One Factor to Bind the Cross-Section of Returns*

Nicola Borri, Denis Chetverikov, Yukun Liu, and Aleh Tsyvinski

April 11, 2024

Abstract

We propose a new non-linear single-factor asset pricing model $r_{it} = h(f_t \lambda_i) + \epsilon_{it}$. Despite its parsimony, this model represents exactly any non-linear model with an arbitrary number of factors and loadings – a consequence of the Kolmogorov-Arnold representation theorem. It features only one pricing component $h(f_t \lambda_i)$, comprising a nonparametric link function of the time-dependent factor and factor loading that we jointly estimate with sieve-based estimators. Using 171 assets across major classes, our model delivers superior cross-sectional performance with a low-dimensional approximation of the link function. Most known finance and macro factors become insignificant controlling for our single-factor.

JEL Codes: G10, G12, C10

Keywords: asset returns, non-linear factor model, Kolmogorov-Arnold, factor zoo

*Nicola Borri is with Luiss University, Rome. Denis Chetverikov is with the University of California, Los Angeles. Yukun Liu is with the University of Rochester, Simon Business School. Aleh Tsyvinski is with Yale University. We thank Yacine Ait-Sahalia, Bryan Kelly, Leonid Kogan, Ralph Koijen, Toby Moskowitz and Semyon Malamud for comments. We thank Andrei Voronin for excellent research assistance.

1 Introduction

Factor models play an important role in finance and economics. The existing literature, however, focuses primarily on linear factor models. In this paper, we propose a new non-linear single factor asset pricing model

$$r_{it} = h(f_t \lambda_i) + \varepsilon_{it}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T, \tag{1}$$

where $r_{it} \in \mathbb{R}$ is the return of asset i at time period t, $f_t > 0$ is the single factor, $\lambda_i > 0$ is the factor loading, h is a continuous function, and $\varepsilon_{it} \in \mathbb{R}$ is an idiosyncratic mean-zero component. Here, the function h, the factor values f_t, and the factor loading values λ_i are all assumed to be unknown, and both N and T are assumed to be large. We refer to our model as the HFL model.

Our model is different from the typical linear factor model. On the one hand, our model is seemingly more restrictive because it allows for only one time-dependent factor and requires both the time factor and the factor loading to be strictly positive, whereas the linear factor model allows for an arbitrary number of factors and allows all factors and factor loadings to take negative values. On the other hand, our model is more flexible because it includes a nonparametric link function h, whereas the linear factor model sets the link function to be identity.

Even though our model has only one factor, the presence of the nonparametric link function h makes our model surprisingly flexible. In particular, using the Kolmogorov-Arnold representation theorem (Kolmogorov (1956); Arnold (1957)), we show that any nonparametric factor model with an arbitrary number of factors and with arbitrary interactions between factors and factor loadings can be reduced to model (1). That is, any arbitrary nonparametric model can be exactly represented by our model. It is thus without loss of generality to focus on our model with just one factor.

We apply our model to characterize risk premia in a large cross-section of assets across major asset classes, and propose a sieve-based least squares estimator of the model. This estimator is defined as a solution to an optimization problem minimizing the sum of squared residuals for the model not only over T potential factor values $\{f_t\}_{t=1}^T$ and N potential factor loading values $\{\lambda_i\}_{i=1}^N$ but also over potential values of the function h in a sieve space, which we choose to consist of all polynomials of a given order that is slowly increasing in the sample dimensions (N,T).

The Kolmogorov-Arnold representation ensures that the function h must be continuous under very mild regularity conditions, and so can be arbitrarily well approximated by a suitable sieve space for estimation purposes. At the same time, readers familiar with the Kolmogorov-Arnold representation theorem might argue that this theorem implies that even though the function h must be continuous, it may nonetheless be rather irregular, and one may need a high-dimensional sieve space in order to approximate it sufficiently well, making estimation of this model difficult. However, whether the function h can be well approximated by a relatively low-dimensional sieve space for the purposes of a particular application, making estimation feasible, is ultimately an empirical question. Our main empirical result is that, for the cross-section of asset returns across major asset classes, the estimated version of our model based on a relatively low-dimensional sieve space outperforms linear factor models with multiple factors in a number of dimensions.

We test the cross-sectional asset pricing predictions of our model using as test assets U.S. equity portfolios, U.S. and international government bonds, commodities, and foreign currency portfolios, for a total of 171 assets. Specifically, we estimate a cross-sectional regression of the average asset realized excess returns on the average asset excess returns predicted by the model and test several implications of our one factor model, such as that the intercept is zero and the significance of the slope coefficient. Furthermore, we evaluate the fit of the model. Our empirical strategy follows the classic two-step Fama and MacBeth (1973)'s procedure. The first step of the Fama and MacBeth (1973)'s procedure, which requires the estimation of the asset specific betas in time-series regressions, is implicit in the algorithm to estimate our single factor model. The second step of the procedure, which requires a cross-sectional regression of average asset excess returns on factor betas, corresponds to our cross-sectional regression.

Our single factor model explains a large fraction of cross-sectional differences in asset returns. We find that a relatively low-dimensional sieve space, with the polynomials of degree as low as four, suffices for the majority of our empirical results. When using all asset classes as test assets, the slope coefficient on the pricing component is significantly different from zero. Furthermore, we cannot reject the null that the slope coefficient is equal to 1, as implied by our model. Moreover, the pricing error, measured by the intercept in the cross-sectional regression, is not significantly different from zero. The HFL model exhibits a cross-sectional regression adjusted R-squared value of 89%.

Next, we compare the HFL model with a number of prominent cross-sectional asset pricing models that have been proposed in the finance and economics literature. First, we consider benchmark equity factor models, such as the CAPM, the Fama and French (1993)'s three factor model, the Fama and French (2015)'s five factor model and the five factor model augmented with the momentum factor of Jegadeesh and Titman (1993). Second, we consider models using factors based on principal component analysis (PCA). Specifically, we consider three estimators of latent factors based on principal components: the standard estimator based on the covariance matrix of asset returns; the RP-PCA estimator, which is constructed to better account for the cross-section of average asset returns; and the kernel PCA estimator, which is a nonlinear form of PCA. Third, we consider higher-order factor models, which account for non-linearities by including higher order factors, such as the squared equity market return of Harvey and Siddique (2000). Specifically, we consider models including the square and cube of equity and PCA factors. Fourth, we consider recent macro-factor models, such as the intermediary capital risk factor model of He, Kelly, and Manela (2017), the downside equity market risk factor model of Lettau, Maggiori, and Weber (2014), and the liquidity factor model of Pástor and Stambaugh (2003). Fifth, we consider the macro factors of Ludvigson and Ng (2009). Our main result is that, once we include our single factor in cross-sectional regressions, the risk prices associated with the alternative factors are all indistinguishable from zero, with the exception of the size factor in Fama and French (1993).

Furthermore, we compare the HFL model with a large set of factors proposed in the last decades to account for asset risk premia, which is commonly referred to as the factor zoo (see, e.g., Feng, Giglio, and Xiu (2020)). We consider 153 established factors, collected by Jensen, Kelly, and Pedersen (2023), and for each of these factors, we first estimate a standard cross-sectional regression of average asset excess returns on the asset factor betas. The betas are measured as slope coefficients in time-series regressions of each asset excess return on the factor. We find that, for a large number of these factors, the risk price, i.e., the slope coefficient in the cross-sectional regression, is statistically significant, although also the estimates for the intercept, a measure of the pricing error of the model, are mostly significantly different from zero. We then repeat the estimation additionally including the predicted values from the HFL model. The results are striking. We find a significant factor risk price estimate for only 3 out of the 153 factors in the factor zoo, while the slope estimates associated with the HFL factor are all statistically significant. Moreover, the pricing error of these models are mostly indistinguishable from zero. Furthermore, we conduct cross-sectional regressions employing both the HFL factor and a diverse set of factors from the factor zoo, employing the double-selection Lasso method. This approach enables simultaneous control over all factors and provides a robust model selection technique for assessing factor contributions in cross-sectional asset pricing regressions (Feng, Giglio, and Xiu (2020)). We find that the HFL component is highly significant, after controlling for the factor zoo, while the majority of the factors in the factor zoo is not.

We then investigate the predictive ability of the HFL model by building portfolios using as signals the predicted asset returns. Specifically, we first estimate the model on all assets and the first 120 months of data. Next, we use the model predicted values to sort assets in five portfolios, and compute the equally-weighted average portfolio returns in the following six months. Finally, we repeat the procedure expanding the estimation window by 6 months each time until the end of the sample. This empirical methodology is designed to capture the strategy of an investor who uses the HFL model to predict future returns rebalancing at a semi-annual frequency. In each period t, we sort assets only using information available up to period t and compute excess returns between t and $t+1$,where each period contains 6 months. The first portfolio groups the assets with, each period, the lowest predicted excess returns by the HFL model. The last portfolio groups the assets with, each period, the highest predicted excess returns by the HFL model. Across the portfolios, we document a sizable, and statistically significant, cross-section of monthly average returns, ranging from 0.08% for the first portfolio to 0.77% for the last one. A zero-cost long-short strategy, wherein assets with the highest predicted returns are bought while those with the lowest predicted returns are sold short, yields a monthly average excess return of 0.69%, with a monthly Sharpe ratio of 0.15. We investigate the risk-adjusted performance of the HFL strategy, by regressing the portfolio returns on the Fama and French (1993)'s three factors, on the Fama and French (2015)'s five factors, and on the five factors augmented by the momentum factor of Jegadeesh and Titman (1993), respectively. The estimates of the intercept from these regressions, a measure of risk-adjusted performance, increase from the first to the last portfolio in all three specifications, and are significant and quantitatively large for portfolio 3 to 5, as well as for the long-short portfolio.

Our empirical results extend to alternative samples. First, we show that the HFL component is statistically significant for each individual asset class using the HFL component estimated with all asset classes. Second, we repeat the cross-sectional asset pricing tests using the first and second half of the sample and find results consistent with those obtained using the full sample. Third, we consider a higher frequency time-variation in the asset pricing test using an estimation based on an expanding window. In this case, we find that the slope coefficient associated with our single factor is significantly different from zero for all months, and we can never reject the null hypothesis that the slope coefficient equals to one as implied by our model. Fourth, we augment the set of test assets with 10 equity momentum portfolios and the momentum factor and report that the slope coefficient associated with the HFL component is statistically different from zero.

Our paper relates to a vast literature on factor models in asset pricing. The majority of the literature focuses on linear factor models, such as characteristic-based models (e.g., Fama and French (1993, 1996); Carhart (1997); Koijen, Lustig, and Van Nieuwerburgh (2017)), statistical-based models (e.g., Connor and Korajczyk (1986)), and macro models (e.g., Adrian, Etula, and Muir (2014); He, Kelly, and Manela (2017)). There is a literature that studies conditional asset pricing models, where the model is conditionally linear (e.g., Jagannathan and Wang (1996);

Lettau and Ludvigson (2001); Kelly, Pruitt, and Su (2019)). There is a literature that shows that the same characteristics-based models are present across a large set of asset classes (e.g., Asness, Moskowitz, and Pedersen (2013); Koijen, Moskowitz, Pedersen, and Vrugt (2018); Bollerslev, Hood, Huss, and Pedersen (2018)). A small subset of papers emphasizes non-linearity in explaining asset returns, such as Bansal and Viswanathan (1993). A growing literature aims at pricing assets with machine learning methods (e.g., Hutchinson, Lo, and Poggio (1994); Gu, Kelly, and Xiu (2020); Kelly, Malamud, and Pedersen (2023); Kelly, Malamud, and Zhou (2024)). Apart for using factor models in understanding asset prices, a recent literature highlights the significance of demand forces (e.g., Koijen and Yogo (2019); Koijen, Richmond, and Yogo (2024)). We contribute to this literature by proposing and testing a parsimonious model that represents exactly any non-linear model with an arbitrary number of factors and factor loadings.

The Kolmogorov-Arnold representation is typically used as a foundation of deep learning and as a benchmark for deriving bounds on neural network approximations; see Schmidt-Hieber (2021), Hecht-Nielsen (1987), Kurkova (1991), and Maiorov and Pinkus (1999). Coppejans (2004) also used this representation for nonparametric estimation in the regression context. We contribute to this literature by applying this representation to factor models.

Finally, our approach complements the literature on non-linear factor extraction. Gu, Kelly, and Xiu (2021) used an auto-encoder approach. Scholkopf, Smola, and Muller (1998) developed a kernel PCA procedure, which estimates a linear factor model for a non-linear transformation of the vectors $(r_{1t}, \ldots, r_{Nt})$. In contrast to these alternative methods, our approach is based on the assumption that the factor affects returns via model (1).

The rest of the paper is organized as follows. Section 2 discusses the motivation and estimation for our one factor model. Section 3 derives the rate of convergence of our sieve-based least squares estimator. Section 4 discusses the data and presents summary statistics. Section 5 presents the cross-sectional asset pricing results and compares the HFL model with alternative models. Section 6 contains additional results and robustness checks. The Online Appendix contains technical derivations and further empirical results.

2 Motivation and Estimation

We start our analysis with a very general factor-type model

$$r_{it} = g(x_{t1}, \ldots, x_{tk}, z_{i1}, \ldots, z_{im}) + \varepsilon_{it}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T, \tag{2}$$

where r_{it} is the excess return of asset i at time period t, $x_{t1}, \ldots, x_{tk} \in [0,1]$ are time-specific effects/factors, $z_{i1}, \ldots, z_{im} \in [0,1]$ are asset-specific effects/factor loadings, g is a continuous function, and ε_{it} is an idiosyncratic mean-zero component. This model is very general as it nests many special cases. For example, it nests any linear factor model

$$r_{it} = \sum_{j=1}^{k} x_{tj} z_{ij} + \varepsilon_{it},$$

any single-index factor model

$$r_{it} = h\left(\sum_{j=1}^{k} x_{tj} z_{ij}\right) + \varepsilon_{it},$$

and any non-linear factor model

$$r_{it} = h(x_{t1} z_{i1}, \ldots, x_{tk} z_{ik}) + \varepsilon_{it}.$$

In fact, model (2) does not even require the number of time-specific effects k to coincide with the number of asset-specific effects m.

Our first result shows that under very mild regularity conditions, namely continuity of the function g, model (2) can be reduced to model (1) with continuous h as a consequence of the Kolmogorov-Arnold theorem (Kolmogorov (1956); Arnold (1957)). Specifically, consider the version of Kolmogorov-Arnold theorem in Schmidt-Hieber (2021), Theorem 2(ii). That representation implies that there exists a continuous function q and monotone functions $\phi_1, \ldots, \phi_k, \psi_1, \ldots, \psi_m$

such that

$$g(x_1,\ldots,x_k,z_1,\ldots,z_m) = q\left(\sum_{j=1}^{k}\phi_j(x_j) + \sum_{j=1}^{m}\psi_j(z_j)\right)$$

for all $x_1,\ldots,x_k,z_1,\ldots,z_m \in [0,1]$. Therefore, denoting $\widetilde{f}_t = \sum_{j=1}^{k}\phi_j(x_{tj})$ and $\widetilde{\lambda}_i = \sum_{j=1}^{m}\psi_j(z_{ij})$, it follows that

$$g(x_{t1},\ldots,x_{tk},z_{i1},\ldots,z_{im}) = q(\widetilde{f}_t + \widetilde{\lambda}_i)$$

for all $i = 1,\ldots,N$ and $t = 1,\ldots,T$. Moreover, by using the identity

$$x + z = \log(\exp(x+z)) = \log(\exp(x)\exp(z)), \quad x,z \in \mathbb{R}$$

and denoting $h(x) = q(\log(x))$, $f_t = \exp(\widetilde{f}_t)$, and $\lambda_i = \exp(\widetilde{\lambda}_i)$, we have

$$g(x_{t1},\ldots,x_{tk},z_{i1},\ldots,z_{im}) = q(\log(f_t\lambda_i)) = h(f_t\lambda_i)$$

for all $t = 1,\ldots,T$ and $i = 1,\ldots,N$. We summarize this result in the following lemma.

Lemma 2.1. *If random variables $\{y_{it}\}_{i,t}^{N,T}$ satisfy model (2) with continuous g, there exist a continuous function h, and strictly positive sequences $\{f_t\}_{t=1}^{T}$ and $\{\lambda_i\}_{i=1}^{N}$ such that*

$$r_{it} = h(f_t\lambda_i) + \varepsilon_{it}, \quad i = 1,\ldots,N,\ t = 1,\ldots,T.$$

This lemma shows that any factor-type model (2) can be reduced to model (1). It is because of this generality, we study the non-linear factor model (1) with just one factor in this paper.

Remark (Alternative Representations). We could consider a more flexible reduction of model (2). Indeed, by applying the Kolmogorov-Arnold representation theorem conditional on $z_1,\ldots,z_m$, we can find a continuous function q and monotone functions $\phi_1,\ldots,\phi_k$ such that

$$g(x_1,\ldots,x_k,z_1,\ldots,z_m) = q\left(\sum_{j=1}^{k}\phi_j(x_j), z_1,\ldots,z_m\right)$$

for all $x_1, \ldots, x_k, z_1, \ldots, z_m \in [0,1]$. Further, by applying the Kolmogorov-Arnold representation theorem conditional on $x_1, \ldots, x_k$, we can find a continuous function u and monotone functions $\psi_1, \ldots, \psi_k$ such that

$$q\left(\sum_{j=1}^{k} \phi_j(x_j), z_1, \ldots, z_m\right) = u\left(\sum_{j=1}^{k} \phi_j(x_j), \sum_{j=1}^{m} \psi_j(z_j)\right)$$

for all $x_1, \ldots, x_k, z_1, \ldots, z_m \in [0,1]$. Thus, denoting $f_t = \sum_{j=1}^{k} \phi_j(x_{tj})$ and $\lambda_i = \sum_{j=1}^{m} \psi_j(z_{ij})$, it follows that model (2) implies that

$$r_{it} = u(f_t, \lambda_i) + \varepsilon_{it}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T. \tag{3}$$

□

In order to estimate model (1), we assume that both the factor f_t and the factor loading λ_i are bounded random variables. By rescaling the function h, it is then without loss of generality to assume that both f_t and λ_i are taking values in the $(0,1]$ interval. Further, since the function h is continuous, it follows from the Weierstrass theorem that it can be approximated arbitrarily well by a finite-degree polynomial:

$$h(x) = \sum_{j=0}^{K} h_j x^j + e_K(x), \quad x > 0,$$

where the residual e_K converges to zero as the degree of the polynomial K gets large. We therefore propose a sieve-based least squares estimator

$$\left(\{\widehat{h}_j\}_{j=0}^{K}, \{\widehat{f}_t\}_{t=1}^{T}, \{\widehat{\lambda}_i\}_{i=1}^{N}\right) = \arg\min \sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{it} - \sum_{j=0}^{K} c_j (\phi_t l_i)^j\right)^2, \tag{4}$$

where the minimum is taken over all $\{c_j\}_{j=0}^{K} \in \mathbb{R}^{K+1}$, $\{\phi_t\}_{t=1}^{T} \in (0,1]^T$, and $\{l_i\}_{i=1}^{N} \in (0,1]^N$. The estimators of the factor values $\{f_t\}_{t=1}^{T}$ and factor loadings $\{\lambda_i\}_{i=1}^{N}$ are then $\{\widehat{f}_t\}_{t=1}^{T}$ and $\{\widehat{\lambda}_i\}_{i=1}^{N}$, respectively, and the estimator of the function h is the Kth degree polynomial

$$\widehat{h}(x) = \sum_{j=0}^{K} \widehat{h}_j x^j, \quad x > 0. \tag{5}$$

Note here that the optimization problem (4) is highly non-convex and may be difficult to solve exactly. To make progress in this direction, we take advantage of the fact that for given values of $\{\phi_t\}_{t=1}^T$ and $\{l_i\}_{i=1}^N$, the optimization over $\{c_j\}_{j=0}^K$ is easy as it can be implemented via OLS. We therefore only need to take care of optimization over $\{\phi_t\}_{t=1}^T \in (0,1]^T$ and $\{l_i\}_{i=1}^N \in (0,1]^N$. In practice, we proceed with optimization over these two sequences as follows. First, we calculate the initial values of $\{\phi_t\}_{t=1}^T$ and $\{l_i\}_{i=1}^N$ as the first left and right singular vectors of the $T \times N$ matrix $\{r_{it}\}_{t,i=1}^{T,N}$ shifted and scaled in a way to make sure that their components are taking values in the $(0,1]$ interval. Second, we perform gradient descent starting from the initial values to find a local minimum of the optimization problem. Third, we reoptimize each component of the sequences $\{\phi_t\}_{t=1}^T$ and $\{l_i\}_{i=1}^N$ in turn multiple times until the change in the criterion function from reoptimization becomes negligible.

Remark (Other Sieve Spaces). We emphasize that although we focus on the polynomial sieves, there exist many other sieve spaces that could be used to approximate the function h arbitrarily well. For example, one could consider splines, wavelets, or neural networks. Our theory in the next section could be extended to cover these alternative sieve spaces but we have decided to focus on the polynomial sieves for conciseness of the derivations. □

3 Rate of Convergence

In this section, we derive the rate of convergence of our sieve-based least squares estimator. To do so, we need the concept of the sub-Gaussian norm, also known as the ψ_2 norm, which is defined as follows. Let ψ_2 be the function defined by $\psi_2(x) = \exp(x^2)$ for all $x \geq 0$. Then for any random variable X, its sub-Gaussian norm $\|X\|_{\psi_2}$ is the smallest constant $C > 0$ such that $E[\psi(|X|/C)] \leq 1$. In addition, we impose the following assumptions.

Assumption 3.1 (Idiosyncratic Noise)**.** *(i) Random variables $\varepsilon_{i,t}$ are mutually indepenent across i and t. (ii)* $\max_{1\le i\le N}\max_{1\le t\le T}\|\varepsilon_{it}\|_{\psi_2}=O(1)$.

Assumption 3.1(i) is fairly standard and can be relaxed to allow weak dependence in exchange for a more complicated analysis. Assumption 3.1(ii) means that the tails of the random variables ε_{it} are not heavier than those of the Gaussian random variables. This assumption is also rather standard in the literature.

Assumption 3.2 (Approximation Error)**.** *There exist constants $\alpha>0$ and $L>0$ such that for each K, there exist constants $h_0^K, h_1^K \dots, h_K^K \in [-L,L]$ such that the polynomial $h_K(x)=\sum_{j=0}^K h_j^K x^j$, $x\in[0,1]$, satisfies* $\sup_{x\in[0,1]}|h_K(x)-h(x)|=O(K^{-\alpha})$.

This assumption means that the function h can be approximated by finite-degree polynomials and specifies how fast the approximation error in the uniform metric converges to zero as we increase the degree of the polynomials.

In light of Assumption 3.2, for the theoretical analysis in this section, we assume that optimization in (4) over slope parameters $\{c_j\}_{j=0}^K$ is performed so that each of these parameters remains bounded: $|c_j|\le L$ for all $j=0,\dots,K$ and the same constant L as that specified in Assumption 3.2. In practice, we find that these constraints are not binding if the constant L is chosen large enough.

The following theorem establishes the main result of this section.

Theorem 3.1 (Rate of Convergence)**.** *Suppose that Assumptions 3.1 and 3.2 are satisfied. Then*

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T(\widehat{h}(\widehat{f_t\lambda_i})-h(f_t\lambda_i))^2=O_P\left(\frac{(T+N+K)\log(TNK)}{NT}+K^{-2\alpha}\right). \tag{6}$$

The proof of this theorem is in Appendix A.1. The theorem shows that our sieve-based least-squares estimator is consistent as long $K\to\infty$ together with N and T in such a way that $K\log(TN)=o(NT)$. Following standard terminology of nonparametric estimation, the first and the second terms on the right-hand side of (6) represent the variance and the bias terms, respectively. The variance term is increasing in the degree of the polynomial sieve K and the bias term is decreasing in K. Moreover, if $K = o(\max(N,T))$, then it follows from the theorem that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t}\widehat{\lambda_i}) - h(f_t\lambda_i))^2 = O_P\left(\frac{\log(NT)}{\min(N,T)} + K^{-2\alpha}\right).$$

Thus, the theorem shows that the variance term converges to zero with a fast rate even if we set K to be rather large.[1]

4 Data and Summary Statistics

We present a description of the data used for the estimation of the HFL model. In order to estimate the HFL model and perform cross-sectional asset pricing tests, we use data for a range of primary asset classes: equities, bonds, commodities, and currencies.

For equities, we use the Fama and French (1993) 25 size and value sorted portfolios and the Fama and French (2015) 25 size and operating profitability sorted portfolios. We also include the size, value and operating profitability equity factors. These equity portfolios and factors are from Ken French's data library. For bonds, we consider U.S. bonds and international bonds. For U.S. bonds, we use eleven maturity-sorted government bond portfolios from CRSP's "Fama Bond Portfolios" file with maturities in six month intervals up to ten years. For international bonds, we use the Refinitiv Datastream Government Bond Indices for 8 advanced economies (Australia, Austria, Canada, France, Germany, Japan, Netherlands and the United Kingdom). Specifically, we consider bond indices with maturity 1-3 years, 2 years, 5 years, 7 years and 10 years. These bond indices include coupon payments and bonds denominated in each country local currency. For commodities, we use returns of 21 commodity spot price indices from Datastream. The commodities are aluminum, Brent oil, cocoa, copper, corn, cotton, crude oil, eggs, gasoline, gold, natural gas, oats, platinum, pork bellies, propane, silver, soybean, sugar, tin, and wheat. For foreign exchange, we

[1]Note also that Theorem 3.1 derives the convergence rate of the HFL estimator $\widehat{h}(\widehat{f_t}\widehat{\lambda_i})$ which we use for all of our empirical results. Another interesting question is about convergence rates of the estimators $\widehat{f_t}$ and $\widehat{\lambda_i}$. It turns out that answering those questions is more difficult. In fact, even deriving conditions under which f_t and λ_i are identified, modulo certain normalizations, is highly non-trivial. In a separate paper (Borri, Chetverikov, Liu, and Tsyvinski (2024)), we establish identification and consistency of f_t and λ_i.

use 46 portfolios from Nucera, Sarno, and Zinna (2024). These portfolios are based on popular investment strategies: carry, short-term and long-term momentum, currency value, net foreign assets and liabilities in domestic currencies, term spread, long-term yields, and output gap.

The selection of advanced economies, for the government bond indices, and commodities is guided by the objective of constructing a balanced long sample, which in our case starts in January 1988. The sample ends in December 2017, when the sample of foreign exchange portfolios ends.

Table 1 reports summary statistics of the test assets and $h(f_t\lambda_i)$ estimated from the sieve-based estimator. We report the mean and standard deviations of the excess returns and $h(f_t\lambda_i)$ for the whole sample and for each asset class. Returns are in excess of the U.S. risk-free rate and reported in percentage. The "All" sample includes 171 assets, each with 360 monthly return observations. The average excess return is 0.32% per month (3.84% annualized). The mean standard deviation is 4.46% (15.44% annualized). The average excess return predicted by the HFL model across all assets exactly matches the sample counterpart, with a monthly volatility which is lower by one order of magnitude. The lower volatility is explained by the fact that realized returns for assets include an idiosyncratic, and volatile, component, which in our model is captured by the volatility of the residuals. Equities and commodities, on average, have higher and more volatile returns than other assets. Bonds, both U.S. and international, have on average lower, and less volatile, returns. Foreign exchange portfolios, in our sample, have average returns not statistically different from zero. For the individual asset classes, as for the whole sample of assets, the mean excess return predicted by the HFL model is very close to the average asset excess return.

Table 1: Summary Statistics of Test Assets

This table reports summary statistics of the tests assets and for $h(f_t\lambda_i)$. For the mean and the standard deviation we report averages across assets. Returns are monthly and in excess of the US risk-free rate, which we take from Ken French's data library. The sample period is January 1988 to December 2017.

	All	Equities	US Bonds	Intl Bonds	Commodities	FX
Mean ($R_{i,t}$, %)	0.320	0.796	0.156	0.217	0.312	-0.097
Std ($R_{i,t}$, %)	4.467	5.210	0.819	1.137	8.810	2.397
Mean ($h(f_t\lambda_i)$, %)	0.320	0.860	0.121	0.172	0.282	-0.110
Std ($h(f_t\lambda_i)$, %)	0.464	0.347	0.029	0.084	0.398	0.186
Assets	171	53	11	40	21	46
Months	360	360	360	360	360	360

5 Cross-Sectional Asset Pricing Tests

5.1 Baseline Results

In this section, we consider cross-sectional asset pricing tests. These tests assess whether the HFL model can account for cross-sectional differences in average excess returns. The HFL model is

$$r_{it} = h(f_t\lambda_i) + \varepsilon_{it} \quad i = 1, \ldots, N \quad 1, \ldots, T, \tag{7}$$

where r_{it} is the time t excess return on asset i and ε_{it} is an idiosyncratic mean zero disturbance. The term $h(f_t\lambda_i)$ includes the single factor f_t, common to all assets, and the asset i factor loading λ_i which are combined by the non-linear link function h by Lemma 2.1. For each asset i, Equation (7) implies $E[r_{it}] = E[h(f_t\lambda_i)]$. We empirically test this prediction by estimating the following linear model

$$E[r_{it}] = \alpha + \beta E[h(f_t\lambda_i)] + \epsilon_i, \tag{8}$$

where the expectation operator is replaced by the sample mean (i.e., $E[r_{it}] = 1/T \sum_{t=1}^{T} r_{it}$), $h(f_t\lambda_i)$'s are replaced by the corresponding estimators $\widehat{h}(\widehat{f_t\lambda_i})$, α and β are the intercept and slope coefficients, and ϵ is a mean zero residual term. Let $E[r_{it}]$ and $E[h(f_t\lambda_i)]$ denote vectors containing, respectively, the averages of the excess returns and predictions from the model.

In the empirical analysis we test several implications of our one factor model, such as that the intercept is zero and that the slope coefficient is equal to 1. Furthermore, we evaluate the fit of the model with regression R-squared and mean absolute pricing error (MAPE). The latter measures the mean absolute residuals in the cross-sectional regression (8). Our empirical strategy follows the classic two-step Fama and MacBeth (1973)'s methodology, where the first step (i.e., the asset level time-series regressions to retrieve the asset beta) is implicit in the algorithm to estimate the HFL model. The second step, that is the cross-sectional regression of mean asset returns on their betas, is replaced by the cross-sectional regression (8).

Remark. In the Appendix, we demonstrate that unless the HFL model is correct, the probability limits α_0 and β_0 of the OLS estimates $\widehat{\alpha}$ and $\widehat{\beta}$ of Equation (7) generally satisfy the restrictions $\alpha_0 = 0$ and $\beta_0 = 1$ only if the factor f_t is constant over time. Since the latter does not hold in practice, it follows that the OLS-based tests considered in this section are indeed consistent. □

Column (1) of Table (2) reports estimates of (8) using all assets for the period of January 1988 to December 2017. We consider estimates of the HFL model based on four degrees of the polynomial function used to approximate the function $h(.)$.[2] The table shows that the slope coefficient $(h(f_t\lambda_i))$ is significantly different from zero at conventional confidence levels. The pricing error, captured be the estimate of the intercept α, is not significantly different from zero. The model provides a good fit with the adjusted R^2 of 89% and the MAPE is small at 0.101%.

[2]We find that using higher order of polynomials generate qualitatively similar results.

Table 2: Cross-Sectional Asset Pricing Tests

This table presents the results of cross-sectional asset pricing tests using all assets for the baseline HFL model, as well as models augmented by, respectively, the US market excess return (CAPM), the Fama and French (1993)'s three factors (FF3), the Fama and French (2015)'s five factors (FF5), and the five factors augmented with the momentum factor of Jegadeesh and Titman (1993) (FF5+MOM). We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i$ and report the estimates for the slope ($h(f_t \lambda_i)$) and intercept (α). Standard errors adjusted with the Fama-MacBeth procedure are in parenthesis. The table additionally reports the regression adjusted R-squared, and the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model.

	(1)	(2)	(3)	(4)	(5)
	Baseline	CAPM	FF3	FF5	FF5+MOM
$h(f_t \lambda_i)$	0.816***	0.715***	0.811***	0.823***	0.821***
	(0.177)	(0.135)	(0.154)	(0.153)	(0.169)
MKTRF		0.147	-0.095	-0.106	-0.104
		(0.260)	(0.259)	(0.261)	(0.281)
SMB			0.331*	0.328*	0.327*
			(0.170)	(0.168)	(0.170)
HML			-0.105	-0.087	-0.085
			(0.181)	(0.177)	(0.175)
CMA				-0.163	-0.162
				(0.199)	(0.203)
RMW				-0.000	-0.001
				(0.217)	(0.218)
MOM					0.111
					(1.109)
α	0.001	0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Adj R^2	0.885	0.900	0.948	0.948	0.948
MAPE, %	0.101	0.098	0.072	0.071	0.071
Assets	171	171	171	171	171
Months	360	360	360	360	360

The good fit of the model is visually confirmed by Figure 1, which plots the average asset excess returns versus the predicted excess returns using the HFL model. All assets line up along the 45 degree line. Figure A.1 in the Appendix reveals that this result is robust to the exclusion of the intercept from the model used to predict average excess returns.

We compare the HFL model with the standard capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), the Fama and French (1993)'s three factor model, the Fama and French (2015)'s five factors model, and the five factors model augmented with the momentum factor of Jegadeesh and Titman (1993).

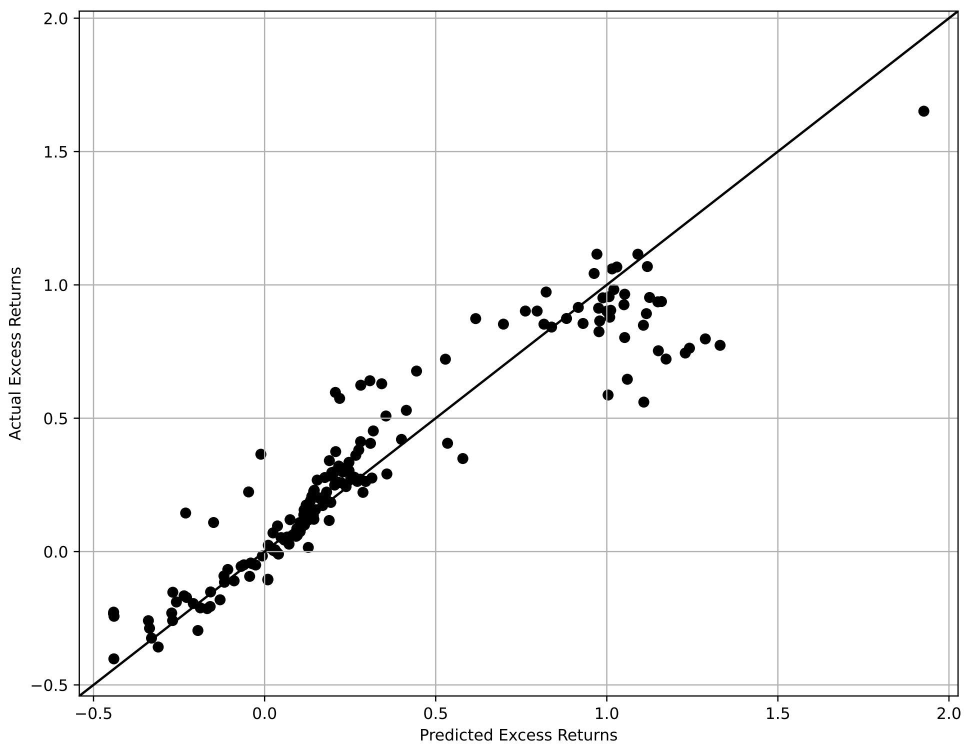
Columns (2) to (5) of Table (2) present the results of cross-sectional asset pricing tests using all assets. In all models, we include the HFL component ($h(f_t\lambda_i)$), and evaluate the effect of including additional factors. First, the slope estimate associated with the HFL component is always significantly different from zero at the 1% confidence level. Second, including the additional factors increases the adjusted R-squareds only marginally, from 89% to a maximum of 95%. In fact, with the exception of the slope coefficient associated with the size factor, which is significant at the 10% level, the slope coefficients associated with all the other factors are never statistically different from zero.

5.2 Comparison with PCA factor models

In this section, we compare the HFL model with factor models based on principal component analysis (see, e.g., Connor and Korajczyk (1986, 1988)). Table 3 presents the results of cross-sectional asset pricing tests using all assets and three estimators of the latent factors related to PCA. The first estimator is based on the standard PCA of the covariance matrix of asset returns. The second estimator, called RP-PCA, is a generalization of PCA, proposed by Lettau and Pelger (2020), which includes a penalty term to account for the pricing errors in the cross-sectional regressions based on the RP-PCA factors. RP-PCA is motivated by the poor performance of standard PCA in identifying factors which are relevant to explain the cross-section of average excess returns (see, e.g., Onatski (2012) and Kozak, Nagel, and Santosh (2020)). The third estimator is the kernel-PCA, proposed by Schölkopf, Smola, and Müller (1998). Kernel-PCA (K-PCA) is a nonlinear form of PCA which allows for the separability of nonlinear data by projecting it onto a higher dimensional space where it becomes linearly separable using kernels.

Figure 1: Actual vs Predicted Expected Returns

The figure plots actual average excess returns on all tested assets ($E[r_{it}]$) versus predicted excess returns using the cross-sectional model based on the HFL model: $E[r_{it}] = \alpha + \beta E[h(\lambda_i f_t)] + \epsilon_i$. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.



In all models, we include the HFL component ($h(f_t\lambda_i)$), and evaluate the effect of including, respectively, the first principal component, the first two principal components, and the first three principal components, separately for standard PCA, RP-PCA and K-PCA. The estimates from the cross-sectional asset pricing regressions reveal that the slope coefficient associated with the HFL component is statistically significant at conventional levels in all models, while the slopes associated with the principal components, for all three estimators we consider, are never statically different from zero. Furthermore, including the principal components increases only marginally the cross-sectional adjusted R-squared, from 89% to a maximum of 91%, while the pricing error

α is never statistically significant at conventional condidence levels.

Table A.3 in the Appendix reveals similar conclusions for asset pricing estimates of PCA models which include up to 6 factors. These estimates confirm the robustness of the results based on PCA models including the first three components.

5.3 Comparison with higher order factor models

In this section, we compare the HFL model with higher order factor models, related to the non-linear version of the CAPM of Kraus and Litzenberger (1976). Harvey and Siddique (2000) develop a three-moment conditional CAPM where skewness is priced. Ang, Hodrick, Xing, and Zhang (2006) develop a model where aggregate market volatility is priced. Table 4 presents the results of cross-sectional asset pricing tests using all assets. In all models, we include the HFL component ($h(f_t\lambda_i)$), and evaluate the effect of including higher order factors. Specifically, we consider both a model specification which additionally includes the U.S. equity market excess returns, the squared U.S. equity market excess returns, and the cubed U.S. equity market excess returns, and the second model specification which additionally includes the first principal component, the squared first principal component, and the cubed first principal component. For the principal components, we consider three estimators: the standard PCA based on the covariance matrix of asset returns; the risk premium PCA (RP-PCA) of Lettau and Pelger (2020), and the kernel PCA (K-PCA) of Schölkopf, Smola, and Müller (1998). In all modes, the slope coefficient associated to the HFL component is statistically significant at conventional confidence levels, while the slopes associated with the higher order factors are never statistically different from zero. Furthermore, including the principal components increases only marginally the cross-sectional R-squared, from 89% to a maximum of 92%. Moreover, the pricing error α is not statistically different from zero in all models, with the exception of the CAPM model augmented with the equity market factor squared and cubed ($CAPM^3$) for which the estimate for the intercept is statistically significant at the 10% level.

Table 3: Comparison with PCA Factor Models

This table presents the results of cross-sectional asset pricing tests using all assets. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i$ augmented by, respectively, the first principal component (PC1), the first two principal components (PC1+PC2) and the first three principal components extracted by the test asset returns. For the principal components, we consider three estimators: the standard PCA based on the covariance matrix of asset returns; the risk premium PCA (RP-PCA) of Lettau and Pelger (2020), and the kernel PCA (K-PCA) of Schölkopf, Smola, and Müller (1998). We report the estimates for the slopes and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression adjusted R-squared, and the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	Standard PCA			RP-PCA			K-PCA		
	PC1	PC1 + PC2	PC1–PC3	PC1	PC1 + PC2	PC1–PC3	PC1	PC1 + PC2	PC1–PC3
$h(f_t \lambda_i)$	0.705***	0.696***	0.654***	0.692***	0.656***	0.635***	0.807***	0.804***	0.777***
	(0.129)	(0.120)	(0.109)	(0.129)	(0.117)	(0.113)	(0.174)	(0.176)	(0.169)
PC1	0.014	0.014	0.015	0.014	0.016	0.016	0.012	0.010	0.009
	(0.020)	(0.019)	(0.019)	(0.020)	(0.019)	(0.019)	(0.018)	(0.017)	(0.017)
PC2		0.001	0.001		0.001	0.002		0.004	0.007
		(0.010)	(0.010)		(0.010)	(0.010)		(0.013)	(0.014)
PC3			-0.005			0.003			-0.008
			(0.010)			(0.010)			(0.014)
α	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Adj R^2	0.907	0.907	0.912	0.908	0.911	0.914	0.888	0.888	0.895
MAPE, %	0.096	0.097	0.090	0.095	0.094	0.089	0.103	0.103	0.104
Assets	171	171	171	171	171	171	171	171	171
Months	360	360	360	360	360	360	360	360	360

Table 4: Comparison with Higher Order Factor Models

This table presents the results of cross-sectional asset pricing tests using all assets. We estimate the cross-sectional linear model $E_t[r_{it}] = \alpha + \beta E_i[h(f_t\lambda_i)] + \epsilon_i$ augmented by the US market excess return and, respectively, by the squared market excess return ($CAPM^2$) and the cubed market excess return ($CAPM^3$); or augmented by the first principal component and, respectively, the squared first principal component ($PC1^2$) and the cubed first principal component ($PC1^3$). For the principal components, we consider three estimators: the standard PCA based on the covariance matrix of asset returns; the risk premium PCA (RP-PCA) of Lettau and Pelger (2020), and the kernel PCA (K-PCA) of Schölkopf, Smola, and Müller (1998). We report the estimates for the slopes and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression adjusted R-squared, and the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	$CAPM^2$	$CAPM^3$	$PC1^2$	$PC1^3$	$RP\text{-}PC1^2$	$RP\text{-}PC1^3$	$K\text{-}PCA1^2$	$K\text{-}PCA1^3$
$h(f_t\lambda_i)$	0.715***	0.711***	0.744***	0.749***	0.725***	0.726***	0.793***	0.672***
	(0.135)	(0.135)	(0.141)	(0.131)	(0.139)	(0.132)	(0.167)	(0.139)
MKTRF	0.126	0.084						
	(0.262)	(0.257)						
$MKTRF^2$	-5.220	-9.051						
	(7.854)	(7.146)						
$MKTRF^3$		8.949						
		(62.955)						
PC1			0.010	0.010				
			(0.020)	(0.020)				
$PC1^2$			0.046	0.050				
			(0.060)	(0.046)				
$PC1^3$				-0.020				
				(0.045)				
RP-PC1					0.011	0.011		
					(0.020)	(0.019)		
$RP\text{-}PC1^2$					0.045	0.045		
					(0.056)	(0.044)		
$RP\text{-}PC1^3$						-0.020		
						(0.041)		
KPC1							0.003	0.001
							(0.021)	(0.021)
$KPC1^2$							0.006	0.007
							(0.011)	(0.011)
$KPC1^3$								-0.005
								(0.007)
α	0.000	0.001*	0.000	0.000	0.000	0.000	0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Adj R^2	0.905	0.921	0.915	0.915	0.917	0.917	0.897	0.923
MAPE, %	0.098	0.086	0.088	0.089	0.103	0.104	0.096	0.086
Assets	171	171	171	171	171	171	171	171
Months	360	360	360	360	360	360	360	360

5.4 Comparison with recent macro-factor models

In this section, we compare the HFL model with factors from recent recent macro-factor models. Specifically, we consider He, Kelly, and Manela (2017)'s intermediary capital risk factor (HKM), the Lettau, Maggiori, and Weber (2014)'s downside equity market risk factor (LMW), and Pástor and Stambaugh (2003)'s liquidity risk factor (PS). Both the HKM and LMW factors are developed to investigate cross-sectional differences in average returns across large cross-section of assets from different asset classes. The PS factor is developed to investigate cross-sectional differences in equity average returns. In all cases, we compare the HFL model with single-index models based on the macro factors.[3]

Table 5 presents the results of cross-sectional asset pricing tests using all assets. In all models, we include the HFL component ($h(f_t\lambda_i)$), and evaluate the effect of including, respectively, the HKM, LMW and PS factors. In all modes, the slope coefficient associated to the HFL component is statistically significant at conventional levels, while the slopes associated with the macro-factor models are never statistically different from zero. Furthermore, including the macro-factors increases only marginally the cross-sectional adjusted R-squared, from 89% to a maximum of 91%. The pricing error α is not statistically different from zero in all specifications.

[3]For the HKM and PS factors, we rely on the data provided by Assaf Manela and Lubos Pastor, respectively. We construct the LMW factor, for the period of analysis, following Lettau, Maggiori, and Weber (2014).

Table 5: Comparison with Recent Macro Factor Models

This table presents the results of cross-sectional asset pricing tests using all assets. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i$ augmented by He, Kelly, and Manela (2017)'s intermediary capital risk factor (HKM), Lettau, Maggiori, and Weber (2014)'s downside equity market risk factor (LMW), and Pástor and Stambaugh (2003)'s liquidity risk factor (PS). We report the estimates for the slopes and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression adjusted R-squared, and the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	HKM	LMW	PS
$h(f_t \lambda_i)$	0.789***	0.714***	0.715***
	(0.152)	(0.135)	(0.141)
MKTRF	0.098	0.132	0.147
	(0.265)	(0.357)	(0.261)
HKM	-0.006		
	(0.007)		
LMW		0.016	
		(0.232)	
PS			0.000
			(0.015)
α	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
Adj R^2	0.918	0.899	0.899
MAPE, %	0.089	0.099	0.098
Assets	171	171	171
Months	360	360	360

Table 6: Comparison with macro factors of Ludvigson and Ng (2009)

This table presents the results of cross-sectional asset pricing tests using all assets. We estimate the cross-sectional linear model $E_i[r_{it}] = \alpha + \beta E_i[h(f_t \lambda_i)] + \epsilon_i$ augmented by the 9 macro factors of Ludvigson and Ng (2009). These factors, denoted by F_{LN}, are eight macro factors extracted from a sample of US government bonds plus the first factor cubed. We report the estimates for the slopes and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression adjusted R-squared. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	Without HFL									With HFL								
	$F_{LN,1}$	$F^3_{LN,1}$	$F_{LN,2}$	$F_{LN,3}$	$F_{LN,4}$	$F_{LN,5}$	F_{LN6}	F_{LN7}	F_{LN8}	$F_{LN,1}$	$F^3_{LN,1}$	$F_{LN,2}$	$F_{LN,3}$	$F_{LN,4}$	$F_{LN,5}$	F_{LN6}	F_{LN7}	F_{LN8}
$F_{LN,1}$	-0.125*									-0.007								
	(-1.830)									(-0.118)								
$F^3_{LN,1}$		-0.289**									-0.011							
		(-2.469)									(-0.113)							
$F_{LN,2}$			0.049									0.023						
			(1.078)									(0.536)						
$F_{LN,3}$				0.062									-0.002					
				(1.621)									(-0.042)					
$F_{LN,4}$					-0.103**									-0.008				
					(-2.128)									(-0.200)				
$F_{LN,5}$						0.037									-0.013			
						(0.989)									(-0.324)			
F_{LN6}							0.166***									0.056		
							(3.560)									(1.524)		
F_{LN7}								0.046***									0.012	
								(2.727)									(0.648)	
F_{LN8}									-0.047***									-0.006
									(-2.939)									(-0.389)
$h(f_t \lambda_i)$										0.810***	0.808***	0.806***	0.817***	0.801***	0.825***	0.774***	0.701***	0.794***
										(4.890)	(5.008)	(4.732)	(4.635)	(5.503)	(4.440)	(4.505)	(5.230)	(4.170)
α	0.003***	0.002***	0.002***	0.003***	0.001*	0.004***	0.005***	0.001**	0.005***	0.001	0.001	0.000	0.001	0.000	0.000	0.001	0.000	0.001*
	(3.473)	(3.246)	(4.026)	(3.388)	(1.775)	(3.991)	(3.865)	(2.032)	(4.555)	(0.854)	(0.859)	(0.364)	(0.886)	(0.686)	(0.851)	(1.562)	(0.700)	(1.770)
Adj R^2	0.107	0.213	0.036	0.058	0.213	0.016	0.189	0.559	0.192	0.885	0.885	0.894	0.885	0.886	0.887	0.904	0.904	0.888
Months	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360

5.5 Comparison with macro factors of Ludvigson and Ng (2009)

In this section, we compare the HFL model with nine macro factors proposed by Ludvigson and Ng (2009) to study bond risk premia.[4] These factors are eight factors extracted using dynamic factor analysis and a panel of US government bonds, and the first factor cubed. For each of these factors (denoted with F_{LN}), separately, we first estimate the cross-sectional regression $E[r_{it}] = \alpha + \beta E[F_{LN}] + \epsilon_i$. Next, we repeat the estimation additionally including the predicted values obtained with the HFL model (i.e., $E[h(f_t\lambda_i)]$).

Table 6 summarizes our results. When taken individually, each of the macro factors is significant, except for the second, third and fifth factors. When we additionally include in the regression specification the HFL component, all of the nine macro factors are not statistically different from zero. In contrast, the HFL component is statistically significant in all models.

5.6 Comparison with the factor zoo

In this section, we compare the HFL model with a large set of factors proposed in the last decades to account for the cross-section of asset returns (see, e.g., Feng, Giglio, and Xiu (2020)). In particular, we consider factors for 153 characteristics in 13 themes, using data from 93 countries and four regions, constructed by Jensen, Kelly, and Pedersen (2023).[5] For each of these factors (denoted with f^{zoo}), separately, we first estimate the cross-sectional regression $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$. Next, we repeat the estimation additionally including the predicted values obtained with the HFL model (i.e., $E[h(f_t\lambda_i)]$). Figure 2 summarizes our results. It plots the average of the absolute values of the t-statistics in a test of the intercepts and slope coefficients in the cross-sectional regressions, respectively, equal to zero. In the left panel of the figure, we consider the regressions which only include the factor asset betas. In the right panel of the figure, we consider regressions which additionally include the average predicted values from the HFL

[4] The macro factors of Ludvigson and Ng (2009) are from Sydney Ludvigson's website.

[5] The data for the factor zoo is available through the Global Factor Data website. We use capped value weighted data for all factors and all countries.

model. The figure reveals that, for the models which do not include the HFL component, the estimates of the slope coefficients are on average significantly different from zero (average absolute t-statistics equal to 2.06). For these models, though, the pricing errors, captured by the estimates of the intercept, are on average statistically significant (average absolute t-statistics equal to 2.91). Furthermore, the figure reveals that, in the models which additionally include the average predicted values from the HFL model, the slope coefficients associated with the factor zoo become not statistically different from zero (average absolute t-statistics equal to 0.85), while those associated with the HFL component are highly significant (average absolute t-statistics equal to 5.29). Moreover, for the models which includes the HFL component, the estimates of the pricing error also become not statistically different from zero (average t-statistics equal to 0.21).

Figure 2: **Comparison with the Factor Zoo**

This figure shows the average of the absolute values of the t-statistics in a test of the intercepts and slope coefficients in the cross-sectional regressions, respectively, equal to zero. Standard errors are adjusted with the Fama-MacBeth procedure. In the left panel, we consider the cross-sectional regressions $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$ for each factor from the factor zoo. In the right panel, we consider cross-sectional regressions which additionally include the average predicted values based on the HFL component (i.e., $E[h(f_t \lambda_i)]$). The vertical bars denote one standard deviation around the mean. The horizontal black dashed-line denote significance at the 10% confidence level. The data for 153 factors from the factor zoo are from Jensen, Kelly, and Pedersen (2023). The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

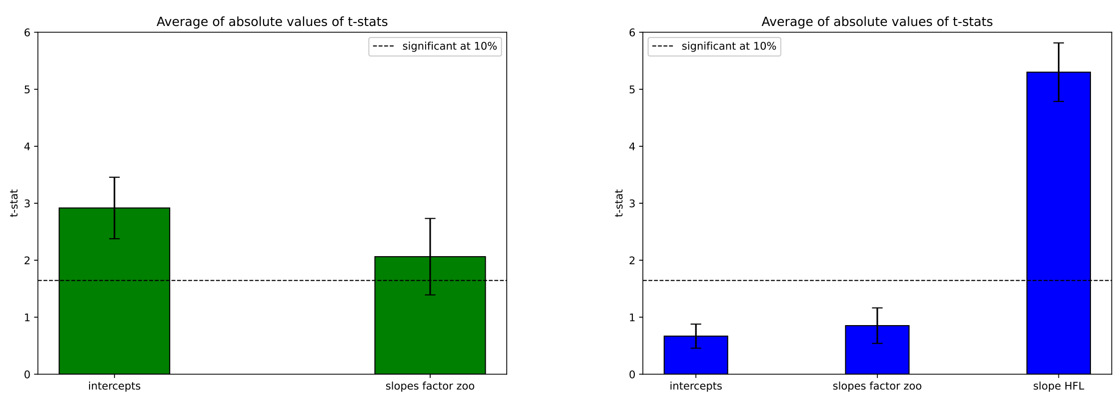


Table 7 provides further details about the comparison with the factor zoo. If the HFL component is not included (Panel A), the fraction of models with a slope coefficient on the factor zoo which is significant at the 10% confidence level is approximately 78%. When we further include the HFL component, this fraction drops to approximately 2%. Moreover, if the HFL component is included (Panel B), all models have a slope on the HFL component which is significant at the 1% confidence level. In these models, the intercept is never statistically different from zero at a confidence level of 10% or lower. Finally, we find that only three of the 153 factors from the factor zoo are significant with a confidence level of 10%. These factors are the capital turnover of Haugen and Baker (1996); the dollar trading volume of Brennan, Chordia, and Subrahmanyam (1998); and the share turnover of Datar, Naik, and Radcliffe (1998).

The double-selection Lasso of Belloni, Chernozhukov, and Hansen (2014) and Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and Robins (2018) allows estimating regressions where the number of right-hand side variables could be large, even larger than the sample size. In practice, double-selection Lasso estimates only one coefficient on the right-hand side at a time by means of a two-step selection method. In the first, factors with a low contribution to the cross-sectional pricing are excluded from the set of controls. In the second step, factors whose covariances with returns are highly correlated in the cross-section with the covariance between returns and a given factor are added to the set of controls. Therefore, the double-selection Lasso achieves dimensionality reduction by eliminating useless and redundant factors. Feng, Giglio, and Xiu (2020) propose double-selection Lasso as model selection method to evaluate the contribution of factors in cross-sectional asset pricing regressions.

The double-selection Lasso is based on two Lasso regressions which contain a regularization parameter. In setting these parameters, we follow a cross-validation, or CV, procedure (see, e.g., Hastie, Tibshirani, Friedman, and Friedman (2009)). We first set a broad range of values for each regularization parameter. Next, we split the data into 5 folds, and train the model on 4 folds with one fold held out as validation set, cycling through all folds. This process is repeated for each value in the range of values for the regularization parameters and the model performance is evaluated on the validation set using the mean squared error as the metric. The parameter value with the lowest average mean squared error across all folds is selected as the chosen regularization parameter.

The HFL component using the double-selection Lasso method is statistically significant at the 1% level (t-statistics equal to 7.34). In this case, the intercept is not statistically different from zero (t-statistics equal to 1.37). Furthermore, Panel C of table 7 reports summary statistics on the statistical significance of slope coefficients on the factor zoo, after controlling for the HFL component, as well as for the remaining factors in the factor zoo. The table reveals that the fraction of significant factors in the factor zoo, with a confidence of 10% or below, is approximately 10%, or 14 factors (see Table A.1 in the Appendix for the list of significant factors).

5.7 Predictability

In this section, we study whether the HFL model can predict returns by investigating the properties of portfolios constructed on the basis of the predictions of the HFL model. We start by estimating the model on all 171 assets and the first 120 months of data.[6] Next, we use the model predicted returns to sort assets in five portfolios, and compute the equally-weighted average portfolio return in the following six months. Finally, we repeat the procedure expanding the estimation window by 6 months each time until the end of the sample (i.e., we repeat the procedure 40 times in order to reach the end of the sample in December, 2017). In each period, the first four portfolios contain 34 assets each, and the last portfolio contains the remaining 35 assets. This empirical methodology is designed to capture the strategy of an investor who uses the model to predict future returns rebalancing at a semi-annual frequency. In each period t, we sort assets only using information available up to period t and compute excess returns between t and $t+1$,where each period contains 6 months. Portfolio 1 groups the assets with, each period, the lowest predicted returns by the HFL model. Portfolio 5 groups the assets with, each period, the highest predicted returns by the HFL model.

[6]The 171 assets contain three equity factors, and the results are qualitatively similar when we exclude them in the predictability exercise.

Table 8 describes the properties of the five portfolios, as well as of a long-short strategy long in portfolio 5 and short in portfolio 1, and their risk-adjusted performance. The analysis of Panel A reveals the predictive ability of the HFL model in forecasting future returns. Across the portfolios, there is a sizable cross-section in average monthly returns, ranging from 0.08% for the first portfolio to 0.77% for the last one. Notably, a zero-cost long-short strategy, wherein assets with the highest predicted returns are bought while those with the lowest predicted returns are sold short, yields a monthly average excess return of 0.69% (8.28% annualized), with a monthly Sharpe ratio of 0.15 (0.51 annualized). Despite the relatively short time-series, the cross-sectional return spread is statistically significant. For instance, the standard error for the long-short strategy is 30 basis points. Consequently, the average excess return is more than two standard errors from zero. Similarly, each individual portfolio, except the first two portfolios, offers average returns which are statistically different from zero. Panel B highlights the significant risk-adjusted performance of the strategy based on the HFL model. For each portfolio, we report the estimates of the intercept in a regression of the portfolio return on the Fama and French (1993)'s three factors (α_{FF3}), on the Fama and French (2015)'s five factors (α_{FF5}), and on the five factors augmented by the momentum factor of Jegadeesh and Titman (1993) ($\alpha_{FF5+MOM}$), respectively. These estimates of the intercept are a measure of risk-adjusted performance. First, the table reveals that the α estimates increase from the first to the last portfolio in all three specifications. Second, we find that the alphas for portfolio 3 to 5, as well as for the long-short portfolio, are statistically different from zero at conventional levels. For the long-short portfolio, the monthly alphas are equal, respectively, to 0.60%, 0.78% and 0.73%, and are statically significant at the 5% level.

Table 7: Comparison with the Factor Zoo

This table provides a comparison of the HFL model and the factor zoo. Panel A refers to cross-sectional regressions $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$ for each factor from the factor zoo. Standard errors are adjusted with the Fama-MacBeth procedure. Panel B refers to cross-sectional regressions which additionally include the average predicted values based on the HFL component (i.e., $E[h(\lambda_i f_t)]$). The two panels report the fraction (in percentage) of intercepts and slope coefficients with p-values less than 1%, between 1 and 5%, between 5 and 10% and larger than 10%. The p-values correspond to tests for the coefficient equal to zero. Panel C refers to the model which includes the HFL component and estimates based on the double-selection Lasso method, and reports the fraction of slope coefficients on the factor zoo with p-values less than 1%, between 1 and 5%, between 5 and 10% and larger than 10%. The HFL component using the double-selection Lasso method is statistically significant at the 1% level. The data for 153 factors from the factor zoo are from Jensen, Kelly, and Pedersen (2023). The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

$p(coeff=0)$	$p \leq 1\%$	$1\% < p \leq 5\%$	$5\% < p \leq 10\%$	$p > 10\%$
	Panel A: Model without HFL			
α	79.08	13.72	2.61	4.57
$E[f_t^{zoo}]$	15.68	54.24	7.84	22.22
	Panel B: Model with HFL			
α	0	0	0	1
$E[f_t^{zoo}]$	0	0	1.96	98.03
$h(f_t\lambda_i)$	1	0	0	0
	Panel C: Model with HFL and double-selection Lasso			
α	98.03	1.96	0	0
$E[f_t^{zoo}]$	1.30	4.57	3.26	90.84

Table 8: HFL Portfolios

This table describes the properties of five portfolios formed by sorting assets by the predicted values of the HFL model and their risk-adjusted performance. Starting in January 1998, we sort all assets by the predicted values from the HFL model estimated using the previous 120 months and compute the equally-weighted average portfolio return in the following 6 months. Finally, we repeat the procedure expanding the estimation window by 6 months each time until the end of the sample (i.e., we repeat the procedure 40 times in order to reach the end of the sample in December, 2017). For each portfolio, and for the long-short portfolio long in the last portfolio and short in the first portfolio, Panel A reports the mean (Mean), standard deviation (SD), and standard error (SE), along the Sharpe ratio (SR) measured as the ratio between mean and standard deviation. The last row reports the number of assets in each portfolio in each period. Panel B reports the estimates of the intercept in a regression of the portfolio return on the Fama and French (1993)'s three factors (α_{FF3}), on the Fama and French (2015)'s five factors (α_{FF5}), and on the five factors augmented by the momentum factor of Jegadeesh and Titman (1993) ($\alpha_{FF5+MOM}$), respectively. Newey and West (1987)'s standard errors with a lag of 3 in parenthesis. *, **, and *** represent significance at the 1%, 5%, and 10% level. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

Portfolio	P1	P2	P3	P4	P5	P5-P1
	Panel A: Portfolio properties					
Mean (%)	0.079	0.084	0.29	0.612	0.769	0.689
SD (%)	3.443	2.29	1.704	2.522	4.123	4.703
SE (%)	0.222	0.148	0.11	0.163	0.266	0.304
SR	0.023	0.037	0.17	0.243	0.186	0.147
Assets	34	34	34	34	35	69
	Panel B: Risk-adjusted performance					
α_{FF3}	-0.098	0.031	0.348***	0.532***	0.505***	0.604**
	(0.181)	(0.124)	(0.11)	(0.108)	(0.176)	(0.281)
α_{FF5}	-0.198	-0.079	0.286***	0.565***	0.583***	0.781***
	(0.179)	(0.113)	(0.108)	(0.113)	(0.189)	(0.285)
$\alpha_{FF5+MOM}$	-0.169	-0.063	0.283**	0.577***	0.56***	0.729**
	(0.181)	(0.112)	(0.112)	(0.117)	(0.192)	(0.291)

6 Additional Results

In this section, we describe four sets of additional results related to our main findings. First, we report the results of asset pricing tests for individual asset classes. Second, we report further asset pricing tests using the first and second half of our sample. Third, we investigate the time-variation in the asset pricing test using an estimation based on an expanding window. Fourth, we consider a larger set of test assets, which further includes the equity momentum portfolios.

6.1 Asset pricing tests for individual asset classes

In this section, we present the result of cross-sectional asset pricing tests for individual asset classes. Table 9 summarizes these results. The slope associated with the HFL component is statistically significant for each asset class. Specifically, it is significant at the 1% level for each asset classes except for the case of commodity, for which the slope is significant at the 10% level. The estimates for the intercept are small in magnitude, and statistically different from zero, at the 5% level, only for US bonds and international bonds. Furthermore, the adjusted R-squareds are above 85% for each asset class, with the exception of equities (47%). The MAPE is the highest for equities (0.13%) and the smallest for US bonds (0.006%). Section B.2 in the Appendix contains scatter plots of average excess returns against predicted excess returns for each individual asset class.

Table 9: Cross-Sectional Asset Pricing Tests: Individual Asset Classes

This table presents the results of cross-sectional asset pricing tests for each asset class separately. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i$ and report the estimates for the slope ($h(f_t \lambda_i)$) and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table columns report estimates for the different asset classes. The table additionally reports the regression adjusted R-squared, the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The KA factor model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	All	Equities	US Bonds	Intl Bonds	Commodities	FX
$h(f_t \lambda_i)$	0.816***	0.441***	2.536***	1.617***	0.813*	0.787***
	(0.177)	(0.152)	(0.767)	(0.275)	(0.490)	(0.101)
α	0.001	0.004	-0.001**	-0.001**	0.001	-0.000
	(0.001)	(0.003)	(0.001)	(0.000)	(0.002)	(0.001)
Adj R^2	0.885	0.472	0.987	0.913	0.898	0.888
MAPE, %	0.101	0.131	0.006	0.032	0.077	0.037
Assets	171	53	11	40	21	46
Months	360	360	360	360	360	360

6.2 Asset-pricing using alternative samples

In this section, we present results of cross-sectional asset pricing tests by asset class using two alternative samples. The first sample starts in January 1988 and ends in December 2003 (first half). The second sample starts in January 2004 and ends in December 2017 (second half). We report the estimates for these two sub-samples, each including 180 monthly dates, in Table 10.

Panel A refers to the estimates using the first half of the sample. In the model with all assets, the slope coefficient associated with $h(f_t \lambda_i)$ is statistically different from zero. Furthermore, the estimate for the α, the model pricing error, is not statistically different from zero and the cross-sectional regression adjusted R-squared is equal to 86%. For the individual asset classes, we reject the null of slope associated with the HFL component equal to zero for all assets, except for commodities (for equities the p-value is equal to 9.6%). The cross-sectional regression adjusted R-squared is highest for US bonds (96%) and lowest for equities (60%).

Panel B refers to the, more recent, second half of the sample. Similarly to the results for the first half of the sample, in the model with all assets, the slope coefficient associated with $h(f_t\lambda_i)$ is statistically different from zero. The slope point estimate is equal to 0.82, whereas it is equal to 0.85 in the first half of the sample. Furthermore, the estimate for the α, the model pricing error, is not statistically different from zero. However, the model fit is lower than in the first half, with a regression adjusted R-squared of 77% and a MAPE of 0.14%. For the individual asset classes, similarly to the estimates for the first half of the sample, we reject the null of slope equal to zero for the HFL component for all asset classes with the exception of commodities. The cross-sectional regression adjusted R-squared is highest for US bonds (98%) and lowest for equities (22%).

Table A.4 in the Appendix presents asset pricing estimates for two samples before and after the global financial crisis of 2007-08, which further confirm the robustness of our results for alternative samples.

6.3 Time variation in asset-pricing tests

In this section, we further investigate time-variation in the cross-sectional asset pricing tests. In section 6.2, we showed that the estimates for the slope associated with the HFL component, on all assets, are similar and statistically significant in both the first and second half of the sample. In this section, we provide further robustness evidence by plotting, in Figure 3, time-varying monthly coefficient estimates. The time-varying estimates are based on an expanding window and starting with a minimum window size of 10 years of data at the monthly frequency following Fama and French (1992). The figure reveals that the slope associated with $E[h(f_t\lambda_i)]$ is statistically different from zero for all months, as highlighted by the shaded region which denotes a two standard error confidence band. Furthermore, we can never reject the hypothesis that the slope is equal to 1, which is one of the cross-sectional predictions of our model.

Table 10: Cross-Sectional Asset Pricing by Asset Class: Alternative Samples

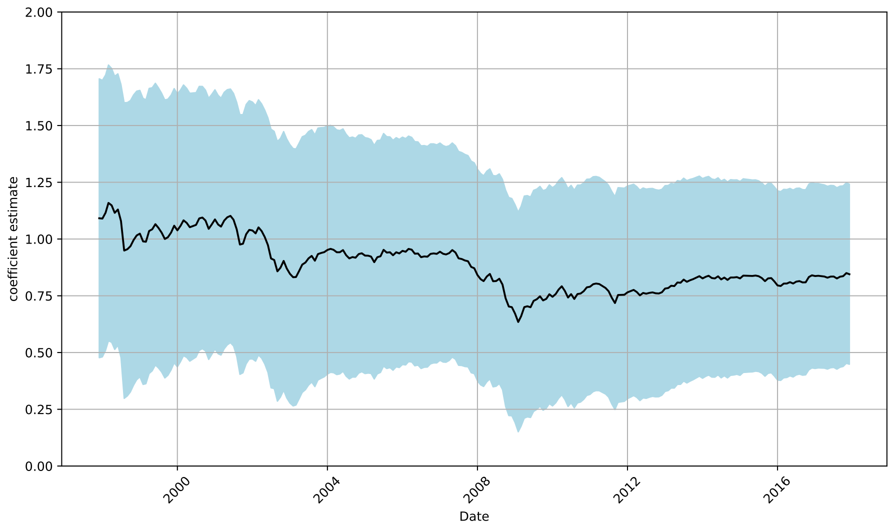
This table presents the results of cross-sectional asset pricing tests for alternative samples. The first sample starts in January 1988 and ends in December 2003 (first half). The second sample starts in January 2004 and ends in December 2017 (second half). We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t\lambda_i)] + \epsilon_i$ and report the estimates for the slope $(h(f_t\lambda_i))$ and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. Panel A reports estimates for the first half of the sample. Panel B reports estimates for the second half of the sample. The table columns report estimates for the different asset classes. The table additionally reports the regression adjusted R-squared, the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

Panel A: First Half	All	Equities	US Bonds	Intl Bonds	Commodities	FX
$h(f_t\lambda_i)$	0.849***	0.546*	2.232**	1.411***	0.782	1.003***
	(0.233)	(0.282)	(0.913)	(0.340)	(0.576)	(0.172)
α	-0.000	0.002	-0.001	-0.000	-0.001	-0.000
	(0.001)	(0.004)	(0.001)	(0.000)	(0.002)	(0.002)
Adj R^2	0.859	0.606	0.961	0.782	0.654	0.789
MAPE, %	0.133	0.127	0.012	0.058	0.218	0.072
Assets	171	53	11	40	21	46
Months	180	180	180	180	180	180

Panel B: Second Half	All	Equities	US Bonds	Intl Bonds	Commodities	FX
$h(f_t\lambda_i)$	0.819***	0.370***	2.984**	1.916***	0.849	0.607***
	(0.277)	(0.126)	(1.342)	(0.462)	(0.810)	(0.117)
α	0.001	0.006*	-0.003**	-0.001***	0.003	0.001
	(0.001)	(0.003)	(0.001)	(0.000)	(0.005)	(0.002)
Adj R^2	0.765	0.222	0.974	0.883	0.495	0.733
MAPE, %	0.140	0.164	0.010	0.038	0.210	0.059
Assets	171	53	11	40	21	46
Months	180	180	180	180	180	180

Figure 3: Time-Varying Coefficient Estimates

The figure plots the estimates of the slope coefficient in the cross-sectional model based on the HFL model: $E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i$. The estimation is based on an expanding window, starting with a minimum window size of 10 years of data at the monthly frequency following Fama and French (1992). The HFL model is estimated using all assets. The colored region denotes a two standard error confidence band. Standard errors are adjusted with the Fama-MacBeth procedure. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.



6.4 Asset-pricing with momentum portfolios

Table 11 presents estimates of cross-sectional asset pricing tests using a set of test assets which additionally includes 10 momentum portfolios and the momentum factor. Both the momentum portfolios and the momentum factor are from Ken French's website and are based on U.S. stocks monthly returns sorted by the 12-2 past return. The table reveals that the slope coefficient associated with the averages of the predicted values from the HFL component is significantly different from zero for all assets as well as for each individual asset class. The pricing error, captured be the estimate of the intercept, is significantly different from zero only for US bonds (significant at the 5%). The model provides the closest fit for US bonds (adjusted-R^2 of 99%), although the model fit is similar across the specifications with the exception of equities with a lower value for the adjusted R-squared of 54%. The MAPE ranges from 0.005% (US bonds) to 0.122% (equities).

Table 11: Cross-Sectional Asset Pricing Tests – With Momentum

This table presents the results of cross-sectional asset pricing tests using all assets and, additionally, 10 momentum portfolios and a momentum factor. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(f_t\lambda_i)] + \epsilon_i$ and report the estimates for the slope ($h(f_t\lambda_i)$) and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression R-squared, the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	All	Equities	US Bonds	Intl Bonds	Commodities	FX
$h(f_t\lambda_i)$	0.795***	0.500***	2.316***	1.529***	0.783*	0.731***
	(0.164)	(0.148)	(0.700)	(0.261)	(0.472)	(0.093)
α	0.001	0.004	-0.001**	-0.000	0.001	-0.000
	(0.001)	(0.003)	(0.000)	(0.000)	(0.002)	(0.001)
Adj R^2	0.877	0.544	0.993	0.918	0.900	0.879
MAPE, %	0.109	0.122	0.005	0.031	0.078	0.036
Assets	182	64	11	40	21	46
Months	180	180	180	180	180	180

7 Conclusion

We propose a parsimonious non-linear single factor asset pricing model $r_{it} = h(f_t\lambda_i) + \epsilon_{it}$. Despite its parsimony, the HFL model represents exactly any non-linear model with an arbitrary number of factors and factor loadings as a consequence of the Kolmogorov-Arnold representation theorem. Empirically, we jointly estimate the link function, the time factor, and the loading with the sieve-based least square estimator.

Using 171 test assets across major asset classes across U.S. equity, U.S. and international bonds, commodities, and currencies, we show that this one factor model delivers superior cross-sectional empirical asset pricing performance with a low-dimensional approximation of the link function. Moreover, controlling for the HFL model, the majority of the established factors for the cross-section of asset returns becomes insignificant. Additionally, we use the HFL model to construct a tradable strategy that generates statistically significant and sizable risk-adjusted long-short excess returns

There are two other broader points we want to make. First, nonlinearity and the higher-order interactions are the key feature of the HFL model. Either as a stand alone tool or incorporating it in other asset pricing settings thus allows a parsimonious, yet general, way to capture the importance of these effects compared to linear asset pricing models. Second, there is significant recent interest in using neural nets and artificial intelligence models in asset-pricing. It is interesting to speculate why these models may deliver superior performance to the existing asset-pricing models. One answer is the Kolmogorov-Arnold representation theorem. This theorem shows that any nonlinear continuous function of many variables can be represented as a continuous outer layer and an inner layer consisting of the sum of functions of single variables. Hecht-Nielsen (1987) argues that this is exactly the structure of a general neural network. Our HFL model thus offers a parsimonious method to capture the performance of neural network models and to evaluate their performance for asset pricing.

References

Adrian, Tobias, Erkko Etula, and Tyler Muir. 2014. "Financial intermediaries and the cross-section of asset returns." *Journal of Finance* 69 (6): 2557–2596.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006. "The cross-section of volatility and expected returns." *Journal of Finance* 61 (1): 259–299.

Arnold, Vladimir. 1957. "On functions of three variables." *Proceedings of the USSR Academy of Sciences* 114 679–681.

Asness, Clifford S, Tobias J Moskowitz, and Lasse Heje Pedersen. 2013. "Value and momentum everywhere." *Journal of Finance* 68 (3): 929–985.

Bansal, Ravi, and Salim Viswanathan. 1993. "No arbitrage and arbitrage pricing: A new approach." *Journal of Finance* 48 (4): 1231–1262.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81 (2): 608–650.

Bollerslev, Tim, Benjamin Hood, John Huss, and Lasse Heje Pedersen. 2018. "Risk everywhere: Modeling and managing volatility." *Review of Financial Studies* 31 (7): 2729–2773.

Borri, Nicola, Denis Chetverikov, Yukun Liu, and Aleh Tsyvinski. 2024. "Factor model with nonparametric link function." *Working Paper*.

Brennan, Michael J, Tarun Chordia, and Avanidhar Subrahmanyam. 1998. "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns." *Journal of Financial Economics* 49 (3): 345–373.

Carhart, Mark M. 1997. "On persistence in mutual fund performance." *Journal of Finance* 52 (1): 57–82.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning." *Econometrics Journal* 21 (1): .

Connor, Gregory, and Robert A Korajczyk. 1986. "Performance measurement with the arbitrage pricing theory: A new framework for analysis." *Journal of Financial Economics* 15 (3): 373–394.

Connor, Gregory, and Robert A Korajczyk. 1988. "Risk and return in an equilibrium APT: Application of a new test methodology." *Journal of Financial Economics* 21 (2): 255–289.

Coppejans, Mark. 2004. "On Kolmogorov's representation of functions of several variables by functions of one variable." *Journal of Econometrics* 123 1–31.

Datar, Vinay T, Narayan Y Naik, and Robert Radcliffe. 1998. "Liquidity and stock returns: An alternative test." *Journal of Financial Markets* 1 (2): 203–219.

Fama, Eugene F., and Kenneth R. French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1): 3–56.

Fama, Eugene F, and Kenneth R French. 1996. "Multifactor explanations of asset pricing anomalies." *Journal of Finance* 51 (1): 55–84.

Fama, Eugene F, and Kenneth R French. 2015. "A five-factor asset pricing model." *Journal of Financial Economics* 116 (1): 1–22.

Fama, Eugene F, and James D MacBeth. 1973. "Risk, return, and equilibrium: Empirical tests." *Journal of Political Economy* 81 (3): 607–636.

Fama, Eugene, and Kenneth French. 1992. "The cross-section of expected stock returns." *Journal of Finance* 47 (2): 427–465.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. "Taming the factor zoo: A test of new factors." *Journal of Finance* 75 (3): 1327–1370.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. "Empirical asset pricing via machine learning." *Review of Financial Studies* 33 (5): 2223–2273.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2021. "Autoencoder asset pricing models." *Journal of Econometrics* 222 (1): 429–450.

Harvey, Campbell R, and Akhtar Siddique. 2000. "Conditional skewness in asset pricing tests." *Journal of Finance* 55 (3): 1263–1295.

Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction*. Volume 2. Springer.

Haugen, Robert A, and Nardin L Baker. 1996. "Commonality in the determinants of expected stock returns." *Journal of Financial Economics* 41 (3): 401–439.

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

Hecht-Nielsen, Robert. 1987. "Kolmogorov's mapping neural network existence theorem." In *Proceedings of the International Conference on Neural Networks*, Volume 3. 11–14, IEEE press New York, NY, USA.

Hutchinson, James M, Andrew W Lo, and Tomaso Poggio. 1994. "A nonparametric approach to pricing and hedging derivative securities via learning networks." *Journal of Finance* 49 (3): 851–889.

Jagannathan, Ravi, and Zhenyu Wang. 1996. "The conditional CAPM and the cross-section of expected returns." *Journal of Finance* 51 (1): 3–53.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *Journal of Finance* 48 (1): 65–91.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen. 2023. "Is there a replication crisis in finance?" *Journal of Finance* 78 (5): 2465–2518.

Kelly, Bryan, Semyon Malamud, and Lasse Heje Pedersen. 2023. "Principal portfolios." *Journal of Finance* 78 (1): 347–387.

Kelly, Bryan, Semyon Malamud, and Kangying Zhou. 2024. "The virtue of complexity in return prediction." *Journal of Finance* 79 (1): 459–503.

Kelly, Bryan T, Seth Pruitt, and Yinan Su. 2019. "Characteristics are covariances: A unified model of risk and return." *Journal of Financial Economics* 134 (3): 501–524.

Koijen, Ralph SJ, Hanno Lustig, and Stijn Van Nieuwerburgh. 2017. "The cross-section and time series of stock and bond returns." *Journal of Monetary Economics* 88 50–69.

Koijen, Ralph SJ, Tobias J Moskowitz, Lasse Heje Pedersen, and Evert B Vrugt. 2018. "Carry." *Journal of Financial Economics* 127 (2): 197–225.

Koijen, Ralph SJ, Robert J Richmond, and Motohiro Yogo. 2024. "Which investors matter for equity valuations and expected returns?" *Review of Economic Studies*.

Koijen, Ralph SJ, and Motohiro Yogo. 2019. "A demand system approach to asset pricing." *Journal of Political Economy* 127 (4): 1475–1515.

Kolmogorov, Andrey. 1956. "On the representation of continuous functions of seversal variables by superpositions of continuous functions of a smaller number of variables." *Proceedings of the USSR Academy of Sciences* 108 179–182.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2020. "Shrinking the cross-section." *Journal of Financial Economics* 135 (2): 271–292.

Kraus, Alan, and Robert H Litzenberger. 1976. "Skewness preference and the valuation of risk assets." *Journal of Finance* 31 (4): 1085–1100.

Kurkova, Vera. 1991. "Kolmogorov's theorem is relevant." *Neural computation* 3 617–622.

Lettau, Martin, and Sydney Ludvigson. 2001. “Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying.” *Journal of Political Economy* 109 (6): 1238–1287.

Lettau, Martin, Matteo Maggiori, and Michael Weber. 2014. “Conditional risk premia in currency markets and other asset classes.” *Journal of Financial Economics* 114 (2): 197–225.

Lettau, Martin, and Markus Pelger. 2020. “Factors that fit the time series and cross-section of stock returns.” *Review of Financial Studies* 33 (5): 2274–2325.

Lintner, John. 1965. “Security prices, risk, and maximal gains from diversification.” *Journal of Finance* 20 (4): 587–615.

Ludvigson, Sydney C, and Serena Ng. 2009. “Macro factors in bond risk premia.” *Review of Financial Studies* 22 (12): 5027–5067.

Maiorov, Vitaly, and Allan Pinkus. 1999. “Lower bounds for approximation by MLP neural networks.” *Neurocomputing* 25 81–91.

Newey, Whitney K, and Kenneth D West. 1987. “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.” *Econometrica* 55 (3): 703–708.

Nucera, Federico, Lucio Sarno, and Gabriele Zinna. 2024. “Currency risk premiums redux.” *Review of Financial Studies* 37 (2): 356–408.

Onatski, Alexei. 2012. “Asymptotics of the principal components estimator of large factor models with weakly influential factors.” *Journal of Econometrics* 168 (2): 244–258.

Pástor, L’uboš, and Robert F Stambaugh. 2003. “Liquidity risk and expected stock returns.” *Journal of Political Economy* 111 (3): 642–685.

Schmidt-Hieber, Johannes. 2021. “The Kolmogorov-Arnold representation theorem revisited.” *Neural Networks* 137 119–126.

Schölkopf, Bernhard, Alexander Smola, and Klaus-Robert Müller. 1998. "Nonlinear component analysis as a kernel eigenvalue problem." *Neural Computation* 10 (5): 1299–1319.

Sharpe, William F. 1964. "Capital asset prices: A theory of market equilibrium under conditions of risk." *Journal of Finance* 19 (3): 425–442.

Online Appendix

A Econometrics Appendix

A.1 Proof of Theorem 3.1

It follows from the definition of the least-squares sieve estimator and Assumption 3.2 that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-\widehat{h}(\widehat{f_t\lambda_i}))^2 \le \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-h_K(f_t\lambda_i))^2.$$

Substituting here $r_{it}=h(f_t\lambda_i)+\varepsilon_{it}$, we obtain

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h(f_t\lambda_i))^2 \\ \le \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(h_K(f_t\lambda_i)-h(f_t\lambda_i))^2+\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i)).\end{aligned}$$

In addition, by the triangle inequality,

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))^2 \\ \le \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h(f_t\lambda_i))^2+\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(h_K(f_t\lambda_i)-h(f_t\lambda_i))^2.\end{aligned}$$

Combining the last two displayed equations, we have

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))^2 \\ \le \frac{4}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(h_K(f_t\lambda_i)-h(f_t\lambda_i))^2+\frac{4}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i)).\end{aligned} \tag{9}$$

The first term on the right-hand side here can be bounded using Assumption 3.2. To bound the second term, we use the empirical process theory. Denote $\mathcal{E} = [-L, L]^{K+1} \times [0,1]^T \times [0,1]^N$ and let d be the metric on $\mathcal{E}$ defined so that for each $\xi_1 = (\{c_{j1}\}_{j=0}^K, \{\phi_{t1}\}_{t=1}^T, \{l_{i1}\}_{i=1}^N)$ and $\xi_2 = (\{c_{j2}\}_{j=0}^K, \{\phi_{t2}\}_{t=1}^T, \{l_{i2}\}_{i=1}^N)$ in $\mathcal{E}$, we have

$$d(\xi_1,\xi_2)^2 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(\sum_{j=0}^K c_{j1}(\phi_{t1}l_{i1})^j - \sum_{j=0}^K c_{j2}(\phi_{t2}l_{i2})^j\right)^2.$$

It is straightforward to show that covering numbers $N(\epsilon, \mathcal{E}, d)$ of the metric space $(\mathcal{E}, d)$ satisfy

$$N(\epsilon,\mathcal{E},d) \leq \begin{cases}(1+2L(K+1)/\epsilon)^{T+N+K+1}, & \text{if } \epsilon \leq 2L(K+1),\\ 1 & \text{if } \epsilon > 2L(K+1).\end{cases} \tag{10}$$

It is also well–known that the corresponding packing numbers $D(\epsilon,\mathcal{E},d)$ satisfy $D(\epsilon,\mathcal{E},d) \leq N(\epsilon/2,\mathcal{E},d)$ for all $\epsilon > 0$; see page 98 of Van der Vaart and Wellner (1996) for the definitions of covering and packing numbers and relations between them.

Further, for each $\xi = (\{c_j\}_{j=0}^K, \{\phi_t\}_{t=1}^T, \{l_i\}_{i=1}^N)$ in $\mathcal{E}$, denote

$$X_\xi = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T \varepsilon_{it}\left(\sum_{j=0}^K c_j(\phi_t l_i)^j\right).$$

Then there exists a constant $C > 0$ such that

$$\|X_{\xi_1} - X_{\xi_2}\|_{\psi_2} \leq C d(\xi_1,\xi_2), \quad \text{for all } \xi_1,\xi_2 \in \mathcal{E}, \tag{11}$$

by Assumption 3.1(ii) and Proposition 2.6.1 in Vershynin (2018). Thus, denoting $\delta = 1/\sqrt{NT}$, it follows from Corollary 2.2.8 in Van der Vaart and Wellner (1996), applied to X_ξ/C instead of X_ξ, that if

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T(\widehat{h}(\widehat{f_t\lambda_i}) - h_K(f_t\lambda_i))^2 \leq \delta^2, \tag{12}$$

then

$$\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))\right| \leq \sup_{\substack{\xi_1,\xi_2\in\mathcal{E}:\\ d(\xi_1,\xi_2)\leq\delta}}|X_{\xi_1}-X_{\xi_2}| \lesssim_P \int_0^\delta \sqrt{\log D(\epsilon,\mathcal{E},d)}d\epsilon$$

Here, using (10), $D(\epsilon,\mathcal{E},d)\leq N(\epsilon/2,\mathcal{E},d)$, and $\delta=1/\sqrt{NT}$, we have

$$\begin{aligned}\int_0^\delta \sqrt{\log D(\epsilon,\mathcal{E},d)}d\epsilon &\lesssim \int_0^\delta \sqrt{(T+N+K+1)\log\left(\frac{8L(K+1)}{\epsilon}\right)}d\epsilon\\ &\lesssim \delta\sqrt{(T+N+K+1)\log\left(\frac{8L(K+1)}{\delta}\right)} \lesssim \sqrt{\frac{(T+N+K)\log(TNK)}{NT}}\end{aligned}$$

by standard integral calculations.

It thus remains to consider the case when (12) does not hold. To cover this case, denote $\xi_0=(\{h_j^K\}_{j=0}^K,\{f_t\}_{t=1}^T,\{\lambda_i\}_{i=1}^N)$. Also, for all $\xi\in\mathcal{E}$, denote $Y_\xi=(X_\xi-X_{\xi_0})/d(\xi,\xi_0)$. Then $\|Y_\xi\|_{\psi_2}\leq C$ by (11). Also, for all $\xi_1,\xi_2\in\mathcal{E}$ such $d(\xi_1,\xi_0)>\delta$,

$$\begin{aligned}\|Y_{\xi_1}-Y_{\xi_2}\|_{\psi_2} &= \left\|\frac{X_{\xi_1}-X_{\xi_0}}{d(\xi_1,\xi_0)}-\frac{X_{\xi_2}-X_{\xi_0}}{d(\xi_2,\xi_0)}\right\|_{\psi_2} \leq \frac{\|X_{\xi_1}-X_{\xi_2}\|_{\psi_2}}{d(\xi_1,\xi_0)}+\|X_{\xi_2}-X_{\xi_0}\|_{\psi_2}\left|\frac{1}{d(\xi_1,\xi_0)}-\frac{1}{d(\xi_2,\xi_0)}\right|\\ &\leq \frac{Cd(\xi_1,\xi_2)}{\delta}+\frac{C|d(\xi_1,\xi_0)-d(\xi_2,\xi_0)|}{\delta}\leq\frac{2Cd(\xi_1,\xi_2)}{\delta},\end{aligned}$$

where the first and the third inequalities follow from the triangle inequality, and the second from (11). Hence, denoting $\mathcal{E}_1=\{\xi\in\mathcal{E}\colon d(\xi,\xi_0)>\delta\}$ and letting d_1 be the metric $\mathcal{E}_1$ such that $d_1(\xi_1,\xi_2)=\|Y_{\xi_1}-Y_{\xi_2}\|_{\psi_2}$ for all $\xi_1,\xi_2\in\mathcal{E}_1$, it follows that the packing numbers $D(\epsilon,\mathcal{E}_1,d_1)$ and the covering numbers $N(\epsilon,\mathcal{E}_1,d_1)$ of the metric space $(\mathcal{E}_1,d_1)$ satisfy

$$D(\epsilon,\mathcal{E}_1,d_1)\leq N(\epsilon/2,\mathcal{E}_1,d_1)\leq N\left(\frac{\delta\epsilon}{4C},\mathcal{E},d\right),\quad \text{for all } \epsilon>0.$$

Thus, it follows from Corollary 2.2.8 in Van der Vaart and Wellner (1996) that if (12) does not hold, then

$$\frac{\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))\right|}{\sqrt{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))^2}} \leq \sup_{\xi\in\mathcal{E}_1}|Y_\xi| \lesssim_P \int_0^\infty \sqrt{\log D(\epsilon,\mathcal{E}_1,d_1)}d\epsilon$$
$$= \int_0^{2C}\sqrt{\log D(\epsilon,\mathcal{E}_1,d_1)}d\epsilon$$
$$\leq \int_0^{2C}\sqrt{\log N\left(\frac{\delta\epsilon}{4C},\mathcal{E},d\right)}d\epsilon$$
$$= \frac{4C}{\delta}\int_0^{\delta/2}\sqrt{\log N(\epsilon,\mathcal{E},d)}d\epsilon \lesssim \sqrt{(T+N+K)\log(TNK)}$$

by standard integral calculations. Therefore, we conclude that

$$\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))\right|$$
$$\lesssim_P \sqrt{\frac{(T+N+K)\log(TNK)}{NT}}\left(\frac{1}{\sqrt{NT}}+\sqrt{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))^2}\right).$$

Combining this bound with (9) and using Assumption 3.2, it follows that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{h}(\widehat{f_t\lambda_i})-h_K(f_t\lambda_i))^2 \lesssim_P \frac{(T+N+K)\log(TNK)}{NT}+K^{-2\alpha}.$$

The asserted claim now follows by combining this bound with the triangle inequality and Assumption 3.2.

A.2 Estimation procedure and cross-sectional asset pricing tests

The HFL model

$$r_{it} = h(f_t\lambda_i)+\varepsilon_{it}, \quad i=1,\dots,N, t=1,\dots,T, \tag{13}$$

implies that the coefficients α and β in the linear regression

$$E[r_{it}] = \alpha + \beta E[h(f_t \lambda_i)] + \epsilon_i, \quad i = 1, \dots, N,$$

should satisfy the following restrictions: $\alpha = 0$ and $\beta = 1$. In Section 5, we tested these restrictions by (i) estimating the HFL model to get $\widehat{h}$, $\{\widehat{f_t}\}_{t=1}^T$, and $\{\widehat{\lambda_i}\}_{i=1}^N$ and (ii) running the OLS of $T^{-1}\sum_{t=1}^T r_{it}$ on the constant one and $T^{-1}\sum_{t=1}^T \widehat{h}(\widehat{f_t}\widehat{\lambda_i})$ to get $\widehat{\alpha}$ and $\widehat{\beta}$ and comparing them with 0 and 1, respecticely. Here, we show that this test indeed has power in the sense that unless model (13) is correct, the probability limits α_0 and β_0 of the OLS estimates $\widehat{\alpha}$ and $\widehat{\beta}$ generally satisfy the restrictions $\alpha_0 = 0$ and $\beta_0 = 1$ only if the factor f_t is constant over time, which is not the case in practice.

Assume, for simplicity, that we approximate the function h with the following polynomial of the second order:

$$h(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2.$$

In this case, the estimator for the HFL model that we propose is:

$$(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \{\hat{f}_t\}, \{\hat{\lambda}_i\}) = \arg \min_{\gamma_0,\gamma_1,\gamma_2,\{\phi_t\},\{l_i\}} \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \left(r_{it} - \gamma_0 - \gamma_1\phi_t l_i - \gamma_2(\phi_t l_i)^2\right)^2.$$

The first order conditions of this optimization problem with respect to $\gamma_0, \gamma_1, \gamma_2$ are

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T (r_{it} - \hat{\gamma}_0 - \hat{\gamma}_1\hat{f}_t\hat{\lambda}_i - \hat{\gamma}_2(\hat{f}_t\hat{\lambda}_i)^2) = 0,$$

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T (r_{it} - \hat{\gamma}_0 - \hat{\gamma}_1\hat{f}_t\hat{\lambda}_i - \hat{\gamma}_2(\hat{f}_t\hat{\lambda}_i)^2)\hat{f}_t\hat{\gamma}_i = 0,$$

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T (r_{it} - \hat{\gamma}_0 - \hat{\gamma}_1\hat{f}_t\hat{\lambda}_i - \hat{\gamma}_2(\hat{f}_t\hat{\lambda}_i)^2)(\hat{f}_t\hat{\gamma}_i)^2 = 0,$$

respectively. Substituting

$$\hat{h}(\hat{f}_t\hat{\lambda}_i) = \hat{\gamma}_0 + \hat{\gamma}_1\hat{f}_t\hat{\lambda}_i + \hat{\gamma}_2(\hat{f}_t\hat{\lambda}_i)^2$$

into these equations, we obtain

$$\frac{1}{NT}\sum_{i=1}^{T}\sum_{t=1}^{T}(r_{it}-\hat{h}(\hat{f}_t\hat{\lambda}_i))=0, \tag{14}$$

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-\hat{h}(\hat{f}_t\hat{\lambda}_i))\hat{f}_t\hat{\lambda}_i=0,$$

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-\hat{h}(\hat{f}_t\hat{\lambda}_i))(\hat{f}_t\hat{\lambda}_i)^2=0.$$

Further, multiplying these equations by $\hat{\gamma}_0$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$, respectively, and taking the sum, we obtain in particular that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-\hat{h}(\hat{f}_t\hat{\lambda}_i))(\hat{\gamma}_0+\hat{\gamma}_1\hat{f}_t\hat{\lambda}_i+\hat{\gamma}_2(\hat{f}_t\hat{\lambda}_i)^2)=0,$$

or, equivalently, that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it}-\hat{h}(\hat{f}_t\hat{\lambda}_i))\hat{h}(\hat{f}_t\hat{\lambda}_i)=0. \tag{15}$$

Taking the probability limits on the left-hand sides of equations (14) and (15) and using Theorem 3.1, it thus follows that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E[r_{it}-h(f_t\lambda_i)]=o(1), \tag{16}$$

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E[(r_{it}-h(f_t\lambda_i))h(f_t\lambda_i)]=o(1). \tag{17}$$

Note that these equations are true regardless of whether the HFL model is correct or not. In case the HFL model is not correct, the quantities $h(f_t\lambda_i)$ should be understood as the probability limits of the estimators $\widehat{h}(\widehat{f_t\lambda_i})$. Intuitively, equations (16) and (17) imply that the mean, across time and assets, of the residuals should be zero and that the covariance between the residuals and the predictor should be zero as well. Note also that even though we used the second-order polynomial approximation to the function h, equations (16) and (17) remain valid for polynomial approximation of any order.

Next, to derive the probability limits α_0 and β_0 of the OLS estimates $\widehat{\alpha}$ and $\widehat{\beta}$, observe that

$$(\hat{\alpha},\hat{\beta}) = \arg\min_{\alpha,\beta} \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T} r_{it} - \alpha - \beta\times\frac{1}{T}\sum_{t}\hat{h}(\hat{f}_t\hat{\lambda}_i)\right)^2.$$

The first-order condition with respect to α is

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T} r_{it} - \hat{\alpha} - \hat{\beta}\times\frac{1}{T}\sum_{t=1}^{T}\hat{h}(\hat{f}_t\hat{\lambda}_i)\right) = 0.$$

The first-order condition with respect to β is

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T} r_{it} - \hat{\alpha} - \hat{\beta}\times\frac{1}{T}\sum_{t=1}^{T}\hat{h}(\hat{f}_t\hat{\lambda}_i)\right)\left(\frac{1}{T}\sum_{t=1}^{T}\hat{h}(\hat{f}_t\hat{\lambda}_i)\right) = 0.$$

These two conditions can be trivially rewritten as

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(r_{it} - \hat{\alpha} - \hat{\beta}\hat{h}(\hat{f}_t\hat{\lambda}_i)) = 0, \tag{18}$$

$$\frac{1}{NT}\sum_{i=1}^{T}\sum_{t=1}^{T}(r_{it} - \hat{\alpha} - \hat{\beta}\hat{h}(\hat{f}_t\hat{\lambda}_i))\widehat{H}_i = 0, \tag{19}$$

where we denoted

$$\widehat{H}_i = \frac{1}{T}\sum_{t}\hat{h}(\hat{f}_t\hat{\lambda}_i), \quad \text{for all } i = 1,\dots,N.$$

Taking the probability limits of these two equations and using Theorem 3.1, it thus follows that α_0 and β_0 satisfy

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E[r_{it} - \alpha_0 - \beta_0 h(f_t\lambda_i)] = o(1), \tag{20}$$

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E[(r_{it} - \alpha_0 - \beta_0 h(f_t\lambda_i))\bar{H}_i] = o(1), \tag{21}$$

where we denoted

$$\bar{H}_i = \frac{1}{T}\sum_{t=1}^{T}E[h(f_t\lambda_i)], \quad i = 1,\dots,N.$$

Comparing Equations (20)-(21) with (16)-(17), it thus follows that unless the HFL model is correct, we generally have $\alpha_0 = 0$ and $\beta_0 = 1$ only if $\overline{H}_i = h(f_t\lambda_i)$ for all $i = 1, \ldots, N$ and $t = 1, \ldots, T$, which is equivalent to $h(f_t\lambda_i) = h(f_s\lambda_i)$ for all $i = 1, \ldots, N$ and $t, s = 1, \ldots, T$, meaning that there is no time-series variation in the factor f_t.

A.3 Related Econometrics Literature

Our model (1) is related to but different from those in the recent literature on non-linear factor and two-way fixed effect models in econometrics. For example, Fernandez-Val and Weidner (2016), Boneva and Linton (2017), Charbonneau (2017), Chen, Fernandez-Val, and Weidner (2021), Wang (2022), and Gao, Liu, Peng, and Yan (2023) considered the model of the form $r_{it} \mid x_{it} \sim h(\cdot \mid x_{it}^\top \beta + f_t^\top \lambda_i)$ and its variants, where x_{it} is a vector of covariates, β is a vector of coefficients, and h is some function. The key difference from our model here is that their function h is assumed to be known. Mugnier and Wang (2022) studied the model of the form $r_{it} \mid x_{it} \sim h(\cdot \mid x_{it}^\top + f_t + \gamma_i)$, where the function $v \mapsto h(\cdot \mid v)$ is unknown but strictly increasing, and assumed the existence of the compensating variable, which is a component of x_{it} that can "undo" the variation in $f_t + \lambda_i$. Chen, Dolado, and Gonzalo (2021) and Ando and Bai (2020) analyzed the quantile factor model, where the τth quantile of the distribution of y_{it} is equal to $f_t^\top \lambda_i$ or to $x_{it}^\top \beta + f_t^\top \lambda_i$, respectively. Chernozhukov, Fernandez-Val, and Weidner (2024) studied the distribution model, where $P(r_{it} \leq \cdot \mid x_{it}) = \Lambda(x_{it}^\top \beta(\cdot) + f_t(\cdot) + \lambda_i(\cdot))$ for some known function Λ.

In addition, our model (1) is related to those in the literature on network models but our analysis is different. For example, Zeleneev (2020) worked with the model $r_{ij} = x_{ij}^\top \beta + h(\xi_i, \xi_j) + e_{ij}$, where $i, j = 1, \ldots, N$, x_{ij} is a pair-specific covariate, ξ_i and ξ_j are individual effects, and h is an unknown function, but his analysis is about identification and estimation of the vector of parameters β, which relies on partialling out $h(\xi_i, \xi_j)$ rather than estimating it. Gao (2020) worked with the model of the form $r_{ij} \mid (x_i, x_j) \sim h(\cdot \mid w(x_i, x_j) + \xi_i + \xi_j)$, where w is an unknown function but $v \mapsto h(\cdot \mid v)$ is *strictly increasing*.

More generally, it appears that non-linear factor models were first developed in the psychology literature; see Bartlett (1953) and McDonald (1962). McDonald (1979) developed an estimator of parametric non-linear factor models. Zhu and Lee (1999) studied a non-linear factor model of the form $r_{it} = \lambda_i^\top h(f_t) + \varepsilon_{it}$, where h is a vector of known functions. Other relevant references to classic results can be found in a review Yalchin and Amemiya (2001). Wang (2022) also provided several references for the analysis of non-linear factor models with small N.

B Empirical Appendix

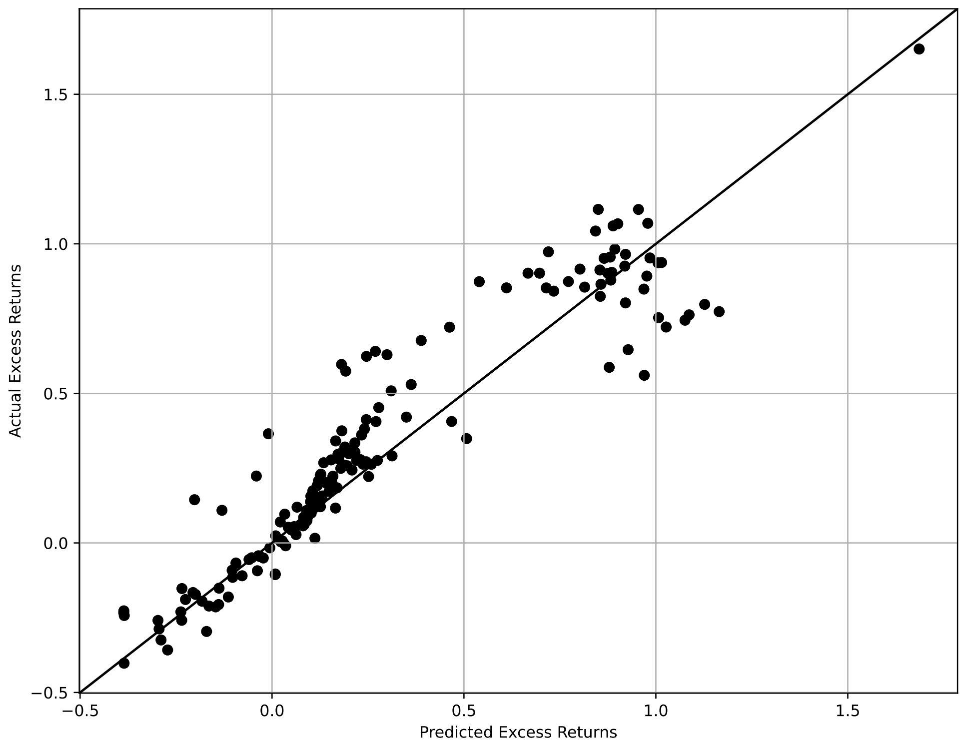
This section presents the results of additional cross-sectional asset pricing tests. Section B.1 presents results of the cross-sectional tests with no intercept. Section B.2 presents results of the cross-sectional tests by asset class. Section B.3 presents further results about the significant factors in the double-selection Lasso estimation. Section B.4 presents a comparison with an alternative factor zoo. Section B.5 presents a comparison with PCA models with up to six factors. Section B.6 presents a further robustness analysis considering the sample before and after the 2008 financial crisis. Section B.7 presents a correlation analysis with benchmark equity factors.

B.1 Cross-sectional asset pricing tests with no constant

Figure A.1 plots the average asset excess returns versus the predicted excess returns using the HFL model with the constraint that the intercept is equal to zero (i.e., $E[r_{it}] = \beta E[h(\lambda_i f_t)] + \epsilon_i$). Note that the latter is one of the implications of our factor model, which we test in the main analysis. The figure reveals that, as for the specification based on the model with a constant, all assets line up along the 45 degree line.

Figure A.1: Actual vs Predicted Expected Returns – No Intercept

The figure plots actual average excess returns on all tested assets ($E[r_{it}]$) versus predicted excess returns using the cross-sectional model based on the HFL model with no intercept: $E[r_{it}] = \beta E[h(\lambda_i f_t)] + \epsilon_i$. The red line denotes the 45 degree line. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.



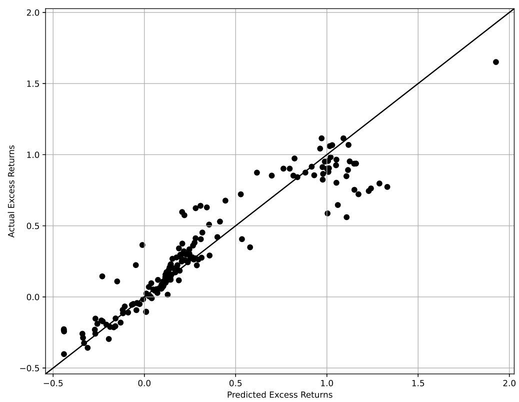
B.2 Cross-sectional asset pricing tests by asset class

In this section we present visual evidence of the cross-sectional performance of the HFL model in asset pricing tests by asset class. Figure A.2 plots, separetely for all assets and for the different asset classes, actual average excess returns ($E[r_{it}]$) versus predicted excess returns using the cross-sectional model based on the HFL model: $E[r_{it}] = \alpha + \beta E[h(\lambda_i f_t)] + \epsilon_i$. The asset classes are: equities, US bonds, international bonds, commodities and foreign exchange. The figures show that, also for the individual asset classes, as for the model tested on all assets, individual assets line up along the 45 degree line.

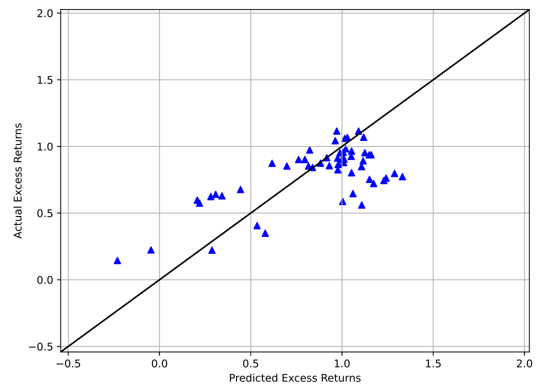
Figure A.2: Actual vs Predicted Expected Returns – By Asset Class

The figure plots, separetely for all assets and for the different asset classes, actual average excess returns ($E[r_{it}]$) versus predicted excess returns using the cross-sectional model based on the HFL model: $E[r_{it}] = \alpha + \beta E[h(\lambda_i f_t)] + \epsilon_i$. The asset classes are: equities, US bonds, international bonds, commodities and foreign exchange. The HFL model is estimated using all assets. The red line denotes the 45 degree line. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

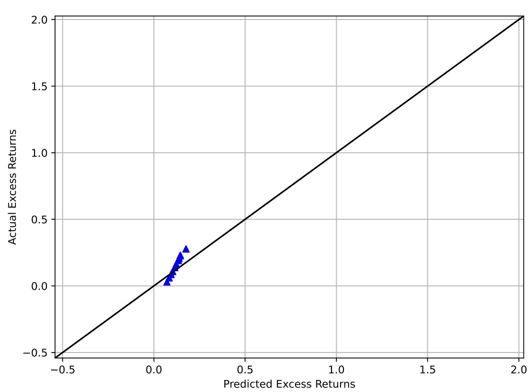
(a) All assets



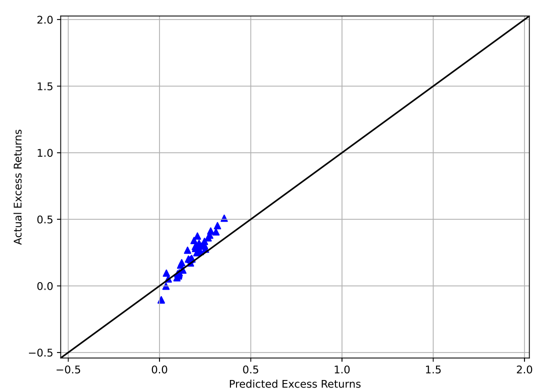
(b) Equities



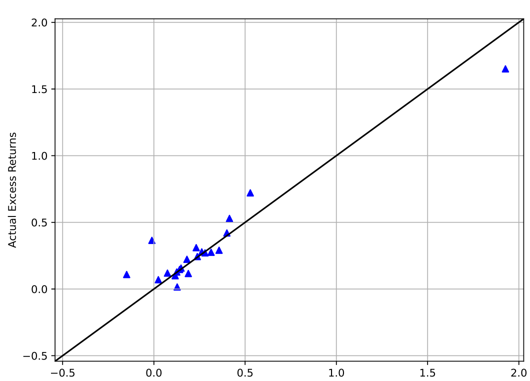
(c) US Bonds



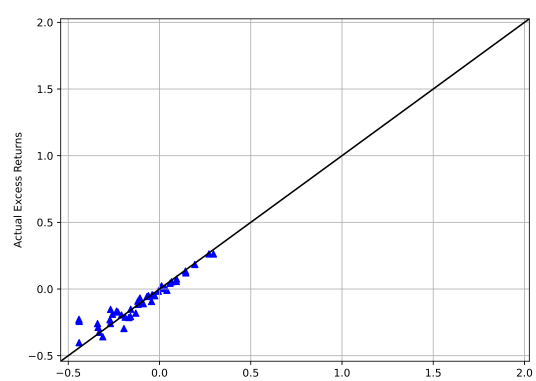
(d) International Bonds



(e) Commodities



(f) Foreign exchange



B.3 Significant factors in the factor zoo using double-selection Lasso

In this section, we report further results of the double-selection Lasso estimation discussed in Section 5.6. Specifically, for each factor in the factor zoo, we estimate the double-selection Lasso model using all the remaining factors from the factor zoo and the HFL component as controls. The factor zoo includes factors for 153 characteristics in 13 themes, using data from 93 countries and four regions, constructed by Jensen, Kelly, and Pedersen (2023). Table A.1 summarizises the additional results of the double-selection estimation, listing all the factors from the factor zoo significant at least at the 10% confidence level, along the corresponding p-value and t-statistics. The factor and factor names are from Jensen, Kelly, and Pedersen (2023).

Table A.1: Significant Factors in the Factor Zoo Using Double-Selection Lasso

This table reports further results from the double-selection Lasso estimation for the factors in the factor zoo. For each factor in the factor zoo, we estimation the double-selection Lasso model using all the remaining factors from the factor zoo and the HFL component as controls. The table lists all the factors from the factor zoo which are significant at least at the 10% confidence level, along the corresponding p-value and t-statistics. Standard errors are adjusted with the Fama-MacBeth procedure. The factor and the factor names are from Jensen, Kelly, and Pedersen (2023).

Factors in Factor Zoo	p-value	t-stat
at_turnover	0	3.969
beta_60m	0.037	2.093
beta_dimson_21d	0.078	1.765
eq_dur	0.037	2.094
inv_gr1a	0.094	1.682
niq_su	0.023	-2.275
nncoa_gr1a	0.021	-2.313
qmj	0.032	-2.158
rd_sale	0.07	1.82
resff3_12_1	0.004	2.894
sale_emp_gr1	0.046	2.005
seas_2_5an	0.017	-2.402
taccruals_ni	0.058	1.905
zero_trades_252d	0.054	-1.937

B.4 Comparison with the factors from Chen and Zimmermann (2022)

In section 5.6 of the main paper, we compared the HFL model with the factors from the factor zoo presented in Jensen, Kelly, and Pedersen (2023) and showed that they become insignificant after controlling for the HFL component. In this section, we show that also the factors from the factor zoo presented in Chen and Zimmermann (2022) become insignificant after controlling for the HFL component.

In particular, we consider 212 factors based on stock return predictors.[7] For each of these factors (denoted with f^{zoo}), separately, we first estimate the cross-sectional regression $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$. Next, we repeat the estimation additionally including the predicted values obtained with the HFL model (i.e., $E[h(\lambda_i f_t)]$). Figure A.3 summarizes our results. It plots the average of the absolute values of the t-statistics in a test of the intercepts and slope coefficients in the cross-sectional regressions, respectively, equal to zero. In the left panel of the figure, we consider the regressions which only include the average predicted values based on the factor zoo. In the right panel of the figure, we consider regressions which additionally include the average predicted values from the HFL model. The figure reveals that, for the models which do not include the HFL component, the estimates of the slope coefficients are on average significantly different from zero (average absolute t-statistics equal to 2.10). For these models, though, also the pricing error, captured by the estimates of the intercept, is on average statistically significant (average absolute t-statistics equal to 2.83). Furthermore, the figure reveals that, in models which additionally include the average predicted values from the HFL model, the slope coefficients associated with the factor zoo become not statistically different from zero (average absolute t-statistics equal to 0.80), while those associated with the HFL component are highly significant (average absolute t-statistics equal to 5.25). Moreover, for the models which includes the HFL component, the estimates of the pricing error also become not statistically different from zero (average t-statistics equal to 0.70).

[7] The data for the factor zoo is available through the Open Source Asset Pricing webpage.

Figure A.3: Comparison with the Factors from Chen and Zimmermann (2022)

This figure shows the average of the absolute values of the t-statistics in a test of the intercepts and slope coefficients in the cross-sectional regressions, respectively, equal to zero. Standard errors are adjusted with the Fama-MacBeth procedure. In the left panel, we consider the cross-sectional regressions $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$ for each factor from the factor zoo. In the right panel, we consider cross-sectional regressions which additionally include the average predicted values based on the HFL component (i.e., $E[h(\lambda_i f_t)]$). The vertical bars denote one standard deviation around the mean. The horizontal black dashed-line denote significance at the 10% confidence level. The data for 212 factors from the factor zoo are from Chen and Zimmermann (2022). The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

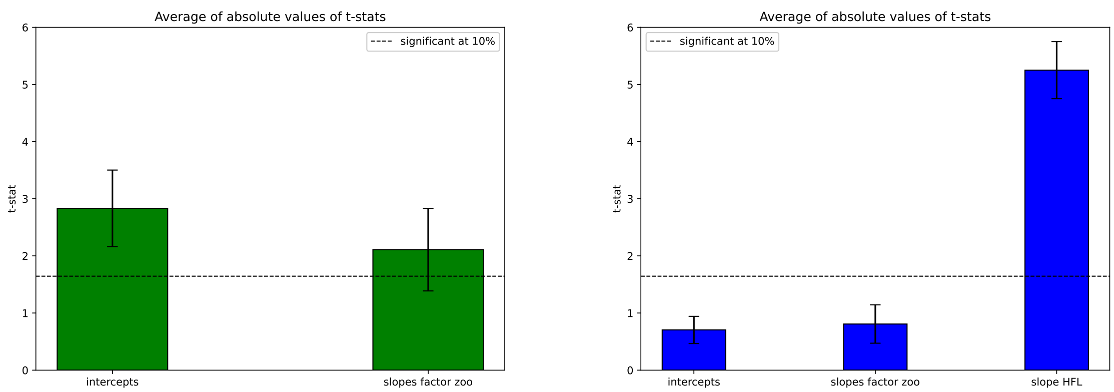


Table A.2 provides further details about the comparison with the factor zoo. If the HFL component is not included (Panel A), the fraction of models with a slope coefficient on the factor zoo which is significant at the 10% confidence level is approximately 80%. When we further include the HFL component, this fraction drops to less than 1%. Moreover, if the HFL component is included (Panel B), the fraction of models with a slope on the HFL component which is significant at the 1% confidence level is about 100%. In these models, the intercept is statisticall different from zero at a confidence level of 10% or below in only 0.5% of the cases. Finally, we find that only one of the 212 factors from the factor zoo is significant with a confidence level of 10%. This factor is the industry return of big firms of Hou (2007).

Furthermore, the table also reports the results of cross-sectional asset pricing regressions based on the double-selection Lasso method of Belloni, Chernozhukov, and Hansen (2014) and Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and Robins (2018) (Panel C). Double-selection Lasso allows estimating regressions where the number of right-hand side variables could be large, even larger than the sample size. In practice, double-selection Lasso estimates only one coefficient on the right-hand side at a time by means of a two-step selection method. In the first, factors with a low contribution to the cross-sectional pricing are excluded from the set of controls. In the second step, factors whose covariances with returns are highly correlated in the cross-section with the covariance between returns and a given factor are added to the set of controls. Feng, Giglio, and Xiu (2020) propose double-selection Lasso as model selection method to evaluate the contribution of factors in cross-sectional asset pricing regressions. The estimates based on the double-selection Lasso method confirm the results based on pairwise comparisons of the factor zoo with the HFL compoment. In particular, the HFL component using the double-selection Lasso method is statistically significant at the 1% level (t-statistics of 7.67). The fraction of slope coefficients on the factor zoo with p-values greater than 10% is 94.33.

B.5 Comparison with PCA models with up to six factors

In this section, we complement the comparison with PCA models (see Table 3). Specifically, we compare the HFL model with PCA models with up to six factors. Including additional principal components is important as some components might explain a large fraction of the time-series variation, but very little of the cross-sectional variation (e.g., the first component which is often only a "level" factor for the cross-section of asset returns). Additional components, on the contrary, might explain a small fraction of the time-series variation, but a large fraction of the cross-sectional variation in returns (see, e.g., Lettau and Pelger (2020), Kozak, Nagel, and Santosh (2020) or Giglio and Xiu (2021) and the review paper by Giglio, Kelly, and Xiu (2022)).

Table A.3 presents the results of cross-sectional asset pricing tests using all assets and three estimators of the latent factors related to PCA. The first estimator is based on the standard PCA of the covariance matrix of asset returns. The second estimator, called RP-PCA, is a generalization of PCA, proposed by Lettau and Pelger (2020), which includes a penalty term to account for the pricing errors in the cross-sectional regressions based on the RP-PCA factors. RP-PCA is motivated by the poor performance of standard PCA in identifying factors which are relevant to explain the cross-section of average excess returns (see, e.g., Onatski (2012) and Kozak, Nagel, and Santosh (2020)). The third estimator is the kernel-PCA, proposed by Schölkopf, Smola, and Müller (1998). Kernel-PCA (K-PCA) is a nonlinear form of PCA which allows for the separability of nonlinear data by projecting it onto a higher dimensional space where it becomes linearly separable using kernels.

In all models, we include the HFL component ($h(f_t\lambda_i)$), and evaluate the effect of including, respectively, the first four principal components, the first five principal components, and the first six principal components, separately for standard PCA, RP-PCA and K-PCA. The estimates from the cross-sectional asset pricing regressions reveal that the slope coefficient associated with the HFL component is statistically significant at conventional levels in all models, while the slopes associated with the principal components, for all three estimators we consider, are never statically different from zero.

Table A.2: Comparison with the Factors from Chen and Zimmermann (2022)

This table provides a comparison of the HFL model and the factor zoo. Panel A refers to cross-sectional regressions $E[r_{it}] = \alpha + \beta^{zoo} E[f_t^{zoo}] + \epsilon_i$ for each factor from the factor zoo. Standard errors are adjusted with the Fama-MacBeth procedure. Panel B refers to cross-sectional regressions which additionally include the average predicted values based on the HFL component (i.e., $E[h(\lambda_i f_t)]$). The two panels report the fraction (in percentage) of intercepts and slope coefficients with p-values less than 1%, between 1 and 5%, between 5 and 10% and larger than 10%. The p-values correspond to tests for the coefficient equal to zero. Panel C refers to the model which includes the HFL component and estimates based on the double-selection Lasso method, and reports the fraction of slope coefficients on the factor zoo with p-values less than 1%, between 1 and 5%, between 5 and 10% and larger than 10%. The HFL component using the double-selection Lasso method is statistically significant at the 1% level. The data for 212 factors from the factor zoo are from Chen and Zimmermann (2022). The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

$p(coeff=0)$	$p \leq 1\%$	$1\% < p \leq 5\%$	$5\% < p \leq 10\%$	$p > 10\%$
	Panel A: Model without HFL			
α	70.28	17.92	3.30	8.49
$E[f_t^{zoo}]$	24.05	45.75	8.49	21.69
	Panel B: Model with HFL			
α	0	0	0.48	99.52
$E[f_t^{zoo}]$	0	0	0.48	99.52
$h(f_t \lambda_i)$	1	0	0	0
	Panel C: Model with HFL and double-selection Lasso			
α	32.54	65.56	1.88	0
$E[f_t^{zoo}]$	0.47	0.94	4.24	94.33

Table A.3: Comparison with PCA Factor Models up to Degree Six

This table presents the results of cross-sectional asset pricing tests using all assets. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(\lambda_i f_t)] + \epsilon_i$ augmented by, respectively, the first four principal components, the first five principal components and the first six principal components, extracted by the test asset returns. For the principal components, we consider three estimators: the standard PCA based on the covariance matrix of asset returns; the risk premium PCA (RP-PCA) of Lettau and Pelger (2020), and the kernel PCA (K-PCA) of Schölkopf, Smola, and Müller (1998). We report the estimates for the slopes and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. The table additionally reports the regression adjusted R-squared, and the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	Standard PCA			RP-PCA			K-PCA		
	PC1–PC4	PC1–PC5	PC1–PC6	PC1–PC4	PC1–PC5	PC1–PC6	PC1–PC4	PC1–PC5	PC1–PC6
$h(f_t\lambda_i)$	0.657***	0.613***	0.595***	0.633***	0.515***	0.612***	0.734***	0.746***	0.757***
	(0.109)	(0.109)	(0.109)	(0.113)	(0.123)	(0.106)	(0.127)	(0.125)	(0.126)
PC1	0.014	0.013	0.012	0.016	0.021	0.014	0.007	0.008	0.006
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.016)	(0.016)	(0.016)
PC2	0.001	0.001	0.001	0.002	0.005	0.003	0.008	0.009	0.008
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
PC3	-0.005	-0.007	-0.009	0.003	0.005	0.005	-0.008	-0.010	-0.005
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.013)	(0.014)
PC4	-0.003	-0.003	-0.004	0.003	0.003	0.003	-0.008	-0.006	-0.002
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.016)	(0.015)	(0.015)
PC5		-0.008	-0.009		0.002	0.001		0.004	0.007
		(0.007)	(0.007)		(0.007)	(0.007)		(0.014)	(0.013)
PC6			-0.005			-0.008			-0.023
			(0.006)			(0.006)			(0.019)
α	0.001	0.001**	0.001***	0.001	0.001**	0.001**	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Adj R^2	0.916	0.936	0.943	0.919	0.932	0.937	0.902	0.903	0.917
MAPE, %	0.091	0.074	0.071	0.090	0.079	0.073	0.103	0.102	0.092
Assets	171	171	171	171	171	171	171	171	171
Months	360	360	360	360	360	360	360	360	360

B.6 Before and after the 2008 financial crisis

In this section, we present the results of cross-sectional asset pricing tests in two alternative samples. The first is a pre-financial crisis sample, which starts in January 1988 and ends in December 2008. The second is a post-financial crisis sample, which starts in January 2009 and ends in December 2017. Table A.4 presents the results for both all assets, and separately for each asset class. When we include all assets, the slope coefficient associated with the HKL component is statistically significant in both sample, and quantitatively very similar. The slope is equal to 0.753 in the pre-crisis sample, and 0.792 in the post-crisis sample. When we consider the different asset classes separately, we find a significant slope for each asset class, except for commodities, in both the pre-crisis and post-crisis samples.

B.7 Correlation with benchmark equity factors

In this section, we discuss the correlation matrix for the HFL component and benchmark equity factors. Specifically, we consider the Fama and French (1993)'s three factors, the additional two factors from the Fama and French (2015)'s five factors model, and the momentum factor of Jegadeesh and Titman (1993). All equity factors are from Ken French's website.

We first estimate $h(\lambda_i f_t)$ for each asset i. Next, we compute the sample correlation coefficient between the asset level HFL components and the asset level factor betas. The latter are slope coefficients in separate time-series regressions of asset excess returns on each benchmark equity factor.

Table A.5 reports the correlation matrix. The HFL component is highly positive correlated with assets' beta exposures to the US equity market excess return (correlation coefficient with MKTRF is 0.74), and positively related with assets' beta exposures to the size factor (correlation coefficient with SMB is 0.43), the value factor (correlation coefficient with HML is 0.48) and the operating profitability factor (correlation coefficient with RMW 0.24). Moreover, the HFL component is somewhat negatively related to the investment factor (correlation coefficient with

CMA is -0.10) and the momentum factor (correlation coefficient with MOM is -0.12). Finally, the HFL component is negatively correlated with assets' beta exposures to $\Delta \log(VIX)$.

Table A.4: Cross-Sectional Asset Pricing by Asset Class: Before and After the 2008 Financial Crisis

This table presents the results of cross-sectional asset pricing tests for alternative samples: before and after the 2008 financial crisis. The pre-financial crisis sample starts in January 1988 and ends in December 2008. The post-financial crisis sample starts in January 2009 and ends in December 2017. We estimate the cross-sectional linear model $E[r_{it}] = \alpha + \beta E[h(\lambda_i f_t)] + \epsilon_i$ and report the estimates for the slope ($h(f_t \lambda_i)$) and intercept (α), and standard errors adjusted with the Fama-MacBeth procedure in parenthesis. Panel A reports estimates for the pre financial crisis sample. Panel B reports estimates for the post financial crisis sample. The table columns report estimates for the different asset classes. The table additionally reports the regression adjusted R-squared, the mean absolute pricing error (MAPE) in percentage terms. *, **, and *** represent significance at the 1%, 5%, and 10% level. We further report the number of assets in the cross-sectional regression and the number of monthly observations for each asset used in the estimation of the averages of the excess returns and predictions from the model. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

Panel A: pre-crisis	All	Equities	US Bonds	Intl Bonds	Commodites	FX
$h(f_t \lambda_i)$	0.753***	0.438**	1.894**	1.271***	0.549	1.079***
	(0.188)	(0.217)	(0.783)	(0.276)	(0.476)	(0.144)
α	0.001	0.004	-0.001	-0.000	0.002	0.002
	(0.001)	(0.004)	(0.000)	(0.000)	(0.002)	(0.001)
Adj R^2	0.835	0.487	0.970	0.795	0.542	0.873
MAPE, %	0.124	0.132	0.009	0.048	0.177	0.057
Assets	171	53	11	40	21	46
Months	240	240	240	240	240	240

Panel B: post-crisis	All	Equities	US Bonds	Intl Bonds	Commodites	FX
$h(f_t \lambda_i)$	0.792**	0.417**	3.829**	2.259***	1.182	0.340**
	(0.374)	(0.169)	(1.673)	(0.620)	(1.080)	(0.131)
α	0.001	0.005	-0.003**	-0.001*	-0.002	-0.000
	(0.001)	(0.004)	(0.002)	(0.000)	(0.006)	(0.002)
Adj R^2	0.697	0.278	0.979	0.827	0.507	0.342
MAPE, %	0.169	0.158	0.011	0.054	0.275	0.075
Assets	171	53	11	40	21	46
Months	180	180	180	180	180	180

Table A.5: Correlation Matrix

This table reports the correlation matrix for the HFL component, the individual asset's beta exposures to the Fama and French (2015)'s five factors, and the momentum factor of Jegadeesh and Titman (1993), and $\Delta\log(VIX)$. For the HFL component, we first estimate $h(\lambda_i f_t)$ for each asset i, and then construct the cross-sectional average. The HFL model is based on a degree of the polynomial used to approximate the function $h(.)$ equal to 4.

	HFL	MKTRF	SMB	HML	RMW	CMA	MOM	$\Delta\log(VIX)$
HFL	1.00							
MKTRF	0.74	1.00						
SMB	0.43	0.64	1.00					
HML	0.48	0.28	0.14	1.00				
RMW	0.24	0.05	0.04	0.13	1.00			
CMA	-0.10	0.14	-0.01	-0.24	0.12	1.00		
MOM	-0.12	-0.30	-0.12	-0.15	-0.09	-0.10	1.00	
$\Delta\log(VIX)$	-0.65	-0.98	-0.68	-0.20	0.07	-0.10	0.32	1.00

Appendix References

Ando, Tomohiro, and Jushan Bai. 2020. "Quantile co-movement in financial markets: a panel quantile model with unobserved heterogeneity." *Journal of American Statistical Association* 115 266–279.

Bartlett, M. 1953. "Factor analysis in psychology as a statistician sees it." *Uppsala Symposium on Phychological Factor Analysis* 23–24.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81 (2): 608–650.

Boneva, Lena, and Oliver Linton. 2017. "A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance." *Journal of Applied Econometrics* 32 1226–1243.

Charbonneau, Karyne. 2017. "Multiple fixed effects in binary response panel data models." *Econometrics Journal* 20 1–13.

Chen, Andrew Y, and Tom Zimmermann. 2022. "Open source cross-sectional asset pricing." *Critical Finance Review* 11 (2): .

Chen, Liang, Juan Dolado, and Jesus Gonzalo. 2021. "Quantile factor models." *Econometrica* 89 875–910.

Chen, Mingli, Ivan Fernandez-Val, and Martin Weidner. 2021. "Nonlinear factor models for network and panel data." *Journal of Econometrics* 220 296–324.

Chernozhukov, Victor, Ivan Fernandez-Val, and Martin Weidner. 2024. "Network and panel quantile effects via distribution regression." *Journal of Econometrics* 240 (2): .

Fama, Eugene F, and Kenneth R French. 1996. "Multifactor explanations of asset pricing anomalies." *Journal of Finance* 51 (1): 55–84.

Fama, Eugene F, and Kenneth R French. 2015. "A five-factor asset pricing model." *Journal of Financial Economics* 116 (1): 1–22.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. "Taming the factor zoo: A test of new factors." *Journal of Finance* 75 (3): 1327–1370.

Fernandez-Val, Ivan, and Martin Weidner. 2016. "Individual and time effects in nonlinear panel models with large N, T." *Journal of Econometrics* 192 291–312.

Gao, Jiti, Fei Liu, Bin Peng, and Yayi Yan. 2023. "Binary response models for heterogeneous panel data with interactive fixed effects." *Journal of Econometrics* 235 1654–1679.

Gao, Wayne. 2020. "Nonparametric identification in index models with link formation." *Journal of Econometrics* 215 399–413.

Giglio, Stefano, Bryan Kelly, and Dacheng Xiu. 2022. "Factor models, machine learning, and asset pricing." *Annual Review of Financial Economics* 14 337–368.

Giglio, Stefano, and Dacheng Xiu. 2021. "Asset pricing with omitted factors." *Journal of Political Economy* 129 (7): 1947–1990.

Hou, Kewei. 2007. "Industry information diffusion and the lead-lag effect in stock returns." *Review of Financial Studies* 20 (4): 1113–1138.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *Journal of Finance* 48 (1): 65–91.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen. 2023. "Is there a replication crisis in finance?" *Journal of Finance* 78 (5): 2465–2518.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2020. "Shrinking the cross-section." *Journal of Financial Economics* 135 (2): 271–292.

Lettau, Martin, and Markus Pelger. 2020. "Factors that fit the time series and cross-section of stock returns." *Review of Financial Studies* 33 (5): 2274–2325.

McDonald, R. 1962. "A general approach to nonlinear factor analysis." *Psychometrika* 27 397–415.

McDonald, R. 1979. "The simultaneous estimation of factor loadings and scores." *British Journal of Mathematical and Statistical Psychology* 32 212–228.

Mugnier, Martin, and Ao Wang. 2022. "Identification and fast estimation of large nonlinear panel models with two-way fixed effects." *Working Paper* 1–64.

Onatski, Alexei. 2012. "Asymptotics of the principal components estimator of large factor models with weakly influential factors." *Journal of Econometrics* 168 (2): 244–258.

Scholkopf, Bernard, Alexander Smola, and Klaus-Robert Muller. 1998. "Nonlinear component analysis as a kernel eigenvalue problem." *Neural Computation* 10 1299–1319.

Van der Vaart, Aad, and Jon Wellner. 1996. *Weak Convergence and Empirical Processes*. Springer Series in Statistics.

Vershynin, Roman. 2018. *High-Dimensional Probability*. Cambridge Series in Statistical and Probabilistic Mathematics.

Wang, Fa. 2022. "Maximum likelihood estimation and inference for high dimensional generalized factor models with application to factor-augmented regressions." *Journal of Econometrics* 229 180–200.

Yalchin, Ilker, and Yasuo Amemiya. 2001. "Nonlinear factor analysis as a statistical method." *Statistical Science* 16 275–294.

Zeleneev, Andrei. 2020. "Identification and estimation of network models with nonparametric unobserved heterogeneity." *Working paper*.

Zhu, Hong-Tu, and Sik-Yum Lee. 1999. "Statistical analysis of nonlinear factor analysis models." *British Journal of Mathematical and Statistical Psychology* 52 225–242.