Revisiting Lettau and Ludvigson (2001): Does *cay* Really Matter for Cross-Sectional Risk Premia?

PAULO MAIO BYOUNG-KYU MIN*

November 2025

Abstract

We revisit the asset pricing tests conducted in Lettau and Ludvigson (2001) (LL). Contrary to LL's claim, their conditional models based on *cay* do not explain the dispersion in risk premia among the size/book-to-market portfolios. The pricing performance is either very poor (conditional CAPM/Consumption-CAPM) or quite modest (conditional Human-capital-CAPM), with all models largely underperforming the Fama–French model. Furthermore, the risk price estimates for the scaled factors are insignificant in several cases. When the zero-beta-rate is unrestricted, we obtain average pricing errors that are similar or above the raw average risk premia. Alternatively, LL's extreme intercept estimates are inconsistent with the equity premium puzzle. Employing alternative test portfolios and evaluating the identification of the risk prices substantially reinforces the negative outlook on their results. LL's models also perform poorly out-of-sample. Overall, LL's wrong conclusions stem from a combination of incorrect empirical choices and a misinterpretation of their results.

*Maio: paulo.maio@hanken.fi, Department of Finance and Economics, Hanken School of Economics; Min: minbk@hanyang.ac.kr, College of Economics and Finance, Hanyang University. We thank two anonymous referees and Ivo Welch (the editor) for valuable comments and suggestions. We also thank Kenneth French, Martin Lettau, and Zhaoguo Zhan for providing part of the data used in this study. This work was supported by the research fund of Hanyang University (HY-202400000003737). A previous version was titled "Resurrecting the (C)CAPM: A Cross-Sectional Test when Risk Premia are Time-Varying: A Comment." All errors are our own.

1 Introduction

In an influential study, Lettau and Ludvigson (2001) (LL henceforth) claim that conditional versions of the CAPM or Consumption-CAPM (CCAPM) based on the lagged log consumption-to-wealth ratio (*cay*) go a long way in pricing the 25 size/book-to-market portfolios, and thus in explaining the value-growth anomaly. Specifically, they write in their abstract: "We demonstrate that such conditional models perform far better than unconditional specifications and about as well as the Fama-French three-factor model on portfolios sorted by size and book-to-market characteristics. The conditional consumption CAPM can account for the difference in returns between low-book-to-market and high-book-to-market portfolios..."

We revisit LL's empirical evidence by conducting a replication (in both a narrow and a broad sense) of their cross-sectional tests. That is, we rely on the same sample period and testing portfolios, but offer a more complete and appropriate picture in terms of empirical methods and empirical choices. Our conclusions are straightforward: Contrary to LL's claim, their three conditional models based on *cay* do not explain the dispersion in risk premia among the 25 portfolios. Indeed, their pricing performance is either very poor (for both the three-factor conditional CAPM and conditional CCAPM) or it is quite modest (for the five-factor conditional version of the human-capital CAPM (HCAPM)). Moreover, all three conditional models underperform to a very large degree the benchmark Fama–French three-factor model (FF3).

Below, we summarize the main issues with LL's empirical asset pricing tests. First, LL do not report the results for the specification with a restricted zero-beta rate, which is consistent with the underlying theoretical model (under the usual assumption that there is a risk-free asset). By estimating the cross-sectional regressions without an intercept, we obtain a negative cross-sectional R^2 for the conditional consumption model (LL3*), while the fit for the scaled CAPM (LL3) is very close to zero. This reveals a very negative pricing performance, that is, these two models do either worse or nearly about the same as a trivial factor model containing solely a constant. In the case of the conditional HCAPM (designated by LL5), the corresponding explanatory ratio is positive, but the magnitude is quite modest (29%), representing less than half the fit obtained for FF3 (80%). Importantly, we cannot reject the null that such a level of fit is equal to zero by relying on a bootstrap simulation that offers good size and power properties.

Second, LL incorrectly interpret the intercept estimates (which reflect severe model misspecification) as part of the risk premium explained by each model rather than as part of the corresponding pricing errors. By including the intercept estimates in the definition of mispricing, we obtain average absolute pricing errors that are actually around or above the average raw risk premium for all three conditional models, with estimates varying between 95% (for LL3) and 168% (for LL5). These estimates of relative mispricing are very large in economic terms and represent a clear rejection of the three scaled models. To have another perspective on the degree of empirical failure of LL's models, those extreme relative average pricing errors arise from the very large intercept estimates, which are above 2% per quarter (or roughly 8% per year) in all three cases. These extreme estimates of the "excess zero-beta rate", if taken seriously, would imply that the equity premium puzzle of Mehra and Prescott (1985) is not really a puzzle, that is the aggregate equity premium would be much smaller than what we find in the data when using the observed risk-free rate.

Third, LL do not provide the individual stochastic discount factors (SDF) coefficient (or equivalently, the covariance risk price) estimates. By estimating the covariance representation of LL5, we find that the risk price estimate for the scaled labor income factor is strongly insignificant. Hence, such a factor is not significantly correlated with the corresponding SDF, and thus it is irrelevant (in statistical terms) to price risk premia.

Fourth, LL actually downplay some of the adverse results reported in their paper. One example is the statistical insignificance of the (beta) risk price estimates for both scaled factors (within their five-factor model) when one relies on the standard errors proposed by Shanken (1992).

To put the previous results into perspective, we employ alternative test portfolios when evaluating the pricing performance of LL's models, which can be interpreted as an "out-of-sample" evaluation of these models at the cross-sectional dimension. Concretely, we consider decile portfolios sorted on three stock characteristics—book-to-market ratio, investment, and operating profitability. We use LL's original sample period, as well as a longer sample period that ends in 2013. Overall, the conditional models from LL are not able to explain portfolios sorted on investment and operating profitability (in addition to the value portfolios). Importantly, LL's models underperform the reference model for this new cross-section (the five-factor model of Fama and French (2015) or FF5) to a large extent. On the other hand, the risk price estimates for the three scaled factors are insignificant in most cases. Thus, if something, the picture looks even more adverse than in the estimation with the baseline 25 portfolios.

Furthermore, we inspect whether the risk prices in LL's models are identified by conducting rank tests on the respective factor loadings. The results indicate that the risk price estimates (and pricing errors) associated with LL's conditional models are largely questionable, including the estimations conducted over the more relevant cross-section of 30 portfolios. This further reinforces the negative outlook on the empirical performance and credibility of LL's models.

We also conduct Monte-Carlo experiments to inspect the relationship between missing factors in the model and the cross-sectional intercept to help us understand the empirical results discussed above. The simulation results show that model misspecification has an important role in terms of driving large intercept estimates, in line with the empirical evidence produced by LL's models.

In the last part of the paper, we evaluate the out-of-sample pricing performance of LL's models over the time-series dimension by comparing predicted risk premia (from the models estimated in sample) with one-quarter ahead realized excess returns. The results show that LL's models have a weak out-of-sample performance, underperforming the reference models (either FF3 or FF5) by a large degree in nearly all cases.

Several studies in the literature have estimated or inspected LL's conditional models. Among these papers, Lewellen et al. (2010) estimate LL's three-factor conditional CCAPM over the 1963–2004 period and assess the statistical significance of the corresponding cross-sectional R^2 (see their Section 4). In related work, Lewellen and Nagel (2006) claim (in their Section 5) that this same model is inconsistent with the underlying consumption CAPM. Yet, Lewellen and Nagel (2006) do not perform any formal estimation and evaluation of the scaled CCAPM, which makes their proposition somewhat speculative and difficult to judge. Importantly, none of these two studies address any of the main issues associated with LL's analysis that are discussed above—e.g., not reporting the results for the restricted zero-beta-rate specification, exclusion of the intercept estimates from the pricing errors, or not reporting the covariance risk price estimates.

The rest of this paper is organized as follows. Section 2 describes the theoretical and empirical backgrounds. Section 3 presents the main asset pricing results, while Section 4 presents the results for the covariance representation of the conditional models. In Section 5, we consider alternative test portfolios, while Section 6 evaluates the identification of the risk prices in LL's models. Section

7 contains a Monte-Carlo analysis, while Section 8 refers to an out-of-sample evaluation of LL's models. Finally, Section 9 concludes.

2 Background

2.1 Theory and Models

Assume that there is a gross-risk free rate between t and $t+1$, denoted by $R_{f,t}$, which is known at the end of period t. The excess return for asset i is defined as $R^e_{i,t+1} \equiv R_{i,t+1} - R_{f,t}$. Working with excess returns represents the norm in the literature and has the advantage of leaving no room for a constant term in the pricing equation, as long as the unobserved zero-beta rate equals the risk-free rate (the common assumption in the literature). Lettau and Ludvigson (2001) (LL henceforth) assume a linear stochastic discount factor (SDF) with time-varying coefficients. To simplify the notation, we consider the following single-factor model with a single instrument,

$$M_{t+1} = a_t + b_t F_{t+1}, \tag{1}$$

where $a_t = \gamma_0 + \gamma_1 z_t$, $b_t = \eta_0 + \eta_1 z_t$, and z_t is the conditioning variable, which is observed at time t. Hence, the SDF can be rewritten as

$$M_{t+1} = \gamma_0 + \gamma_1 z_t + \eta_0 F_{t+1} + \eta_1 F_{t+1} z_t. \tag{2}$$

By taking unconditional expectations to both sides of the conditional SDF representation, we obtain an equivalent unconditional representation,

$$0 = \mathrm{E}_t(R^e_{i,t+1} M_{t+1}) \Rightarrow 0 = \mathrm{E}(R^e_{i,t+1} M_{t+1}), \tag{3}$$

by using the law of iterated expectations.

The equivalent covariance pricing equation is given by

$$\mathrm{E}(R^e_{i,t+1}) = -\,\mathrm{E}(R_{f,t})\,\mathrm{cov}(R^e_{i,t+1}, M_{t+1}), \tag{4}$$

where we use the equality between the average gross risk-free rate and the reciprocal of the mean SDF, $\mathrm{E}(R_{f,t}) = 1/\,\mathrm{E}(M_{t+1})$. By substituting the SDF above in this general pricing equation, we obtain the following three-factor model,

$$\mathrm{E}(R^e_{i,t+1}) = \gamma_z \operatorname{cov}(R^e_{i,t+1}, z_t) + \gamma_f \operatorname{cov}(R^e_{i,t+1}, F_{t+1}) + \gamma_{fz} \operatorname{cov}(R^e_{i,t+1}, F_{t+1} z_t), \tag{5}$$

where $\gamma_z \equiv -\,\mathrm{E}(R_{f,t})\gamma_1$, $\gamma_f \equiv -\,\mathrm{E}(R_{f,t})\eta_0$, and $\gamma_{fz} \equiv -\,\mathrm{E}(R_{f,t})\eta_1$ denote the covariance risk prices. The risk prices coincide with (minus) the SDF coefficients up to a constant close to one, $\mathrm{E}(R_{f,t}) \approx 1$.

Equivalently, we can define the pricing equation in terms of multivariate betas, which is the representation estimated by LL. First, we define the covariance representation as follows,

$$\mathrm{E}(R^e_{i,t+1}) = \boldsymbol{\gamma}' \boldsymbol{\sigma}_i, \tag{6}$$

where $\boldsymbol{\gamma} \equiv (\gamma_z, \gamma_f, \gamma_{fz})'$ is the vector of risk prices and

$$\boldsymbol{\sigma}_i \equiv [\operatorname{cov}(R^e_{i,t+1}, z_t), \operatorname{cov}(R^e_{i,t+1}, F_{t+1}), \operatorname{cov}(R^e_{i,t+1}, F_{t+1} z_t)]', \tag{7}$$

denotes the vector of factor covariances for asset i. Denoting the (3×3) factor variance-covariance matrix by $\boldsymbol{\Sigma}_f$, it follows that the pricing equation can be written as

$$\mathrm{E}(R^e_{i,t+1}) = \boldsymbol{\gamma}' \boldsymbol{\Sigma}_f \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\sigma}_i = \boldsymbol{\lambda}' \boldsymbol{\beta}_i, \tag{8}$$

where $\boldsymbol{\lambda} \equiv \boldsymbol{\Sigma}_f \boldsymbol{\gamma}$ is the (3×1) vector of (beta) risk prices and $\boldsymbol{\beta}_i \equiv \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\sigma}_i$ is the (3×1) vector of factor betas. This pricing equation can be rewritten as

$$\mathrm{E}(R^e_{i,t+1}) = \lambda_z \beta_{zi} + \lambda_f \beta_{fi} + \lambda_{fz} \beta_{fzi}, \tag{9}$$

where the factor betas are obtained from the following multiple time-series regression:

$$R^e_{i,t+1} = \delta_i + \beta_{zi} z_t + \beta_{fi} F_{t+1} + \beta_{fzi} F_{t+1} z_t + \varepsilon_{i,t+1}. \tag{10}$$

This derivation makes clear that the factors used in the computation of the multivariate betas need to be the very same factors included in the pricing equation. Otherwise, we have an inconsistency between the model being estimated and the underlying "theoretical" pricing equation.

In the empirical analysis conducted in the following sections, we consider the three conditional factor models estimated by LL. The first model is the three-factor conditional CAPM (LL3 henceforth),

$$\mathrm{E}(R^e_{i,t+1}) = \lambda_z \beta_{zi} + \lambda_{vw} \beta_{vwi} + \lambda_{vwz} \beta_{vwzi}, \tag{11}$$

where λ_{vw} and λ_{vwz} denote the risk prices for the raw and scaled market factors, respectively. β_{vwi} and β_{vwzi} represent the corresponding factor loadings. λ_z and β_{zi} denotes the risk price and beta for the lagged instrument as a separate factor. The baseline CAPM is obtained by setting $\lambda_z = \lambda_{vwz} = 0$.

The second conditional model is the five-factor conditional human-capital-CAPM (LL5 henceforth),

$$\mathrm{E}(R^e_{i,t+1}) = \lambda_z \beta_{zi} + \lambda_{vw} \beta_{vwi} + \lambda_{vwz} \beta_{vwzi} + \lambda_{\Delta y} \beta_{\Delta yi} + \lambda_{\Delta yz} \beta_{\Delta yzi}, \tag{12}$$

where $\lambda_{\Delta y}$ and $\lambda_{\Delta yz}$ denote the risk prices for the original and scaled labor income growth factors, respectively. $\beta_{\Delta yi}$ and $\beta_{\Delta yzi}$ represent the corresponding factor loadings. The baseline two-factor human-capital-CAPM (HCAPM) is obtained by imposing $\lambda_z = \lambda_{vwz} = \lambda_{\Delta yz} = 0$.

The last conditional model represents a scaled version of the Consumption-CAPM (LL3* henceforth),

$$\mathrm{E}(R^e_{i,t+1}) = \lambda_z \beta_{zi} + \lambda_{\Delta c} \beta_{\Delta ci} + \lambda_{\Delta cz} \beta_{\Delta czi}, \tag{13}$$

where $\lambda_{\Delta c}$ and $\lambda_{\Delta cz}$ denote the risk prices for the raw and scaled consumption growth factors, respectively. The respective betas are $\beta_{\Delta ci}$ and $\beta_{\Delta czi}$. The baseline CCAPM is obtained by setting $\lambda_z = \lambda_{\Delta cz} = 0$.

Finally, and following LL, we also estimate the three-factor model of Fama and French (1993) (henceforth, FF3), which acts as the reference model.

2.2 Methodology

Following LL, we use the two-pass ordinary least squares (OLS) regression approach to estimate the different factor models (e.g., Brennan et al. , 2004, Cochrane , 2005, among many others).

Taking LL3 as an example, in the first step the factor betas are estimated from the time-series regressions for each testing portfolio,

$$R^e_{i,t+1} = \delta_i + \beta_{zi} z_t + \beta_{vwi} R_{vw,t+1} + \beta_{vwzi} R_{vw,t+1} z_t + \varepsilon_{i,t+1}, \tag{14}$$

where R_{vw} denotes the baseline market factor (excess market return). In the second step, the beta representation is estimated from an OLS cross-sectional regression,

$$\mu_i = \lambda_0 + \lambda_z \widehat{\beta}_{zi} + \lambda_{vw} \widehat{\beta}_{vwi} + \lambda_{vwz} \widehat{\beta}_{vwzi} + \alpha_i, \tag{15}$$

where μ_i represents the average (time-series) excess return for asset i, α_i denotes the respective pricing error, and $\widehat{\beta}_{zi}$, $\widehat{\beta}_{vwi}$, and $\widehat{\beta}_{vwzi}$ represent the factor loading estimates for asset i.[1] We estimate the cross-sectional regression both with and without the intercept (λ_0). The t-statistics associated with the risk price estimates are based on Shanken's standard errors (Shanken , 1992), which account for the estimation error in the factor loadings.[2]

Following LL, we compute the cross-sectional OLS coefficient of determination,

$$R^2 = 1 - \frac{\text{Var}_N(\widehat{\alpha}_i)}{\text{Var}_N(\mu_i)}, \tag{16}$$

where $\text{Var}_N(\cdot)$ stands for the cross-sectional sample variance.[3] R^2 represents the fraction of the cross-sectional variation in the average excess returns (on the test assets) that is explained by the factor loadings associated with a given model.[4] Following Lewellen et al. (2010) and Guo and

[1] Throughout the paper, we use a "hat" to denote a sample estimate.

[2] LL also consider the t-ratios based on the standard errors from Fama and MacBeth (1973). We do not consider these standard errors, as it is well known that the Shanken's standard errors tend to be more conservative (e.g., Kan et al. , 2013).

[3] $\widehat{\alpha}_i$ represents the residual from the cross-sectional regression, albeit we often designate it by "pricing error."

[4] Following LL, we do not consider the case in which the second-step cross-sectional regression is estimated by GLS (rather than OLS), an approach that is advocated by Kandel and Stambaugh (1995) and Lewellen et al. (2010), among others. The reason is that, similarly to LL, we are interested in assessing the capacity of the conditional models in explaining the original 25 portfolios, which are associated with the size and value anomalies, rather than explaining a mean-variance transformation of those same portfolios ("repackaged portfolios"). Importantly, it is often the case that the repackaged portfolios involve extreme weights on the original portfolios, which makes the former difficult to be implemented in practice. For these reasons, and despite Lewellen et al. (2010), the bulk of the literature still focuses nowadays on the OLS (rather than the GLS) cross-sectional regression when estimating and evaluating linear models that contain non-traded factors. As discussed in Cochrane (2005), if all the factors in the model are traded, it follows that the GLS risk price estimates (obtained from a cross-section that includes the factors as test assets) coincide with the sample factor means (that is, the implied risk price estimates when using the popular time-series approach). Yet, the corresponding GLS cross-sectional pricing errors do not measure the capacity of the

Maio (2020), to address the statistical uncertainty associated with the sample R^2, we compute empirical p-values based on a bootstrap simulation. The empirical p-value represents the fraction of artificial samples in which the pseudo explanatory ratio is higher than the corresponding sample R^2. Our simulation extends the bootstrap employed in Guo and Maio (2020) in two aspects. First, we impose the restriction that the risk prices (but not the intercept) of a given model are zero, something that is consistent with the null of no pricing ability, when simulating the (realized) excess portfolio returns. Second, we impose a common resampling on all portfolio returns and factors, to allow for a possible correlation among both sets of variables. Appendix A.1 contains the complete description of the bootstrap simulation.

Despite representing an intuitive goodness-of-fit measure, the cross-sectional R^2 might be misleading. For example, we can have a situation in which the pricing errors are large (in magnitude), but close to each other, thus originating a small value of $\mathrm{Var}_N(\widehat{\alpha}_i)$, and thus a high R^2. On the other hand, we can have a situation in which the risk premia does not vary much across the testing assets (i.e., a low value of $\mathrm{Var}_N(\mu_i)$), thus originating a low R^2 despite the small magnitudes of the pricing errors. Moreover, R^2 might be sensitive to outlier test assets, that is, assets having pricing errors and risk premia that deviate substantially from those of the other portfolios. Therefore, we also compute the mean absolute pricing error (MAE),

$$MAE = \mathrm{E}_N(|\widehat{\alpha}_i|), \tag{17}$$

where $\mathrm{E}_N(\cdot)$ stands for the cross-sectional sample mean.[5] MAE does not suffer from some of the problems associated with R^2 described above. However, one key disadvantage of MAE is that it represents a measure of absolute pricing performance. Thus, it could be the case that the average pricing error seems economically significant (i.e., assumes a large magnitude), yet such an estimate is small in comparison to the cross-sectional average risk premium that we seek to explain in the first place. To address such an important limitation, we also compute the MAE/MAR ratio, where MAR is the mean absolute raw risk premium, $MAR = \mathrm{E}_N(|\mu_i|)$. This metric measures the mispricing as a proportion of the mean raw (absolute) risk premia to be explained, and (similar to

model in terms of explaining the original test portfolios of economic interest. Importantly, the three conditional models employed by LL contain non-traded factors.

[5] LL report the root mean squared pricing error in their Table 4. Yet, the two average pricing error metrics are very similar.

R^2) represents a measure of relative pricing performance.[6]

When the cross-sectional regression contains the intercept, we include the "excess zero-beta" rate estimate ($\widehat{\lambda}_0$) in the definition of the "pricing error" (e.g., Burnside , 2011 and Maio , 2024). This choice comes from acknowledging that the intercept should be interpreted as a specification error ("common pricing error"), and thus handled in the same way as the cross-sectional residual. This implies that the mean absolute pricing error becomes:

$$MAE = \mathrm{E}_N(|\widehat{\alpha}_i + \widehat{\lambda}_0|). \tag{18}$$

Since $\mathrm{E}_N(\widehat{\alpha}_i) = 0$, the intercept estimate represents the mean pricing error. In comparison, R^2 is not affected by the inclusion of the intercept in the pricing errors, as it is based on centered (rather than uncentered) cross-sectional second moments.

LL's empirical approach has several key differences relative to the standard two-step approach described above. First, they introduce a decoupling between the factors used in the time-series regressions and those employed in the cross-sectional regression, a procedure that is theoretically invalid. Specifically, in the estimation of their two-factor conditional CAPM that excludes the lagged instrument, both $\widehat{\beta}_{vwi}$ and $\widehat{\beta}_{vwzi}$ are obtained from a regression that also includes the lagged instrument,

$$R^e_{i,t+1} = \delta_i + \beta_{zi} z_t + \beta_{vwi} R_{vw,t+1} + \beta_{vwzi} R_{vw,t+1} z_t + \varepsilon_{i,t+1}, \tag{19}$$

instead of a bivariate regression that excludes such a variable. Given that both $R_{vw,t+1}$ and $R_{vw,t+1} z_t$ have non-zero correlations with z_t, it follows that both $\widehat{\beta}_{vwi}$ and $\widehat{\beta}_{vwzi}$ differ from the "correct" estimates, which in turn has an impact on the second-pass risk price estimates and associated pricing errors. Such a procedure is theoretically incorrect and difficult to accept (even by the standards of the asset pricing literature back in 2001).[7] LL argue in their footnote 25 that the reason for including *cay* in the time-series regression is that the corresponding SDF coefficient (b) may be non-zero despite the beta risk price (λ) being zero. This argument makes no sense. In

[6] In our empirical analysis in the next section, it turns out that MAR is equal to the mean risk premium, $MAR = \mathrm{E}_N(\mu_i)$, as all testing portfolios have positive risk premia (i.e., positive average excess returns).

[7] Note that the pricing errors (and thus the aggregate performance metrics) are the same under the multivariate- and single-beta representations of a given linear asset pricing model. However, such an equivalence does not hold when the multivariate betas are estimated from time-series regressions that incorporate "outside" factors, that is, factors that are not in the model.

multifactor models, there can exist a discrepancy between the covariance and beta prices of risk (and also in their statistical significance) due to the non-zero correlations among the factors (and this is the reason why we also estimate covariance risk prices in our empirical analysis). However, this issue has no connection with the (wrong) practice of including outside factors (factors that do not belong to the model) when computing the (multivariate) betas for the relevant factors (those that are in the model). In fact, we are not aware of another study in the literature that adopts such a decoupling in the factors used in the time-series and cross-sectional regressions (when estimating the multivariate beta representation of a given model). Moreover, the decoupling between the factors included in the time-series and cross-sectional regressions has an impact on the standard errors of the risk price estimates.[8] In light of this discussion, we do not attempt to replicate the results for LL's scaled models that exclude the lagged instrument as a factor, a choice that is also consistent with the theoretical derivations above.

A second major issue with LL's empirical approach is that they do not report the estimation results when the "excess zero-beta rate" is constrained to zero, $\lambda_0 = 0$. As noted above, restricting the intercept to zero is consistent with the underlying theoretical model under the assumption that there is a risk-free asset, thus implying that a risky asset with zero factor loadings earns a zero risk premium. There are several other important advantages of restricting the intercept to zero. First, it allows a proper comparison with APT-based equity models (containing exclusively traded factors), such as FF3. These models are traditionally estimated under the time-series method in which the intercept (in the corresponding cross-sectional pricing equation obtained from the time-series regressions of the testing realized excess returns onto the factors) is implicitly restricted to zero (e.g., Fama and French , 1993, 1996, Cochrane , 2005, among many others).[9] Second, the presence of the intercept in the cross-sectional regression often creates (near perfect) multicollinearity, which is driven by the small cross-sectional dispersion in the betas for several factors in the literature (with most macro factors and the market factor being notable examples). This often leads to largely implausible factor risk price estimates from an economic perspective, such as negative

[8]Concretely, the Shanken's standard errors are derived under the (usual) assumption that the decoupling documented above does not exist. Indeed, the GMM-based derivation of the Shanken's correction presented in Cochrane (2005) (Chapter 12) makes clear that such a decoupling is not possible. Thus, what LL label as "Shanken's standard errors" do not actually represent Shanken's standard errors. Relatedly, several studies have shown that Shanken's standard errors are invalid in the context of misspecified models (e.g., Kan et al. , 2013 and Raponi et al. , 2020).

[9]In other words, the intercepts from the time-series regressions (alphas) represent the pricing errors.

market prices of risk, which clearly represents an economic rejection of the models under analysis. Third, the cross-sectional OLS R^2 defined above can assume negative values when $\lambda_0 = 0$, which has an intuitive interpretation: A negative estimate indicates that the cross-sectional regression (including the factor betas as regressors) does worse than a trivial regression containing only the intercept, something that indicates a very negative performance. In contrast, when the cross-sectional regression includes an intercept, the R^2 always assumes positive values (by construction). Yet, such positive estimates are often misleading, as these arise exclusively from the ability of the model in matching (partially) the average risk premium rather than from explaining any cross-sectional dispersion in risk premia. However, the latter represents the main goal of cross-sectional asset pricing tests.

A third key problem with LL's empirical approach is a clear misinterpretation of their cross-sectional regressions. As discussed in the next section, the estimates of the excess zero-beta rate are very large in economic terms. Transaction costs or imperfections in money markets, such as bid-ask spreads or liquidity premiums, do not represent credible explanations for such very large estimates (e.g., Burnside , 2011). In fact, the most credible argument for including the intercept is that it captures the cross-sectional average risk premium (risk price times average loading) associated with missing risk factors. Importantly, such an argument reflects misspecification of the model, which means that the intercept should be part of the pricing error (rather than included in the explained part of the model, as done by LL). On the other hand, a statistically insignificant intercept estimate provides a statistical validation of the zero-intercept restriction, especially when such a specification is consistent with theory, as discussed above.

To see why large intercept estimates reflect model's misspecification, consider the case of the baseline CCAPM as an example. The OLS estimate of the intercept is given by

$$\widehat{\lambda}_0 = \mathrm{E}_N(\mu_i) - \widehat{\lambda}_{\Delta c}\, \mathrm{E}_N(\widehat{\beta}_{\Delta ci}), \tag{20}$$

where $\widehat{\lambda}_{\Delta c}$ denotes the risk price estimate for the baseline consumption factor. If the average portfolio consumption beta, $\mathrm{E}_N(\widehat{\beta}_{\Delta ci})$, has a small magnitude, it follows that the average portfolio fitted risk premium, $\widehat{\lambda}_{\Delta c}\, \mathrm{E}_N(\widehat{\beta}_{\Delta ci})$, clearly falls below the average portfolio raw risk premium, $\mathrm{E}_N(\mu_i)$, thus originating a large $\widehat{\lambda}_0$. In fact, if the average beta is around zero, $\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) \approx 0$, it follows that the intercept estimate roughly equals the average raw risk premium, which is positive by construction, $\widehat{\lambda}_0 \approx \mathrm{E}_N(\mu_i) > 0$. Hence, large intercept estimates represent a situation in which the factors are nearly "useless" (very small or near zero correlations with all the testing returns). In turn, this implies that there are important missing factors in the model. In the CCAPM example, suppose the "true" model includes an additional factor (G), which is assumed to be orthogonal to Δc, without loss of generalization.[10] G is assumed to drive nearly all of the model's pricing performance for equity risk premia. The intercept estimate in the cross-sectional regression associated with this two-factor model is given by

$$\widehat{\theta}_0 = \mathrm{E}_N(\mu_i) - \widehat{\theta}_{\Delta c}\,\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) - \widehat{\theta}_G\,\mathrm{E}_N(\widehat{\beta}_{Gi}), \tag{21}$$

where $\widehat{\theta}_{\Delta c}$ and $\widehat{\theta}_G$ denotes the risk price estimate for Δc and G, respectively. $\widehat{\beta}_{\Delta ci}$ and $\widehat{\beta}_G$ represent the respective estimated factor loadings. Given the factor orthogonality assumption, it follows that the consumption beta estimated from the following multiple time-series regression,

$$R^e_{i,t+1} = \delta_i + \beta_{\Delta ci}\Delta c_{t+1} + \beta_{Gi}G_{t+1} + \varepsilon_{i,t+1},$$

is equal to the corresponding univariate consumption beta. Given the assumption that the two-factor model is correctly specified, it follows that $\widehat{\theta}_0 \approx 0$, which in turn implies that $\mathrm{E}_N(\mu_i) \approx \widehat{\theta}_{\Delta c}\,\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) + \widehat{\theta}_G\,\mathrm{E}_N(\widehat{\beta}_{Gi})$. Consequently, the intercept estimate in the single-factor model becomes:

$$\widehat{\lambda}_0 \approx (\widehat{\theta}_{\Delta c} - \widehat{\lambda}_{\Delta c})\,\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) + \widehat{\theta}_G\,\mathrm{E}_N(\widehat{\beta}_{Gi}). \tag{22}$$

Under the assumption that $\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) \approx 0$, the intercept estimate equals the average risk premium of the missing factor:[11]

$$\widehat{\lambda}_0 \approx \widehat{\theta}_G\,\mathrm{E}_N(\widehat{\beta}_{Gi}). \tag{23}$$

We note that the existence of missing risk factors in the empirical conditional models tested by LL might be generically consistent with the underlying theoretical background of those models.

[10] Gagliardini et al. (2019) propose a method to determine the approximate number of factors consistent with a given cross-section of asset returns.

[11] Note that $\mathrm{E}_N(\widehat{\beta}_{\Delta ci}) \approx 0$ also implies $\widehat{\theta}_{\Delta c} \approx \widehat{\lambda}_{\Delta c}$ from the algebra of least squares.

This is because the "theoretical" conditional CAPM (CCAPM) does not specify the identity of the instruments that drive variation in the SDF coefficients over time. Indeed, the choice of the instruments used in the empirical specifications is largely an arbitrary one, and the assumption that *cay* is the sole driver of the SDF coefficients can be overly restrictive. Consequently, there might be missing scaled factors in the empirical conditional models, but, as noted above, the average risk premium associated with those omitted factors (captured by the estimate of λ_0) represents misspecification error of the model being tested.

2.3 Data and Variables

This section offers a concise description for the factors' construction, eliminating the need for readers to refer back to Lettau and Ludvigson (2001). All the factor and portfolio return data used in this study are obtained from Zhaoguo Zhan and correspond to the data employed in the replication of LL conducted by Kleibergen and Zhan (2020).

Consumption growth is measured as the log growth in real aggregate per capita consumption, where consumption data is restricted to nondurables and services, excluding shoes and clothing, and is presented in 1992 chain weighted dollars. To compute per capita real consumption, nominal consumption is divided by both total population and the Personal Consumption Expenditure deflator. The standard "end-of-period" timing convention, matching consumption growth from time t to $t+1$ with asset returns calculated over the same time interval, is used. Labor income is measured as total personal income minus dividend income from NIPA. A one-month lag in labor income is adopted to account for any reporting delays in official reports of aggregate income. Labor income growth is calculated using the same approach as for consumption growth, by representing the log growth in real per capita labor income.

The lagged instrument used in LL is the log consumption-to-wealth ratio, denoted as *cay*.[12] *cay* is demeaned, following the convention in the conditional asset pricing literature. The test portfolios are 25 size/book-to-market sorted portfolios over the period from 1963:Q4 to 1998:Q3. Accordingly, *cay* is used from 1963:Q3 to 1998:Q2. Excess portfolio returns are computed by subtracting the

[12]It should be noted that LL re-scale *cay* by multiplying it by 10 in some specifications, while using the original scale in other specifications. To ensure consistency and eliminate any ambiguity in the interpretation of the asset pricing results, we employ the original scale of *cay* in all models. As a result, the magnitudes of some of the estimated risk prices reported in our study differ from those reported by LL.

quarterly three-month Treasury-bill return from the raw quarterly returns.

3 Main Results

This section contains a replication (both in a narrow and in a broad sense) of the cross-sectional tests conducted in LL. That is, we rely on the same sample period and test portfolios, but offer a more complete and appropriate picture in terms of empirical methods and empirical choices.

3.1 Conditional CAPM

We start with the empirical tests for the three-factor conditional CAPM.[13] We note that our intercept estimates represent proxies of the "excess zero-beta rate" (relative to the T-bill rate), whereas LL's estimates represent proxies of the "zero-beta rate".

The results presented in Table 1 (Panel A) show that the scaled market factor, $R_{vw,t+1}cay_t$, is priced (at the 5% level). However, the R^2 is somewhat modest, representing less than half the corresponding fit obtained for FF3 (31% versus 80%), and such a fit is strongly statistically insignificant based on the bootstrap. Importantly, the intercept estimate is 2.14% per quarter, compared to only 0.30% for FF3, which translates into an annual estimate as large as 8.56%. This represents a significant misspecification of the model, which is reflected in a very large average pricing error (2.14% per quarter or around 8.6% per year). Such level of mispricing is very close to the raw average risk premia that we want to explain in the first place, as indicated by the MAE/MAR ratio of 0.95 (compared to a ratio of only 0.17 for FF3), something that reveals a very negative performance. This fact is, however, ignored by LL, as they (erroneously so) include the intercept as part of the portfolio risk premia being explained by the model.

[Table 1 here]

In fact, another way to assess the severe implausibility of such an intercept estimate is by looking at the implied, and counterfactual, aggregate equity premium. Given the average risk-free rate of

[13] Our risk price estimates have a different scaling than LL's counterparts in their Table 1: While we use the standard procedure in the literature of multiplying all (beta) risk price estimates by 100, LL's adopt an asymmetric, and unusual, scaling across factors: The risk prices associated with cay_t and the scaled factor are both multiplied by 1000, while the baseline market risk price is multiplied by 100.

1.56% (per quarter) over LL's sample period, an intercept of 2.14% translates into a zero-beta rate as large as 3.70% per quarter (or 14.80% per year). With an average market return of 4.79% per quarter, the corresponding equity premium would be as low as 1.09% per quarter (or 4.36% per year), which compares to the observed quarterly equity premium in LL's sample of 3.23% (or 12.92% per year).

Critically, the simple fact that the intercept estimate is largely insignificant in statistical terms (t-ratio of 1.51) would indeed legitimate focusing on the specification with a restricted zero-beta rate, especially when such a specification is consistent with the theoretical model. Yet, when we estimate such a specification, the results in Panel B of Table 1 show a R^2 very close to zero (7%) for LL3, compared to a fit as large as 80% for FF3.[14] This also illustrates that the three-factor conditional CAPM does not perform much better than the baseline (single-factor) CAPM in terms of tracking cross-sectional dispersion in portfolio risk premia. The poor fit also makes it irrelevant the fact that $\lambda_{vw,cay}$ remains statistically significant when the intercept is restricted to zero. In other words, why would one be interested in the finding that the two factors in such a model are priced (in the specification without an intercept) when these two factors combined deliver a fit close to zero? Actually, this represents a very good example that significant factor risk price(s) does not imply that the respective model explains the cross-section of equity risk premia (that is, the factor betas are closely aligned cross-sectionally with the raw portfolio risk premia).[15]

Next, we turn attention to the five-factor conditional human capital-CAPM, which incorporates both the baseline and scaled human capital factors.[16] The results in Panel A of Table 1 show that the risk price estimates of both scaled factors are strongly statistically insignificant (in fact, there is no significance even at the 10% level), something that is in line with the evidence provided in LL's

[14] We note that the appropriate method to estimate FF3 is the time-series method (Fama and French 1993, 1996) rather than the two-step procedure. This is because all of the factors in the model are traded (excess returns), which implies that the OLS risk price estimates should coincide with the corresponding factor sample means. Yet, the sample means of the size (0.50%) and value (1.27%) factors do not differ substantially from the respective OLS risk price estimates. The reason is that these two factors are strongly correlated with the 25 portfolios by construction. Hence, the R^2 of 80% does not significantly overstate the fit of the three-factor model in this specific empirical setting.

[15] The classic example is the baseline CAPM: Typically, the market price of risk is statistically significant when the intercept is restricted to zero and the test assets represent equity portfolios. Yet, the corresponding cross-sectional OLS R^2 is often negative, which means that the single-factor model does worse than a trivial constant factor (that is, the market betas do not align with the portfolio risk premia and can actually go in the wrong direction). The evidence provided in Panel B of Table 1 represents another example of such a pattern.

[16] By comparing with our results, it looks like LL multiply the risk price for cay_t within LL5 by 100 (rather than 1000). This lack of consistency in scaling the risk price estimates for either different factors in the same model or the same factor in different models is clearly non-standard in the literature and makes it difficult to replicate LL's results in a narrow sense.

Table 1. Thus, each of these factors is irrelevant, in statistical terms, to explain risk premia. This very negative result is, however, "downplayed" by LL (see their page 1259). Indeed, they argue that the Shanken's standard errors are inevitably large for macro risk prices, something that is not true, as evidenced by the subsequent voluminous empirical asset pricing literature.

The five-factor model generates a large R^2, which is only slightly below the fit obtained for FF3 (77% versus 80%), being also identical to the estimate reported in LL. However, such apparent large fit is illusory and mainly arises from the inclusion of the intercept in the pricing equation. In fact, we obtain an even more extreme intercept estimate than in the three-factor conditional CAPM, as the "excess zero-beta rate" becomes as large as 3.77% per quarter (or 15.08% per year). Such an estimate represents almost 13 times (!) the magnitude of the corresponding estimate for the benchmark model (FF3). Naturally, this represents an extreme level of misspecification and a strong economic rejection of the model. To have an idea of the degree of model misspecification, the associated MAE of 3.77% per quarter is approximately 70% above (!) the observed average risk premium among the 25 portfolios (that is, a MAE/MAR ratio as large as 1.68), something that represents evidence of a very negative pricing performance. In other words, the seemingly large fit of 77% is entirely illusory (even if statistically above zero based on the empirical p-value of 2%), and this represents a great example of how misleading the cross-sectional R^2 can be when the intercept is not restricted to zero. To add to the negative picture, we see that the cross-sectional R^2 associated with LL5 is only moderately above the corresponding fit achieved by the simple two-factor human capital model (77% versus 58%), which indicates that the two scaled factors combined play a secondary role.

To have another idea on the very large misspecification of the five-factor model, the intercept estimate of 3.77% implies a zero-beta rate of 5.33% per quarter, which translates into a negative equity premium (-0.54% per quarter or -2.16% per year) compared to a sample (annual) equity premium of nearly 13%, as reported above. Hence, if we would take the estimation results of the model at face value, the equity premium puzzle of Mehra and Prescott (1985)—one of the classic findings in asset pricing—would simply not represent a puzzle in the first place. In other words, the five-factor model is strongly rejected on economic grounds.

The results in Panel B of Table 1 show that the five-factor model has a modest performance when $\lambda_0 = 0$, as indicated by the R^2 of only 29% (representing about a third the fit obtained for the benchmark model), an estimate that is largely statistically insignificant according to the bootstrap. This shows the very perverse effect of the intercept in terms of inflating the cross-sectional fit of the conditional model (from 29% to 77%). Moreover, the risk price estimates associated with all factors, except the baseline market factor, remain insignificant (at the 5% level) when we drop the intercept. This represents "bad news" for LL, as both scaled factors, arguably the key factors in the conditional HCAPM, are not priced irrespective of whether the intercept is omitted or not.

We can also see from Table 1 that the risk price estimate associated with the lagged instrument is largely insignificant in both LL3 and LL5, irrespective of whether we include or not the intercept. In other words, the lagged instrument is irrelevant on a separate basis.

The results above show that the conditional model has either a very poor (CAPM) or a rather modest (HCAPM) pricing performance for the 25 portfolios.[17] Moreover, both models are clearly dominated by the benchmark model (FF3). Since the size premium has become insignificant since the 1980s, our results suggest that the two conditional models are not able to explain the value premium anomaly. To inspect this further, we conduct an "accounting" decomposition of the value premium (the average return of the high-minus-low BM portfolio) within each size quintile for LL3, where the risk price estimates are obtained from the zero-intercept specification.

The results in Table 2 (Panel A) show that the risk premium (risk price times the beta of the long-short BM portfolio) associated with the baseline market factor is negative within all five size quintiles, something that only confirms the well known failure of the simple CAPM in explaining the value premium anomaly. In comparison, the risk premium associated with the scaled market factor has the right sign (positive) within the first four size quintiles. However, the respective magnitudes are too small to compensate for the negative risk premia associated with the baseline market factor, and thus, too small to match the original raw value premia (even if added to the smaller positive risk premia attached to *cay*). Consequently, the pricing errors for the value-growth portfolio vary between 1.61% and 2.48% within the first three size quintiles, estimates that are either around or close to the corresponding original value premia. The situation is even worse in the case of the last

[17] We compute another widely used goodness-of-fit metric, $R^{2*} = 1 - \mathrm{E}_N(\widehat{\alpha}_i^2)/\,\mathrm{Var}_N(\mu_i)$, which is based on the cross-sectional second moment (rather than variance) of the pricing errors. This metric accounts for the possibility that the pricing errors have a large magnitude in average, but a small cross-sectional dispersion. The distinction between R^{2*} and the baseline metric (R^2) only arises in the specification with $\lambda_0 = 0$. When there is an intercept, the two metrics coincide, since $\mathrm{E}_N(\widehat{\alpha}_i) = 0$ by construction. Untabulated results show that R^{2*} produces identical estimates to the baseline metric. Specifically, the explanatory ratios associated with LL3 and LL5 are 0.06 and 0.28, respectively. The empirical p-values are also very similar to those associated with the main metric.

size quintile, as both the scaled factor and *cay* risk premia become negative (−0.15% and −0.07%, respectively), leading the model further away (compared to the simple CAPM) in terms of matching the raw value premium of 0.59% per quarter (as indicated by the pricing error of 1.24%). The sole exception to this very negative pattern holds in the case of the fourth size quintile, in which case the risk premium for the scaled factor (1.42%) has an important contribution in terms of producing a small pricing error for the respective value-growth portfolio (−0.21%).

[Table 2 here]

We observe a similar pattern for the scaled market risk premia within the five-factor conditional model (LL5), as illustrated in Panel B of Table 2: With the sole exception of the fourth size quintile, the scaled market risk premia are either very small with respect to the raw value premia to be explained (first three size quintiles) or have the wrong sign (last size quintile). The key difference relative to the three-factor conditional model is that both the original and scaled labor income factors help the model in explaining the value premium, as the corresponding factor risk premia are positive in nine out of 10 cases (two factors times five size quintiles). Nonetheless, the magnitudes of the risk premia associated with the two labor income factors are still modest (in comparison to the raw value premia) across most size quintiles. Moreover, the *cay* risk premium assumes the wrong sign (negative) both in the first and last size quintiles. Consequently, the model does not succeed in explaining the value-growth portfolios within the first three size quintiles (with pricing errors varying between 0.78% and 2.40% per quarter) and also within the last size quintile (pricing error of 0.48% compared to a raw value premium of 0.59%). As in the case of the three-factor model, only within the fourth size quintile does the model do better in terms of matching the value premium, with a pricing error of only −0.49% (compared to a raw value premium of 1.47%). Yet, such a level of mispricing is larger (in magnitude) than the corresponding estimate from the three-factor model (−0.21%).

Obviously, as with any linear asset pricing model, the failure of these two conditional models in explaining the risk premia associated with the 25 portfolios stems from the absence of enough dispersion in betas across the extreme BM portfolios (which are associated with the bulk of the observed value premia). These facts can be confirmed in Figure 1, which plots the factor betas associated with LL3 alongside the portfolio risk premia (to save space, we do not plot the betas for *cay*). We can see that the scaled factor betas go in the "wrong direction" within the last size quintile, while showing little dispersion (with the right direction) within the first three size quintiles. With the exception of the fourth size quintile, such dispersion in betas is too small to offset the systematic wrong pattern in the loadings of the original market factor. Consequently, the three-factor conditional model cannot explain the value premia embedded in the 25 portfolios. Critically, Table 3 shows that nearly all the 25 betas associated with the scaled market factor are statistically insignificant, the sole exception holding for the S4B5 portfolio (with significance at the 5% level). In other words, $R_{vw,t+1}cay_t$ largely represents an "useless factor" (Kan and Zhang , 1999), and a similar pattern holds for cay_t. In comparison, all the 25 betas associated with the baseline market factor are strongly significant (1% level).

[Table 3 here]

[Figure 1 here]

The results in Figure 2 show a similar pattern for the scaled market betas within the conditional HCAPM. Turning to the scaled labor income factor, the pattern in betas (across the extreme BM portfolios) is relatively flat within the first two and last two size quintiles. Hence, the negative (high-minus-low) spreads in betas multiplied by the respective negative risk price are not large enough (in magnitude) to generate sizable positive risk premia for such a scaled factor. On the other hand, the direction of the betas for the baseline labor income factor is always the correct one, yet the respective dispersion is also not large enough to match the corresponding raw value premia. Therefore, the conclusion is the same as that obtained for LL3: With the exception of the fourth size quintile, there is not enough cross-sectional dispersion (with the right direction) in the betas of both scaled factors, as well as in the betas of the baseline labor income factor, to compensate for the large dispersion (in the wrong direction) in betas for the raw market factor. Consequently, the model produces average pricing errors for the value-growth portfolios that, despite being generally smaller than the corresponding mispricing obtained under the three-factor model, are still large relative to the raw value premia (at least within four of the five size quintiles). Importantly, according to Table 3, only seven of the 25 betas for the scaled labor income factor are statistically significant (5% level), while only two portfolios have significant loadings on the scaled market factor

(S4B4 and S4B5) under LL5. This provides another perspective on the (statistical) irrelevance of the scaled factors within the conditional HCAPM.

[Figure 2 here]

3.2 Conditional CCAPM

In this section, we shift attention to the three-factor conditional CCAPM.

The results in Table 4 (Panel A) show that the risk price estimate for $\Delta c_{t+1} cay_t$ is significant (5% level) and the fit is large (70%) and statistically significant. However, this apparent positive pricing performance is entirely illusory. Indeed, the excess zero-beta rate estimate is as large as 2.76% per quarter (with significance at the 1% level) and the corresponding average pricing error as large as 2.8% per quarter, which is above the average raw average risk premia by 23% (MAE/MAR of 1.23). These results point to a strong empirical rejection of the model. In other words, the model's performance is quite negative if we look beyond the cross-sectional R^2 and correctly interpret the estimate of λ_0. In the same vein as the case of LL5 discussed above, this example illustrates well the danger of assessing the performance of linear asset pricing models by relying exclusively on R^2, something that is especially relevant when the intercept is not restricted to zero. The reason is that R^2 is not penalized by large intercepts, which in turn indicate a large model's misspecification. One aspect that represents particularly bad news for LL is that their three-factor consumption model generates an average pricing error clearly above that associated with the baseline CCAPM (2.76% versus 1.58% per quarter), when the pricing error is correctly computed. All of these facts are ignored by LL, as they (incorrectly) include the intercept in the risk premium explained by the model rather than as part of the pricing error, while also failing to report the average pricing error as a fraction of the raw risk premia.[18]

[Table 4 here]

When we restrict the cross-sectional intercept to zero, it turns out that the scaled consumption factor remains priced (at the 1% level), as shown in Panel B of Table 4. However, the model's

[18] See their Table 4. They also report the formal χ^2 specification test on the joint significance of the pricing errors. However, such a statistic is often irrelevant in empirical applications given the poor inversion of the covariance matrix of the pricing errors. For this reason, many studies do not even report such a statistic.

explanatory ratio becomes negative (−4%), which reveals a very poor performance: As discussed in the previous section, a negative R^2 means that the model has a very negative performance, that is, it does worse than a trivial model containing a constant as the sole risk factor.[19] Indeed, this example illustrates well the very perverse effect driven by the inclusion of the intercept, as the fit moves from being negative to 70%. This result also means that the model does not do better than the baseline consumption model (on qualitative terms) when it comes to pricing the 25 portfolios (as both models produce a negative fit). Naturally, these results could not have been inferred by LL, as they do not report results for the zero-intercept specification. Similarly to the cases of the conditional CAPM and conditional HCAPM, it turns out that the risk price estimate for the lagged instrument is insignificant (at the 5% level) irrespective of whether we include or not the intercept, albeit in the latter case there is marginal significance (10% level).

The accounting decomposition of value risk premia in Panel C of Table 2 helps us understand the very poor performance of the three-factor conditional CCAPM. We can see that the risk premium associated with the scaled consumption factor is positive within all size quintiles, thus helping the model in terms of explaining the value premium. However, given the large negative premia associated with the lagged instrument and nearly zero premia for the unscaled consumption factor, it follows that the magnitudes of the scaled factor premia are not large enough to produce small pricing errors for the value-growth portfolios. Indeed, the pricing error for the high-minus-low BM portfolio varies between 0.77% and 2.32% within the first four size quintiles (which compares with raw value premia between 1.47% and 2.46%). On the other hand, the risk premium associated with the scaled factor is too large ("overshooting") in the last size quintile (1.65% compared to a raw value premium of 0.59%), thus producing a somewhat large negative pricing error (−0.34%). As shown in Figure 3, there is not enough dispersion in the betas of the scaled consumption factor across the extreme BM portfolios within most size quintiles. Importantly, as shown in Table 3, none of the 25 betas associated with the scaled consumption factor is significant (at the 5% level), which suggests that such a factor is irrelevant to price these portfolios.

[Figure 3 here]

[19] The estimate of R^{2*} associated with LL3* is −0.08, and thus, very close to the corresponding R^2 counterpart (when there is no intercept in the cross-sectional regression).

To summarize, the punchline from the results above is straightforward: Contrary to LL's claim, their three conditional models based on *cay* do not explain the dispersion in risk premia among the 25 portfolios. If we consider the specification with no intercept and rely on the cross-sectional R^2, it turns out that the pricing performance is either very poor (LL3 or conditional CCAPM) or it is quite modest (LL5). On the other hand, if we consider the specification with an intercept, it turns out that all three conditional models produce a very large mispricing, generating an average pricing error that is around or above the average risk premia that we want to explain in the first place. Moreover, all three models underperform to a very large degree the benchmark Fama–French model.

Therefore, LL's conclusion that "Although the (C)CCAPM can explain a substantial fraction of the cross-sectional variation in these 25 portfolio returns, this result suggests that the conditional models do a poor job of simultaneously pricing the hypothetical zero-beta portfolio" does not provide a minimally correct picture on the performance of their models. Importantly, the very large intercept estimates discussed above cannot be credibly interpreted as valid estimates of the "excess zero-beta rate", as money market frictions cannot (even remotely) explain those extreme magnitudes. Instead, these large intercept estimates indicate a very large misspecification of these models in terms of pricing the 25 portfolios, that is, there are important missing risk factors, which in turn represents a clear empirical failure of LL's models.[20]

3.3 Inspecting the Bootstrap

We conduct Monte-Carlo experiments to assess the validity of the bootstrap simulation associated with the cross-sectional R^2. We employ the 25 portfolios as test assets and rely on FF3. The corresponding cross-sectional regression,

$$\mu_i = \lambda_0 + \lambda_{vw}\widehat{\beta}_{vwi} + \lambda_{SMB}\widehat{\beta}_{SMBi} + \lambda_{HML}\widehat{\beta}_{HMLi} + \alpha_i, i = 1, .., 25, \tag{24}$$

is estimated over the original sample (with and without intercept).

In each of the Monte-Carlo pseudo samples, we implement the bootstrap simulation (with

[20] We note that, even by the standards of the asset pricing literature back in 2001, imposing the restriction on the zero-beta rate was a frequent procedure used in empirical tests of asset pricing models. A self-contained list includes Fama and French (1993, 1996, 1997), Cochrane (1996), Carhart (1997), Grauer (1999), Lewellen (1999), and Davis et al. (2000).

1000 replications) and compute the empirical p-value for the simulated OLS R^2. Based on the distribution of 2000 pseudo bootstrap p-values, we compute the fraction of the 2000 Monte-Carlo draws in which the null, $R^2 = 0$, is rejected at the 5% level (that is, the empirical p-value is below 5%).[21]

The experiment consists of several steps. First, we compute the residuals from the following auxiliary time-series regressions for portfolio i,

$$R_{i,t}^{e} = \mu_i + \varsigma_{i,t}, \tag{25}$$

which decomposes the total excess return into the expected (μ_i) and unexpected excess return ($\varsigma_{i,t}$). Then, we simulate (jointly) the factors and the unexpected returns (with 2000 draws) by using the multivariate normal distribution and the first two sample moments of all the variables. We use the same sample size as the original data ($T = 140$), thus producing pseudo series for these variables, $\left\{R_{vw,t}^{m}, SMB_t^m, HML_t^m, \varsigma_{i,t}^{m}\right\}, i = 1, ..., 25, m = 1, ..., 2000$.

We use different return generating processes when evaluating the power and size of the bootstrap. To assess power, the pseudo return for portfolio i (in each Monte-Carlo draw) is defined as

$$R_{i,t}^{e,m} \equiv \widehat{\lambda}_{vw}\widehat{\beta}_{vwi} + \widehat{\lambda}_{SMB}\widehat{\beta}_{SMBi} + \widehat{\lambda}_{HML}\widehat{\beta}_{HMLi} + \varsigma_{i,t}^{m}, \tag{26}$$

where the risk price estimates $(\widehat{\lambda})$ are obtained from (24). This return generating process is consistent with a scenario in which FF3 is true in the population. We obtain a rejection rate of 92%, irrespective of whether we include or not the cross-sectional intercept, which indicates a high power of the bootstrap.

To assess the size of the bootstrap, the pseudo return for portfolio i is defined as

$$R_{i,t}^{e,m} \equiv \widehat{\lambda}_0 + \widehat{\alpha}_i + \varsigma_{i,t}^{m}, \tag{27}$$

where $\widehat{\lambda}_0$ and $\widehat{\alpha}_i$ denotes respectively the intercept and the pricing error from (24). Naturally, when the cross-sectional regression does not include the intercept, the pseudo return for portfolio

[21] We use the terminology "pseudo bootstrap p-value" to make it clear that the empirical p-value at Monte-Carlo draw m comes from applying the bootstrap simulation to an artificial (rather than the original) sample.

i equals:

$$R^{e,m}_{i,t} \equiv \widehat{\alpha}_i + \varsigma^m_{i,t}. \tag{28}$$

This return generating process is consistent with the null that the factors in FF3 have no pricing power in the population (that is, all risk prices are zero). We obtain a rejection rate (across the 2000 Monte-Carlo draws) of 6.8% and 6.3% in the specifications without and with intercept, respectively. These rejection rates indicate only a slightly over-rejection of the null. Overall, these results show good size and power properties for the bootstrap simulation associated with the sample R^2.

4 Covariance Representation

In this section, we estimate the covariance representation of the conditional factor models (e.g., Eq. (5)), which is identical to the representation based on single-regression betas. This enables us to obtain direct estimates of the SDF coefficients, which represent the covariance prices of risk (γs). Unless the factors in a given model are uncorrelated, we can have a scenario in which the beta prices of risk (λ) are statistically significant, but the corresponding SDF coefficients (γ) are insignificant, and vice versa. In fact, we can even face a situation where the sign differs among the two types of risk price (see Cochrane , 2005 and Kan et al. , 2013 for a discussion).

The estimation of the covariance representation is by first-stage GMM (based on the identity matrix), which is equivalent to an OLS cross-sectional regression of risk premia onto factor covariances (Cochrane , 2005). We use heteroskedasticity-robust standard errors (White , 1980). Importantly, the GMM standard errors relax some of the restrictive assumptions underlying the Shanken's standard errors for the beta risk prices, namely the assumption of independence between the factors and the errors (from the time-series regressions) and the assumption that the model is true (that is, correctly specified). It is important to note that, contrary to the beta models estimated in the previous section, we can compare nested models on statistical terms. For example, if the risk price estimates associated with the two scaled factors within LL5 are both insignificant, we can conclude that such a model is equivalent (from a statistical viewpoint) to the baseline HCAPM.

The risk price estimates for the conditional CAPM and conditional HCAPM are presented in Table 5. To be consistent with the previous section, we report the results for both the restricted and unrestricted zero-beta rate.[22] Since the pricing errors are the same as those from the corresponding beta specifications, we do not report the aggregate performance metrics. The results show that the estimates of γ for both scaled factors have roughly the same sign and the same qualitative significance as the corresponding estimates of λ. In particular, both scaled factors within LL5 do not significantly influence the SDF in the specification with an intercept (even at the 10% level), and the same holds for cay_t. This means that the model that excludes those three factors (i.e., the baseline two-factor HCAPM) is statistically equivalent to the conditional HCAPM. In other words, both scaled factors are irrelevant when the intercept is included in the pricing equation. When the intercept is excluded from the estimation, it follows that the conditional HCAPM is equivalent (in statistical terms) to the three-factor conditional CAPM, as the risk price estimates associated with both labor income factors are largely insignificant. Another relevant finding from Table 5 is that the risk price estimates for cay_t are strongly insignificant, and this pattern holds irrespective of whether the intercept is included or not in the pricing equation. This result is in line with the evidence obtained from the beta representation of these models in Table 1 and confirms the irrelevance of the lagged instrument on an individual basis.

[Table 5 here]

Table 6 presents the risk price estimates for the conditional CCAPM. It turns out that the scaled consumption risk price is statistically significant (5% level) in both specifications, something that is consistent with the corresponding beta risk price estimates in Table 4. However, the risk price estimate for cay_t is strongly insignificant (even at the 10% level) in both cases, which makes LL3* invalid from a statistical viewpoint.

[Table 6 here]

The findings from this section could not have been inferred from the evidence provided by LL: In their Section 5, LL conduct a GMM estimation of the SDF representation of their models (which is equivalent to the covariance representation described above) by using equally-weighted moments (see their Table 8). However, they do not report the estimates (and respective significance) of

[22] Naturally, the estimates of the excess zero-beta rate are numerically equal to those in the beta representations.

the individual SDF coefficients. Instead, they report the joint significance of the SDF coefficient estimates associated with a given model. However, such a statistics is useless when it comes to assessing the statistical significance of the impact of each individual factor into the SDF.

5 Alternative Test Assets

In this section, we employ alternative test portfolios when evaluating the pricing performance of LL's models. Roughly speaking, this analysis can be interpreted as an "out-of-sample" evaluation of the conditional models for cross-sectional equity risk premia. To keep the focus, and given the shortcomings associated with the cross-sectional intercept, as discussed in the previous sections, we restrict the analysis to the zero-intercept specification.

Concretely, we consider decile portfolios sorted on three stock characteristics—book-to-market ratio (BM10), investment or asset growth (INV10, Titman et al. , 2004, Fama and French , 2008, and Cooper et al. , 2008), and operating profitability (OP10, Novy-Marx , 2013 and Fama and French , 2015). The choice of these portfolios is straightforward. First, these portfolio groups correspond to prominent patterns in cross-sectional risk premia left unexplained by the baseline CAPM (designated by "CAPM anomalies"), offering a significant cross-sectional dispersion in equity risk premia. Second, these portfolios are closely connected to the non-market factors in the five-factor model of Fama and French (2015) (FF5 henceforth), which represents one of the most popular models in the literature. The reason is that the underlying equity characteristics are the same for both the factors and portfolios (with the latter being constructed from a finer sorting). We ignore the portfolios sorted on the market value of equity (size), as it is well known that the size premium has declined substantially since the 1980s. We also note that the BM deciles are strongly correlated with the 25 portfolios by construction. Hence, INV10 and OP10 represent the "truly" new portfolios to be priced. In sum, these 30 portfolios can be interpreted as a "minimum cross-section" to judge the pricing performance of LL's models.[23] The monthly data on the new decile portfolio returns and factors are obtained from Kenneth French's data library. We compound the monthly

[23]We note that a positive pricing performance for the three portfolio groups (on a joint basis) does not imply a positive pricing performance for each group of deciles on a stand-alone basis, as discussed in Maio (2024). Yet, we avoid estimating the conditional models on each individual decile group for two main reasons. First, these models do poorly in terms of explaining jointly the 30 portfolios, as discussed below. Second, there is a severe risk of overfitting for the model containing more factors (LL5).

returns and monthly factors to get the quarterly counterparts. To obtain quarterly excess portfolio returns, we subtract the quarterly risk-free rate from the quarterly returns.

The estimation results for the conditional CAPM and conditional HCAPM are displayed in Table 7 (Panel A). The reference model is now FF5 given the well known failure of FF3 in terms of explaining both INV10 and OP10 (e.g., Fama and French , 2015). The corresponding cross-sectional regression is given by

$$\mu_i = \lambda_0 + \lambda_{vw}\widehat{\beta}_{vwi} + \lambda_{SMB}\widehat{\beta}_{SMBi} + \lambda_{HML}\widehat{\beta}_{HMLi} + \lambda_{RMW}\widehat{\beta}_{RMWi} + \lambda_{CMA}\widehat{\beta}_{CMAi} + \alpha_i, i = 1, .., N, \quad (29)$$

where RMW and CMA denote respectively the profitability and investment factors.

[Table 7 here]

We can see that the cross-sectional R^2 is negative for LL3. On the other hand, LL5 delivers a positive explanatory ratio, yet the level of the fit is modest and far below that of FF5 (29% versus 74%), being also strongly insignificant (according to the empirical p-value of 0.28). Turning to the risk prices, it follows that the estimates of λ associated with the scaled factors, including $cay_t R_{vw,t+1}$, are strongly insignificant (even at the 10% level) in all cases. By comparing these results with those associated with the 25 portfolios in Table 1 (Panel B), the evidence against the conditional models looks even stronger.

Table 8 (Panel A) presents the estimation results for the conditional CCAPM. In contrast to the estimation with the 25 portfolios, the R^2 associated with the conditional consumption model appears positive, yet the magnitude is very modest (14%) and the estimate is strongly insignificant in statistical terms. On the other hand, the risk price estimate associated with the scaled consumption factor is strongly insignificant, something that is clearly at odds with the largely significant estimate reported in Panel B of Table 4.

[Table 8 here]

Similarly to the baseline empirical setting in Section 3, we report the covariance risk prices obtained from the new test portfolios. Table 9 (Panel A) reports the risk price estimates for the conditional CAPM and conditional HCAPM models, whereas Table 10 (Panel A) does the same for the conditional CCAPM. We can see that the estimates of γ associated with all scaled factors in these models are insignificant at the 5% level, something that is in line with the evidence for the beta prices of risk discussed above. Again, this represents a more negative scenario as in the corresponding estimations with the 25 portfolios (displayed in Panel B of Tables 5 and 6), thus indicating that the scaled factors are irrelevant on statistical grounds: A given conditional model is equivalent to a nested model that excludes the scaled factor(s).

[Table 9 here]

[Table 10 here]

Next, we inspect the capacity of the conditional models in terms of pricing the same 30 portfolios, but using an extended sample that ends in 2013, which corresponds to the sample period employed in Fama and French (2015). The construction of the factors for the "out-of-sample period" is as follows. The *cay* series, obtained from Martin Lettau's website, corresponds to the latest version estimated over the sample period from 1952:Q1 to 2019:Q3 and is used for the period from 1963:Q3 to 2013:Q3.[24] The log growth in real aggregate per capita consumption (Δc) and labor income (Δy) is constructed in the same way as described in Section 3.[25] All data used to compute consumption and labor income growth are obtained from the NIPA, published by the Bureau of Economic Analysis. The construction of both the quarterly factors and quarterly portfolio excess returns is identical to the procedure employed for the early sample period.

The estimation results for the scaled CAPM and scaled HCAPM are displayed in Table 7 (Panel B). At a first glance, the picture looks more favorable for LL3, as the corresponding explanatory ratio becomes positive. However, the level of fit is quite modest for both LL3 and LL5 (8% and 28%, respectively) and we cannot reject the null of no explanatory power for both models, as the empirical p-values are quite large. Indeed, the main novelty relative to the early sample estimation (with the same portfolios) is that the risk prices for the scaled market factor are now significantly negative in both conditional models. As shown in Panel B of Table 9, the same pattern holds for

[24] This version of *cay* is constructed using Personal Consumption Expenditures (PCE), whereas the version used in the original study is based on nondurables and services. See Martin Lettau's website for further details. Naturally, for our purposes, *cay* is demeaned over the sample period from 1963:Q3 to 2013:Q3.

[25] Note that the series for consumption and labor income used to construct the factors differ from those used in estimating *cay*. Specifically, the consumption factor is based on nondurables and services, whereas *cay* is constructed using PCE consumption. Labor income for estimating *cay* is defined as wages and salaries plus transfer payments plus other labor income minus personal contributions for social insurance minus taxes.

the corresponding covariance risk price estimates. The negative scaled market risk prices represent a key difference relative to the corresponding positive estimates documented in Panel B of Tables 1 and 5 and raises obvious concerns about the economic plausibility of the conditional CAPM and conditional HCAPM when forced to explain the 30 portfolios.

Panel B in Table 8 displays the estimation results for the conditional CCAPM. The picture is even worse than in the estimation over the short sample, as the model produces a very negative explanatory ratio. Moreover, the estimate of λ for the scaled consumption factor remains strongly insignificant (even at the 10% level), and the same holds for the respective covariance risk price, as shown in Panel B of Table 10. Thus, similarly to the baseline sample period, we cannot reject the null that LL3* is statistically equivalent to a nested model that excludes the scaled consumption factor.

Finally, one relevant finding is that both the beta and covariance risk price estimates associated with cay_t are insignificant in most estimations of the conditional models with the 30 portfolios. The sole exception arises for LL5 over the early sample period, where the estimate of γ is significantly negative (5% level). This implies that the scaled models are not correctly specified from a statistical viewpoint.

Overall, according to the evidence in this section, the conditional models from LL are not able to explain portfolios sorted on investment and operating profitability (in addition to the BM portfolios). If something, the picture looks even more adverse than in the estimation with the baseline 25 portfolios.

6 Risk Price Identification

In this section, we inspect whether the risk prices in LL's models are identified. Burnside (2011) and Kleibergen and Zhan (2015, 2020), among others, argue that the risk price estimates, and associated pricing errors, are unreliable when the risk prices are not identified. This happens when the $N \times K$ factor loading matrix ($\boldsymbol{\beta}$) does not have full column rank (considering the case with $\lambda_0 = 0$), where K denotes the number of factors. In the specification with an intercept, the condition for identified risk prices is that the matrix containing the factor loadings and a vector of ones ($\boldsymbol{\beta}^* \equiv [\mathbf{1} \vdots \boldsymbol{\beta}]$) has full column rank.

We employ the rank tests on the factor loadings proposed by Kleibergen and Zhan (2020). The null hypothesis is that there is a reduced rank for either $\boldsymbol{\beta}$ or $\boldsymbol{\beta}^*$, which means that the risk prices are not identified. When $\lambda_0 = 0$, the statistic is distributed as $F(N+1-K, T-N-1)$. When there is an intercept, the statistic is distributed as $F(N-K, T-N)$. Full details are available in Kleibergen and Zhan (2020), including their internet appendix.

The p-values associated with the rank tests are displayed in Table 11. We consider four different setups. The first two columns correspond to the estimations in which the test assets are the 25 portfolios (with and without intercept). The last two columns correspond to the estimations in which the test assets are the 30 portfolios (over both the early and more recent sample periods).

[Table 11 here]

Overall, the evidence is very negative for LL's models. Indeed, in none of the 12 estimations of their conditional models (three models and four estimations for each model), is the null of a reduced rank (in $\boldsymbol{\beta}$ or $\boldsymbol{\beta}^*$) rejected (at the 5% level). These results generalize the evidence offered in Kleibergen and Zhan (2020) exclusively for LL3* when estimated with an intercept over the 25 portfolios (see their Section II). Importantly, the rank-test results are consistent with the evidence provided in Table 3 above showing that most factor loadings in these three models (in particular for the scaled factors) are statistically insignificant.

For comparison purposes, the null is strongly rejected (1% level) for the CAPM as well as the reference multifactor models (FF3 and FF5) in most cases. The sole exception holds for FF3 when estimated with an intercept over the 25 portfolios, in which case the rejection is at the 5% level. This example also shows that including the intercept in the cross-sectional regression makes it more difficult to pass the identification test (and thus, makes it more likely that the risk price estimates are not credible). The reason is that it is no longer enough that the betas associated with a given factor are jointly different from zero, as there must exist relevant dispersion in betas (among the test assets) for the rank condition to be satisfied.[26]

The punch line of this section is clear: The risk price estimates (and pricing errors) associated with LL's conditional models, which have been the focus of the analyses conducted in the previous

[26]Often, the violation of the rank condition under $\lambda_0 \neq 0$ represents a situation where the cross-sectional regression exhibits (near) multicollinearity: The betas associated with a given factor are approximately constant, thus being proportional to a vector of ones.

sections, are questionable (at best) in most cases, including the estimations conducted over the more relevant cross-section of 30 portfolios. This further reinforces the negative outlook on the empirical performance and credibility of LL's models.[27]

7 Monte-Carlo Experiments

In this section, we conduct Monte-Carlo experiments to illustrate the relationship between model misspecification and the cross-sectional intercept.[28]

We use the 25 portfolios as the test assets and either FF3 or the CAPM as the reference model. In all three experiments, we simulate (jointly) the factors and the unexpected portfolio returns in Eq. (25) (with 10000 draws) by using the multivariate normal distribution and the first two sample moments of the factors and residual returns. We use the same sample size as the original data (over the early sample period), thus producing the following pseudo series:

$$\left\{R_{vw,t}^m, SMB_t^m, HML_t^m, \varsigma_{i,t}^m\right\},$$
$$\left\{R_{vw,t}^m, \varsigma_{i,t}^{*m}\right\}, m = 1, ..., 10000.$$

Armed with the artificial data, we estimate the models (by using the two-step approach) and obtain a pseudo intercept estimate for each of the 10000 draws, $\left\{\widehat{\lambda}_0^m\right\}, m = 1, ..., 10000$. Finally, we compute the average pseudo intercept estimate across the 10000 replications:

$$\mathrm{E}^*(\widehat{\lambda}_0) \equiv \frac{1}{10000} \sum_{m=1}^{10000} \widehat{\lambda}_0^m. \tag{30}$$

The first two experiments are based on the CAPM. The corresponding cross-sectional regression is given by

$$\mu_i = \lambda_0 + \lambda_{vw} \widehat{\beta}_{vwi} + \alpha_i, i = 1, .., 25, \tag{31}$$

which is estimated over the original sample.

First, we assume that the CAPM is the true model in the population. Thus, the pseudo return

[27] There are concerns about the low power of the F-tests (see Kroencke , 2025 and Kleibergen and Zhan , 2025 for a discussion). However, we note that the p-values associated with LL's conditional models are very large (above 10%) in most cases.

[28] We thank an anonymous referee for suggesting this analysis.

for portfolio i (in each Monte-Carlo draw) is defined as

$$R_{i,t}^{e,m} \equiv \widehat{\lambda}_{vw}\widehat{\beta}_{vwi} + \varsigma_{i,t}^{*m}, \tag{32}$$

where the beta and risk price estimates are obtained from the original sample. It follows that the average CAPM intercept estimate (across the 10000 replications) is around zero (−0.02%).

In the second experiment, we assume that the CAPM is misspecified, thus implying the following return generating process,

$$R_{i,t}^{e,m} \equiv \widehat{\lambda}_0 + \widehat{\alpha}_i + \varsigma_{i,t}^{*m}, \tag{33}$$

where the intercept and pricing error estimates are also obtained from the original sample. In this case, the average pseudo intercept estimate is as large as 2.62% per quarter, as a result of the CAPM being misspecified in the population.

The last experiment is based on FF3, with cross-sectional regression given in Eq. (24). We assume that FF3 holds in the population, with the return generating process given in Eq. (26). The intercepts associated with both FF3 and the CAPM are estimated for each of the 10000 replications. We obtain average intercept estimates of 0.62% and 2.32% for FF3 and CAPM, respectively. In other words, the misspecified model (CAPM) produces an average intercept that is nearly four times as large as the corresponding estimate for the "true" model.

Overall, these Monte-Carlo experiments suggest that model misspecification has an important role in terms of driving large cross-sectional intercept estimates.

8 Out-of-Sample Asset Pricing Tests

In this section, we evaluate the out-of-sample pricing performance of LL's models over the time-series dimension.[29]

Following Simin (2008) and Cooper et al. (2021), we form one-quarter ahead forecasts of portfolio excess returns using only information available at the forecast date, and compare these forecasts with realized excess returns next quarter. A key ingredient in our analysis is in employing a real-time estimate of cay_t: At each forecast date t, we re-estimate the dynamic least squares

[29] We thank an anonymous referee for suggesting this analysis.

specification of Lettau and Ludvigson (2001) on a 20-year rolling window (ending at t) to construct the trend deviation from consumption, asset wealth, and labor income. In addition to the baseline 25 portfolios, we consider the 30 portfolios employed in Section 5 as test assets. The whole sample period spans 1963:Q4–2013:Q4. To keep the focus, and given the discussion in the previous sections, this analysis is restricted to the specification without an intercept.

The first 20 years in the sample (1963:Q4–1983:Q3) serve as the initial training window for in-sample estimation of factor betas and risk prices.[30] The first one-quarter ahead portfolio excess return forecast is computed for 1983:Q4. The forecasts for the remaining 120 quarters (1984:Q1–2013:Q4) are constructed by using a 20-year rolling window (i.e., moving one quarter ahead at a time), based on the in-sample betas and risk prices estimated with data available up to each forecast date (i.e., the end of each window). At the end, this procedure yields 121 equity premium forecasts (from 1983:Q4 to 2013:Q4) for each model and test portfolio.

The estimation procedure for the in-sample factor betas and risk prices is as follows. Taking the three-factor scaled CAPM (LL3) as an example, for each portfolio i and for each quarter t, the in-sample betas are estimated by running the following time-series regression over the rolling window:

$$R^e_{i,\tau} = \delta_i + \beta_{zi} z_{\tau-1} + \beta_{vwi} R_{vw,\tau} + \beta_{vwzi} R_{vw,\tau} z_{\tau-1} + \varepsilon_{i,\tau}, \quad \tau \in \{t-79, \ldots, t\}. \tag{34}$$

Next, the in-sample factor risk prices are estimated by running an OLS cross-sectional regression of the (time-series) average excess portfolio returns on the in-sample betas ($\widetilde{\beta}$),

$$\widetilde{\mu}_i = \lambda_z \widetilde{\beta}_{zi} + \lambda_{vw} \widetilde{\beta}_{vwi} + \lambda_{vwz} \widetilde{\beta}_{vwzi} + \alpha_i, \tag{35}$$

where $\widetilde{\mu}_i$ denotes the sample risk premium (for asset i) computed over the rolling window. The estimates $\widetilde{\lambda}$ represent the in-sample factor risk prices (for the window ending at t). Finally, the one-quarter ahead (i.e., for $t+1$) predicted portfolio risk premium is computed as the sum of the risk premia (products of in-sample betas and risk prices) across all factors,

$$\widetilde{R}^e_{i,t+1} \equiv \widetilde{\lambda}_z \widetilde{\beta}_{zi} + \widetilde{\lambda}_{vw} \widetilde{\beta}_{vwi} + \widetilde{\lambda}_{vwz} \widetilde{\beta}_{vwzi} \tag{36}$$

[30]The window used to estimate *cay* matches the window used to obtain in-sample betas and risk prices.

which relies exclusively on information up to quarter t.

Following Cooper et al. (2021), we define the out-of-sample one-period ahead forecasting error as

$$\varpi_{i,t+1} \equiv R^e_{i,t+1} - \widetilde{R}^e_{i,t+1}, \tag{37}$$

i.e., the realized excess returns minus the fitted risk premia implied by the model (estimated at t). Iterating this process forward (from $t+1$ to $T-1$) and averaging over the evaluation (or out-of-sample) window yields,

$$\widehat{\mathrm{E}}(R^e_{i,s}) = \widehat{\mathrm{E}}(\widetilde{R}^e_{i,s}) + \widehat{\mathrm{E}}(\varpi_{i,s}), \tag{38}$$

where $\widehat{\mathrm{E}}(.)$ denotes a time-series sample mean computed over $s = t+1, ..., T$.

The out-of-sample cross-sectional R^2, which is analogous to the in-sample cross-sectional R^2, is given by

$$R^2_{OS} = 1 - \frac{\mathrm{Var}_N[\widehat{\mathrm{E}}(\varpi_{i,s})]}{\mathrm{Var}_N[\widehat{\mathrm{E}}(R^e_{i,s})]}. \tag{39}$$

Untabulated results show that the estimates of R^2_{OS} for FF3, LL3, LL5, and LL3* in the estimation with SBM25 are 0.23, 0.07, 0.40, and −0.09, respectively. On the other hand, when the test assets are the 30 portfolios the estimates associated with FF5, LL3, LL5, and LL3* are 0.66, −0.24, 0.37, and −0.31, respectively. Thus, apart from the case of LL5 in the estimation with the 25 portfolios, it follows that the conditional models substantially underperform the respective reference model (either FF3 or FF5). Indeed, the scaled models produce a negative explanatory ratio in three of the six cases (LL3 in the estimation with the 30 portfolios and LL3* in both estimations), evidence that is not very different from the corresponding in-sample fit documented in Sections 3 and 5.

The out-of-sample analysis has the advantage of imposing that the data used in the estimation and evaluation of a given model do not overlap. Nonetheless, these results should be interpreted with some caution. This is because the one-period realized excess return is a very noisy proxy for "true" risk premia, implying that the out-of-sample forecasting errors defined in Eq. (37) represent a poor proxy for the model's pricing errors. Consistent with these concerns, we see that the out-of-sample fit produced by FF3 is much smaller than the in-sample counterpart reported in Table 1 (23% versus 80%), albeit we obtain more similar (out-of-sample versus in-sample) levels of fit in the case of FF5 (66% versus 74%).

9 Conclusion

We revisit LL's empirical evidence by conducting a replication of the cross-sectional tests conducted in their study. Our conclusions are straightforward: Contrary to LL's claim, their three conditional models based on *cay* do not explain the dispersion in risk premia among the 25 size-value portfolios. Their pricing performance is either very poor (for the scaled CAPM and the scaled CCAPM) or it is quite modest (for the scaled HCAPM). Moreover, all three conditional models underperform to a very large degree the benchmark Fama–French model.

When the zero-beta rate is unrestricted, we obtain average absolute pricing errors that are actually around or above the average raw risk premium (among the 25 portfolios) for all three conditional models, with estimates varying between 95% and 168% (thus indicating a very negative performance). Furthermore, the SDF coefficient estimate associated with the scaled labor income factor is strongly insignificant.

To summarize, LL's misleading conclusions stem from a combination of (i) incorrect empirical choices—e.g., not reporting the results when the excess zero-beta rate is restricted to zero and not reporting the SDF coefficient estimates (and respective statistical significance)—and (ii) incorrect interpretation of their results—not including the intercept estimates as part of the model's pricing errors and downplaying some of the adverse empirical evidence they actually report. Those problems reinforce each other in producing LL's wrong conclusions regarding the pricing ability of their conditional models.

Employing alternative test assets, namely portfolios sorted on investment and operating profitability (over both the original sample period and an extended period), and evaluating the identification of LL's risk prices reinforces the negative outlook on their models to a substantial degree. Additionally, LL's models perform poorly in an out-of-sample evaluation, substantially underperforming the Fama–French models in most cases.

Naturally, the evidence in this paper does not preclude LL's conditional models from explaining other dimensions of the cross-section of equity risk premia (that is, other CAPM anomalies beyond the baseline book-to-market effect or the profitability and investment portfolios) or other asset classes (e.g., bond risk premia). This is left for future research.

Lastly, our critique does not question in any way the historical importance of LL from a methodological viewpoint. Indeed, LL represents one of the first studies that conducted cross-sectional tests of linear conditional asset pricing models by specifying a pricing kernel with time-varying coefficients (as being affine in a lagged instrument), thus emphasizing the importance of scaled risk factors for asset pricing. Therefore, their study significantly helped to establish a vast subsequent literature.[31]

[31] Cochrane (1996) represents another early example of a study that specifies the SDF coefficients as being affine in a conditioning variable. Yet, he estimates the SDF (rather than the beta) representation of his models by GMM (on the cross-section of stock returns). Ferson and Harvey (1999) and Lewellen (1999) represent two other examples of early studies that relied on scaled factors. Yet, unlike LL, they do not employ the two-step regression approach when testing the beta representation of their models, using instead the time-series approach (given that their factors are traded). Moreover, the scaled factors in these two studies arise from assuming time-varying betas rather than time-varying risk prices.

References

Brennan, M. J., A. W. Wang, and Y. Xia. 2004. "Estimation and Test of a Simple Model of Intertemporal Capital Asset Pricing." *Journal of Finance.* 59: 1743–1775.

Burnside, C. 2011. "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk: Comment." *American Economic Review.* 101: 3456–3476.

Carhart, M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance.* 52: 57–82.

Cochrane, J. H. 1996. "A Cross-Sectional Test of an Investment-Based Asset Pricing Model." *Journal of Political Economy.* 104: 572–621.

Cochrane, J. H. 2005. *Asset Pricing* (Princeton University Press, Princeton, NJ).

Cooper, M. J., H. Gulen, and M. J. Schill. 2008. "Asset Growth and the Cross-Section of Stock Returns." *Journal of Finance.* 63: 1609–1651.

Cooper, I., L. Ma, P. Maio, and D. Philip. 2021. "Multifactor Models and Their Consistency with the APT." *Review of Asset Pricing Studies.* 11: 402–444.

Davis, J. L., E. F. Fama, and K. R. French. 2000. "Characteristics, Covariances, and Average Returns: 1929 to 1997." *Journal of Finance.* 55: 389–406.

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

Fama, E. F., and K. R. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance.* 51: 55–84.

Fama, E. F., and K. R. French. 1997. "Industry Costs of Equity." *Journal of Financial Economics.* 43: 153–193.

Fama, E. F., and K. R. French. 2008. "Dissecting Anomalies." *Journal of Finance.* 63: 1653–1678.

Fama, E. F., and K. R. French. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics.* 116: 1–22.

Fama, E. F., and J. D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy.* 81: 607–636.

Ferson, W. E., and C. R. Harvey. 1999. "Conditioning Variables and the Cross Section of Stock Returns." *Journal of Finance.* 54: 1325–1360.

Gagliardini, P., E. Ossola, and O. Scaillet. 2019. "A Diagnostic Criterion for Approximate Factor Structure." *Journal of Econometrics.* 212: 503–521.

Grauer, R.R. 1999. "On the cross-Sectional Relation between Expected Returns, Betas, and Size." *Journal of Finance.* 54: 773–789.

Guo, H., and P. Maio. 2020. "ICAPM and the Accruals Anomaly." *Quarterly Journal of Finance.* 10: 2050014.

Kan, R., and C. Zhang. 1999. "Two-Pass Tests of Asset Pricing Models with Useless Factors." *Journal of Finance.* 54: 203–235.

Kan, R., C. Robotti, and J. Shanken. 2013. "Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology." *Journal of Finance.* 68: 2617–2649.

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

Kleibergen, F., and Z. Zhan. 2015. "Unexplained Factors and Their Effects on Second Pass R-Squared's." *Journal of Econometrics.* 189: 101–116.

Kleibergen, F., and Z. Zhan. 2020. "Robust Inference for Consumption-Based Asset Pricing." *Journal of Finance.* 75: 507–550.

Kleibergen, F., and Z. Zhan. 2025. "A Powerful Test Needs to Be Size-Correct: Response to "Robust Inference for Consumption-Based Asset Pricing with Power"." *Critical Finance Review.* 14: 179–185.

Kroencke, T. A. 2025. "Robust Inference for Consumption-Based Asset Pricing with Power." *Critical Finance Review.* 14: 129–178.

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

Lewellen, J. 1999. "The Time-Series Relations among Expected Return, Risk, and Book-to-Market." *Journal of Financial Economics.* 54: 5–43.

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

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

Maio, P. 2024. "Is Long-Run Risk Really Priced? Revisiting Liu and Matthies (2022)." *Journal of Finance.* 79: 2885–2900.

Novy-Marx, R. 2013. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics.* 108: 1–28.

Mehra, R., and E. C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics.* 15: 145–161.

Raponi, V., C. Robotti, and P. Zaffaroni. 2020. "Testing Beta-Pricing Models Using Large Cross-Sections." *Review of Financial Studies.* 33: 2796–2842.

Shanken, J. 1992. "On the Estimation of Beta Pricing Models." *Review of Financial Studies.* 5: 1–34.

Simin, T. 2008. "The Poor Predictive Performance of Asset Pricing Models." *Journal of Financial and Quantitative Analysis.* 43: 355–380.

Titman, S., K. J. Wei, and F. Xie. 2004. "Capital Investments and Stock Returns." *Journal of Financial and Quantitative Analysis.* 39: 677–700.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica.* 48: 817–838.

Table 1: Pricing Performance for Conditional CAPM and HCAPM

This table presents the estimation and evaluation results for the conditional CAPM and conditional human-capital CAPM. The estimation procedure is the two-pass regression approach. The test portfolios are the 25 size/book-to-market sorted portfolios. R_{vw}, Δy, SMB, and HML represents the excess market return, labor income growth, size, and value factor, respectively. $cay_t R_{vw,t+1}$ and $cay_t \Delta y_{t+1}$ denotes respectively the scaled market and scaled labor income growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the risk price estimates, while the second line reports t-statistics based on Shanken's standard errors (in parenthesis). Panel A and B presents the results from estimating the cross-sectional regression with and without an intercept (λ_0), respectively. R^2 denotes the cross-sectional OLS R^2, with the respective empirical p-value obtained from a bootstrap simulation reported in parentheses. $MAE(\%)$ is the average absolute pricing error (in %). MAE/MAR represents the ratio of MAE to MAR, where MAR denotes the mean absolute raw portfolio risk premium. The sample period is from 1963:Q4 to 1998:Q3. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	$\lambda_0(\%)$	cay_t	$R_{vw,t+1}$	Δy_{t+1}	SMB_{t+1}	HML_{t+1}	$cay_t R_{vw,t+1}$	$cay_t \Delta y_{t+1}$	R^2	$MAE(\%)$	$\frac{MAE}{MAR}$
					Panel A (Intercept)						
CAPM	2.62		−0.32						0.01	2.62	1.17
	(2.78**)		(−0.27)						(0.80)		
HCAPM	3.22		−1.10	1.26					0.58	3.22	1.43
	(1.89)		(−0.61)	(1.92)					(0.06)		
FF3	0.30		1.33		0.47	1.46			0.80	0.38	0.17
	(0.19)		(0.78)		(0.94)	(3.22**)			(0.01)		
LL3	2.14	−0.05	−0.06				0.11		0.31	2.14	0.95
	(1.51)	(−0.15)	(−0.03)				(2.47*)		(0.47)		
LL5	3.77	−0.44	−1.99	0.56			0.03	−0.02	0.77	3.77	1.68
	(2.38*)	(−0.98)	(−1.17)	(1.29)			(1.06)	(−1.45)	(0.02)		
					Panel B (No Intercept)						
CAPM			1.84						−0.41	0.70	0.31
			(2.49*)						(0.99)		
HCAPM			1.61	0.99					−0.04	0.61	0.27
			(2.14*)	(1.75)					(0.72)		
FF3			1.62		0.47	1.46			0.80	0.22	0.10
			(2.32*)		(0.94)	(3.22**)			(0.01)		
LL3		−0.50	1.74				0.11		0.07	0.53	0.23
		(−0.97)	(2.23*)				(2.11*)		(0.72)		
LL5		−0.45	1.72	0.94			0.08	−0.00	0.29	0.46	0.20
		(−0.90)	(2.38*)	(1.72)			(1.83)	(−0.21)	(0.55)		

Table 2: Accounting of Value Risk Premia

This table reports a decomposition of risk premia for the value-minus-growth portfolio within each size quintile associated with each of the three conditional factor models. cay, R_{vw}, Δy, and Δc represents the risk premium (beta times risk price) associated with the cay, excess market return, labor income growth, and consumption growth factor, respectively. $cayR_{vw}$, $cay\Delta_y$, and $cay\Delta c$ represents the risk premium for the scaled market factor, scaled labor income factor, and scaled consumption factor, respectively. $\overline{R}$ denotes the average realized excess return for a given value-growth portfolio, $\mathrm{E}(R)$ represents the respective total risk premium (the sum of the factor risk premia), while $\alpha = \overline{R} - \mathrm{E}(R)$ represents the respective pricing error. All the values are presented in percentage points. The sample period is from 1963:Q4 to 1998:Q3. S1 (S5) represents the first (last) size quintile.

Panel A (Conditional CAPM)								
	$\overline{R}$	cay	R_{vw}	$cayR_{vw}$			$\mathrm{E}(R)$	α
S1	2.46	0.13	−0.57	0.41			−0.02	2.48
S2	1.95	0.35	−0.63	0.62			0.35	1.61
S3	1.82	0.32	−0.58	0.33			0.07	1.75
S4	1.47	0.50	−0.24	1.42			1.68	−0.21
S5	0.59	−0.07	−0.43	−0.15			−0.65	1.24
Panel B (Conditional HCAPM)								
	$\overline{R}$	cay	R_{vw}	Δ_y	$cayR_{vw}$	$cay\Delta_y$	$E(R)$	α
S1	2.46	−0.20	−0.57	0.41	0.31	0.12	0.06	2.40
S2	1.95	0.19	−0.63	1.06	0.52	0.04	1.18	0.78
S3	1.82	0.44	−0.56	0.63	0.29	−0.07	0.74	1.09
S4	1.47	0.43	−0.24	0.68	1.08	0.00	1.95	−0.49
S5	0.59	−0.51	−0.43	0.98	−0.08	0.16	0.11	0.48
Panel C (Conditional CCAPM)								
	$\overline{R}$	cay	Δc	$cay\Delta c$			$E(R)$	α
S1	2.46	−1.14	−0.04	1.32			0.14	2.32
S2	1.95	−1.70	0.04	2.01			0.35	1.61
S3	1.82	−1.69	0.08	2.23			0.62	1.20
S4	1.47	−1.39	0.14	1.95			0.70	0.77
S5	0.59	−0.84	0.12	1.65			0.93	−0.34

Table 3: Factor Loadings

This table reports the beta estimates associated with the factors in the scaled factor models. The test portfolios are 25 size/book-to-market portfolios. cay, R_{vw}, Δy, and Δc denotes the cay, excess market return, labor income growth, and consumption growth factor, respectively. $cay_t R_{vw,t+1}$ and $cay_t \Delta y_{t+1}$ denotes respectively the scaled market and scaled labor income growth factor. $cay_t \Delta c_{t+1}$ denotes the scaled consumption growth factor. The conditioning variable is the lagged consumption-wealth ratio (cay_t). The sample period is from 1963:Q4 to 1998:Q3. Betas marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

	Conditional CAPM			**Conditional HCAPM**					**Conditional CCAPM**		
Portfolio	cay_t	$R_{vw,t+1}$	$cay_t R_{vw,t+1}$	cay_t	$R_{vw,t+1}$	Δy_{t+1}	$cay_t R_{vw,t+1}$	$cay_t \Delta y_{t+1}$	cay_t	Δc_{t+1}	$cay_t \Delta c_{t+1}$
S1B1	0.16	1.61**	1.94	1.58	1.60**	0.57	1.81	−84.41	4.28	6.81*	−55.35
S1B2	0.24	1.43**	2.52	2.35**	1.42**	0.65	2.16	−126.92*	3.69	6.76**	12.73
S1B3	0.12	1.33**	2.94	2.18*	1.31**	0.74	2.68	−123.54*	3.25	5.57*	22.39
S1B4	0.12	1.26**	0.56	1.80*	1.24**	0.79	0.50	−99.53	2.99	5.96*	41.39
S1B5	−0.10	1.29**	5.78	2.03	1.27**	1.00	5.69	−126.57	2.68	6.24*	103.86
S2B1	0.25	1.55**	0.52	1.55	1.54**	0.17	0.12	−78.90	4.58**	4.84	−167.93
S2B2	0.11	1.36**	−2.24	2.11**	1.35**	0.92	−2.33	−118.36*	3.28	4.08	−9.56
S2B3	0.07	1.25**	−0.19	1.45*	1.24**	0.36	−0.47	−83.24	3.10*	4.37*	−20.61
S2B4	0.20	1.14**	2.68	1.67**	1.13**	0.96	2.85	−85.67*	2.69	3.98	59.55
S2B5	−0.46	1.19**	6.32	1.12	1.18**	1.30	6.70	−91.97	2.19	5.28*	75.06
S3B1	0.37	1.43**	−3.36	1.38*	1.43**	0.28	−3.56	−60.65	4.56**	3.16	−231.80
S3B2	0.68	1.22**	−0.05	1.82**	1.21**	0.15	−0.40	−68.86*	3.67**	3.22	−43.26
S3B3	0.23	1.11**	−0.01	1.36**	1.10**	0.39	−0.17	−67.89*	2.66*	3.37	34.21
S3B4	0.22	1.04**	4.53	1.47**	1.03**	0.51	4.42	−74.59**	2.76**	3.03	−9.60
S3B5	−0.27	1.10**	−0.29	0.39	1.10**	0.95	0.20	−36.12	2.18	4.16	38.03
S4B1	0.55	1.26**	−2.95	1.37**	1.26**	−0.09	−3.35	−50.30	4.25**	2.42	−212.54
S4B2	0.21	1.17**	−3.55	1.03**	1.16**	0.29	−3.65	−48.86	3.31**	3.07	−112.21
S4B3	−0.12	1.06**	2.52	0.54	1.05**	0.28	2.47	−39.56	2.60**	2.40	−59.62
S4B4	−0.19	1.03**	7.30	0.60	1.02**	0.55	7.41*	−46.13	2.35*	2.96	−3.74
S4B5	−0.46	1.12**	10.31*	0.41	1.12**	0.63	10.45*	−50.34	2.29	4.24	23.08
S5B1	−0.58*	1.05**	2.79	−1.33**	1.06**	−1.12**	2.21	40.78	2.80**	2.06	−222.01
S5B2	0.14	0.96**	1.71	0.29	0.96**	−0.47	1.27	−11.22	2.56**	1.61	−51.76
S5B3	0.31	0.78**	−2.75	−0.45	0.79**	−0.16	−2.56	46.08	2.49**	2.78*	−92.51
S5B4	−0.14	0.84**	0.78	−0.11	0.84**	−0.07	0.72	−2.08	2.27**	1.69	−114.70
S5B5	−0.43	0.81**	1.37	−0.18	0.80**	−0.08	1.20	−15.64	1.62*	3.53**	−22.31

Table 4: Pricing Performance for Conditional CCAPM

This table presents the estimation and evaluation results for the conditional consumption-CAPM. The estimation procedure is the two-pass regression approach. The test portfolios are the 25 size/book-to-market sorted portfolios. Δc represents the consumption growth factor. $cay_t \Delta c_{t+1}$ denotes the scaled consumption growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the risk price estimates, while the second line reports t-statistics based on Shanken's standard errors (in parenthesis). Panel A and B presents the results from estimating the cross-sectional regression with and without an intercept (λ_0), respectively. R^2 denotes the cross-sectional OLS R^2, with the respective empirical p-value obtained from a bootstrap simulation reported in parentheses. $MAE(\%)$ is the average absolute pricing error (in %). MAE/MAR represents the ratio of MAE to MAR, where MAR denotes the mean absolute raw portfolio risk premium. The sample period is from 1963:Q4 to 1998:Q3. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	$\lambda_0(\%)$	cay_t	Δc_{t+1}	$cay_t \Delta c_{t+1}$	R^2	$MAE(\%)$	$\frac{MAE}{MAR}$
			Panel A (Intercept)				
CCAPM	1.57		0.22		0.16	1.58	0.70
	(2.19*)		(1.16)		(0.27)		
LL3*	2.76	−0.13	0.02	0.01	0.70	2.76	1.23
	(2.66**)	(−0.31)	(0.15)	(2.25*)	(0.02)		
			Panel B (No Intercept)				
CCAPM			0.64		−0.47	0.70	0.31
			(1.52)		(0.84)		
LL3*		0.71	0.08	0.01	−0.04	0.59	0.26
		(1.90)	(0.49)	(3.01**)	(0.70)		

Table 5: Covariance Representation for Conditional CAPM and HCAPM

This table presents the estimation results for the covariance representation of the conditional CAPM and conditional human-capital CAPM. The estimation procedure is first-stage GMM with equally-weighted moment conditions. The test portfolios are 25 size/book-to-market sorted portfolios. R_{vw}, Δy, SMB, and HML represents the excess market return, labor income growth, size, and value factor, respectively. $cay_t R_{vw,t+1}$ and $cay_t \Delta y_{t+1}$ denotes respectively the scaled market and scaled labor income growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the covariance risk price estimates, while the second line reports t-statistics (in parenthesis). Panel A and B presents the results from estimating the model with and without an intercept (γ_0), respectively. The sample period is from 1963:Q4 to 1998:Q3. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	γ_0(%)	cay_t	$R_{vw,t+1}$	Δy_{t+1}	SMB_{t+1}	HML_{t+1}	$cay_t R_{vw,t+1}$	$cay_t \Delta y_{t+1}$
Panel A (Intercept)								
CAPM	2.62		−0.47					
	(2.77**)		(−0.27)					
HCAPM	3.22		−0.95	176.48				
	(1.93)		(−0.28)	(2.05*)				
FF3	0.30		4.36		−0.71	8.15		
	(0.19)		(1.09)		(−0.23)	(3.13**)		
LL3	2.14	−26.84	1.28				1062.13	
	(1.56)	(−0.62)	(0.36)				(2.26*)	
LL5	3.77	136.72	−2.45	94.86			385.08	−10446.25
	(2.46*)	(1.28)	(−0.79)	(1.18)			(1.19)	(−1.20)
Panel B (No Intercept)								
CAPM			2.75					
			(2.31*)					
HCAPM			2.96	139.75				
			(1.62)	(1.99*)				
FF3			5.05		−1.15	8.52		
			(3.43**)		(−0.66)	(4.78**)		
LL3		−73.36	5.86				1108.52	
		(−1.30)	(2.08*)				(2.18*)	
LL5		−116.06	6.01	133.74			961.39	3814.53
		(−0.96)	(2.12*)	(1.91)			(2.05*)	(0.51)

Table 6: Covariance Representation for Conditional CCAPM

This table presents the estimation results for the covariance representation of the conditional consumption-CAPM. The estimation procedure is first-stage GMM with equally-weighted moment conditions. The test portfolios are 25 size/book-to-market sorted portfolios. Δc represents the consumption growth factor. $cay_t \Delta c_{t+1}$ denotes the scaled consumption growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the covariance risk price estimates, while the second line reports t-statistics (in parenthesis). Panel A and B presents the results from estimating the model with and without an intercept (γ_0), respectively. The sample period is from 1963:Q4 to 1998:Q3. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	γ_0(%)	cay_t	Δc_{t+1}	$cay_t \Delta c_{t+1}$
Panel A (Intercept)				
CCAPM	1.57		98.13	
	(1.92)		(1.12)	
LL3*	2.76	−79.21	27.61	16772.23
	(3.07**)	(−1.59)	(0.28)	(2.00*)
Panel B (No Intercept)				
CCAPM			292.54	
			(1.32)	
LL3*		−0.62	70.50	14551.27
		(−0.01)	(0.85)	(2.13*)

Table 7: Pricing Performance for Conditional CAPM and HCAPM: Alternative Test Assets

This table presents the estimation and evaluation results for the conditional CAPM and conditional human-capital CAPM. The estimation procedure is the two-pass regression approach. The test portfolios are deciles sorted on book-to-market ratio, investment, and operating profitability (for a total of 30 portfolios). R_{vw}, Δy, SMB, HML, RMW, and CMA represents the excess market return, labor income growth, size, value, profitability, and investment factor, respectively. $cay_t R_{vw,t+1}$ and $cay_t \Delta y_{t+1}$ denotes respectively the scaled market and scaled labor income growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the risk price estimates, while the second line reports t-statistics based on Shanken's standard errors (in parenthesis). R^2 denotes the cross-sectional OLS R^2, with the respective empirical p-value obtained from a bootstrap simulation reported in parentheses. $MAE(\%)$ is the average absolute pricing error (in %). MAE/MAR represents the ratio of MAE to MAR, where MAR denotes the mean absolute raw portfolio risk premium. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	cay_t	$R_{vw,t+1}$	Δy_{t+1}	SMB_{t+1}	HML_{t+1}	RMW_{t+1}	CMA_{t+1}	$cay_t R_{vw,t+1}$	$cay_t \Delta y_{t+1}$	R^2	$MAE(\%)$	$\frac{MAE}{MAR}$
Panel A (1963:Q4–1998:Q3)												
CAPM		1.72								−0.37	0.39	0.22
		(2.47*)								(0.99)		
HCAPM		1.84	0.59							−0.15	0.35	0.20
		(2.59**)	(1.40)							(0.95)		
FF5		1.66		0.92	1.05	0.53	0.48			0.74	0.18	0.10
		(2.39*)		(1.41)	(2.27*)	(1.81)	(1.35)			(0.01)		
LL3	0.46	1.67						0.08		−0.27	0.38	0.21
	(1.41)	(2.39*)						(1.57)		(0.97)		
LL5	−0.17	2.02	1.19					0.02	0.01	0.29	0.29	0.16
	(−0.35)	(2.68**)	(1.40)					(0.30)	(1.13)	(0.28)		
Panel B (1963:Q4–2013:Q4)												
CAPM		1.76								−0.58	0.38	0.21
		(2.87**)								(1.00)		
HCAPM		1.76	0.37							−0.51	0.37	0.21
		(2.87**)	(1.66)							(1.00)		
FF5		1.56		0.92	0.99	0.68	0.55			0.73	0.16	0.09
		(2.55*)		(1.60)	(2.25*)	(2.10*)	(1.71)			(0.01)		
LL3	0.21	1.60						−0.16		0.08	0.27	0.15
	(0.31)	(2.61**)						(−2.81**)		(0.63)		
LL5	0.49	1.61	0.36					−0.12	0.01	0.28	0.24	0.13
	(0.62)	(2.63**)	(1.17)					(−2.25*)	(1.35)	(0.36)		

Table 8: Pricing Performance for Conditional CCAPM: Alternative Test Assets

This table presents the estimation and evaluation results for the conditional consumption-CAPM. The estimation procedure is the two-pass regression approach. The test portfolios are deciles sorted on book-to-market ratio, investment, and operating profitability (for a total of 30 portfolios). Δc represents the consumption growth factor. $cay_t \Delta c_{t+1}$ denotes the scaled consumption growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the risk price estimates, while the second line reports t-statistics based on Shanken's standard errors (in parenthesis). R^2 denotes the cross-sectional OLS R^2, with the respective empirical p-value obtained from a bootstrap simulation reported in parentheses. $MAE(\%)$ is the average absolute pricing error (in %). MAE/MAR represents the ratio of MAE to MAR, where MAR denotes the mean absolute raw portfolio risk premium. t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	cay_t	Δc_{t+1}	$cay_t \Delta c_{t+1}$	R^2	$MAE(\%)$	$\frac{MAE}{MAR}$
			Panel A (1963:Q4–1998:Q3)			
CCAPM		0.77		−0.31	0.43	0.24
		(1.36)		(0.76)		
LL3*	0.41	0.36	0.00	0.14	0.33	0.19
	(1.19)	(1.89)	(1.43)	(0.41)		
			Panel B (1963:Q4–2013:Q4)			
CCAPM		0.60		−0.47	0.37	0.21
		(1.77)		(0.88)		
LL3*	0.89	0.50	0.01	−0.30	0.37	0.21
	(1.10)	(1.68)	(0.94)	(0.94)		

Table 9: Covariance Representation for Conditional CAPM and HCAPM: Alternative Test Assets

This table presents the estimation results for the covariance representation of the conditional CAPM and conditional human-capital CAPM. The estimation procedure is first-stage GMM with equally-weighted moment conditions. The test portfolios are deciles sorted on book-to-market ratio, investment, and operating profitability (for a total of 30 portfolios). R_{vw}, Δy, SMB, HML, RMW, and CMA represents the excess market return, labor income growth, size, value, profitability, and investment factor, respectively. $cay_t R_{vw,t+1}$ and $cay_t \Delta y_{t+1}$ denotes respectively the scaled market and scaled labor income growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the covariance risk price estimates, while the second line reports t-statistics (in parenthesis). t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	cay_t	$R_{vw,t+1}$	Δy_{t+1}	SMB_{t+1}	HML_{t+1}	RMW_{t+1}	CMA_{t+1}	$cay_t R_{vw,t+1}$	$cay_t \Delta y_{t+1}$
Panel A (1963:Q4–1998:Q3)									
CAPM		2.57							
		(2.22^{*})							
HCAPM		3.09	84.05						
		(2.15^{*})	(1.54)						
FF5		4.83		2.57	7.12	17.81	9.57		
		(2.79^{**})		(1.01)	(2.05^{*})	(3.66^{**})	(1.78)		
LL3	20.23	1.85						666.28	
	(0.65)	(0.85)						(1.52)	
LL5	−235.93	6.09	152.16					463.38	13460.92
	(-2.00^{*})	(2.29^{*})	(1.51)					(0.73)	(1.80)
Panel B (1963:Q4–2013:Q4)									
CAPM		2.36							
		(2.62^{**})							
HCAPM		2.50	66.51						
		(1.97^{*})	(1.31)						
FF5		3.72		1.78	1.71	6.31	5.11		
		(2.83^{**})		(0.80)	(0.69)	(2.67^{**})	(1.44)		
LL3	11.37	1.45						−525.18	
	(0.51)	(0.93)						(-3.09^{**})	
LL5	−12.43	1.19	97.43					−383.55	5425.04
	(−0.28)	(0.62)	(1.16)					(-2.59^{**})	(1.06)

Table 10: Covariance Representation for Conditional CCAPM: Alternative Test Assets

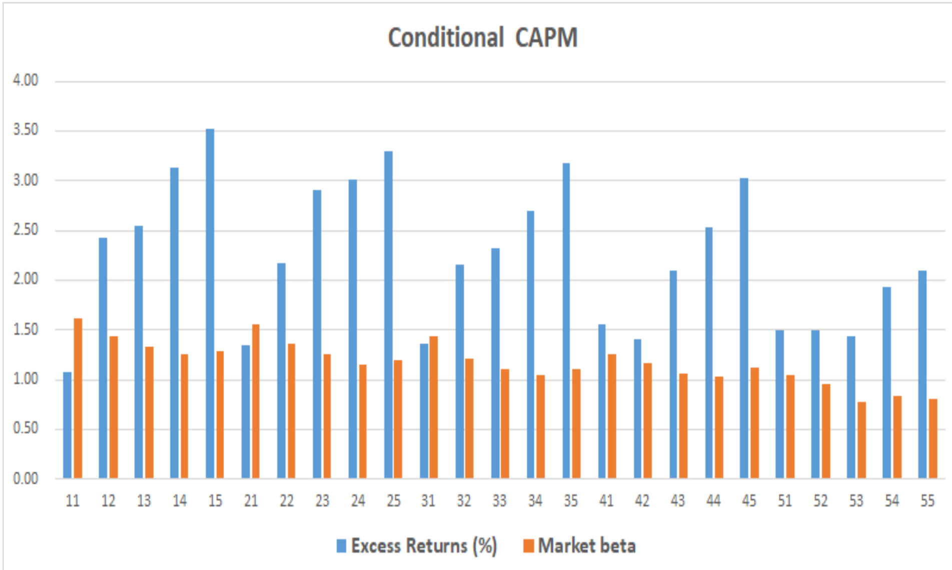
This table presents the estimation results for the covariance representation of the conditional consumption-CAPM. The estimation procedure is first-stage GMM with equally-weighted moment conditions. The test portfolios are deciles sorted on book-to-market ratio, investment, and operating profitability (for a total of 30 portfolios). $cay_t \Delta c_{t+1}$ denotes the scaled consumption growth factor, where the conditioning variable is the lagged consumption-wealth ratio (cay_t). The first line associated with each model contains the covariance risk price estimates, while the second line reports t-statistics (in parenthesis). t-ratios marked with * and ** denote statistical significance at the 5% and 1% level, respectively.

Model	cay_t	Δc_{t+1}	$cay_t \Delta c_{t+1}$
Panel A (1963:Q4–1998:Q3)			
CCAPM		349.00	
		(1.20)	
LL3*	18.99	182.05	5524.54
	(0.37)	(2.08*)	(0.80)
Panel B (1963:Q4–2013:Q4)			
CCAPM		290.45	
		(1.47)	
LL3*	5.39	275.69	4654.17
	(0.15)	(1.42)	(0.76)

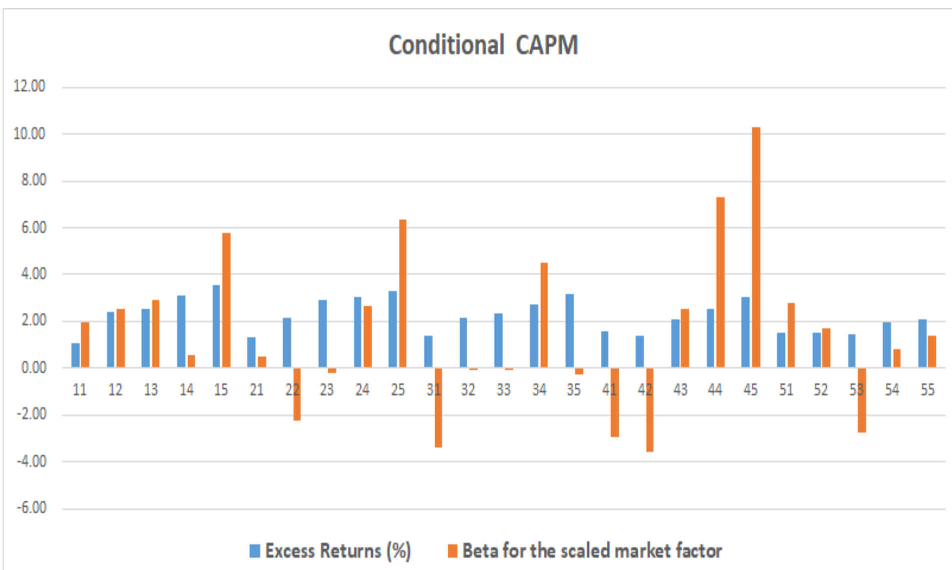
Table 11: Rank Tests on Factor Loadings

This table presents the p-values associated with the F-test on the factor loading estimates. For a description of the different models, see Tables 1 and 4. The test portfolios are the 25 size/book-to-market portfolios (SBM25) and deciles sorted on book-to-market ratio, investment, and operating profitability (BM10+INV10+OP10). "SBM25 (int.)" refers to the case in which the cross-sectional regression includes an intercept. In this case, the null hypothesis is that the matrix containing a vector of ones and the factor loadings has a reduced rank. In the remaining three columns, the intercept is excluded from the cross-sectional regression, which means that the null hypothesis is that the matrix containing the factor loadings has a reduced rank. The sample period is from 1963:Q4 to 1998:Q3, except in the last column where the sample is 1963:Q4 to 2013:Q4.

	SBM25 (int.)	SBM25	BM10+INV10+OP10	BM10+INV10+OP10 (1963–2013)
Panel A (Conditional CAPM and HCAPM)				
CAPM	0.00	0.00	0.00	0.00
HCAPM	0.54	0.56	0.52	0.77
FF3	0.03	0.00		
FF5			0.00	0.00
LL3	0.13	0.16	0.89	0.57
LL5	0.95	0.96	0.99	0.91
Panel B (Conditional CCAPM)				
CCAPM	0.15	0.12	0.31	0.39
LL3*	0.27	0.10	0.25	0.74



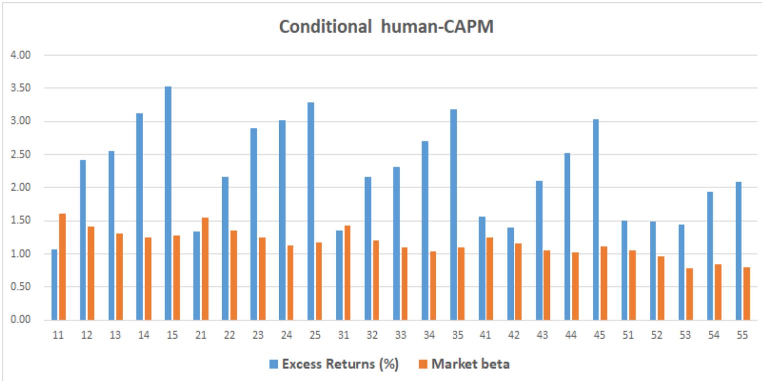
Panel A: Market Factor



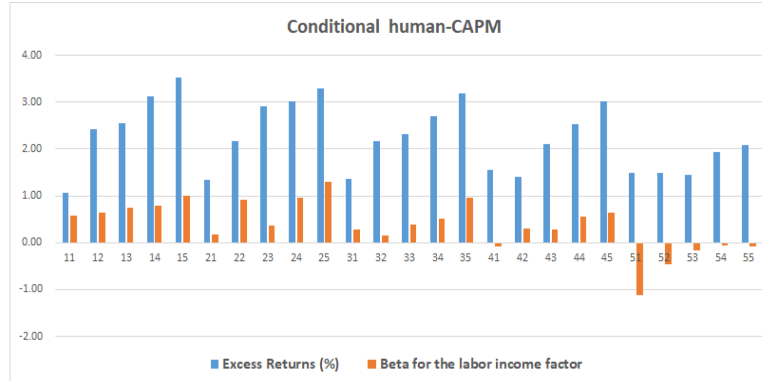
Panel B: Scaled Market Factor

Figure 1: Factor Betas: Conditional CAPM

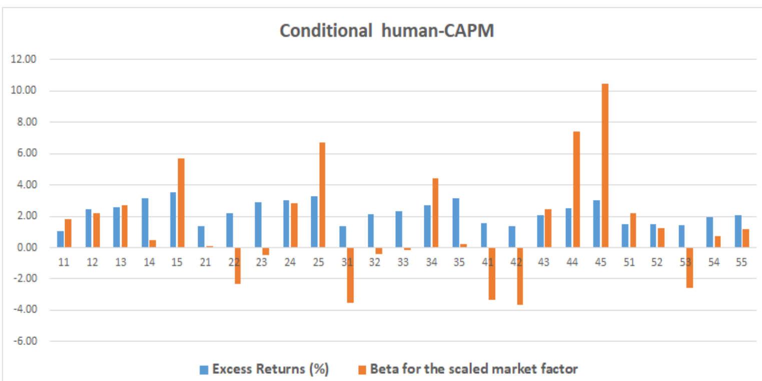
This figure plots the beta estimates for the factors in the Conditional CAPM, along with the average excess returns of the 25 size/book-to-market portfolios. Each portfolio is represented by a two-digit number, where the first digit refers to the size quintile (1-5, smallest to largest), and the second digit refers to the book-to-market quintile (1-5, lowest to highest). The sample period is from 1963:Q4 to 1998:Q3.



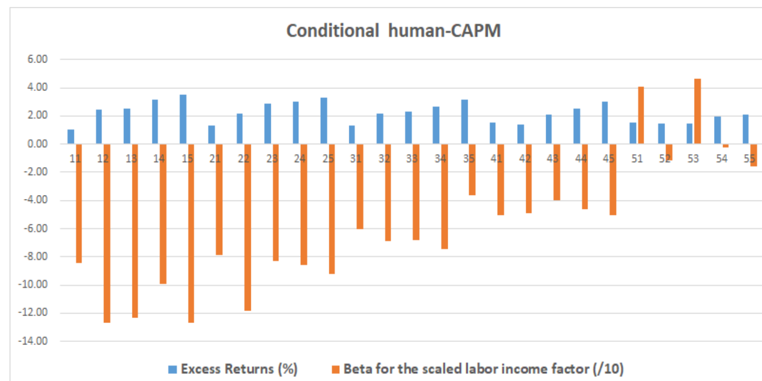
Panel A: Market Factor



Panel B: Labor Income Factor



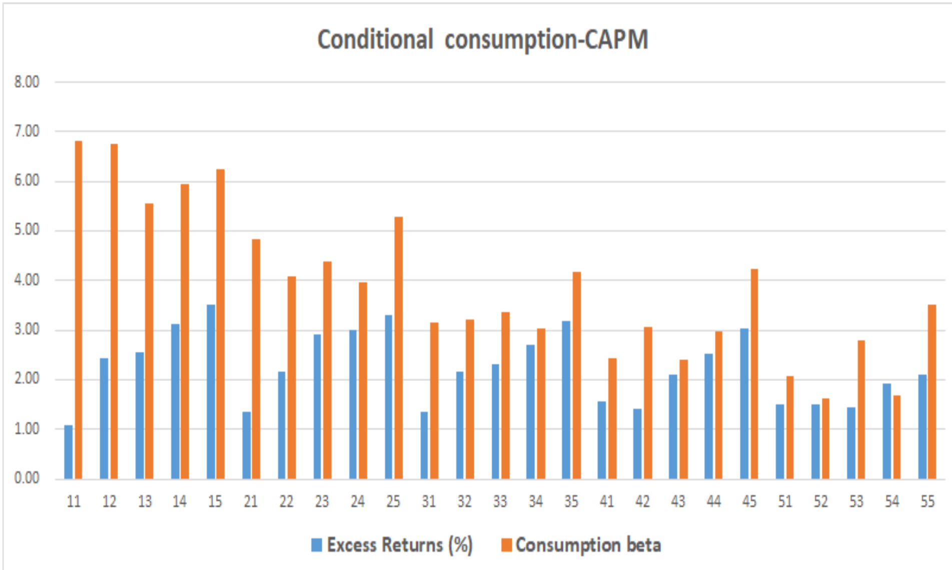
Panel C: Scaled Market Factor



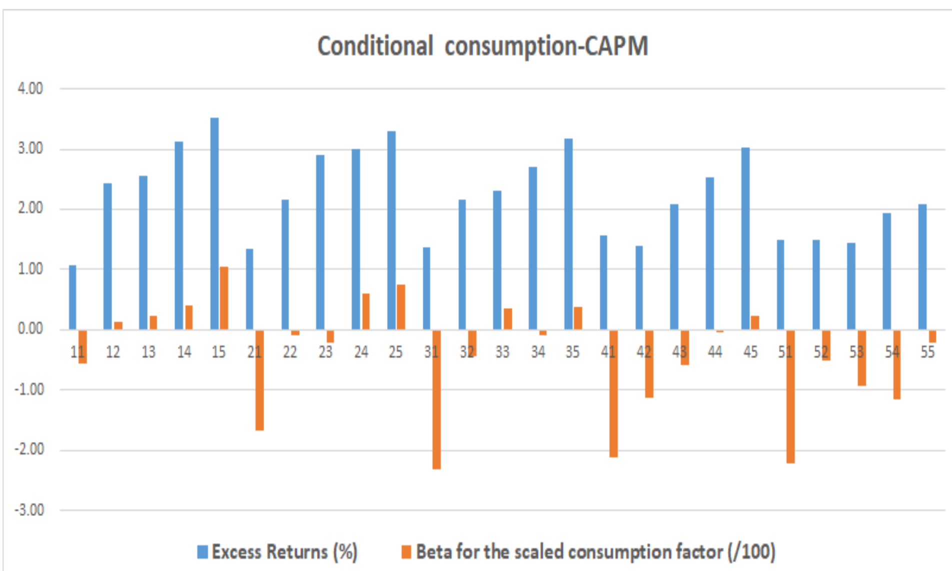
Panel D: Scaled Labor Income Factor

Figure 2: Factor Betas: Conditional HCAPM

This figure plots the beta estimates for the factors in the Conditional human capital-CAPM, along with the average excess returns of the 25 size/book-to-market portfolios. Each portfolio is represented by a two-digit number, where the first digit refers to the size quintile (1-5, smallest to largest), and the second digit refers to the book-to-market quintile (1-5, lowest to highest). The sample period is from 1963:Q4 to 1998:Q3.



Panel A: Consumption Factor



Panel B: Scaled Consumption Factor

Figure 3: Factor Betas: Conditional CCAPM

This figure plots the beta estimates for the factors in the Conditional consumption-CAPM, along with the average excess returns of the 25 size/book-to-market portfolios. Each portfolio is represented by a two-digit number, where the first digit refers to the size quintile (1-5, smallest to largest), and the second digit refers to the book-to-market quintile (1-5, lowest to highest). The sample period is from 1963:Q4 to 1998:Q3.

A Online Appendix

A.1 Bootstrap Simulation

The bootstrap simulation used to assess the statistical significance of the cross-sectional OLS R^2 consists of the following steps. We use the simple two-factor HCAPM to illustrate the procedure. Moreover, we consider the case in which the intercept is included in the cross-sectional regression. Naturally, the specification with no intercept holds by setting $\lambda_0 = 0$.

1. First, we estimate the time-series regressions to obtain the factor loadings,

$$R^e_{i,t} = \delta_i + \beta_{vwi} R_{vw,t} + \beta_{\Delta yi} \Delta y_t + \varepsilon_{i,t},$$

   and in a second step, the expected return-beta representation is estimated by an OLS cross-sectional regression,

$$\mu_i = \lambda_0 + \lambda_{vw} \widehat{\beta}_{vwi} + \lambda_{\Delta y} \widehat{\beta}_{\Delta yi} + \alpha_i.$$

   We compute and save the sample explanatory ratio (R^2), the intercept estimate ($\widehat{\lambda}_0$), and the series of the pricing errors ($\widehat{\alpha}_i$).

   We also run the following auxiliary time-series regressions,

$$R^e_{i,t} = \mu_i + \varsigma_{i,t},$$

   and save the series of the corresponding residuals, $\varsigma_{i,t}$.

2. In each replication $b = 1, ..., 10000$, we construct pseudo-samples for the time-series residuals of each test asset and for each factor (of size T) by drawing with replacement,

$$\{\varsigma^b_{i,t}\}, i = 1, ..., N,$$
$$\{R^b_{vw,t}, \Delta y^b_t\}, t = s^b_1, s^b_2, ..., s^b_T,$$

   where the time indices $s^b_1, s^b_2, ..., s^b_T$ are created randomly from the original time sequence $1, ..., T$. Notice that the time sequence is the same for all variables, thus we allow for a possible correlation between the test returns and the factors.

3. For each replication, the pseudo asset excess returns are constructed by imposing the null

that the factor risk prices are zero:

$$R^{e,b}_{i,t} \equiv \widehat{\lambda}_0 + \widehat{\alpha}_i + \varsigma^b_{i,t}.$$

4. In each replication, we estimate the CAPM by the two-step procedure, but using the artificial data rather than the original data:

$$\begin{aligned} R^{e,b}_{i,t} &= \delta^b_i + \beta^b_{vwi} R^b_{vw,t} + \beta^b_{\Delta yi} \Delta y^b_t + \varepsilon^b_{i,t}, \\ \mu^b_i &= \lambda^b_0 + \lambda^b_{vw} \widehat{\beta}^b_{vwi} + \lambda^b_{\Delta y} \widehat{\beta}^b_{\Delta yi} + \alpha^b_i. \end{aligned}$$

   We compute and save the pseudo cross-sectional R^2, leading to an empirical distribution of this statistic, $R^2_b, b = 1, ..., 10000$.

5. Finally, the empirical p-value associated with R^2 is computed as

$$p(R^2) = \#\left\{R^2_b \geq R^2\right\}/10000,$$

   where $\#\left\{R^2_b \geq R^2\right\}$ denotes the number of pseudo samples in which the condition, $R^2_b \geq R^2$, is true.