilon_t\right]\to N(0,\Omega)$. We also have

$$\frac{1}{N}\sum_{i=1}^{N}\left[D_{iT}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\varepsilon_t\right]=O_p(N^{-1/2})\rightarrow 0\,. \tag{A12}$$

To see (A12) is $O_p(N^{-1/2})$, we write

$$D_{iT}=\frac{1}{T}\sum_{t=1}^{T}x_t z_{it}'=\frac{1}{T}\sum_{t=1}^{T}\left(x_t^*+v_t\right)\left(x_t^{*\prime}\beta_i+e_{it}\right)=\frac{1}{T}\sum_{t=1}^{T}x_t^*x_t^{*\prime}\beta_i+O_p\left(T^{-1/2}\right),$$

and $E\left(D_{it}\right)=G\beta_i$, where $G=E\left(\frac{1}{T}\sum_{t=1}^{T}x_t^*x_t^{*\prime}\right)$. Adding and subtracting $E\left(D_{iT}\right)$, then equation (A12) can be rewritten as

$$\frac{1}{N}\sum_{i=1}^{N}\left[D_{iT}-E\left(D_{iT}\right)\right]\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\varepsilon_t+GN^{-1/2}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i e_{it}\varepsilon_t=I+II\,. \tag{A13}$$

Since $D_{iT}-E\left(D_{iT}\right)=O_p\left(T^{-1/2}\right)$, term *I* is dominated by *II*. From

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i e_{it}\varepsilon_t=O_p\left(1\right), \tag{A14}$$

we have $II=O_p\left(N^{-1/2}\right)$.

If N is fixed, (A12) is $O_p\left(1\right)$ and is not negligible. This term will contribute to the limiting distribution; but the $\sqrt{T}$ consistency and the asymptotic normality still hold.

Appendix B. Proof of Proposition 2

The proof of Proposition 1 remains valid up to (A11). We show (A12) is still asymptotically negligible if $\sqrt{T}/N\rightarrow 0$. It is sufficient to consider *II* in (A13). Let $\gamma_i=E\left(e_{it}\varepsilon_t\right)$, with $\gamma_i\neq 0$, equation (A14) will no longer hold. But it can be rewritten as

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i e_{it}\varepsilon_t=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i\left[e_{it}\varepsilon_t-E\left(e_{it}\varepsilon_t\right)\right]+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i\gamma_i\,. \tag{B1}$$

The first term on the right hand side is $O_p(1)$. Assuming $\|\beta_i\| \leq M$ for all *i*, the second term is bounded by $M\left(\sqrt{T/N}\right)\sum_{i=1}^{N}|\gamma_i| = O\left(\sqrt{T/N}\right)$ because $\sum_{i=1}^{N}|\gamma_i| = O(1)$ by assumption (9). Thus (B1) is $O_p(1) + O_p(\sqrt{T/N})$. This implies that, noting the extra term $N^{-1/2}$, *II* in (A13) is equal to $O_p\left(N^{-1/2}\right) + O_p(\sqrt{T}/N)$, which converges to zero if $\sqrt{T}/N \to 0$.

Appendix C. Derivations for Joint GMM

We derive the formulae for w_{jk}^{ih}, the (j,k)th element of the block off-diagonal matrix W_{ih}, in each of the following four possible cases.

Case 1: $j \neq h \text{ and } k \neq i$.

$$\begin{aligned} w_{jk}^{ih} &= E(u_{ijt}u_{hkt}) = E\left[y_{jt}(e_{it} - v_t{}'B_i)(e_{ht} - B_h{}'v_t)y_{kt} \right] \\ &= E\left[y_{jt}\left(B_i{}'\mathrm{var}(v_t)B_h\right)y_{kt} \right] \\ &= E\left[\left(B_j{}'x_t^* + e_{jt}\right)\left(x_t^*{}'B_k + e_{kt}\right)\left(B_i{}'\mathrm{var}(v_t)B_h\right) \right] \\ &= E\left[\left(B_j{}'\mathrm{var}(x_t^*)B_k + \delta_{jk}\,\mathrm{var}(e_{jt})\right)\left(B_i{}'\mathrm{var}(v_t)B_h\right) \right] \\ &= \left(B_j{}'\mathrm{var}(x_t^*)B_k + \delta_{jk}\,\mathrm{var}(e_{jt})\right)\left(B_i{}'\mathrm{var}(v_t)B_h\right) \end{aligned} \quad , \tag{C1}$$

where $\delta_{jk} = 1$ if $j = k$, and zero otherwise.

Case 2: $j = h \text{ and } k \neq i$.

$$
\begin{aligned}
w_{jk}^{ih} &= E(u_{ijt}u_{hkt}) = E\left[y_{ht}(e_{it} - v_t{}'B_i)(e_{ht} - B_h{}'v_t)y_{kt} \right] \\
&= E\left[\left(B_h{}'x_t^* + e_{ht} \right)\left(e_{ht} - B_h{}'v_t \right)\left(e_{it} - v_t{}'B_i \right)\left(x_t^*{}'B_k + e_{kt} \right) \right] \\
&= E\left[\left(B_h{}'x_t^* e_{ht} - B_h{}'x_t^* v_t{}'B_h + e_{ht}^2 - B_h{}'v_t e_{ht} \right)\left(B_k{}'x_t^* e_{it} - B_k{}'x_t^* v_t{}'B_i + e_{it}e_{kt} - B_i{}'v_t e_{kt} \right) \right] \\
&= E\begin{bmatrix} B_h{}'x_t^* x_t^*{}'B_k \bullet e_{ht}e_{it} - B_h{}'x_t^* x_t^*{}'B_k \bullet B_i{}'v_t \bullet e_{ht} + B_h{}'x_t^* e_{ht}e_{it}e_{kt} - B_h{}'x_t^* v_t{}'B_i e_{ht}e_{kt} \\ -B_h{}'x_t^* x_t^*{}'B_k \bullet B_h{}'v_t e_{it} + B_h{}'x_t^* x_t^*{}'B_k \bullet B_h{}'v_t v_t{}'B_i - B_h{}'x_t^* v_t{}'B_h \bullet e_{it}e_{kt} + B_h{}'v_t v_t{}'B_i \bullet B_h{}'x_t^* e_{kt} \\ +B_k{}'x_t^* \bullet e_{it}e_{ht}^2 - B_k{}'x_t^* v_t{}'B_i \bullet e_{ht}^2 + e_{it}e_{kt}e_{ht}^2 - B_i{}'v_t e_{kt}e_{ht}^2 \\ -B_k{}'x_t^* v_t{}'B_h \bullet e_{ht}e_{it} + B_h{}'v_t v_t{}'B_i \bullet B_k{}'x_t^* \bullet e_{ht} - B_h{}'v_t e_{ht}e_{it}e_{kt} + B_h{}'v_t v_t{}'B_i \bullet e_{ht}e_{kt} \end{bmatrix} \qquad \text{(C2)} \\
&= E\left[B_h{}'x_t^* x_t^*{}'B_k \bullet B_h{}'v_t v_t{}'B_i \right] \\
&= B_h{}'\operatorname{var}(x_t^*)B_k \bullet B_h{}'\operatorname{var}(v_t)B_i
\end{aligned}
$$

Case 3, $j \neq h$ and $k = i$. This is the mirror case of Case 2. In short, the formula is:

$$
\begin{aligned}
w_{jk}^{ih} &= E(u_{ijt}u_{hkt}) = E\left[y_{jt}(e_{it} - v_t{}'B_i)(e_{ht} - B_h{}'v_t)y_{it} \right] \\
&= B_i{}'\operatorname{var}(x_t^*)B_j \bullet B_h{}'\operatorname{var}(v_t)B_i
\end{aligned}. \qquad \text{(C3)}
$$

Case 4: $j = h \; and \; k = i$.

$$
\begin{aligned}
w_{jk}^{ih} &= E(u_{ijt}u_{hkt}) = E\left[y_{ht}(e_{it} - v_t{}'B_i)(e_{ht} - B_h{}'v_t)y_{it} \right] \\
&= E\left[\left(B_h{}'x_t^* + e_{ht} \right)\left(e_{it} - v_t{}'B_i \right)\left(e_{ht} - B_h{}'v_t \right)\left(x_t^*{}'B_i + e_{it} \right) \right] \\
&= E\left[\left(B_h{}'x_t^* e_{it} - B_h{}'x_t^* v_t{}'B_i + e_{ht}e_{it} - B_i{}'v_t e_{ht} \right)\left(B_i{}'x_t^* e_{ht} + e_{ht}e_{it} - B_h{}'v_t x_t^*{}'B_i - B_h{}'v_t e_{it} \right) \right] \\
&= E\begin{bmatrix} B_h{}'x_t^* x_t^*{}'B_i \bullet e_{ht}e_{it} + B_h{}'x_t^* \bullet e_{ht}e_{it}^2 - B_h{}'x_t^* x_t^*{}'B_i \bullet B_h{}'v_t \bullet e_{it} - B_h{}'x_t^* v_t{}'B_h \bullet e_{it}^2 \\ -B_h{}'x_t^* x_t^*{}'B_i \bullet B_i{}'v_t \bullet e_{ht} - B_h{}'x_t^* v_t{}'B_i \bullet e_{ht}e_{it} + B_h{}'x_t^* x_t^*{}'B_i \bullet B_h{}'v_t v_t{}'B_i + B_h{}'v_t v_t{}'B_i \bullet B_h{}'x_t^* \bullet e_{it} \\ +B_i x_t^* \bullet e_{ht}^2 e_{it} + e_{ht}^2 e_{it}^2 - B_h{}'v_t x_t^*{}'B_i \bullet e_{ht}e_{it} - B_h{}'v_t \bullet e_{ht}e_{it}^2 \\ -B_i{}'v_t x_t^*{}'B_i \bullet e_{ht}^2 - B_i{}'v_t \bullet e_{ht}^2 e_{it} + B_h{}'v_t v_t{}'B_i \bullet B_i{}'x_t^* \bullet e_{ht} + B_h{}'v_t v_t{}'B_i \bullet e_{ht}e_{it} \end{bmatrix} \qquad \text{(C4)} \\
&= E\left[B_h{}'x_t^* x_t^*{}'B_i \bullet B_h{}'v_t v_t{}'B_i + e_{ht}^2 e_{it}^2 \right] \\
&= B_h{}'\operatorname{var}(x_t^*)B_i \bullet B_h{}'\operatorname{var}(v_t)B_i + \operatorname{var}(e_{it})^2
\end{aligned}
$$

References

Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica* 71, 135-171.

Barnes, M. , and A. Hughes, 2002 A quantile regression analysis of the cross section of stock market returns, Working Paper, Federal Reserve Bank of Boston.

Black, F., M. Jensen, and M. Scholes, 1972, The Capital Asset Pricing Model: some empirical tests, in M. Jensen, eds.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Campbell, J., and J. Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

Campbell, J., A. Lo, and C. MacKinlay, 1997, *The econometrics of financial markets* (Princeton University Press).

Chamberlain, G., and M. Rothschild, 1983, Arbitrage, factor structure and mean-variance analysis in large asset markets, *Econometrica* 51, 1305-1324.

Chao, J, and N. Swanson, 2005, Consistent Estimation with a Large Number of Weak Instruments, *Econometrica*, 73, 1673-1692.

Chen, N., R. Roll, and S. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383-403.

Cochrane, J., 1996, A cross-sectional test of an investment-base asset pricing model, *Journal of Political Economy* 104, 572-621.

Cochrane, J., 2001, *Asset pricing* (Princeton University Press).

Coën, A., and F. Racicotv, 2007, Capital asset pricing models revisited: evidence from errors in variables, *Economic Letters* 95, 443-450.

Connor, G., and R. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory, *Journal of Financial Economics* 15, 373-394.

Connor, G., and R. Korajczyk, 1991, The attributes, behavior and performance of U.S. mutual funds, *Review of Quantitative Finance and Accounting* 1, 5-26.

Connor, G., and R. Korajczyk, 1993, A test for the number of factors in an approximate factor model, *Journal of Finance* 48, 1263-1291.

Davidson, R., and J. MacKinnon, 1993, *Estimation and inference in econometrics* (Oxford University Press, New York).

Donald, S., and W. Newey, 2001, Choosing the number of instruments, *Econometrica* 69, 1161-1191.

Doran, H., and P. Schmidt, 2006, GMM estimators with improved finite sample properties using principal components of the weighting matrix, with an application to the dynamic panel data model, *Journal of Econometrics* 133, 387 -409.

Durbin, J., 1954, Errors in variables, *International Statistical Review* 22, 23-32.

Fama, E., and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Fama, E., and K. French, 1992, The cross-section of expected returns, *Journal of Finance* 47, 427-465.

Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Ferson, W., and S. Foerster, 1994, Finite sample properties of the Generalized Methods of Moments tests of conditional asset pricing models, *Journal of Financial Economics* 36, 29-55.

Ferson, W., and C. Harvey, 1999, Conditioning variables and the cross-section of stock returns, *Journal of Finance* 54, 1325-1360.

Ferson, W., S. Sarkissian, and T. Simin, 2008, Asset pricing models with conditional betas and alphas: The effects of data snooping and spurious regression, *Journal of Financial and Quantitative Analysis* 43, 331-354.

Fuller, W., 1977, Some properties of a modification of the limited information estimator, *Econometrica* 45, 939-954.

Hahn, J., and J. Hausman, 2002, A new specification test for the validity of instrumental variables, *Econometrica* 70, 163-189.

Hahn, J., J. Hausman, and G. Kuersteiner, 2004, Estimation with weak instruments: accuracy of higher-order bias and MSE approximations, *Econometrics Journal* 7, 272-306.

Hansen, L., J. Heaton, and A. Yaron, 1996, Finite-sample properties of some alternative GMM estimators, *Journal of Business and Economic Statistics* 14, 262-280.

Hussman, J., 1993, A note on the interpretation of cross-sectional evidence against the beta-expected return relationship, Working Paper, University of Michigan.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-54.

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

Jones, C., 2001, Extracting factors from heteroskedastic asset returns, *Journal of Financial Economics* 62, 293-325.

Kim, D., 1995, The errors in the variables problem in the cross-section of expected stock returns, *Journal of Finance* 50, 1605-1634.

Lettau, M., and S. Ludvigson, 2001a, Consumption, aggregate wealth and expected stock returns, *Journal of Finance* 56, 815-849.

Lettau, M., and S. Ludvigson, 2001b, Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109,1238-1287.

Miller, M., and M. Scholes, 1972, Rates of return in relation to risk: A re-examination of some recent findings, in M. Jensen, eds.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Newey, W., and R. Smith, 2004, Higher order properties of GMM and generalized empirical likelihood estimators, *Econometrica* 72, 219-255.

Ng, S., 2006, Testing cross-section correlation in panel data using spacings, *Journal of Business and Economic Statistics* 24, 12–23.

Pal, M., 1980, Consistent moment estimators of regression coefficients in the presence of errors in variables, *Journal of Econometrics* 14, 349-364.

Pesaran, M., 2004, General diagnostic tests for cross section dependence in panels, Working Paper, CESifo.

Roll, R., 1977, A critique of asset pricing theory's test: Part 1, *Journal of Financial Economics* 4, 129-176.

Rothenberg, T., 1983, Asymptotic properties of some estimators in structural models, in *Studies in Econometrics, Time Series, and Multivariate Statistics.*

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

Wansbeek, T., and E. Meijer, 2000, Measurement error and latent variables in econometrics, in *Advanced textbooks in economics* (North-Holland, Amsterdam).

Wei, J., C. Lee, and A. Chen, 1991, Multivariate regression tests of the arbitrage pricing theory: the instrumental-variables approach, *Review of Quantitative Finance and Accounting* 1, 191-208.

Windmeijer, F., 2005, A finite sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics* 126, 25-51.

Wyhowski, D., 1998, Monte Carlo evidence for dynamic panel data models, Working Paper, Australian National University.

TABLE 1. Simulation Results.

TABLE 1. PANEL A. Medium Measurement Error, Medium Instruments. σ_v=0.1, σ_e=0.1.

K	Estimator	Mean Bias	Mean AD	SQRT. MSE	IQR	Dec. Rge	Cov. Rate
2	OLS	-0.3300	0.3300	0.3440	0.1251	0.2499	1.0000
	2SLS	0.0028	0.1236	0.1625	0.1999	0.3831	1.0000
	LIML	0.0104	0.1378	0.2607	0.1997	0.3865	0.9980
	B2SLS	0.0028	0.1236	0.1625	0.1999	0.3831	1.0000
	FULLER1	0.0009	0.1227	0.1620	0.1962	0.3769	1.0000
	FULLER4	-0.0336	0.1159	0.1473	0.1823	0.3444	1.0000
	OLIVE	0.0111	0.1249	0.1641	0.2024	0.3854	1.0000
	2GMM	0.0025	0.1239	0.1615	0.2038	0.3780	1.0000
10	OLS	-0.3305	0.3305	0.3444	0.1306	0.2416	1.0000
	2SLS	-0.0657	0.1161	0.1425	0.1641	0.3097	1.0000
	LIML	0.0092	0.1117	0.1442	0.1820	0.3503	1.0000
	B2SLS	-0.0008	0.1137	0.1466	0.1836	0.3550	1.0000
	FULLER1	0.0000	0.1100	0.1408	0.1773	0.3427	1.0000
	FULLER4	-0.0265	0.1076	0.1354	0.1693	0.3175	1.0000
	OLIVE	0.0055	0.1073	0.1385	0.1818	0.3369	1.0000
	2GMM	-0.0050	0.2826	2.1224	0.1822	0.3814	0.9900
45	OLS	-0.3300	0.3300	0.3432	0.1271	0.2384	1.0000
	2SLS	-0.2672	0.2676	0.2851	0.1318	0.2497	1.0000
	LIML	0.0297	0.1530	0.2170	0.2414	0.4661	0.9990
	B2SLS	-0.0114	0.1829	0.2529	0.2657	0.5385	0.9990
	FULLER1	0.0186	0.1479	0.2044	0.2337	0.4514	1.0000
	FULLER4	-0.0121	0.1372	0.1804	0.2132	0.4182	1.0000
	OLIVE	0.0061	0.1036	0.1325	0.1760	0.3320	1.0000
	2GMM	-0.4270	0.5082	3.9644	0.2098	0.4392	0.9880
150	OLS	-0.3317	0.3317	0.3453	0.1330	0.2443	1.0000
	OLIVE	0.0040	0.1032	0.1315	0.1773	0.3239	1.0000
	2GMM	-0.3492	0.5876	1.3594	0.2679	0.6326	0.9760
600	OLS	-0.3336	0.3336	0.3469	0.1268	0.2483	1.0000
	OLIVE	0.0099	0.1055	0.1318	0.1806	0.3353	1.0000
	2GMM	-0.9429	1.1228	5.4738	0.3163	0.8479	0.9570

TABLE 1. PANEL B. Large Measurement Error, Weak Instruments. σ_v=1, σ_e=1.

K	Estimator	Mean Bias	Mean AD	SQRT. MSE	IQR	Dec. Rge	Cov. Rate
2	OLS	-0.9874	0.9874	0.9965	0.1788	0.3459	1.0000
	2SLS	-1.1503	1.5473	6.1335	1.0882	2.5046	0.8440
	LIML	-9.8045	12.5131	276.1521	1.9171	5.9127	0.6640
	B2SLS	-1.1503	1.5473	6.1335	1.0882	2.5046	0.8440
	FULLER1	-0.9446	0.9604	1.0804	0.6494	1.3348	0.9750
	FULLER4	-0.9695	0.9695	1.0025	0.3166	0.6293	0.9990
	OLIVE	-1.1474	1.5594	6.4277	1.0981	2.6147	0.8410
	2GMM	-0.7191	1.7035	4.7551	1.1113	2.6944	0.8230
10	OLS	-0.9839	0.9839	0.9918	0.1685	0.3182	1.0000
	2SLS	-0.9523	0.9548	1.0224	0.4443	0.9014	0.9910
	LIML	-1.2332	4.5724	21.8788	2.1373	6.5960	0.6440
	B2SLS	-0.0264	4.1301	28.0351	1.3482	4.3566	0.7460
	FULLER1	-0.8979	0.9889	1.2033	1.1304	2.1220	0.8960
	FULLER4	-0.9425	0.9437	1.0214	0.5661	1.0293	0.9970
	OLIVE	-0.9447	0.9495	1.0284	0.4911	0.9903	0.9890
	2GMM	-1.0307	1.1889	4.8548	0.5351	1.0949	0.9720
45	OLS	-0.9755	0.9755	0.9840	0.1672	0.3238	1.0000
	2SLS	-0.9660	0.9660	0.9770	0.1964	0.3701	1.0000
	LIML	-1.3414	4.5150	18.5643	2.0296	6.7210	0.6470
	B2SLS	0.8161	6.2762	63.5363	1.3027	3.9060	0.7640
	FULLER1	-0.9182	1.2109	1.5134	1.5227	3.1658	0.7920
	FULLER4	-0.9311	0.9555	1.1331	1.0281	1.6975	0.9430
	OLIVE	-0.9371	0.9371	0.9605	0.2548	0.5220	1.0000
	2GMM	-0.9600	1.0554	1.9130	0.2843	0.5908	0.9780
150	OLS	-0.9747	0.9747	0.9835	0.1716	0.3191	1.0000
	OLIVE	-0.9376	0.9376	0.9537	0.2304	0.4401	1.0000
	2GMM	-0.8838	1.4939	6.8534	0.2629	0.6783	0.9540
600	OLS	-0.9724	0.9724	0.9814	0.1746	0.3246	1.0000
	OLIVE	-0.9338	0.9338	0.9468	0.1939	0.3728	1.0000
	2GMM	-0.9629	1.8064	9.3302	0.3039	1.0251	0.9230

Note: This table presents the simulation results. We compare OLIVE with other IV estimators including 2SLS, LIML, B2SLS, FULLER1 and FULLER4 (the choice of the α parameter is either 1 or 4), as well as the two-step equation-by-equation GMM estimator (2GMM). We first generate a security with a true beta of one, which is to be estimated using the above estimators. Then we generate $K = N$-1 other securities which serve as instruments using the data generating process (DGP) detailed in Section IV. We use K = (2, 10, 45, 150, 600), T = 60, π = 0.1, σ_x = 0.1, σ_β = 1, and 1,000 replications. Without loss of generality, we set J, the number of explanatory variables to be 1, which makes the model specification equivalent to the CAPM for the excess return. We allow x and β to be normally generated. One advantage of OLIVE is that when K is larger than T, it still works while most other IV estimators no longer do. This is why we can only compare the performance of OLS, OLIVE, and 2GMM for K = (150, 600). The statistics reported include mean bias (Mean Bias), mean absolute deviation (Mean AD), squared-root of mean squared error (SQRT. MSE), inter-quartile range (IQR), the difference between the 1st and 9th deciles (Dec. Rge), and the coverage rate of a nominal 95% confidence interval (Cov. Rate). In Panel A we set $\sigma_v = 0.1$ and $\sigma_e = 0.1$ (medium measurement error and medium instruments). In Panel B we set $\sigma_v = 1$ and $\sigma_e = 1$ (large measurement error and weak instruments).

TABLE 2. Simulation Results, Allowing for Weak Cross-Sectional Correlation.

K	Estimator	Mean Bias	Mean AD	SQRT. MSE	IQR	Dec. Rge	Cov. Rate
2	OLS	-0.3296	0.3296	0.3446	0.1368	0.2562	1.0000
	2SLS	0.0106	0.1335	0.2009	0.2065	0.4021	0.9990
	LIML	0.0254	0.1414	0.2278	0.2141	0.4186	0.9980
	B2SLS	-0.0106	0.1335	0.2009	0.2065	0.4021	0.9990
	FULLER1	0.0062	0.1276	0.1698	0.2047	0.4059	1.0000
	FULLER4	-0.0323	0.1190	0.1508	0.1902	0.3683	1.0000
	OLIVE	0.0190	0.1355	0.2016	0.2096	0.4090	0.9990
	2GMM	0.0120	0.1379	0.2116	0.2102	0.4088	0.9990
10	OLS	-0.3300	0.3300	0.3433	0.1307	0.2386	1.0000
	2SLS	-0.0669	0.1153	0.1406	0.1461	0.3099	1.0000
	LIML	0.0067	0.1103	0.1401	0.1884	0.3434	1.0000
	B2SLS	-0.0023	0.1124	0.1430	0.1896	0.3500	1.0000
	FULLER1	-0.0025	0.1086	0.1371	0.1835	0.3345	1.0000
	FULLER4	-0.0288	0.1066	0.1327	0.1715	0.3205	1.0000
	OLIVE	0.0051	0.1074	0.1356	0.1810	0.3401	1.0000
	2GMM	-0.0540	0.1755	0.6182	0.1919	0.3629	0.9930
45	OLS	-0.3329	0.3329	0.3462	0.1255	0.2384	1.0000
	2SLS	-0.2701	0.2701	0.2879	0.1360	0.2557	1.0000
	LIML	0.0316	0.1588	0.2209	0.2508	0.4924	0.9990
	B2SLS	-0.0130	0.1805	0.2380	0.2830	0.5421	1.000
	FULLER1	0.0198	0.1528	0.2049	0.2424	0.4823	1.0000
	FULLER4	-0.0117	0.1403	0.1817	0.2254	0.4482	1.0000
	OLIVE	0.0144	0.1071	0.1377	0.1760	0.3361	1.0000
	2GMM	-0.2680	0.3699	2.7643	0.2040	0.4195	0.9850
150	OLS	-0.3339	0.3339	0.3475	0.1273	0.2515	1.0000
	OLIVE	0.0026	0.1091	0.1400	0.1822	0.3351	1.0000
	2GMM	-0.3814	0.5515	1.0523	0.2612	0.6299	0.9760
600	OLS	-0.3303	0.3303	0.3435	0.1253	0.2394	1.0000
	OLIVE	0.0133	0.1056	0.1336	0.1805	0.3375	1.0000
	2GMM	0.9979	2.5784	50.2473	0.2964	0.8726	0.9550

Note: This table presents the simulation results, allowing for weak cross-sectional correlation among securities. We compare OLIVE with other IV estimators including 2SLS, LIML, B2SLS, FULLER1 and FULLER4 (the choice of the α parameter is either 1 or 4), as well as the two-step equation-by-equation GMM estimator (2GMM). We first generate a security with a true beta of one, which is to be estimated using the above estimators. Then we generate $K = N$-1 other securities which serve as instruments using the data generating process (DGP) detailed in Section IV. We use $K = (2, 10, 45, 150, 600)$, $T = 60$, $\pi = 0.1$, $\sigma_x = 0.1$, $\sigma_\beta = 1$, and 1,000 replications. Without loss of generality, we set J, the number of explanatory variables to be 1, which makes the model specification equivalent to the CAPM for the excess return. We allow x and β to be normally generated. One advantage of OLIVE is that when K is larger than T, it still works while most other IV estimators no longer do. This is why we can only compare the performance of OLS, OLIVE, and 2GMM for K = (150, 600). The statistics reported include mean bias (Mean Bias), mean absolute deviation (Mean AD), squared-root of mean squared error (SQRT. MSE), inter-quartile range (IQR), the difference between the 1st and 9th deciles (Dec. Rge), and the coverage rate of a nominal 95% confidence interval (Cov. Rate).

TABLE 3. Fama-MacBeth Regressions Using 25 Fama-French Portfolios: λ_j Coefficient Estimates on Betas in Cross-Sectional Regression

			Factors$_{t+1}$				$\widehat{cay}_t \cdot$ Factors$_{t+1}$		
Row	Constant	$\widehat{cay}_t$	R_{vw}	Δy	SMB	HML	R_{vw}	Δy	R^2
1-OLS	4.18 (4.47) (4.47)		-0.32 (-0.27) (-0.27)						0.01
1-OLIVE	3.48 (4.71) (4.70)		0.25 (0.30) (0.30)						0.01
2-OLS	4.47 (4.78) (2.66)		-1.10 (-0.95) (-0.53)	1.26 (3.40) (1.89)					0.58
2-OLIVE	4.42 (5.76) (5.69)		-1.22 (-1.32) (-1.29)	0.06 (3.55) (4.49)					0.78
3-OLS	1.76 (1.22) (1.12)		1.44 (0.89) (0.82)		0.49 (0.97) (0.89)	1.49 (3.31) (3.03)			0.81
3-OLIVE	1.79 (1.63) (1.49)		1.41 (1.11) (1.01)		0.48 (0.96) (0.88)	1.53 (3.39) (3.10)			0.83
4-OLS	3.69 (3.90) (2.62)	-0.04 (-0.19) (-0.13)	-0.07 (-0.06) (-0.04)				1.14 (3.60) (2.41)		0.31
4-OLIVE	3.09 (4.20) (4.18)	-0.02 (-0.60) (-0.58)	0.13 (0.15) (0.15)				0.10 (3.54) (3.51)		0.80
5-OLS	3.70 (3.86) (2.61)		-0.09 (-0.07) (-0.05)				1.17 (3.59) (2.41)		0.31
5-OLIVE	3.54 (2.84) (2.82)		-0.43 (-0.26) (-0.25)				0.11 (2.93) (2.90)		0.78
5'-OLS	3.83 (4.06) (2.62)		-0.23 (-0.19) (-0.12)				1.27 (3.60) (2.31)		0.30
5'-OLIVE	3.15 (4.20) (4.18)		0.06 (0.07) (0.07)				0.10 (3.65) (3.62)		0.80

TABLE 3 Continued

Row	Constant	$\widehat{cay}_t$	Factors$_{t+1}$ R_{vw}	Δy	SMB	HML	$\widehat{cay}_t \cdot$ Factors$_{t+1}$ R_{vw}	Δy	R^2
	5.19	-0.44	-2.00	0.56			0.35	-0.17	
6-OLS	(5.60)	(-1.59)	(-1.74)	(2.11)			(1.69)	(-2.39)	0.77
	(3.33)	(-0.94)	(-1.03)	(1.25)			(1.00)	(-1.42)	
	1.91	0.01	1.33	0.02			0.10	-0.0005	
6-OLIVE	(1.80)	(0.16)	(1.07)	(0.67)			(2.00)	(-0.05)	0.83
	(1.76)	(0.15)	(1.04)	(0.66)			(1.95)	(-0.05)	
	5.14		-1.98	0.59			0.60	-0.08	
7-OLS	(5.59)		(-1.73)	(2.20)			(2.71)	(-2.52)	0.75
	(3.85)		(-1.19)	(1.51)			(1.86)	(-1.73)	
	2.00		1.26	0.02			0.10	-0.002	
7-OLIVE	(2.30)		(1.25)	(0.66)			(1.89)	(-1.00)	0.83
	(2.26)		(1.22)	(0.64)			(1.85)	(-0.98)	
	4.26		-0.97	0.91			0.43	-0.10	
7'-OLS	(4.58)		(-0.84)	(2.96)			(2.10)	(-1.65)	0.71
	(3.40)		(-0.51)	(1.80)			(1.28)	(-1.00)	
	2.14		1.06	0.01			0.15	0.002	
7'-OLIVE	(2.17)		(0.91)	(0.24)			(2.60)	(0.20)	0.83
	(2.13)		(0.89)	(0.23)			(2.54)	(0.20)	

Note: This table corresponds to Table 1 in Lettau and Ludvigson (2001b). The table presents $\boldsymbol{\lambda}$ estimates from cross-sectional Fama-MacBeth regressions using returns of 25 Fama-French portfolios:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_i{}'\lambda .$$

The individual λ_j estimates (from the second-pass cross-sectional regression) for the beta of the factor listed in the column heading are reported. In the first pass, the time-series betas β_i are computed in one multiple regression of the portfolio returns on the factors, using either OLS or OLIVE as noted in each row. R_{vw} is the CRSP value-weighted index return, Δy is labor income growth, and SMB and HML are the Fama-French mimicking portfolios related to size and book-to-market equity ratios. The scaling variable is $\widehat{cay}$. The table reports the Fama-MacBeth cross-sectional regression coefficients, with two t-statistics in parentheses for each coefficient estimate. The top t-statistic uses uncorrected Fama-MacBeth standard errors, and the bottom t-statistic uses the Shanken (1992) correction. The cross-sectional R^2 is reported. The model is estimated using data from 1963:Q3 to 1998:Q3. The coefficient estimates of the factors are multiplied by 100, and the estimates of the scaled terms are multiplied by 1,000.

TABLE 4. Consumption CAPM, Fama-MacBeth Regressions Using 25 Fama-French Portfolios: λ_j Coefficient Estimates on Betas in Cross-Sectional Regression

Row	Constant	$\widehat{cay}_t$	Δc_{t+1}	$\widehat{cay}_t \cdot \Delta c_{t+1}$	R^2
1-OLS	3.25 (4.94) (4.50)		0.22 (1.26) (1.14)		0.16
1-OLIVE	3.34 (4.74) (0.88)		0.003 (0.46) (0.45)		0.03
2-OLS	4.27 (6.10) (4.24)	-0.12 (-0.42) (-0.29)	0.03 (0.22) (0.15)	0.06 (3.14) (2.18)	0.70
2-OLIVE	4.72 (3.70) (3.67)	0.04 (0.90) (0.88)	-0.004 (-0.27) (-0.27)	0.009 (3.44) (3.40)	0.82
3-OLS	4.09 (6.81) (5.13)		-0.02 (-0.12) (-0.09)	0.07 (3.22) (2.41)	0.69
3-OLIVE	3.49 (5.85) (5.83)		0.006 (0.42) (0.41)	0.007 (3.36) (3.34)	0.81
3'-OLS	2.77 (4.37) (3.77)		0.01 (0.09) (0.07)	0.04 (2.36) (2.03)	0.27
3'-OLIVE	5.40 (6.07) (6.05)		0.04 (2.95) (2.92)	-0.005 (-2.29) (-2.27)	0.34

Note: This table corresponds to Table 3 in Lettau and Ludvigson (2001b). The table presents λ estimates from cross-sectional Fama-MacBeth regressions using returns of 25 Fama-French portfolios:

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_i{}'\lambda .$$

The individual λ_j estimates (from the second-pass cross-sectional regression) for the beta of the factor listed in the column heading are reported. In the first pass, the time-series betas β_i are computed in one multiple regression of the portfolio returns on the factors, using either OLS or OLIVE as noted in each row. Δc denotes consumption growth (log difference in consumption). The scaling variable is $\widehat{cay}$. The table reports the Fama-MacBeth cross-sectional regression coefficients, with two t-statistics in parentheses for each coefficient estimate. The top t-statistic uses uncorrected Fama-MacBeth standard errors, and the bottom t-statistic uses the Shanken (1992) correction. The cross-sectional R^2 is reported. The model is estimated using data from 1963:Q3 to 1998:Q3. The coefficient estimates of the factors are multiplied by 100, and the estimates of the scaled terms are multiplied by 1,000.