Capital Share Risk in U.S. Asset Pricing: A Reappraisal

NIKOLAY GOSPODINOV and CESARE ROBOTTI*

ABSTRACT

Using long-horizon beta estimates, Lettau, Ludvigson, and Ma (2019) document striking pricing ability and cross-sectional explanatory power for a capital share growth factor across major asset classes. We revisit their findings and show that the statistical significance of their results is likely due to the interaction between the lack of identification of the proposed single-factor model and the persistence induced by overlapping the data to obtain long-horizon beta estimates. This casts doubts on whether capital share betas are truly priced in the cross-section of expected returns and calls for alternative methods to validate the existence of capital share risk in U.S. asset prices.

Keywords: Capital share risk; Long-horizon betas; Model identification; Overlapping data.

JEL classification: G12; G20; C12; C14; C15.

*Gospodinov (corresponding author, nikolay.gospodinov@atl.frb.org) is at the Federal Reserve Bank of Atlanta and Robotti is at the University of Warwick (Warwick Business School). We would like to thank the Editor (Stefan Nagel) and three anonymous referees for comments and suggestions. We also thank Martin Lettau, Sydney Ludvigson, and Sai Ma for sharing their data and code and for providing clarifications. We have read the *Journal of Finance*'s disclosure policy and have no conflicts of interest to disclose. The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

Lettau, Ludvigson, and Ma (2019, LLM) provide evidence that the capital share growth of aggregate income carries a positive and statistically significant risk premium for a wide range of assets. The main results are based on long-horizon betas that are obtained by regressing H-period compounded test asset returns on the H-period growth of the capital share factor. The authors argue that this H-period aggregation can mitigate the effect of measurement error in the data as well as the impact of model misspecification arising from omission of additional risk factors. The estimated H-horizon betas are then shown to be useful for explaining one-period expected return premia. Given the nonstandard setup with *H-period* test asset returns (and factors) in the first pass but *one-period* returns in the second pass, statistical inference on the risk premia parameters was performed using a bootstrap method.

Table I reproduces LLM's results for equity and nonequity portfolio returns. In addition to horizons $H = 4$ and 8, we report results for $H = 1$. The table presents the estimate for the risk premium parameter, $\hat{\lambda}_H$, (multiplied by 100) for the capital share factor, the corresponding 95% bootstrap confidence interval based on 10,000 replications (in square brackets below the estimate), and the cross-sectional adjusted R^2 (at the bottom in curly brackets).[1]

Table I about here

Several remarks on the results in Table I are in order. While the confidence intervals are relatively wide for $H = 1$, they tighten substantially for $H = 4$ and even further for $H = 8$. At these longer horizons, the risk premium is positive and highly significant at the 5% nominal level. Similarly, for equity portfolios, the cross-sectional adjusted R^2 increases from zero and negative values to values as high as 80% to 86%, suggesting that a single nontraded factor can explain a large fraction of the cross-sectional variation in equity portfolio returns.[2] The pricing ability of LLM's single-factor model is not limited to equity portfolios and it appears to extend to nonequity asset classes. The high

[1] Slight differences in the bootstrap confidence bounds between our Table I and LLM's Table III arise from the different number of bootstrap replications and the inherent randomness embedded in the resampling of the data.

[2] While LLM (p. 1757) acknowledge that this does not hold for industry portfolios, FX and commodity returns, in Appendix B (Table B.I) we show that the positive and statistically significant risk premium on the capital share factor is not robust across a wider set of equity and nonequity portfolio returns. The only commonality across all of these portfolios is the identification failure that we discuss below.

cross-sectional regression R^2s are even more remarkable for these nonequity test assets with values for $H = 4$ of 0.86, 0.79, and 0.95 for bonds, sovereign bonds, and options, respectively.[3] Overall, these results seem to support – at least for the specific equity and nonequity portfolio returns considered in Table I – LLM's argument that aggregating over H periods may sharpen the signal in the data and produce more informative inference. We return to these issues later in this note.

This evidence raises two questions. First, is the capital share factor unique in producing this highly significant risk premium? If the answer to this question is affirmative, this would elevate the status of the capital share factor in enhancing our understanding about the underlying drivers of the cross-section of asset returns. As a reminder, the empirical asset pricing literature has had difficulty identifying priced nontraded macroeconomic factors that pass statistical scrutiny and robust evaluation. Second, is the proposed multi-period estimation and inference approach indeed a more informative method for uncovering priced factors in the cross-section of asset returns? Again, an affirmative answer to this question could expand our tools for characterizing and spanning the factor space of asset returns. We provide evidence that sheds light on these questions.

I. Is the Capital Share Factor Unique in Producing Highly Significant Risk Premia Estimates?

To address this question, it is instructive to subject other nontraded factors to the same estimation and inference procedure that is used in LLM and Table I. In Table II below, we report results for four series from the FRED quarterly database for macroeconomic research: new private housing units authorized by building permits in the Northeast, Midwest, South, and West Census regions (denoted by PerN, PerMW, PerS, and PerW, respectively).[4] In general, the level of these series does not share

[3]For context, the number of time-series observations used for computing these R^2s is 143 (with 20 bonds), 60 (with six sovereign bonds), and 98 (with 18 options).

[4]The results reported below hold for all 11 of the nonprice series from the housing group, as well as other macroeconomic variables, in the FRED quarterly database for macroeconomic research. All of these series are available at https://research.stlouisfed.org/econ/mccracken/fred-databases/. To avoid data-mining and conserve space, we focus on the regional building permit series. The mnemonics for the four regional permit variables are PERMITNE, PERMITMW, PERMITS, and PERMITW, while those for the unreported housing variables are HOUST, HOUST5F, PERMIT, HOUSTMW, HOUSTNE, HOUSTNS, and HOUSTNW.

the business cycle or lower-frequency dynamics of the capital share factor and their growth rates are only weakly correlated with the growth rate of the capital share factor, with correlations ranging from 8% to 11%. The correlations among the regional housing variables are also not extremely high, with the correlation between the growth rates of building permits in the Northeast and West, for example, being 50.6%. Importantly, the regional nature of the building permits undermines the theoretical arguments that underlie the usefulness of these series as common risk factors in asset pricing models.[5]

We construct the housing factors exactly the same way the capital share factor is constructed in LLM (by summing up the log differences over H periods and exponentiating) and scale them to have the same standard deviation as the capital share factor to facilitate comparisons of the risk premium estimate across factors. We apply LLM's bootstrap method to these four nontraded factors in exactly the same way as in Table I. The results for $H = 8$ (two-pass risk premium estimates, 95% bootstrap confidence intervals constructed using LLM's method, and cross-sectional adjusted R^2s) are presented in Table II.

Table II about here

The results are rather surprising. All of these variables appear to possess the same pricing ability as the capital share factor, with positive, highly significant risk premia and substantial explanatory power. For example, for "All Equities," the 95% confidence interval for the risk premium on building permits in the Northeast is [0.69, 0.89] with a cross-sectional adjusted R^2, denoted by $\bar{R}^2$, of 0.73. For options, the $\bar{R}^2$s for the Midwest, South, and West regions are 0.99, 0.98, and 0.99, respectively. Similar results, with a positive and significant risk premium, also obtain for $H = 4$.[6] We should note that this evidence does not invalidate the capital share as a genuine priced risk factor. The evidence in Table II simply demonstrates that, provided LLM's bootstrap method is theoretically valid, there are many nontraded priced factors. However, given the limited sample size, the tight confidence intervals

[5]Of course, this does not rule out the possibility that housing, and its empirical proxies, is a priced risk factor (e.g., Piazzesi, Schneider, and Tuzel (2007)).

[6]For robustness, we explored various factors from the FRED database and other sources (the results are not reported here to conserve space) and found that many of them exhibit similar characteristics when the inference is conducted with H-period returns and factors using LLM's bootstrap method.

and extremely high $\bar{R}^2$s (for asset classes known to be difficult to price) raise some doubts about the validity of the statistical procedure.

II. Is the Multi-Period Estimation and Inference Approach a More Informative Method for Uncovering Priced Factors in the Cross-Section of Asset Returns?

The evidence presented so far runs against the consensus in the empirical asset pricing literature that only a few nontraded factors are priced. Accordingly, we now turn to the second question that we raised above and scrutinize the implicit assumptions that underpin the validity of the statistical inference procedure that produced the results in LLM and Tables I and II. More specifically, the validity of the statistical inference depends critically on the full column rank of the matrix $\mathbf{X} = [\mathbf{1}_N,\ \boldsymbol{\beta}_H]$, where $\boldsymbol{\beta}_H = [\beta_{1,H}, \ldots, \beta_{N,H}]'$ is the H-period vector of risk exposures on the N test assets and $\mathbf{1}_N$ denotes an N-vector of ones. This identification condition follows immediately from defining (in the ordinary least squares (OLS) framework) the zero-beta rate (λ_0) and risk premium (λ_H) parameters as $\boldsymbol{\lambda}_H \equiv [\lambda_0, \lambda_H]' = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\mu}_R$, where $\boldsymbol{\mu}_R$ is the N-vector of expected one-period test asset returns. In single-factor models such as the capital share factor model, rank failure of $\mathbf{X}$ can arise because the risk factor's variation does not induce changes in test asset returns (Pukthuanthong, Roll, and Subrahmanyam (2019)) or because $\boldsymbol{\beta}_H$ is a constant vector.[7]

Before we subject the identification condition to formal statistical testing, we want to highlight the constrained nature of the problem in recovering information about the true risk premia. In the one-period setting, a large cross-section N of test assets relative to the time-series observations T can severely impair the accuracy of standard inference procedures. Since in LLM's setting the two-pass procedure can be rewritten as a moment condition problem with $3N$ moment conditions, the effective number of time-series observations per moment condition (test asset) could be quite limited for large N. As a reminder, for "All Equities" in Table I, $N = 85$ and $T = 200$ for $H = 1$. For nonequity asset classes, N is an even larger fraction of T, with the most extreme case being CDS ($N = 20$, $T = 38$). The H-period setting affects the quality of inference along two dimensions. First,

[7] In multifactor models, linear combinations of the columns of $\boldsymbol{\beta}_H$ represent another potential source of identification failure.

H-period data overlapping reduces the effective number of time-series observations as it trims the independent information in the sample. (See Richardson and Stock (1989), Valkanov (2003).) Second, the overlapping structure induces strong persistence in the series that may obscure true relationships and lead to spurious comovements. To be clear, averaging can potentially reduce the noise and sharpen the low-frequency signal, but it should be done without overlapping data for which identification-robust inference (Kleibergen and Zhan (2020)) is readily applicable.

Given the small T, large N/T, and possibly large H/T in LLM's analysis, standard asymptotic rank tests for testing the reduced rank of $\mathbf{X}$, $H_0 : \text{rank}(\mathbf{X}) = 1$, would provide an extremely inaccurate approximation of the sampling distributions of these statistics. Simulation results for $H = 4$ and 8 (reported in Appendix A) show that under H_0, the asymptotic rank tests exhibit empirical rejection rates close to 100% at the 5% nominal level, that is, the tests suggest that the model is well identified when the null of a reduced rank is true. Instead, in Appendix A we propose a new bootstrap-based rank test that is characterized by excellent size and power properties. Table III reports the results (p-values) from the bootstrap rank test for single-factor models based on the capital share factor and the four regional building permits that we explored earlier. For comparison, we also include one traded factor – the popular small-minus-big (SMB) factor – that is known to have a nontrivial correlation with the equity portfolio returns.[8] This traded factor serves as a natural benchmark in evaluating the identification properties of the model.

Table III about here

We start with the capital share (KS) factor. Based on a 5% significance level, we can never reject the null hypothesis of a reduced rank. (In fact, this is true for all $H = 1, \ldots, 8$.)[9] This preliminary

[8] In unreported results, the risk premium estimates for SMB are also statistically significant with 95% confidence intervals, based on LLM's bootstrap, being somewhat wider and not exhibiting as much tightening with H as those for the nontraded factors in Tables I and II. It should be noted that the block bootstrap in LLM does not properly mimic the persistence in H-period returns and factors, which suppresses the true sampling uncertainty and results in artificially tighter confidence intervals.

[9] While the identification-robust method of Kleibergen and Zhan (2020) cannot be readily applied to the H-period setup, we used it to construct confidence sets for the capital share factor at $H = 1$. These confidence intervals are either unbounded, disjoint, or very wide, suggesting lack of identification. In an earlier version of this note, we proposed a

test has crucial implications for the subsequent analysis on the risk premium parameter λ_H, including inconsistency of the OLS estimator of λ_H, highly nonstandard large-sample behavior, and invalidity of the bootstrap method as implemented in LLM. The reader may be skeptical that the rank test has low power in rejecting the null, so that the resulting inference is overly conservative. These concerns are valid as the power of the test depends on the distance from the null (for example, how far $\hat{\boldsymbol{\beta}}_H$ is from a vector of zeros or ones), which is a function of the effective number of observations. As argued above, large H and N relative to T reduce the amount of independent information in the sample, which is accompanied by elevated sample uncertainty that could overwhelm seemingly large and dispersed estimates of $\hat{\boldsymbol{\beta}}_H$. Following a request by a referee, we evaluated the power of the rank test in simulations, calibrated to the application setup, with true values for $\boldsymbol{\beta}_H$ set to their sample estimates for $H = 8$. Despite the small sample size, the empirical power of the bootstrap rank test in this case is close to 100% at the 5% nominal level for all equity portfolios. Similar results – indicating that the power of the bootstrap rank test is quite high in empirically relevant settings – obtain using an alternative simulation design reported in Appendix A. Having said that, it is important to highlight the extremely limited information in the sample, especially for nonequity asset classes, that is insufficient to identify the parameters of interest. So it could well be the case that the capital share factor is a genuine risk factor but the limited sample information does not allow the researcher to identify the risk premia and possible pricing ability of the factor. Nevertheless, the empirical failure of the rank test to reject the null of reduced rank implies that standard methods (asymptotic or bootstrap) would lead to unwarranted conclusions and false inference. This is the main takeaway of our note.[10]

nonparametric, identification-robust bootstrap method for the H-period setting that did not lend empirical support to LLM's claim that capital share growth is a priced risk factor.

[10]It is important to stress that significance tests on individual betas or beta spreads (as in Table VII in LLM), while informative for some purposes, would not address the issue of identification failure. Statistical significance of individual betas or spreads – as is the case in this application – is consistent with the possibility of identification failure, that is, that all betas are jointly statistically indistinguishable from a vector of zeros. This occurs when the betas across the different test assets are highly correlated, as reflected in the large off-diagonal elements of the covariance matrix of $\hat{\boldsymbol{\beta}}_H$. An informal diagnostic to gauge how far the covariance matrix of $\hat{\boldsymbol{\beta}}_H$, $\mathbf{V}_{\hat{\boldsymbol{\beta}}_H}$, is from a diagonal matrix with diag$(\mathbf{V}_{\hat{\boldsymbol{\beta}}_H})$ on the main diagonal, which is used for the individual tests, is the ratio between the largest eigenvalue of these two matrices. With one being the value of this ratio when the two matrices have identical largest eigenvalues, the ratio for the 25 Size/BM equity portfolios ranges between 7.0 and 9.4 for $H = 1, \ldots, 8$. This suggests that significance tests on

The rank test results for the regional housing factors are largely similar, indicating that their risk premia cannot be identified from sample information, and the high statistical significance in Table II is due to an inadequate statistical methodology. While there are some borderline cases (for example, PerW for sovereign bonds),[11] the large p-values suggest that the rank failure of $\mathbf{X}$ – for whatever reason, irrelevance of the factor or lack of sufficient information in the sample – invalidates the high statistical significance reported in Table II. With this in mind, there may be other methods to evaluate the usefulness and pricing ability of a particular factor. We discuss alternative approaches in the next section.

Finally, we observe in Table III that based on a one-factor model with a traditional traded factor such as SMB, we can convincingly reject the null hypothesis of a reduced rank for the equity portfolios. For nonequity asset classes, the results for SMB in Table III suggest that the rank condition can be compromised even for traded factors if these factors are only weakly correlated with the test asset returns or the sample size is small. The reduced effective number of observations in the H-period two-pass methodology further inflates the sampling uncertainty, and the failure of the full rank condition becomes more pronounced, at larger horizons H. Some interesting results for the market factor (not reported here but available upon request) shed further light on the potential statistical problems underlying the H-period two-pass methodology. Despite its high correlation with equity returns, the rank condition for the H-period CAPM also appears to be violated – not because $\boldsymbol{\beta}_H = \mathbf{0}_N$, but because $\hat{\boldsymbol{\beta}}_H$ cannot be statistically distinguished from a column of ones given the heightened sampling uncertainty for large overlapping horizons H. Thus, overlapping the data may not necessarily lead to a higher signal-to-noise ratio and may, in fact, distort the standard inference procedures.

---

the individual betas or beta spreads would not be appropriate for evaluating the validity of the identification condition. Clearly, individual tests based on beta spreads would deliver qualitatively similar conclusions as those of the joint test if the off-diagonal elements of the covariance matrix of $\hat{\boldsymbol{\beta}}_H$ were small. This is not the case for all asset classes considered in this application.

[11]Note that Table III reports 56 rank tests for the four housing factors, and it is natural to observe p-values below the 0.1 or 0.05 thresholds.

III. Alternative Empirical Approaches

If the risk premium parameter is poorly identified in the two-pass regression framework, are there alternative empirical approaches that would prove more informative about the pricing ability of LLM's candidate risk factor? While the failure of the identification condition suggests that the sample does not contain sufficient information to estimate consistently the parameter of interest, it is possible that alternative methods that exploit other characteristics of the data (time variability of the betas, out-of-sample evaluation, etc.) may be better suited for the H-period setup. It is important, however, to recognize that a fair comparison of various competing methods requires that these procedures account for all sources of statistical uncertainty in the construction of the factor's risk premium.

One such method, suggested to us by a referee, is based on sorting test asset portfolio returns on the H-period, rolling-window estimates of the betas for the capital factor. The cumulative (out-of-sample) return on a long/short portfolio obtained from high/low capital share betas would then inform the researcher if capital share is a priced risk factor or not. Since this approach can be cast in a nonparametric estimation setting (see Cattaneo et al. (2020)), the statistical evaluation of the cumulative return should be adjusted for sampling uncertainty in the portfolio sorting, estimation uncertainty in the betas used for sorting, the short time span used for out-of-sample evaluation, possible serial correlation, etc. A well-designed bootstrap method that incorporates all of these sources of uncertainty can produce a sampling distribution for evaluating the statistical significance of the cumulative return.

Second, if other risk factors (Fama-French factors, for example) load significantly on the capital share factor, it is prudent to conduct the portfolio sorting on the capital share betas by controlling for these factors in the long-horizon beta regressions. This will ensure that the capital share factor exhibits genuine pricing ability, in addition to that contained in the traditional risk factors. The inference method for the first approach can be readily modified to account for the additional controls. Finally, a third evaluation approach can be based on the out-of-sample Sharpe ratio of a mimicking portfolio for the capital share factor, which employs the test assets and other traded factors.

Given the large number of tuning choices in these alternative methods (size of the rolling window, determination of the optimal number of portfolios, test/basis asset returns used in the formation of the mimicking portfolio, sample and frequency of the data, etc.), it is difficult to argue with a high degree of confidence for or against the pricing ability of a particular factor. Nevertheless, we applied these alternative methods to the value-, profitability-, and investment-sorted portfolios with $H = 8$ and betas computed over a five-year rolling window. The results (not reported here to conserve space) can be summarized as follows. The cumulative return for the first method (single-factor model with capital share as the only risk factor) is relatively large and positive in the evaluation period but its sampling distribution (accounting for the different sources of uncertainty) is very wide and does not suggest statistical significance. The second method, which controls for the four Fama-French factors (size, value, operating profitability, and investment), produces an essentially flat and statistically insignificant cumulative return. It is important to stress that the pricing ability of the capital share factor further deteriorates when considering shorter compounding horizons such as $H = 1$ and $H = 4$. The third method also produces an out-of-sample Sharpe ratio for the mimicking portfolio of the capital share factor that is essentially zero and statistically different and dominated by the market portfolio. While these results do not lend support to the capital share as a priced risk factor in the cross-section of these equity returns, they do not rule out the possibility that other sorting choices, test assets, or mimicking portfolios may provide more favorable evidence for the pricing ability of the capital share factor.

IV. Concluding Remarks

In summary, the results in this note highlight some problems arising from identification failure and H-period compounding of the data. We present evidence that questions the robustness and reliability of LLM's findings of high statistical significance and explanatory power for a capital share factor. The validity of these results relies critically on the identification (full-rank) condition in their H-period model. We find no empirical evidence that this condition holds in the sample. We also demonstrate that LLM's inference method produces pervasive evidence of high statistical significance and explanatory power for other nontraded factors that also fail the full-rank condition. This identification failure is more subtle than in the standard setup as it is partially obscured by the persistence induced by H-period compounding of the data. This suggests the need for extra caution in conducting inference in a multi-period setting as the proposed statistical method should account for the possibility of spurious commonality arising from overlapping data over long horizons. Performing the analysis at a lower frequency (without overlapping data) or using alternative empirical methods may be better suited for robust evaluation of nontraded macroeconomic factors.

Appendix A. Bootstrap Rank Test

The identification condition for the second-pass risk premium in single-factor models with a zero-beta rate is that the $N \times 2$ matrix $\mathbf{X} = [\mathbf{1}_N, \boldsymbol{\beta}_H]$ is of full column rank. Let $\mathbf{I}_{N-1}$ be an $(N-1)\times(N-1)$ identity matrix and $\mathbf{P}$ denote an $N \times (N-1)$ orthonormal matrix ($\mathbf{P}'\mathbf{P} = \mathbf{I}_{N-1}$) whose columns are orthogonal to $\mathbf{1}_N$ such that

$$\mathbf{P}\mathbf{P}' = \mathbf{I}_N - \mathbf{1}_N(\mathbf{1}_N'\mathbf{1}_N)^{-1}\mathbf{1}_N'. \tag{A1}$$

Using this notation, the null of reduced column rank, $H_0:\ \text{rank}(\mathbf{X}) = 1$, can be expressed as $H_0:$ $\mathbf{P}'\boldsymbol{\beta}_H = \mathbf{0}_{N-1}$, where $\mathbf{0}_{N-1}$ is an $(N-1)$-vector of zeros. A simple Wald test of $H_0:\ \mathbf{P}'\boldsymbol{\beta}_H = \mathbf{0}_{N-1}$ can be performed using the test statistic

$$\mathcal{W}_T = (T-H)\hat{\boldsymbol{\beta}}_H'\mathbf{P}\hat{\mathbf{V}}^{-1}_{\mathbf{P}'\hat{\boldsymbol{\beta}}_H}\mathbf{P}'\hat{\boldsymbol{\beta}}_H, \tag{A2}$$

where $\hat{\boldsymbol{\beta}}_H$ is the OLS estimate of $\boldsymbol{\beta}_H$ (from a regression of H-period test asset excess returns $\mathbf{R}^e_{t+H,t}$ on a constant and the H-period factor $f_{t+H,t}$) and $\hat{\mathbf{V}}_{\mathbf{P}'\hat{\boldsymbol{\beta}}_H}$ is a consistent estimator of the long-run covariance matrix

$$\mathbf{V}_{\mathbf{P}'\hat{\boldsymbol{\beta}}_H} = \sum\nolimits_{j=-\infty}^{\infty} \mathbb{E}\left[\mathbf{m}_{t,H}\mathbf{m}_{t+j,H}'\right], \tag{A3}$$

with $\mathbf{m}_{t,H} = \frac{(f_{t+H,t}-\mu_{f_H})}{\sigma^2_{f_H}}\mathbf{P}'\boldsymbol{\epsilon}_{t+H,t}$, $\mu_{f_H} = \mathbb{E}\left[f_{t+H,t}\right]$, and $\sigma^2_{f_H} = \text{Var}\left[f_{t+H,t}\right]$. In the numerical implementation of the test, we use the Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) estimator with a bandwidth set equal to H.

While under some regularity conditions the test $\mathcal{W}_T$ is asymptotically chi-squared distributed with $N-1$ degrees of freedom, this approximation will likely provide a very poor approximation of the finite-sample distribution for the reasons discussed in this note: small T and large N and H (both relative to T) that further reduce the effective number of time-series observations.[12] Before describing the bootstrap procedure for approximating the finite-sample distribution of the test $\mathcal{W}_T$, it is convenient

[12] The H-period overlapping data also induces strong serial correlation of a telescoping sum pattern. It is widely documented that the Newey and West (1987) HAC estimator is not well suited to capture this type of serial dependence. We also experimented with the Hansen and Hodrick (1980) HAC estimator, but this estimator is not guaranteed to be positive semi-definite. This is the case in LLM's empirical application given the large N and the relatively small T.

to pre-multiply the first-pass regression model by $\mathbf{P}'$, which yields

$$\mathbf{P}'\mathbf{R}^e_{t+H,t} = \mathbf{P}'\boldsymbol{\alpha} + \mathbf{P}'\boldsymbol{\beta}_H f_{t+H,t} + \mathbf{P}'\boldsymbol{\epsilon}_{t+H,t}. \tag{A4}$$

This model also facilitates imposing the null hypothesis of reduced rank $H_0: \ \mathbf{P}'\boldsymbol{\beta}_H = \mathbf{0}_{N-1}$ in the bootstrap sample. Under the null, we have

$$\mathbf{P}'\mathbf{R}^e_{t+H,t} = \boldsymbol{\mu}_{\mathbf{P}'\mathbf{R}^e} + \mathbf{P}'\boldsymbol{\epsilon}_{t+H,t}, \tag{A5}$$

where $\boldsymbol{\mu}_{\mathbf{P}'\mathbf{R}^e} = \mathbb{E}\left[\mathbf{P}'\mathbf{R}^e_{t+H,t}\right]$. Let $\mathbf{P}'\hat{\boldsymbol{\epsilon}}_{t+H,t}$ denote the OLS estimate of $\mathbf{P}'\boldsymbol{\epsilon}_{t+H,t}$ and $\hat{\boldsymbol{\mu}}_{\mathbf{P}'\mathbf{R}^e}$ be the sample estimate of $\boldsymbol{\mu}_{\mathbf{P}'\mathbf{R}^e}$. Stack the H-period factor $f_{t+H,t}$ and the $(N-1)$-vector $\tilde{\mathbf{R}}^e_{t+H,t} = \hat{\boldsymbol{\mu}}_{\mathbf{P}'\mathbf{R}^e} + \mathbf{P}'\hat{\boldsymbol{\epsilon}}_{t+H,t}$ in a $(T-H)\times N$ matrix $\mathbf{Z}$ with rows $\mathbf{z}_t = [f_{t+H,t}, (\tilde{\mathbf{R}}^e_{t+H,t})']$ for $t = 1, \ldots, T-H$. The bootstrap samples are constructed by drawing with replacement blocks of M $(1 \leq M < T-H)$ observations from matrix $\mathbf{Z}$, denoted by $\mathbf{Z}^* = \{(\mathbf{z}^*_1, \mathbf{z}^*_2, \ldots, \mathbf{z}^*_M), (\mathbf{z}^*_{M+1}, \mathbf{z}^*_{M+2}, \ldots, \mathbf{z}^*_{2M}), \ldots, (\mathbf{z}^*_{T-M-H}, \mathbf{z}^*_{T-M+1-H}, \ldots, \mathbf{z}^*_{T-H})\}$, with $\mathbf{z}^*_t = [f^*_{t+H,t}, (\tilde{\mathbf{R}}^{e*}_{t+H,t})']$ being the resampled analog of the original data $\mathbf{z}_t = [f_{t+H,t}, (\tilde{\mathbf{R}}^e_{t+H,t})']$. Using the bootstrap sample, we obtain the estimated quantities $\mathbf{P}'\hat{\boldsymbol{\beta}}^*_H$ and $\mathbf{P}'\hat{\boldsymbol{\epsilon}}^*_{t+H,t}$ by running an OLS regression of $\tilde{\mathbf{R}}^{e*}_{t+H,t}$ on $f^*_{t+H,t}$ (and a constant). Then, the bootstrap analog of $\mathcal{W}_T$ for the j-th bootstrap sample is constructed as

$$\mathcal{W}^*_{T,j} = (T-H)\hat{\boldsymbol{\beta}}^{*\prime}_H \mathbf{P}\hat{\mathbf{V}}^{*-1}_{\mathbf{P}'\hat{\boldsymbol{\beta}}^*_H}\mathbf{P}'\hat{\boldsymbol{\beta}}^*_H, \tag{A6}$$

where $\hat{\mathbf{V}}^*_{\mathbf{P}'\hat{\boldsymbol{\beta}}^*_H}$ denotes the HAC estimator of $\mathbf{V}_{\mathbf{P}'\hat{\boldsymbol{\beta}}_H}$, with the bootstrap sample analog of $\mathbf{m}_{t,H}$ being $\hat{\mathbf{m}}^*_{t,H} = \frac{(f^*_{t+H,t} - \hat{\mu}^*_{f^*_H})}{\hat{\sigma}^{*2}_{f^*_H}}\mathbf{P}'\hat{\boldsymbol{\epsilon}}^*_{t+H,t}$, $\hat{\mu}^*_{f^*_H} = \frac{1}{T-H}\sum_{t=1}^{T-H} f^*_{t+H,t}$, and $\hat{\sigma}^{*2}_{f^*_H} = \frac{1}{T-H}\sum_{t=1}^{T-H}(f^*_{t+H,t} - \hat{\mu}^*_{f^*_H})^2$. With B bootstrap replications, the bootstrap p-value of the rank test is computed as $\frac{1}{B}\sum_{j=1}^{B}\mathbb{I}\left\{\mathcal{W}^*_{T,j} > \mathcal{W}_T\right\}$, where $\mathbb{I}\{\cdot\}$ is the indicator function.

A similar bootstrap test can be constructed for the null $H_0: \ \boldsymbol{\beta}_H = \mathbf{0}_N$ but without pre-multiplying by the matrix $\mathbf{P}'$. While the two tests yield almost identical results for models with a single spurious factor, differences emerge in the presence of useful factors in single- or multi-factor models. For example, since the identification condition is concerned with the matrix $\mathbf{X} = [\mathbf{1}_N, \ \boldsymbol{\beta}_H]$, the rank of $\mathbf{X}$ can be compromised if $\boldsymbol{\beta}_H = \mathbf{c}$ for some $\mathbf{c} \neq \mathbf{0}_N$. Furthermore, in multi-factor models, $\boldsymbol{\beta}_H$ is a matrix and rank failure can also occur if two or more of its columns are linear combinations of each other (even if, individually, they are different than a zero vector, that is, the factors are not spurious).

Below we report the results from a Monte Carlo simulation study that evaluates the size and power properties of the asymptotic and bootstrap versions of the rank test of $\mathbf{X} = [\mathbf{1}_N, \boldsymbol{\beta}_H]$. To assess power, we simulate one-period data jointly for the SMB factor and the test portfolio returns, using the sample means and sample covariance matrix. For the size of the test, we impose the null of rank failure by setting the covariance between the factor and the test asset returns equal to a zero vector. Since there is evidence of non-Gaussianity in the data, SMB and portfolio returns are generated from a multivariate-t distribution with six degrees of freedom. The rank test is then performed on the betas estimated from H-period data by compounding returns as a moving product over a sliding window of length H. For the bootstrap version of the test, we impose the null of rank deficiency on the compounded returns (as explained above), and we use the block bootstrap to compute the rank test statistic.[13]

We consider two sample sizes: $T = 202$, the (before-transformation) sample size in LLM, and $T = 1,000$, a sufficiently large sample size to determine how quickly the empirical power of the test improves as T increases. The number of Monte Carlo runs is set to 10,000. The chosen horizons are $H = 1$, 4, and 8, and the number of bootstrap replications for the bootstrap rank test is set to $B = 399$. Finally, the test portfolios are the 10 long-run reversal portfolios ($N = 10$) and the 25 size and book-to-market sorted portfolios ($N = 25$), respectively.

Table A.I presents our simulations results, where Panels A and B report results for the size and power of the asymptotic rank test (which is robust to serial correlation and conditional heteroskedasticity) while Panels C and D report results for the size and power of the bootstrap test.

Table A.I about here

The size distortions of the asymptotic rank test for nonoverlapping data ($H = 1$) arise from the large dimension of N relative to T but tend to vanish as T grows. For the overlapping setting ($H > 1$), however, the size distortions of the asymptotic rank test are massive and remain large for $T = 1,000$.

[13]We employ a block size of $M = H$. Moreover, we use the Newey and West (1987) HAC estimator with a bandwidth set to H in the computation of the asymptotic and bootstrap versions of the tests (see also footnote 11).

For example, for $N = 25$, $T = 202$, and $H = 4, 8$, the asymptotic rank test (Panel A) leads to empirical rejection rates close to 100% at a 5% nominal level. Even for $T = 1,000$, these rejection rates exceed 65% and 86% for $H = 4$ and $H = 8$, respectively. These results highlight the hidden challenges of using standard inference with overlapping data. These tests would suggest, erroneously, that a model with a spurious factor is well identified. Certainly, a smaller N and a larger T help, but the overrejections of the asymptotic versions of these tests are still substantial, as emphasized in Panel A for the 10 long-run reversal portfolios. Panel C displays a dramatic size improvement when considering the bootstrap implementation of the rank test. The size properties of the test are now very good regardless of the chosen overlapping horizon H. The bootstrap rank test slightly underrejects for $T = 202$, but its empirical size approaches the nominal level as T increases.[14] Importantly, the empirical power of the test is high even for the sample size in the empirical analysis. While the power of the test deteriorates for $H = 8$ and $T = 202$, this is to be expected because, as argued in the main text, the number of effective independent observations decreases with the overlapping horizon H. As T increases, the empirical power improves materially. Another notable feature is that the power of the rank test improves as, for fixed H and T, the number of test assets N increases. In summary, this simulation evidence suggests that the bootstrap-based identification test used in this note should be fairly reliable for the sample sizes and compounding horizons considered by LLM.

[14]We attribute the slight size distortions of the bootstrap to several sources. First, our choice of block size, $M = H$, may be partly responsible for this. A more judicious or data-driven selection of M would likely eliminate these distortions. Second, as argued earlier, Newey-West estimator does not provide a reliable approximation of the asymptotic variance for H-period overlapping. Third, the choice of the multivariate-t distribution with six degrees of freedom may be too extreme as some higher-moment regularity conditions, required for asymptotic validity, may not be satisfied. Nevertheless, it is instructive to assess the robustness of the proposed bootstrap method to possible deviations from standard regularity conditions.

Appendix B. Additional Empirical Evidence

This appendix presents results for additional test assets that were not explicitly considered by LLM. The test asset portfolios are, in order, the 25 size and momentum (Size/MOM), 10 short term reversal (ShREV), 17 industry (Industries), and 10 dividend yield (DIVYLD) sorted portfolios from Kenneth French's website. We also consider these equity portfolios together in the "All Equities" column. The additional test asset portfolios are the 23 commodity (Commodities) and 12 foreign exchange (FX) portfolios from He, Kelly, and Manela (2017). The sample period is the same as in LLM (1963:Q3 to 2013:Q4) except for Commodities (1986:Q4 to 2012:Q4) and FX (1976:Q2 to 2009:Q4). For each set of test assets, we run a battery of one-factor models based either on capital share or on individual regional permits as pricing factors. For each one-factor model and compounding horizon $H = 4$ and 8, Table B.I reports LLM's bootstrap confidence interval for the risk premium estimate (in square brackets), the cross-sectional adjusted R^2 ($\bar{R}^2$, in curly brackets), and the p-value of the bootstrap rank test (in round brackets). The number of bootstrap replications for the calculation of the confidence intervals and the p-values of the bootstrap rank test is 10,000.

Table B.I about here

Several observations emerge from the table. First, the pricing ability of the capital share factor vanishes when considering these additional asset classes. The capital share risk premium estimates are often negative (in seven cases out of 14) and the overall magnitude of these estimates is much smaller than that documented by LLM. Second, the confidence intervals for the KS factor are generally wide and contain zero in 11 out of 14 cases. The three cases in which the confidence intervals do not contain zero are those in which the risk premium estimates are negative. Moreover, the KS one-factor model does not violate the full-rank condition in only three cases in which the risk premium estimates are statistically insignificant. Third, the capital share model cross-sectional adjusted R^2s are small and negative in eight cases out of 14. Finally, when considering one-factor models with individual regional permits and LLM's bootstrap method, the risk premium estimates are often statistically significant. The overall picture that emerges from this analysis is that the results documented by LLM are not robust across test assets returns and are driven by an erroneous identification scheme.

REFERENCES

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Ernst Schaumburg, 2020, Characteristic-sorted portfolios: Estimation and inference, *Review of Economics and Statistics* 102, 531–551.

Hansen, Lars P., and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

He, Zhiguo, Brian Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

Kleibergen, Frank, and Zhaoguo Zhan, 2020, Robust inference for consumption-based asset pricing, *Journal of Finance* 75, 507–550.

Lettau, Martin, Sydney C. Ludvigson, and Sai Ma, 2019, Capital share risk in U.S. asset pricing, *Journal of Finance* 74, 1753–1792.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2007, Housing, consumption, and asset pricing, *Journal of Financial Economics* 83, 531–569.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam, 2019, A protocol for factor identification, *Review of Financial Studies* 32, 1573–1607.

Richardson, Matthew, and James H. Stock, 1989, Drawing inferences from statistics based on multiyear asset returns, *Journal of Financial Economics* 25, 323–348.

Valkanov, Rossen, 2003, Long-run regressions: Theoretical results and applications, *Journal of Financial Economics* 68, 201–232.

Table I
Two-Pass Cross-Sectional Regressions with Capital Share

The table presents the two-pass risk premium estimate in one-factor models with capital share as the only pricing factor. The test asset portfolios are, in order, the 25 size and book-to-market (Size/BM), 10 long-term reversal (REV), 25 size and investment (Size/INV), and 25 size and operating profitability (Size/OP) sorted portfolios from Kenneth French's website. LLM also consider these equity portfolios together in the "All Equities" column. The additional test asset portfolios are the 20 corporate and government bond (Bonds), six sovereign bond (Sov. Bonds), and 18 option (Options) portfolios from He, Kelly, and Manela (2017). The sample period is 1963:Q3 to 2013:Q4 except for Bonds (1975:Q1 to 2011:Q4), Sov. Bonds (1995:Q1 to 2011:Q1), and Options (1986:Q2 to 2011:Q4). For each model and compounding horizon ($H = 1$, 4, and 8), we report LLM's confidence interval for the risk premium estimate (in square brackets) and the cross-sectional adjusted R^2 ($\bar{R}^2$, in curly brackets). The number of bootstrap replications for the calculation of the confidence intervals is 10,000.

	Size/BM	REV	Size/INV	Size/OP	All Equities	Bonds	Sov. Bonds	Options
$H = 1$	0.53	0.40	−0.91	1.36	0.58	1.17	2.69	5.59
	[−0.49, 1.51]	[−0.92, 1.78]	[−2.06, 0.19]	[0.60, 2.09]	[0.08, 1.08]	[0.07, 2.21]	[2.15, 3.24]	[4.86, 6.26]
	{$\bar{R}^2$=−0.02}	{$\bar{R}^2$=−0.11}	{$\bar{R}^2$=0.00}	{$\bar{R}^2$=0.21}	{$\bar{R}^2$=0.02}	{$\bar{R}^2$=0.18}	{$\bar{R}^2$=0.91}	{$\bar{R}^2$=0.96}
$H = 4$	0.74	0.63	0.61	0.70	0.68	0.82	1.41	1.87
	[0.42,1.08]	[0.33, 0.92]	[0.27, 0.96]	[0.53, 0.87]	[0.53, 0.83]	[0.59, 1.03]	[0.88, 1.96]	[1.41, 2.32]
	{$\bar{R}^2$=0.51}	{$\bar{R}^2$=0.70}	{$\bar{R}^2$=0.39}	{$\bar{R}^2$=0.78}	{$\bar{R}^2$=0.58}	{$\bar{R}^2$=0.86}	{$\bar{R}^2$=0.79}	{$\bar{R}^2$=0.95}
$H = 8$	0.68	0.41	0.55	0.57	0.57	0.57	1.18	1.80
	[0.52, 0.83]	[0.30, 0.50]	[0.37, 0.74]	[0.42, 0.71]	[0.49, 0.66]	[0.40, 0.71]	[0.17, 2.19]	[0.82, 2.78]
	{$\bar{R}^2$=0.80}	{$\bar{R}^2$=0.86}	{$\bar{R}^2$=0.62}	{$\bar{R}^2$=0.76}	{$\bar{R}^2$=0.74}	{$\bar{R}^2$=0.89}	{$\bar{R}^2$=0.32}	{$\bar{R}^2$=0.81}

Table II
Two-Pass Cross-Sectional Regressions with Building Permits

The table presents the two-pass risk premium estimates in one-factor models with building permits in the North East (PerNE), Mid West (PerMW), South (PerS), and West (PerW) Census regions as pricing factors. The test asset portfolios are, in order, the 25 size and book-to-market (Size/BM), 10 long-term reversal (REV), 25 size and investment (Size/INV), and 25 size and operating profitability (Size/OP) sorted portfolios from Kenneth French's website. LLM also consider these equity portfolios together in the "All Equities" column. The additional test asset portfolios are the 20 corporate and government bond (Bonds), six sovereign bond (Sov. Bonds), and 18 option (Options) portfolios from He, Kelly, and Manela (2017). The sample period is 1963:Q3 to 2013:Q4 except for Bonds (1975:Q1 to 2011:Q4), Sov. Bonds (1995:Q1 to 2011:Q1), and Options (1986:Q2 to 2011:Q4). For each model and a compounding horizon of $H = 8$, we report LLM's confidence interval for the risk premium estimate (in square brackets) and the cross-sectional adjusted R^2 ($\bar{R}^2$, in curly brackets). The number of bootstrap replications for the calculation of the confidence intervals is 10,000.

	Size/BM	REV	Size/INV	Size/OP	All Equities	Bonds	Sov. Bonds	Options
PerNE	0.77	1.13	0.86	0.81	0.79	0.49	0.61	3.04
	[0.62, 0.90]	[0.72, 1.49]	[0.66, 1.04]	[0.56, 1.04]	[0.69, 0.89]	[0.38, 0.61]	[0.29, 0.95]	[2.39, 3.76]
	{$\bar{R}^2$=0.83}	{$\bar{R}^2$=0.59}	{$\bar{R}^2$=0.74}	{$\bar{R}^2$=0.60}	{$\bar{R}^2$=0.73}	{$\bar{R}^2$=0.76}	{$\bar{R}^2$=0.66}	{$\bar{R}^2$=0.86}
PerMW	0.84	0.85	0.85	0.80	0.82	0.42	1.31	1.46
	[0.65, 1.01]	[0.63, 1.05]	[0.63, 1.05]	[0.51, 1.09]	[0.70, 0.93]	[0.29, 0.56]	[0.84, 1.78]	[1.38, 1.54]
	{$\bar{R}^2$=0.75}	{$\bar{R}^2$=0.80}	{$\bar{R}^2$=0.69}	{$\bar{R}^2$=0.48}	{$\bar{R}^2$=0.68}	{$\bar{R}^2$=0.64}	{$\bar{R}^2$=0.80}	{$\bar{R}^2$=0.99}
PerS	0.70	1.03	0.64	0.70	0.68	0.83	1.50	2.38
	[0.44, 0.93]	[0.87, 1.19]	[0.40, 0.87]	[0.49, 0.91]	[0.56, 0.80]	[0.65, 1.01]	[0.91, 2.07]	[2.22, 2.54]
	{$\bar{R}^2$=0.51}	{$\bar{R}^2$=0.89}	{$\bar{R}^2$=0.46}	{$\bar{R}^2$=0.57}	{$\bar{R}^2$=0.55}	{$\bar{R}^2$=0.75}	{$\bar{R}^2$=0.69}	{$\bar{R}^2$=0.98}
PerW	0.78	0.99	0.76	0.66	0.75	0.54	1.09	1.92
	[0.58, 0.96]	[0.81, 1.15]	[0.57, 0.93]	[0.37, 0.96]	[0.62, 0.85]	[0.40, 0.69]	[0.62, 1.56]	[1.84, 2.00]
	{$\bar{R}^2$=0.67}	{$\bar{R}^2$=0.87}	{$\bar{R}^2$=0.67}	{$\bar{R}^2$=0.37}	{$\bar{R}^2$=0.62}	{$\bar{R}^2$=0.68}	{$\bar{R}^2$=0.69}	{$\bar{R}^2$=0.99}

Table III
Bootstrap Rank Tests in One-Factor Models

The table presents the p-values of the proposed bootstrap rank tests in one-factor models with capital share (KS), building permits in the North East (PerNE), Mid West (PerMW), South (PerS), and West (PerW) Census regions, and SMB as pricing factors. The test asset portfolios are, in order, the 25 size and book-to-market (Size/BM), 10 long-term reversal (REV), 25 size and investment (Size/INV), and 25 size and operating profitability (Size/OP) sorted portfolios from Kenneth French's website. The additional test asset portfolios are the 20 corporate and government bond (Bonds), six sovereign bond (Sov. Bonds), and 18 option (Options) portfolios from He, Kelly, and Manela (2017). The sample period is 1963:Q3 to 2013:Q4 except for Bonds (1975:Q1 to 2011:Q4), Sov. Bonds (1995:Q1 to 2011:Q1), and Options (1986:Q2 to 2011:Q4). The number of bootstrap replications for the calculation of the p-values is 10,000.

	Size/BM		REV		Size/INV		Size/OP		Bonds		Sov. Bonds		Options	
Factor/H	4	8	4	8	4	8	4	8	4	8	4	8	4	8
KS	0.217	0.758	0.358	0.323	0.433	0.466	0.177	0.442	0.886	0.654	0.337	0.703	0.673	0.833
PerNE	0.274	0.424	0.079	0.670	0.131	0.228	0.134	0.494	0.277	0.400	0.154	0.044	0.276	0.602
PerMW	0.488	0.331	0.035	0.414	0.113	0.122	0.117	0.332	0.377	0.443	0.249	0.119	0.082	0.105
PerS	0.295	0.480	0.131	0.280	0.259	0.212	0.267	0.330	0.481	0.507	0.140	0.055	0.357	0.255
PerW	0.504	0.467	0.240	0.721	0.122	0.197	0.365	0.479	0.261	0.434	0.068	0.034	0.483	0.242
SMB	0.000	0.000	0.023	0.013	0.000	0.000	0.000	0.000	0.321	0.337	0.281	0.516	0.374	0.475

Table A.I

Size and Power of Asymptotic and Bootstrap Rank Tests with H-period data

The table presents Monte Carlo simulation results for the empirical size and power of the asymptotic and bootstrap versions of the rank test. We consider two sample sizes ($T = 202$, the length of LLM's original sample, and $T = 1,000$) and three compounding horizons ($H = 1$, 4, and 8). The number of simulated paths for the returns and factor is 10,000. For each of the 10,000 Monte Carlo iterations, the p-values for the bootstrap versions of the tests are computed based on 399 replications. The test portfolios are the 10 long-run reversal portfolios ($N = 10$) and the 25 size and book-to-market sorted portfolios ($N = 25$).

Panel A: Size of Asymptotic Rank Test							
		$N = 10$			$N = 25$		
	T	10%	5%	1%	10%	5%	1%
$H = 1$	202	0.352	0.255	0.124	0.847	0.784	0.637
	1,000	0.165	0.095	0.030	0.344	0.234	0.099
$H = 4$	202	0.684	0.597	0.431	0.998	0.997	0.992
	1,000	0.359	0.256	0.119	0.753	0.659	0.463
$H = 8$	202	0.881	0.835	0.728	1.000	1.000	1.000
	1,000	0.467	0.358	0.193	0.913	0.865	0.744
Panel B: Power of Asymptotic Rank Test							
		$N = 10$			$N = 25$		
	T	10%	5%	1%	10%	5%	1%
$H = 1$	202	1.000	1.000	1.000	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000	1.000
$H = 4$	202	1.000	1.000	1.000	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000	1.000
$H = 8$	202	1.000	1.000	0.999	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000	1.000
Panel C: Size of Bootstrap Rank Test							
		$N = 10$			$N = 25$		
	T	10%	5%	1%	10%	5%	1%
$H = 1$	202	0.058	0.023	0.004	0.016	0.004	0.000
	1,000	0.075	0.035	0.005	0.044	0.015	0.002
$H = 4$	202	0.113	0.051	0.006	0.054	0.017	0.001
	1,000	0.131	0.066	0.012	0.126	0.060	0.011
$H = 8$	202	0.087	0.035	0.004	0.024	0.005	0.000
	1,000	0.125	0.064	0.012	0.093	0.038	0.006
Panel D: Power of Bootstrap Rank Test							
		$N = 10$			$N = 25$		
	T	10%	5%	1%	10%	5%	1%
$H = 1$	202	1.000	0.998	0.995	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000	1.000
$H = 4$	202	0.994	0.979	0.884	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000	1.000
$H = 8$	202	0.811	0.662	0.311	1.000	0.999	0.994
	1,000	1.000	1.000	0.998	1.000	1.000	1.000

Table B.I
Additional Empirical Evidence

The table presents the two-pass risk premium estimates in one-factor models with capital share (KS) and building permits in the North East (PerNE), Mid West (PerMW), South (PerS), and West (PerW) Census regions as pricing factors. The test asset portfolios are, in order, the 25 size and momentum (Size/MOM), 10 short-term reversal (ShREV), 17 industry (Industries), and 10 dividend yield (DIVYLD) sorted portfolios from Kenneth French's website. We also consider these equity portfolios together in the "All Equities" column. The additional test asset portfolios are the 23 commodity (Commodities) and 12 foreign exchange (FX) portfolios from He, Kelly, and Manela (2017). The sample period is 1963:Q3 to 2013:Q4 except for Commodities (1986:Q4 to 2012:Q4) and FX (1976:Q2 to 2009:Q4). For each one-factor model and compounding horizon $H = 4$ and 8, we report LLM's confidence interval for the risk premium estimate (in square brackets), the cross-sectional adjusted R^2 ($\bar{R}^2$, in curly brackets), and the p-value of the bootstrap rank test (in round brackets). The number of bootstrap replications for the calculation of the confidence intervals and the p-values of the bootstrap rank test is 10,000.

	Size/MOM	ShREV	Industries	DIVYLD	All Equities	Commodities	FX
				$H = 4$			
KS	−0.48	0.43	0.05	0.08	−0.05	−0.22	0.64
	[−0.96, −0.01]	[−0.16, 1.03]	[−0.12, 0.22]	[−0.20, 0.37]	[−0.28, 0.19]	[−0.63, 0.18]	[−0.51, 1.79]
	{$\bar{R}^2$=0.18}	{$\bar{R}^2$=0.22}	{$\bar{R}^2$=−0.04}	{$\bar{R}^2$=−0.09}	{$\bar{R}^2$=−0.01}	{$\bar{R}^2$=−0.01}	{$\bar{R}^2$=−0.00}
	(0.866)	(0.018)	(0.200)	(0.026)	(0.793)	(0.700)	(0.239)
PerNE	0.71	0.90	0.02	−0.20	0.38	0.33	−0.79
	[0.10, 1.31]	[0.32, 1.43]	[−0.15, 0.19]	[−0.58, 0.18]	[0.15, 0.61]	[−0.08, 0.73]	[−1.95, 0.49]
	{$\bar{R}^2$=0.13}	{$\bar{R}^2$=0.56}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=−0.00}	{$\bar{R}^2$=0.12}	{$\bar{R}^2$=0.03}	{$\bar{R}^2$=0.15}
	(0.197)	(0.741)	(0.064)	(0.623)	(0.675)	(0.621)	(0.331)
PerMW	0.73	0.82	0.05	−0.27	0.39	0.24	−0.68
	[0.12, 1.33]	[0.25, 1.32]	[−0.11, 0.20]	[−0.87, 0.32]	[0.16, 0.61]	[−0.33, 0.79]	[−2.94, 1.61]
	{$\bar{R}^2$=0.17}	{$\bar{R}^2$=0.40}	{$\bar{R}^2$=−0.04}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=0.15}	{$\bar{R}^2$=−0.02}	{$\bar{R}^2$=−0.05}
	(0.124)	(0.441)	(0.005)	(0.825)	(0.469)	(0.270)	(0.661)
PerS	0.94	0.94	0.06	−0.30	0.52	0.33	−1.93
	[0.34, 1.52]	[0.14, 1.66]	[−0.12, 0.24]	[−0.88, 0.25]	[0.28, 0.76]	[−0.39, 1.06]	[−4.22, 0.49]
	{$\bar{R}^2$=0.24}	{$\bar{R}^2$=0.22}	{$\bar{R}^2$=−0.04}	{$\bar{R}^2$=−0.02}	{$\bar{R}^2$=0.19}	{$\bar{R}^2$=−0.02}	{$\bar{R}^2$=0.21}
	(0.311)	(0.790)	(0.009)	(0.911)	(0.673)	(0.325)	(0.547)
PerW	0.75	0.72	0.04	0.19	0.45	0.69	0.18
	[0.14, 1.36]	[0.28, 1.12]	[−0.14, 0.21]	[−0.39, 0.74]	[0.21, 0.69]	[0.14, 1.24]	[−2.25, 2.55]
	{$\bar{R}^2$=0.14}	{$\bar{R}^2$=0.46}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=−0.08}	{$\bar{R}^2$=0.15}	{$\bar{R}^2$=0.10}	{$\bar{R}^2$=−0.10}
	(0.063)	(0.373)	(0.001)	(0.811)	(0.292)	(0.680)	(0.501)
				$H = 8$			
KS	−0.38	0.34	−0.00	0.09	−0.07	−0.26	0.28
	[−0.73, −0.04]	[−0.12, 0.81]	[−0.12, 0.12]	[−0.03, 0.23]	[−0.25, 0.10]	[−0.52, −0.01]	[−0.42, 0.98]
	{$\bar{R}^2$=0.21}	{$\bar{R}^2$=0.20}	{$\bar{R}^2$=−0.07}	{$\bar{R}^2$=0.15}	{$\bar{R}^2$=−0.00}	{$\bar{R}^2$=0.13}	{$\bar{R}^2$=−0.03}
	(0.674)	(0.014)	(0.052)	(0.136)	(0.833)	(0.275)	(0.460)
PerNE	0.80	0.73	0.01	0.10	0.30	0.07	−0.35
	[0.34, 1.25]	[0.48, 0.99]	[−0.10, 0.11]	[−0.13, 0.32]	[0.13, 0.47]	[−0.24, 0.39]	[−0.96, 0.30]
	{$\bar{R}^2$=0.27}	{$\bar{R}^2$=0.77}	{$\bar{R}^2$=−0.07}	{$\bar{R}^2$=−0.04}	{$\bar{R}^2$=0.13}	{$\bar{R}^2$=−0.04}	{$\bar{R}^2$=0.03}
	(0.131)	(0.253)	(0.317)	(0.568)	(0.626)	(0.298)	(0.489)
PerMW	0.75	0.71	0.02	0.16	0.29	0.17	0.06
	[0.27, 1.23]	[0.39, 1.04]	[−0.09, 0.12]	[−0.05, 0.38]	[0.11, 0.46]	[−0.21, 0.56]	[−1.21, 1.35]
	{$\bar{R}^2$=0.22}	{$\bar{R}^2$=0.61}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=0.11}	{$\bar{R}^2$=0.12}	{$\bar{R}^2$=−0.02}	{$\bar{R}^2$=−0.10}
	(0.090)	(0.448)	(0.186)	(0.605)	(0.784)	(0.467)	(0.711)
PerS	0.98	0.95	0.03	0.17	0.50	0.33	−0.70
	[0.56, 1.40]	[0.28, 1.57]	[−0.11, 0.18]	[−0.19, 0.52]	[0.31, 0.68]	[−0.24, 0.92]	[−2.10, 0.69]
	{$\bar{R}^2$=0.39}	{$\bar{R}^2$=0.34}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=−0.05}	{$\bar{R}^2$=0.24}	{$\bar{R}^2$=−0.01}	{$\bar{R}^2$=0.02}
	(0.182)	(0.712)	(0.096)	(0.691)	(0.469)	(0.426)	(0.732)
PerW	0.64	0.64	0.01	0.19	0.35	0.50	0.81
	[0.16, 1.13]	[0.38, 0.90]	[−0.12, 0.14]	[−0.02, 0.39]	[0.16, 0.54]	[0.07, 0.90]	[−0.68, 2.46]
	{$\bar{R}^2$=0.14}	{$\bar{R}^2$=0.65}	{$\bar{R}^2$=−0.06}	{$\bar{R}^2$=0.19}	{$\bar{R}^2$=0.13}	{$\bar{R}^2$=0.12}	{$\bar{R}^2$=−0.02}
	(0.191)	(0.535)	(0.222)	(0.462)	(0.637)	(0.414)	(0.464)