Empirical Crypto Asset Pricing using Factor Models with High-Dimensional Characteristics

Adam Baybutt

Last updated: November 9, 2023

We study the dynamics of crypto asset returns through the lens of factor models. Given the limited number of tradable assets and years of data and the rich set of available asset characteristics, we develop novel estimation procedures with supporting econometric theory for a dynamic latent-factor model with high-dimensional asset characteristics, that is, the number of characteristics is on the order of the sample size. Utilizing the Double Selection Lasso estimator, our procedure employs regularization to eliminate characteristics with low signal-to-noise ratios yet maintains asymptotically valid inference for our asset pricing tests. In our empirical panel, we find the new estimator obtains comparable out-of-sample pricing ability and risk-adjusted returns to benchmark methods. We provide an inference procedure for measuring the risk premium of an observable nontradable factor, and employs this to find crypto's inflation mimicking-portfolio has positive risk compensation. Finally, specifying a factor model with nonparametric loadings and factors, we utilize recent methods in deep-learning to maximize out-of-sample risk-adjusted returns in an hourly panel, which yields economically significant alphas even after a detailed accounting of transaction costs.

Contact: adam baybutt at protonmail dot com. *Acknowledgments*: I am grateful to Denis Chetverikov for guidance, support, and uniquely valuable mathematical insights. I thank the following people for valuable comments and suggestions: Jesper Böjeryd, Mikhail Chernov, Valentin Haddad, Basil Halperin, Jinyong Hahn, Niklas Henke, Zhipeng Liao, Manu Navjeevan, Matthew Reid, Andres Santos, Pierre-Olivier Weill, and participants in UCLA's Econometrics Proseminar. I thank the following individuals for their research assistance: Jacob Brophy, Chuhan Guo, and, Tiffany Zho. Thanks to Pascal Michaillat for the formatting of this paper. Thank you to Coin Metrics, CoinMarketcap, and Glassnode for providing academic research discounts on data purchases. All code available at https://github.com/adambaybutt/crypto_asset_pricing.

Contents

1. Introduction

In this manuscript, we investigate the dynamics of crypto asset returns through the lens of factor models. After presenting a set of motivating empirical facts, we develop novel estimation procedures with supporting econometric theory for a dynamic latent-factor model with high-dimensional asset characteristics. We then provide empirical results to investigate why different crypto assets earn different average returns; conduct inference for crypto's inflation risk premium; and, estimate risk premia of crypto asset excess returns with classic factor models, our new dynamic latent-factor model, and a deep-learning-based factor model.

Why Crypto. Nakamoto (2008) gifted a novel mechanism design known as Proof-of-Work, enabling a set of adversarial entities to reach consensus on the current state of a digital database using cryptography, often framed as a solution to the Byzantine Generals' Problem (Lamport, Shostak, and Pease 1982). The Bitcoin blockchain launched in early 2009 employing Proof-of-Work to pioneer a censorship-resistant digital transaction ledger. This innovation introduced a permissionless payment network for transferring bitcoin, its native digital asset. The emergence of Nakamoto Consensus, along with other blockchain-based consensus mechanisms that followed, enabled the scarcity of digital information, particularly in the form of digital or crypto assets, and thus introduces a new area of economic research.

An Attractive and Independent Asset Class. We motivate research into the return dynamics of these digital assets by establishing the following empirical facts. The advent of Bitcoin sparked a Cambrian explosion of other crypto assets, evolving from initial valuations as collectibles into a trillion dollar asset class. Bitcoin has matured into a substantial payment network, settling hundreds of billions of dollars annually, with the large majority of transactions settling for a cost of less than one dollar, thereby offering monetary functions with distributed consensus.

Bitcoin exhibited superior risk-adjusted returns when compared to traditional asset classes over our study period of 2018-2022, inclusive. With regard to independence, bitcoin has lower correlations with the Nasdaq and the S&P500, at 0.23 and 0.21 respectively, as compared to gold's correlations with these indices at 0.26 and 0.28. Moreover, bitcoin's correlation with other assets exhibits significant temporal variance, including several quarters of zero or negative correlation with the Nasdaq; their high correlation

(> 0.3) is only a recent phenomenon seen in 2022. While these measures are suggestive of an independent asset class, a possibly sufficient statistic is whether there are risk-adjusted return gains from including crypto assets in one's portfolio. From diversifying a risk portfolio of holding 100% Nasdaq to instead hold 60% Nasdaq and 40% the crypto market, one would obtain a Sharpe Ratio gain of 0.53 (from 0.43 to 0.96).

Crypto Signals. The emergence of thousands of crypto assets expands to a new asset class the central focus of empirical asset pricing: the search for explanations of why different assets earn different average returns. A fundamentally unique aspect of the crypto asset class is open state: the state of the digital ledger is readable. This is termed onchain data where one has access to the (onchain) economy's full history of transactions. For instance, we directly observe the hodling time of all Bitcoin wallets to discover a majority of wallets utilize bitcoin as a store of value rather than for speculatively trading.

In this manuscript, we formulate several novel crypto asset characteristics in addition to investigating the signal content of numerous characteristics previously studied in the literature. An additional distinction of this study is to build a panel of tradable crypto asset excess return data with more realistic inclusion criteria than previously studied in the literature. In examining the signal content across this rich set of asset characteristics, although there are some redundant characteristics and signal decays over the study years, we observe numerous sources of signal for the cross-section of one-week ahead expected returns. We now turn to modeling asset excess returns using these characteristics.

Factor Models. Our empirical search centers on factor models: the cornerstone cross-sectional asset pricing model. The aforementioned empirical observations motivate the adoption of a factor model that accommodates time-varying relationships between assets and characteristics, while capable of incorporating and compressing the signal across a large number of characteristics. We now provide the setting for the major contribution of this research, namely, developing econometric estimation methods for a factor model with high-dimensional asset characteristics, such that, the number of asset characteristics is on the order of the number of assets and time periods.

We assume a statistical model where asset excess returns $r_{i,t+1} \in \mathbb{R}$ are a function of common time-varying latent (unobserved) factors, $f_{t+1} \in \mathbb{R}^k$, as dictated by time-varying asset-specific factor loadings $\beta_{i,t} \in \mathbb{R}^k$, that is, for all assets $i \in \{1, \ldots, N\}$ and time $t \in \{1, \ldots, T\}$

$$\begin{aligned} r_{i,t+1} &= \alpha_{i,t} + \beta_{i,t}^{\top} f_{t+1} + \epsilon_{i,t+1}^{r} \\ \beta_{i,t} &= \Gamma_{\beta}^{\top} z_{i,t} + \epsilon_{i,t}^{\beta} \end{aligned} \tag{1}$$

where $\alpha_{i,t} \in \mathbb{R}$ are average pricing errors of the factor model; $\epsilon_{i,t+1}^{r} \in \mathbb{R}$ are uncorrelated idiosyncratic errors, i.e. $\mathbb{E}_t[\epsilon_{i,t+1}^{r} f_{t+1}] = 0$; $\Gamma_{\beta} \in \mathbb{R}^{p \times k}$ is a static latent loading parameter; and, $z_{i,t} \in \mathbb{R}^{p}$ are time-varying asset-specific characteristics where p is high-dimensional, i.e. on the order of N and T. Crucially, we follow the established practice in the literature of assuming the number of factors k is low-dimensional, i.e., $N, T, p >> k \in \{1, 2, 3, \ldots\}$. This model has theoretical underpinnings motivated by a structural model for asset excess returns or by the assumption of no arbitrage, as discussed in our literature review.

Double Selection Lasso Factor Model. The main contribution of this research is to develop new estimation and inference procedures, termed the the Double Selection Lasso Factor Model (DSLFM), to fit the latent factors f_{t+1} and loadings Γ_{β} in (1) and to conduct standard asset pricing tests under the novel setting of high-dimensional asset characteristics.

The DSLFM remains consistent with the equilibrium asset pricing principle that risk premia are solely determined by risk exposures and specifies a linear loading mapping Γ_{β} between characteristics and dynamic factor loadings $\beta_{i,t}$. We have two novel assumptions for Γ_{β}. First, we develop estimation procedures and large-sample theory that allows $p, T, N \to \infty$. Given our focus is studying the cross-section of crypto assets, this assumption is particularly relevant given the numerous asset characteristics available, as previously discussed, as well as the existence of only a small number of tradable assets and years of relevant data such that p, T, N are of similar order. Second, we assume exact row sparsity in Γ_{β}; that is, only a small number of the p asset characteristics determine the content of the factor loading, which matches empirical findings in cross-sectional asset pricing (Babiak and Bianchi (2021) and Bianchi, Guidolin, and Pedio (2022)). These novel assumptions within a dynamic latent-factor model require novel estimation procedures and supporting asymptotic theory.

The DSLFM aims to jointly and consistently estimate the loading matrix Γ_{β} and latent factors f_{t+1}. If we were to utilize the MSE objective function to minimize over the p–dimensional choice vector $\Gamma_{\beta} f_{t+1}$, for each t, the mean-squared error of $\sum_i (r_{i,t+1} -$

$z_{i,t}^{\top} \Gamma_{\beta} f_{t+1})^2$, we will have not only a noisily estimated design matrix when $p \sim N$, (or, at worst, a nonsingular design matrix when $p > N$) but also a non-convex objective function given the interaction between minimization arguments f_{t+1} and Γ_{β}. The next logical step would be to introduce sparsity in Γ_{β}, which would amount to adding a regularization parameter to the aforementioned objective function to combat the curse of dimensionality from $z_{i,t}$. However, although potentially helpful for minimizing MSE by decreasing the variance of the estimator, this regularization introduces a bias in estimation, which would lead to invalid asymptotic inference for asset pricing tests, defeating a goal of this research.

We therefore adapt for our purpose the Double Selection Lasso (DSL) estimator developed by Belloni, Chernozhukov, and Hansen (2014). The key insight from their work was to introduce an orthogonality wherein, assuming Γ_{β} is row sparse, the regularization bias from the LASSO first-stage estimation does not pass through to the target parameter of interest when conducting inference. We first estimate for each time period t and each characteristic j the scalar $\Gamma_{\beta,j}^{\top} f_{t+1}$ using DSL; then, stacking these estimates into a $T \times p$ matrix, we use PCA to obtain separate estimates for latent loadings $\widehat{\Gamma}_{\beta}$ and factors $\{\widehat{f}_{t+1}\}_{t=1}^{T}$; and, finally, we soft-threshold $\widehat{\Gamma}_{\beta}$ to set numerous rows to zero given the assumption of sparsity for Γ_{β}.

This procedure has several additional benefits. Given the period-by-period cross-sectional regression—mirroring Fama-MacBeth regressions–our estimation procedure accommodates unbalanced panels. This first DSL step does require running $T \times p$ cross-sectional regressions, however this can be done in parallel and is on the order of minutes in practice as each is computationally light. In the second step, the high dimensionality of the PCA procedure, given we have a $p \times T$ matrix, is adapted from the established theory for $N \times T$ excess return matrices (Bai 2003). The final soft-thresholding step exploits the sparsity in Γ_{β} to remove noise from characteristics with low signal-to-noise ratios. [1]

Under standard DSL assumptions adapted to our setting (Belloni, Chernozhukov, and Hansen 2014), high-dimensional PCA assumptions (Bai 2003), and assuming we observe the true number of latent factors (Bai and Ng 2002), we develop the asymptotic consistency of the latent factors $\widehat{f}_{t+1}$ and loading matrix $\check{\Gamma}_{\beta}$ for the latent factors f_{t+1} and loadings Γ_{β}, respectively. Monte Carlo simulations corroborate with finite-sample

[1] Finally, just as DSL laid the groundwork for the more general Debiased Machine Learning (DML) theory, this work sets up future research to extend the framework with a semi-parametric specification to utilize the rich set of available machine learning estimators that have been shown to handle well the nonlinearities in cross-sectional asset returns (Gu, Kelly, and Xiu 2020).

evidence that the performance of the DSLFM is comparable to or surpasses relevant benchmarks. As is standard in this setting, without further restrictions outlined in Bai and Ng (2013), F^0 and Γ_β^0 are not separately identifiable; hence, the $k \times k$ invertible matrix transformation H appears in each asymptotic result. However, in many cases, knowing $F^0 H$ is equivalent to knowing F^0; for example, using the regressor F^0 will give the same predicted values as using the regressor $F^0 H$ given they have the same column space. Similarly, in the coming inference result, our target parameter is rotation invariant to H.

To show the generality of these estimation procedures, we enrich our model—with one of several possible extensions—to address the common question in asset pricing research. We ask whether an observable factor $g_{t+1} \in \mathbb{R}$ carries a risk premium: compensation for exposure to the risk factor holding constant exposure to all other sources of risk, i.e., variation with other factors. In the subsequent empirical applications, we investigate a common hypothesis for the crypto asset class: exposure to inflation offers crypto investors a positive risk premium.

Following a recent approach in the literature (Giglio and Xiu (2021) and Giglio, Xiu, and Zhang (2021)), we assume the "true" latent factors f_{t+1} can be decomposed into the latent-factor risk premia $\gamma \in \mathbb{R}^k$ and latent-factor innovations $v_{t+1} \in \mathbb{R}^k$, that is, $f_{t+1} := \gamma + v_{t+1}$. Then, we specify the observable factor g_{t+1} as potentially linearly correlated with the latent factors through

$$g_{t+1} = \eta v_{t+1} + \epsilon_{t+1}^g,$$

where $\eta \in \mathbb{R}^k$ is an unknown parameter mapping the relation between the latent-factor innovations and the observable factors, and $\epsilon_{t+1}^g \in \mathbb{R}$ is measurement error in g_{t+1}.

The risk premium of an observable factor—our target parameter in this extension—is defined to be the expected excess return of a portfolio with loading (i.e., beta) of 1 with respect to this factor in g_{t+1} and zero loadings on all other factors; in this model, that parameter is $\gamma_g := \eta^\top \gamma = \eta^{0\top} H H^{-1} \gamma^0 = \eta^{0\top} \gamma^0$, which utilizes the rotation invariant result of Giglio and Xiu (2021).

We thus extend with our estimation procedure to a dynamic latent factor model with high-dimensional characteristics, instead of a static latent factor model–the estimation procedure of Giglio and Xiu (2021)to measure the risk premium γ_g. We additionally develop our estimator's large-sample distribution and variance to conduct asset pricing tests on the sign of the observable factor risk premium. We apply this test for the risk premium of the observable inflation factor within the crypto asset panel studied herein.

Empirical Setting. To assemble a weekly panel of tradable crypto assets, we prospectively identify, at the start of each month from 2018 to 2022, inclusive, tradable crypto assets on US centralized crypto exchanges with sufficient trading volume and market capitalization, which results in the number of assets growing from 9 on January 7, 2018 to 168 on December 1, 2022. There are 210 unique assets in the panel.

Motivated by a 1% threshold on an order book's volume, the most restrictive inclusion criteria applied each month is for each asset to have a median weekly volume across US exchanges of $500k over the trailing three months. Using this strict set of inclusion criteria to study tradable assets without bias, our panel has a challenge we have not seen in the literature wherein many assets repeatedly enter and leave the panel over time as they rise above and fall below the inclusion criteria. We thus have to reform the panel monthly when fitting models. For example, an asset may not meet the January 2018 criteria even though we have data for it, but then the asset will be included in February 2018, for which, we would want to use its historical data to inform estimation.

This panel of weekly crypto asset excess returns is not only novel to the literature based on the inclusion criteria but also given its contains several novel asset characteristics across the 63 characteristics studied within the following categories: onchain, social, financial, momentum, exchange, and microstructure. We do note that five years of data is limited, but this the reality for empirical crypto asset pricing.

We then provide empirical results to investigate why different crypto assets earn different average returns; conduct inference for crypto's inflation risk premium; and, estimate risk premia of crypto asset excess returns with classic factor models, our new dynamic latent-factor model, and a deep-learning-based factor model.

Empirical Applications: Factor Models with Low-Dimensional Characteristics. We begin our study of estimating risk premia of crypto asset excess returns with classic factor models. In studying 63 univariate factors formed as the long-short quintile portfolio sorted on each asset characteristic, we find financial factors (e.g., momentum, beta, idiosyncratic skewness, and 5% shortfall) are the only ones associated with significant differences in average one-week ahead excess returns between the top and bottom quintiles. Given the significant factors were those that are functions of only previous returns, this is suggestive that crypto asset returns are not driven by fundamental factors.

To build a set of benchmarks for later results, we compare the out-of-sample predictive power in the Q3-Q4 2022 data of three models, namely, a three-factor model of size, crypto market, and momentum; a latent three-factor model fit with PCA; and a dynamic latent-factor model fit with IPCA using a subset of the characteristics selected based on in-sample performance. We find IPCA outperforms the other models, suggestive of the signal in the characteristics. Its predictive pricing signal outperforms a random walk and it provides economically and statistically significant risk-adjusted returns, whereas the other models underperform a random walk yet provide modest risk-adjusted returns.

Empirical Applications: Factor Models with High-Dimensional Characteristics. Utilizing the broader set of asset characteristics, we first establish the comparable out-of-sample predictive ability of the DSLFM compared to the benchmark methods, with supporting bootstrapped characteristic importance measures to elucidate the drivers of returns. Exchanges inflows and outflows were significant characteristics, showing the importance of these onchain measures. While DSLFM achieved a maximum, with one latent factor, out-of-sample Sharpe of $\sim$ 3.3, this underperforms ICPA's maximum, with one latent factor, Sharpe of $\sim$4.

Additionally, we implement our testing procedure to find that the crypto asset class provides investors a positive inflation risk premium. Early proponents of Bitcoin and other cryptocurrencies framed these as an outside option or hedge against traditional fiat currencies. To study this question, we use our extended model to recover the 10-year inflation expectation mimicking portfolio and measure its risk premium. This inflation risk premium was estimated at a statistically significant 1.4 bps with a standard error of 0.0097 bps. This translate to a 7.3% annual excess return, suggestive of positive compensation for investors holding an inflation-hedge crypto portfolio, ceteris paribus.

Finally, to explore further the predictability of crypto asset expected returns, we expand our panel to the hourly frequency and extend a recent deep-learning model nested within a factor-model structure. This hypothesizes there are nonlinearities in the DGP for returns that this nonparametric specification can recover. Although we utilize the feed-forward network architecture of Gu, Kelly, and Xiu (2021) for both the factors and factor loadings, we find several notable adjustments improved performance in validation samples, including a nonlinear factor autoencoder, employing more recent optimization methods, using a weighted loss function to focus on asset-hours with high trade volume, among other adjustments. We employ an iterative step-forward cross-validation procedure to optimize hyperparameters by sequentially predicting on validation data for each quarter in Q1 2021 through Q2 2022 before a single prediction on Q3-Q4 2022 test period data.

The out-of-sample empirical results are impressive. Ventile portfolios are nearly monotonic with a pre-transaction cost arithmetic hourly average spread of $\sim$16 bps between the top and bottom ventiles. After a detailed accounting of transaction costs (i.e. price impact, bid-ask spread, trading fee, and margin), the model yields out-of-sample hourly alphas of 5 bps. This portfolio is significantly correlated with the market, given the loss function over-weights large volume assets. Also, it failed to achieve positive out-of-sample predictive R^2.

To account for the long memory present across several asset characteristics, we nest a Transformer architecture into a factor-model to find improved pricing ability with a statistically significant out-of-sample predictive R^2 of 3.6%, with associated positive quintile spread returns. However, these returns were not robust in the Q3-Q4 2022 data to transaction costs, which motivates further research with 2023 data, which I have yet to obtain.

Overview. Section 2 frames this study in relation to the relevant literature. Section 3 provides an overview of the panel data studied herein. Section 4 presents motivating stylized facts frame the research herein and encourage academic research in this new asset class. Section 5 develops the econometric theory of the DSLFM with supporting simulation evidence. Section 6 covers numerous empirical applications. Section 7 concludes. The appendix contains supporting tables and figures, algorithmic details, and mathematical proofs.

2. Literature Review

This paper builds on two literatures: empirical crypto asset pricing and the econometric theory of factor models.

I will complete the drafted empirical crypto asset pricing literature review in November, once I get a few day break from interviews!

2.1. Econometric Theory of Factor Models

This manuscript contributes to the literature developing the econometric theory of factor models (1), particularly those with high-dimensional asset characteristics. Giglio, Kelly, and Xiu (2022) provide an excellent review of recent machine-learning based factor model applications and relevant econometric theory, including the common asymptotic frameworks of fixed N and $T \to \infty$, fixed T and $N \to \infty$, and $T, N \to \infty$. We will, nevertheless, attempt a brief summary of the closely related econometric methods.

Starting from either a structural model for asset excess returns in the style of the Capital Asset Pricing Model (Sharpe 1964), or the assumption of no arbitrage, as in Arbitrage Pricing Theory (Ross 1976), a stochastic discount factor $m_{t+1} \in \mathbb{R}$ exists and an Euler equation, termed the Law of One Price, holds for asset excess returns $r_{i,t+1} \in \mathbb{R}$ for assets $i \in \{1, 2, \ldots, N\}$ in time periods $t \in \{1, 2, \ldots, T\}$

$$\mathbb{E}_t[m_{t+1} r_{i,t+1}] = 0,$$

which by the definition of the variance and covariance operators,

$$\mathbb{E}_t[m_{t+1} r_{i,t+1}] = \underbrace{\frac{Cov_t(m_{t+1}, r_{i,t+1})}{Var_t\,(m_{t+1})}}_{\beta_{i,t}} \underbrace{\frac{-Var_t\,(m_{t+1})}{E_t[m_{t+1}]}}_{\lambda_t}.$$

As discussed in Section 1 we directly assume the statistical model

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t}^{\top} f_{t+1} + \epsilon_{i,t+1}^{r}.$$

To map this model to its theoretical underpinnings in the Law of One Price, one can assume for all i and t: mean zero unobserved idiosyncratic errors $\mathbb{E}_t[\epsilon_{i,t+1}^{r}] = 0$, uncorrelated errors $\mathbb{E}_t[\epsilon_{i,t+1}^{r} f_{t+1}] = 0$, the price of risk associated with the factors to be defined as $\lambda_t := \mathbb{E}_t[f_{t+1}]$, and zero average pricing errors $\alpha_{i,t} = 0$ (Cochrane 2009).

The factor model posits that asset excess returns $r_{i,t+1}$ signify compensation for asset-specific, time-varying exposure $\beta_{i,t} \in \mathbb{R}^k$ to systematic risk factors $f_{t+1} \in \mathbb{R}^k$.

The classic factor model is a static loading observable factor model—in the style of Fama and French—where pricing errors and static factor loadings β_i in (1) using exogenously-defined factors are estimated via the "Fama-MacBeth" two-step procedure (Fama and MacBeth 1973), which has rich supporting inferential theory (Shanken 1992). This procedure relies on ex ante declaration of observable factors f_{t+1} formed as a convex combination (i.e., a portfolio) of sorted asset returns r_{t+1} based on asset characteristics $z_t \in \mathbb{R}^p$. This is likely an incomplete model of the relationship between z_t and r_{t+1} and prone to overfit.

Factor Zoo. Recognizing the proliferation of factors, a "Factor Zoo" (Cochrane 2011), explaining the cross section of expected returns, Feng, Giglio, and Xiu (2020) propose the use of Double Selection Lasso Belloni, Chernozhukov, and Hansen (2014) combined with two-pass Fama-MacBeth regressions to evaluate the contribution of a new factor, g_{t+1}, explaining asset expected returns above and beyond an existing high-dimensional set of factors. However, as the recent empirical literature has shown (e.g., Kelly, Pruitt, and Su (2019); Chen, Pelger, and Zhu (2020)), allowing the data to construct the relevant latent factors offers superior explanatory and predictive power as compared to using a set of selected observable factors from the literature.

PCA. Factor models (1) with latent factors have been a focal point since the development of APT (Ross 1976) and early empirical efforts (Chamberlain 1983). PCA is the common estimation framework. Bai (2003) laid the groundwork for PCA estimation and inference under joint $N, T \to \infty$ high-dimensional asymptotics, with Bai and Ng (2002) introducing a novel Information Criterion penalization to ascertain the true number of latent factors. Assumptions for the joint identification of factor loadings and factors are summarized in Bai and Ng (2013). This study leverages these contributions yet employs instead a dynamic latent-factor model with $p, N, T \to \infty$ high-dimensional asymptotics. In the asset pricing context, the static model can still perform well for describing portfolios over time given the dynamics are captured by the dynamic factors; however, this has not performed as well in describing individual asset returns, as noted by Ang, Liu, and Schwarz (2009). Although this model allows the data to statistically inform the factor structure, it fails to incorporate rich data available in asset characteristics and it assumes a static factor loading that maps systematic risks to excess returns.

IPCA. A recent and significant methodological advance, instrumented PCA (IPCA) from Kelly, Pruitt, and Su (2019), utilizes asset characteristics in a linear model of dynamic factor loadings to parameterize the more general semi-parametric method studied in Connor and Linton (2007). That is,

$$\beta_{i,t} = \Gamma_\beta^\top z_{i,t} + \epsilon_{i,t}^\beta$$

where $\Gamma_\beta \in \mathbb{R}^{p \times k}$ is a loading matrix mapping asset characteristics to the factor loading. IPCA has several benefits, including compressing the $N \times T$ factor loading matrix β to a lower dimensional $p \times k$ loading matrix Γ_β, as well as specifying a time-varying relationship $\beta_{i,t}$ between characteristics and returns, which as stated previously appears to be the empirical reality in crypto cross-sectional asset pricing.

Setting MSE as the objective function, the authors develop an iterated least squares estimation procedure where factor loadings Γ_β are the coefficient from regressing the panel of returns on the factors interacted with firm-specific characteristics. Factor realizations are period-by-period cross-sectional regression coefficients of r_{t+1} on the latent loadings. This system of first-order conditions must be solved numerically by alternating between the two objectives. In a separate theory paper, the authors develop the asymptotic distributional theory—following $N, T \to \infty$ asymptotics from Bai (2003)—for the factor realizations and loadings under quite general identifying restrictions on loadings and factors (Kelly, Pruitt, and Su 2020).

The IPCA procedure benefits from the following: the efficiency gains from using asset characteristics for estimating the latent factors and their loadings; accommodation of unbalanced panels; maintaining an expected return factor model structure to ascertain the economic relationships among factors and assets via the observable characteristics; a parametric model with inference procedures for asset pricing tests; among other reasons. The IPCA estimation procedure, however, is not possible under high-dimensional asset characteristics (i.e., $p > T, N$), or, if regularization is used, produces biased inference for asset pricing tests.

Similar to IPCA, the DSLFM procedure is a reduced form statistical model, in the spirit of APT, and thus leaves open structural model research to connect to the economic mechanism driving returns. We nevertheless in empirical applications offer results to guide this theoretical work for crypto assets.

Testing Observable Factor Premium. Giglio and Xiu (2021) develop a three-step procedure combining estimation of latent-factor model via PCA with standard two-pass regressions to recover the risk premium of an observable nontradable factor $g_{t+1}\mathbb{R}$, which are potentially correlated with the latent factors:

$$g_{t+1} = \eta^\top v_{t+1} + \epsilon_{t+1}^{g}$$

where v_{t+1} are (mean-zero) latent-factor innovations (i.e., $f_{t+1} = \gamma + v_{t+1}$); $\eta \in \mathbb{R}^k$ is a linear mapping of the true latent factors to the observed factors; and, ϵ_{t+1}^{g} is measurement error. This allows the observable factors to either be some component of the latent factors or just correlated with v_{t+1} and therefore still carry a risk premium. The target parameter is the risk premium associated with the observable factors $\gamma_g \coloneqq \eta\gamma$.

To demonstrate the generality of the DSLFM, we extend our estimation procedure by adding the procedure of Giglio and Xiu (2021) to address this classic asset pricing test of the recovering the risk premium of an observable factor. The DSLFM theory presented in Section 5 extends Giglio and Xiu (2021) by incorporating not only dynamic factor loadings, but also high-dimensional asset characteristics.

Machine Learning Factor Models. More recent literature has incorporated a wide array of machine learning-based estimation approaches within the factor model structure. Gu, Kelly, and Xiu (2020) study a set of machine learning estimation procedures for measuring the equity risk premium to find that deep learning methods outperform out-of-sample. Gu, Kelly, and Xiu (2021) develop a factor model in the structure of IPCA, but allows for non-linear mappings to the factor loadings and the factors through two feed-forward neural networks. Although, in practice they only use linear mappings to the factors, they still obtain out-of-sample predictive R^2 and Sharpe ratio gains in relation to benchmark methods. Other notable uses of deep learning within factor model structures are Chen, Pelger, and Zhu (2020); Feng et al. (2018); Guijarro-Ordonez, Pelger, and Zanotti (2021); among others.

A fundamental result in machine learning is the no free lunch theorem (NFLT): all optimization problem strategies perform equally well when averaged over all possible problems. However, there is a convergence in several domains toward the use of deep learning as the dominant ML paradigm. A recent paper, Gómez and Rojas (2016), evaluates the performance of several popular ML methods to a diverse set of real-world problems to empirically reject the NFLT. Their conclusion is that the empirical settings of interest do not have a common DGP but rather are sufficiently heterogeneous such that certain ML algorithms are more amenable in the finite sample.

A fundamental difference between common empirical settings for machine learning applications and their use in empirical asset pricing is the uniquely low signal-to-noise DGP. Theory suggests market efficiency pushes the signal into the unforecastable idiosyncratic error. This critical issue significantly compounds the curse-of-dimensionality of using high-dimensional characteristics, which further motivates the parsimonious specification of the factor model. An important open question in empirical asset pricing is whether deep learning scaling laws will apply in financial forecasting.

3. Description of Data

See Appendix A and Appendix D Figures A1-A5 for an overview of the panel. This drafted section will be completed when I have a few day break from interviews...

4. Motivating Empirical Facts

> "We have always had bad money because private enterprise was not permitted to give us a better one...The important truth to keep in mind is that we cannot count on intelligence or understanding but only on sheer self-interest to give us the institutions we need."
>
> —Friedrich A. Hayek (1976) *The Denationalization of Money*.

See appendix figures A6-A24 for these motivating empirical facts. This drafted section will be completed when I have a few day break from interviews...

The following are motivating empirical facts observed in the panel of crypto asset returns during the study time period of 2018-2022, inclusive.

(i) From a market value of zero in 2009, Bitcoin and a Cambrian explosion of thousands of other crypto assets have rose to trillion dollar market capitalization at the end of 2022, with several multi billion dollar sub-industries.

(ii) Bitcoin achieved a superior risk-adjusted returns (i.e. four-year rolling Sharpe ratio) for nearly the entire study time period as compared to the traditional asset classes.

(iii) Bitcoin has lower Pearson correlations to the Nasdaq and S&P500 (at 0.23 and 0.21) as compared to Gold (at 0.26 and 0.28), and numerous other asset classes.

(iv) Bitcoin's correlation with other assets is highly time varying, including several quarters of zero or negative correlation with the Nasdaq; their high correlation (> 0.3) is only observed recently, i.e. 2022.

(v) From diversifying a risk portfolio of holding 100% Nasdaq to instead hold 60% Nasdaq and 40% the crypto market, one would obtain a Sharpe Ratio gain of 0.53 (from 0.43 to 0.96).

(vi) The crypto market offers a positive inflation risk premium of 31 bps.

(vii) Buying (at the time they are tradable on a CEX) and holding (until the end of the study period) the panel assets that ICO'ed offers a near zero buy-and-hold average return, far from a significant loss of capital.

(viii) Bitcoin is used to store value by a majority of wallets, rather than speculatively trading.

(ix) Bitcoin is a payment network settling hundreds of billions of dollars annually where the large majority of transactions cost less then one USD, offering monetary functions with distributed consensus.

(x) Efforts to fork (i.e. copy) the Bitcoin blockchain have had immaterial adverse effects; an event study of forks observes significant positive effects on price, trading volume, active addresses, and social activity.

(xi) In studying the signal content across a rich set of asset characteristics, although decaying over the years, we observe numerous sources of signal for the cross-section of one-week ahead expected returns as well as for returns at several week and month horizons.

(xii) The asset characteristics do contain redundant information; however, the variation cannot be captured by just a few principal components.

5. Double Selection Lasso Factor Model

Our goal is to consistently estimate a dynamic latent-factor model and conduct asset pricing tests with valid inference under the novel setting of high-dimensional asset characteristics, which is of particular relevance in the crypto asset class given the open state of digital ledgers.

5.1. Setup

Setting and Observable Random Variables. Assume for time periods $t = 1, 2, \ldots, T$ and assets $i = 1, 2, \ldots, N$, that we observe realizations of random variables for asset excess returns $r_{i,t+1} \in \mathbb{R}$ and asset characteristics $z_{i,t} \in \mathbb{R}^p$. An asset's excess return is the simple return of asset i from time t to $t+1$ net the assumed simple return of the risk-free rate (e.g., one-month US Treasury Bill). An asset characteristic of asset i is known at time t : for example, the total fees for a crypto protocol between time $t-1$ and t. Note that asset characteristics are information from the previous period to follow the established convention in the literature and to be able to use this model for prediction. Importantly, we will introduce the novel asymptotic assumption for dynamic latent-factor models to let p grow to infinity simultaneously with N and T.

Model. Given the highly nonlinear data-generating process observed in empirical asset returns (Gu, Kelly, and Xiu (2020); Chen, Pelger, and Zhu (2020); Bianchi, Büchner, and Tamoni (2021)), we specify a semi-parametric factor model—where $\beta_{i,t}$ is a function of asset characteristics $z_{i,t}$—studied in recent literature (Connor and Linton (2007); Connor, Hagmann, and Linton (2012); and Fan, Liao, and Wang (2016)). We assume a dynamic latent-factor model where

$$r_{i,t+1} = \beta_{i,t}^{\top} f_{t+1} + \epsilon_{i,t+1}^{r}.$$

However, given we are interested in conducting inference, we specify a linear model—to build on the foundational work of Kelly, Pruitt, and Su (2019)—for the factor loadings

$$\beta_{i,t} = \Gamma_{\beta}^{\top} z_{i,t} + \epsilon_{i,t}^{\beta}.$$[2]

[2] Although we are working with this parametric specification for the factor loadings, we can in our setting, nevertheless, employ feature engineering to generate many different functional forms of our asset characteristics, given the coming dimensionality reduction from LASSO.

Parameters and Unobserved Random Variables. $f_{t+1} \in \mathbb{R}^k$ are low-dimensional latent factors; $\beta_{i,t} \in \mathbb{R}^k$ are latent-factor loadings; $\Gamma_\beta \in \mathbb{R}^{p \times k}$ is an unknown factor loading parameter matrix; and, $\epsilon^r_{i,t+1} \in \mathbb{R}$ and $\epsilon^\beta_{i,t} \in \mathbb{R}^k$ are unobserved idiosyncratic errors.

The latent factors f_{t+1} should be interpreted as purely statistical in nature. That is, these risk factors do not necessarily capture fundamental shocks to productive technologies as modeled in canonical theoretical models. Nevertheless, the latent factors capture systemic risk or covariance among asset returns that is non-diversifiable. We follow the literature in restricting k to be a small finite constant (i.e., $k \in \{1, 2, 3, 4, 5\}$) that, in our asymptotic theory, does not grow with p, T, N. It should be noted that, although ubiquitous, it is nevertheless a strong assumption: the empirical content of asset returns can be captured by a small number of strictly time-varying systematic risk factors.

The specification of $\beta_{i,t}$ provides several benefits. First, we enable the use of a dynamic loading to model a changing relationship (e.g., regime changes) between the cross section of returns and systematic risk. Yet, we reduce the parameter space from a $N \times T$ loading matrix β to the $p \times k$ loading matrix Γ_β. Second, we incorporate additional data from the large number of asset characteristics to influence the factor model for returns through the loading matrix Γ_β. This addresses a challenge of migrating assets wherein an asset-specific but static β_i would not be able to capture an asset moving from, for example, a crypto asset earning low fees to one with high protocol fees. The classic way to handle this issue was to sort assets into portfolios of similar characteristics to form test assets, which then compresses the dimensionality of the cross-section. Thus, as discussed in Kelly, Pruitt, and Su (2019), this model specification skips ad hoc test asset formation to instead accommodate working directly with the high-dimensional system of individual assets. Finally, we assume exact row sparsity in Γ_β—precisely stated in the coming Assumption 1(ii)—a novel assumption to the literature; that is, only a small number of the p asset characteristics determine the content of the factor loading, which matches empirical findings in cross-sectional asset pricing (Kelly, Pruitt, and Su (2019); Babiak and Bianchi (2021); Bianchi, Guidolin, and Pedio (2022)).

Extended Model. In order to show the generality of this approach, we enrich the model—with one of several possible extensions—to address the common question in asset pricing research of whether an observable factor $g_{t+1} \in \mathbb{R}$ carries a risk premium: compensation for exposure to the risk factor holding constant exposure to all other sources of risk, i.e., variation with other factors.

In the context of asset pricing, a factor can be either tradable or nontradable. A tradable factor is a portfolio, that is, a convex combination of tradable asset returns. The risk premium is straightforward to calculate for tradable factors: it is the time series average excess return of the factor. However, many risk factors are nontradable, e.g., inflation expectations, consumption, liquidity, etc. Thus, we must estimate the risk premia of nontradable observable risk factors as the risk premia associated with their tradable portfolio.

Following a recent approach in the literature (Giglio and Xiu (2021) and Giglio, Xiu, and Zhang (2021)), first, we assume the aforementioned model for returns is a function of the "true" latent factors f_{t+1}, [3] and, second, we assume these true latent factors can be decomposed into the latent-factor risk premia $\gamma \in \mathbb{R}^k$ (i.e., unknown parameters of the long-run average excess return) and latent-factor innovations $v_{t+1} \in \mathbb{R}^k$ (i.e., mean zero risk factor random variable), that is, $f_{t+1} := \gamma + v_{t+1}$. Then, we specify the observable factor g_{t+1} as potentially linearly correlated with the latent factors through

$$g_{t+1} = \eta v_{t+1} + \epsilon^g_{t+1},$$

where $\eta \in \mathbb{R}^k$ is an unknown parameter mapping the relation between the latent-factor innovations and the observable factors, and $\epsilon^g_{t+1} \in \mathbb{R}$ is measurement error in g_{t+1}. This specification allows the observable factors to be either simply some component(s) of the true latent factors (e.g., setting ϵ^g_{t+1} to zero with η set to $(1, 0, 0, \ldots, 0)$ or, more generally, some unknown linear function η of the true latent factors and thus still carry a risk premium. To recover the tradable portfolio representing the nontradable observable risk factor, we map g_{t+1} through η onto the column space of the true latent factors.

Precisely, the risk premium of an observable factor—our target parameter in this extension—is defined to be the expected excess return of a portfolio with loading (i.e., beta) of 1 with respect to this factor in g_{t+1} and zero loadings on all other factors; in this model, that parameter is $\gamma_g := \eta^\top \gamma$.

Goal. Under the novel asymptotic assumption for this setting of $p, N, T \to \infty$, we aim to develop estimation procedures for the latent loadings Γ_β and factors f_{t+1}, $\forall t$; in addition, we aim to estimate and conduct inference on γ_g under the novel use of a dynamic latent-factor model and the aforementioned novel high-dimensional asset characteristics.

[3] To be precise, by true latent factors, we mean we can consistently estimate the finite constant of the number of latent factors that span the cross section of returns.

5.2. Estimation

Motivating DSLFM Estimation. The goal is to jointly estimate the loading matrix Γ_β and latent factors f_{t+1}, which are not separately identifiable without further restrictions, to be discussed (Bai and Ng 2013). However, to begin, given the model takes the form

$$r_{i,t+1} = z_{i,t}^\top \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}$$

where $\epsilon_{i,t+1} = (\epsilon_{i,t}^\beta)^\top f_{t+1} + \epsilon_{i,t+1}^r$ is the composite idiosyncratic error, we observe that our setting is high-dimensional panel data where we project $\{r_{i,t+1}\}_{i=1,t=1}^{i=N,t=T}$ onto the column space of $\{z_{i,t}\}_{i=1}^N$ for each time period to estimate each p dimensional time-varying vector $\{\widehat{\Gamma_\beta f_{t+1}}\}_{t=1}^T$ where we have to address $p \sim \max(N, T)$ or $p >> \max(N, T)$.

Thus, if we utilize the objective function to minimize over the p–dimensional choice vector $\Gamma_\beta f_{t+1}$ the mean-squared error of $\sum_i (r_{i,t+1} - z_{i,t}^\top \Gamma_\beta f_{t+1})^2$, we will not only have a noisily estimated design matrix when $p \sim \max(N, T)$, (or, at worst, a nonsingular design matrix when $p > \max(N, T)$) but also a non-convex objective given the interaction between minimization arguments f_{t+1} and Γ_β. This rules out implementing low-dimensional (in p) methods.

One potential solution would be to introduce sparsity in $z_{i,t}$, given that empirical estimates show, ex post, few covariates contribute the vast majority of the signal, as stated earlier.[4] This would amount to adding a regularization parameter to the aforementioned objective to combat the curse of dimensionality from $z_{i,t}$ [5] However, although potentially helpful for minimizing MSE by decreasing the variance of the estimator, this regularization introduces a bias in estimation, which will lead to invalid asymptotic inference for asset pricing tests, defeating a goal of this work.

We therefore adapt for our purpose the Double Selection Lasso estimator introduced by Belloni, Chernozhukov, and Hansen (2014). The key insight from their work was to introduce an orthogonality wherein, assuming Γ_β is row sparse, the regularization bias from the LASSO first-stage estimation does not pass through to the target parameter of interest when conducting inference.[6] First, we estimate for each time period t and

[4] The online implementation of IPCA Kelly, Pruitt, and Su (2019) does indeed offer an ℓ_1 regularization to their MSE objective, which is not discussed in the econometric theory of Kelly, Pruitt, and Su (2020).

[5] One of several ways to interpret the curse of dimensionality is that as the number of covariates increases linearly, the volume of the parameter space to estimate grows nonlinearly; hence the density of the data falls.

[6] The ideas developed in the Double Selection Lasso paper for inference in partially linear models with high-dimensional controls were the basis for the generalization of this idea in the DML procedures

each characteristic j the scalar $\Gamma_{\beta,j}^{\top} f_{t+1}$ using DSL; then stacking these estimates into a $T \times p$ matrix, we use PCA to obtain separate estimates for latent loadings $\widehat{\Gamma}_{\beta}$ and factors $\{\widehat{f}_{t+1}\}_{t=1}^{T}$; and, finally, soft-threshold $\widehat{\Gamma}_{\beta}$ to set the majority of the rows to zero given the assumption of sparsity.

We rewrite the DSLFM model and introduce a first-stage:

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^{\top} c_{t+1,-j} + \epsilon_{i,t+1}, \qquad && \mathbb{E}[\epsilon_{i,t+1} | z_{i,t}] = 0, \\
z_{i,t,j} &= z_{i,t,-j}^{\top} \delta_{t,j} + \epsilon_{i,t,j}^{z}, && \mathbb{E}[\epsilon_{i,t,j}^{z} | z_{i,t,-j}] = 0,
\end{aligned} \tag{2}
$$

where $c_{t+1,j}$ refers to the $j \in \{1, \ldots, p\}$ component of $c_{t+1} := \Gamma_{\beta} f_{t+1}$ while $-j$ refers to the remaining $p - 1$ elements of the vector; $\delta_{t,j} \in \mathbb{R}^{(p-1)}$ is an unknown, possibly time-varying, parameter; and, $\epsilon_{i,t,j}^{z}$ is an unknown scalar random error. $c_{t+1,j}$ is an asset return when its j–th characteristic is set to 1 and all other characteristics are set to zero, less its idiosyncratic return $\epsilon_{i,t+1}$.

There are several ways to interpret and justify the first-stage equation as discussed in Belloni, Chernozhukov, and Hansen (2014). Intuitively, the procedure does not rely on perfect model selection for valid inference as instead we not only recover controls $z_{i,t,-j}$ in the second-stage equation for their pricing ability in the cross section of returns but also recover controls with high correlation to our target variable $z_{i,t,j}$. From a theoretical perspective, the first-stage equation accounts for potential omitted variable bias if one estimated only the second-stage equation. That is, the set of potentially relevant asset covariates is enormous (Chen, Pelger, and Zhu (2020) and Bianchi, Guidolin, and Pedio (2022)), and thus a researcher may be motivated to select their preferred subset to ameliorate the curse of dimensionality, which could introduce model selection mistakes. Moreover, it is known LASSO can have poor finite sample model selection performance (Chernozhukov, Hansen, and Spindler 2015). Thus, selecting covariates with only the second-stage equation could fail to include relevant controls.

as developed in Chernozhukov et al. (2018). It would likely be closer to the empirical reality to maintain a nonparametric loading (Fan, Liao, and Wang 2016). As aforementioned, there is thus a natural extension of the work herein to use DML wherein the target variable $c_{t,j}$ is linear while the controls are nonparametrically estimated via a machine learning method, which would require the development of a Neyman Orthogonal score for this panel data setting, perhaps in a similar fashion to Semenova and Chernozhukov (2021). However, the econometric theory is unknown for inference in the nonparametric setting. Moreover, given the high-dimension asset characteristics, using a conditional independence assumption to obtain a causal parameter may be the most fruitful path toward a causal factor model, a major area of future work (e.g., Lopez de Prado (2022)). We explore the out-of-sample predictive ability of a non-parametric model in the last section of this manuscript.

DSLFM Estimation Procedure. Our estimation procedure for $\{f_{t+1}\}_{t=1}^{T}$ and Γ_β has three steps: Double Selection Lasso, PCA, and soft-thresholded step.

(i) *DSL*: To estimate $\hat{c}_{t+1,j}$, run $T \times p$ Double Selection Lasso cross-sectional regressions. [7]

- Run LASSO of $\{r_{i,t+1}\}_{i=1}^{N}$ on $\{z_{i,t}\}_{i=1}^{N}$ for $\hat{c}_{t+1,j}$ and $\hat{c}_{t+1,-j}$.
  - Let $\hat{I}_1$ denote the nonzero elements of $\hat{c}_{t+1,-j}$.
- Run LASSO $\{z_{i,t,j}\}_{i=1}^{N}$ on $\{z_{i,t,-j}\}_{i=1}^{N}$ for $\hat{\delta}_{t,j}$.
  - Let $\hat{I}_2$ denote the nonzero elements of $\hat{\delta}_{t,j}$.
- Define the set $\hat{I} = \hat{I}_1 \cup \hat{I}_2 \cup \hat{I}_3$ where $\hat{I}_3$ is the set of controls in $z_{i,t,-j}$ not included in the first two LASSOs that the econometrician thinks are important for ensuring robustness, termed the amelioration set.
- Run OLS of $\{r_{i,t+1}\}_{i=1}^{N}$ on $\{z_{i,t,j}, \tilde{z}_{i,t,-j}\}_{i=1}^{N}$ where $\tilde{z}_{i,t,-j}$ includes only elements of $z_{i,t,-j}$ in $\hat{I}$. That is,

$$(\hat{c}_{t+1,j}, \hat{c}_{t+1,-j}) \coloneqq \arg\min_{c_j, c_{-j}} \{\mathbb{E}_N[(r_{i,t+1} - z_{i,t,j}c_{t+1,j} - z_{i,t,-j}^{\top} c_{t+1,-j})] \ : \ c_{t+1,-j,l} = 0, \forall l \notin \hat{I}\}.$$

(ii) *PCA*: To estimate Γ_β and f_{t+1}, run PCA on $\hat{C} = \widehat{F}\widehat{\Gamma}_\beta^{\top}$—a $T \times p$ matrix—to decompose it into $p \times k$ and $T \times k$ matrices $\hat{\Gamma}_\beta$ and $\widehat{F}$.

(iii) *Soft-thresholded*: Given the assumed exact row sparsity of Γ_β, we set to zero all rows of $\hat{\Gamma}_\beta$ whose row ℓ_1 norm is below a cross-validated hyperparameter λ. That is,

$$\check{\Gamma}_{\beta,j} = \left(\left\|\widehat{\Gamma}_{\beta,j}\right\|_1 - \lambda\right)_+ \operatorname{sign}\left(\left\|\widehat{\Gamma}_{\beta,j}\right\|_1\right), \;\; j \in \{1, \ldots, p\}. \tag{3}$$

This does require running $T \times p$ versions of the cross-sectional Double Selection Lasso regressions, which can be in the thousands in empirical settings; however, these regressions are all computationally light and can be run in parallel. Moreover, this allows for unbalanced panels as each cross-section can have a different number of assets. [8] Additionally, these cross-sectional regressions, followed by estimations with

[7] To set the penalty parameter in the LASSO implementations, one can follow the analytic method developed for heteroskedastic, non-Gaussian settings detailed in Appendix A, Algorithm 1 of Belloni, Chernozhukov, and Hansen (2014). For a more modern approach, one can use the bootstrap-after-cross-validation method of Chetverikov and Sorensen (2021). In practice, we use cross validation.

[8] In empirical practice, we find the entire estimation procedure is on the order of ten minutes.

the entire panel, mirror the effort of the most commonly used estimation procedure in the factor model setting, namely two-pass Fama-MacBeth regressions.

The high dimensionality of the PCA procedure, given we have a $p \times T$ matrix, is adapted from the existing literature using $N \times T$ matrices (Bai 2003). The estimated factor matrix $\widehat{F}$ is the product of $\sqrt{T}$ and the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrix $(T\,p)^{-1}\widehat{C}\widehat{C}^\top$. The estimated factors are normalized such that $\widehat{F}^\top\widehat{F} = I_{k\times k}$, a standard approach. The estimated loadings are $\widehat{\Gamma}_\beta = T^{-1}\widehat{C}^\top\widehat{F}$. We thus see the main challenge in deriving the large-sample theory will be handling the estimation error in using $\widehat{C}$ instead of the unobserved C.

The final soft-thresholded step 3 in our estimation procedure exploits the sparsity in Γ_β to not only reduce the dimensionality of the characteristic space s ($<<\ p$) but also remove noise from the characteristics that have low signal-to-noise ratios. Belloni et al. (2018) discuss the general theoretical properties of the soft-thresholded estimator with theory-based hyperparameter selection, and its close relation, the better known LASSO and Dantzig selector estimators. We further discuss constraints and selection of the hyperparameter in Appendix C.

Estimating the Risk Premium of an Observable Factor. Under the richer setting that includes the observable factor g_{t+1}, our model has an additional specification and moment conditions.

$$\begin{aligned} r_{i,t+1} &= z_{i,t}^\top\Gamma_\beta(\gamma + \nu_{t+1}) + \epsilon_{i,t+1}, \quad \mathbb{E}[\nu_{t+1}] = \mathbb{E}[\epsilon_{i,t+1}] = 0, \quad \mathbb{E}[\nu_{t+1}\epsilon_{i,t+1}] = 0, \\ g_{t+1} &= \eta\nu_{t+1} + \epsilon^g_{t+1}, \qquad\qquad\qquad\qquad\qquad \mathbb{E}[\epsilon^g_{t+1}] = 0, \quad \mathbb{E}[\nu_{t+1}\epsilon^g_{t+1}] = 0. \end{aligned} \tag{4}$$

Our goal is to estimate and conduct inference on $\gamma_g := \eta^\top\gamma$. Given the latent factors are unobserved, we cannot jointly estimate ν_{t+1} and η without further restrictions. We would have to invoke one of the three classic identification approaches of Bai and Ng (2013); however, by using the key rotation invariance result of Giglio and Xiu (2021), we can estimate the latent factors up to an invertible rotation matrix $H \in \mathbb{R}^{k\times k}$ and still maintain identification of our target parameter γ_g. That is, both of the underlying parameters will be identified up to this rotation matrix: $\eta^\top = \eta_0^\top H^{-1}$ and $\gamma = H\gamma_0$. Thus, the target parameter is rotationally invariant to H : $\gamma_g = \eta_0^\top H^{-1}H\gamma_0 = \eta^\top\gamma$.

For our estimation procedure, we replace the first PCA step of Giglio and Xiu (2021) with our above procedure—augmented to use return innovations—to estimate the latent loadings $\check{\Gamma}_\beta$ and factor innovations $\hat{\nu}_{t+1}$ for all t. We then proceed with the latter two steps of the authors' procedure to obtain our target estimator.

(i) To estimate latent-factor risk premia $\hat{\gamma}$, run cross-sectional OLS of average returns $\bar{r} \in \mathbb{R}^N$ on averaged estimated latent-factor loadings $\bar{\hat{\beta}} = \bar{Z}^\top \hat{\Gamma}_\beta \in \mathbb{R}^{N \times k}$.

(ii) To estimate latent to observable factor mapping $\hat{\eta}$, run a time series OLS regression of $\{g_{t+1}\}_{t=1}^T$ on factor innovations $\hat{V} \in \mathbb{R}^{T \times k}$.

We can thus form our estimator of the risk premium for the observable factors g_{t+1} by combining these estimators into $\hat{\gamma}_g = \hat{\eta}^\top \hat{\gamma}$.

This procedure extends the estimation in Giglio, Xiu, and Zhang (2021) to dynamic loadings and high-dimensional asset characteristics, while inheriting the rotation invariance and the specification consistent with two-pass estimators in this literature. Again, simply applying IPCA instead of PCA in the first step of Giglio and Xiu (2021) would not be feasible with $p > \max\{N, T\}$ or would yield biased inference if an ℓ_1 penalty were simply added to the IPCA objective. The cross-sectional OLS of average returns on the estimated latent-factor loadings is the standard second step in two-pass Fama-MacBeth regressions, which could be replaced with generalized least squares or weighted least squares to explore asymptotic efficiency gains. The final time series regression is critical to translate the uninterpretable risk premia of latent factors to those of factors proposed by economic theory. Moreover, this procedure handles omitted variable bias which we now briefly discuss.

To illustrate, assume we have a scalar observable factor g_{t+1}, which is the first component of a two-dimensional latent-factor innovation vector: $v_{t+1} = (g_{t+1}, v_{2,t+1})^\top$ (i.e. $\eta = (1, 0)$). The vector-version of our model is thus

$$r_{t+1} = z_t \Gamma_{\beta,1}(\gamma_g + g_{t+1}) + z_t \Gamma_{\beta,2}(\gamma_2 + v_{2,t+1}) + \epsilon_{t+1}.$$

Using the standard Fama-MacBeth two-pass regressions (Fama and MacBeth 1973) will produce bias in estimating γ_g if $v_{2,t+1}$ is omitted. The first step of a time series regression of asset excess returns on g_{t+1} will give a biased estimate of $\hat{\beta}_1$ as long as $v_{2,t+1}$ is correlated with both g_{t+1} and r_{t+1}, per the standard OVB term: the covariance between between the outcome and the excluded regressor times the covariance between the included and excluded regressor, up to scale. Moreover, in the second step of a cross-sectional regression of average returns on estimated loadings, a second omitted variable bias is introduced if the loading of the omitted factor $\hat{\beta}_2$ is correlated with both $\hat{\beta}_1$ and $\bar{r}_{t+1}$.

Estimating the latent factors via the DSLFM procedure resolves this issue of omitting a potentially relevant factor given one can utilize a consistent estimator of the true number of latent factors, which we assume spans the true factor space.[9]

5.3. Asymptotic Theory

In this section, we present the asymptotic results for consistent estimation of the latent factors and loadings and the large sample distribution of the nontradable observable factor risk premium estimator under the assumed setting discussed in Section 5.1 and using estimation procedures discussed in Section 5.2 for models (2) and (4). We first provide the regularity conditions sufficient for the validity of the estimation and inference results. For clarity of exposition, we focus on motivating the assumptions and interpreting the results, while theoretical details and mathematical proofs are provided in Appendix C.

Throughout, let $\|A\| = [tr(A^\top A)]^{1/2}$ denote the Frobenius norm of matrix A or $\|x\| = \left(\sum_i x_i^2\right)^{1/2}$ for the ℓ_2 norm of a vector x. Let $\|x\|_0$ and $\|x\|_1$ be the usual ℓ_0 and ℓ_1 norms, respectively. All limits are simultaneous where we will restrict the rates among p, T, N, to allow $p \to \infty$, as discussed below.

5.3.1. Regularity Conditions

Consistent Estimators for the Latent-Factor Model. The following assumptions enable the consistent estimation of the factors $\{f_{t+1}\}_{t=1}^T$ and the loadings Γ_β. Let f_{t+1}^0 and Γ_β^0 be the true factors and loadings such that $f_{t+1} = Hf_{t+1}^0$ and $\Gamma_\beta = \Gamma_\beta^0 H^{-1}$ where H is an unobserved $k \times k$ invertible rotation matrix.

In regard to identification, our results do not require the identification of the true factors f_{t+1}^0 and loadings Γ_β^0 but rather simply factors (loadings) that span the true factors (loadings) up to the rotation matrix H. Bai and Ng (2013) show identification results for PCA under three different sets of assumptions to pin down the $k \times k$ elements in H, which requires pinning down the covariance matrices of the factor loadings and factors to be diagonal matrices or identity matrices to provide $k(k-1)/2 + k(k+1)/2 = k^2$ restrictions. The researcher can choose which asymptotic covariance matrix to restrict. As we will discuss, we will additionally not need these identification restrictions for the observable factor risk premia given the aforementioned rotation invariance result of the target parameter.

[9] The DSLFM could be further extended to estimate the zero-beta rate (i.e., alpha) using a very similar approach to that discussed in Online appendix I.2 of Giglio and Xiu (2021).

Assumption 1 (Consistency of DSL). (i) *Bounded Characteristic Portfolios: For a finite absolute constant M and $\forall t, j$, $\left|c_{t+1,j}\right| = \left|\Gamma_{\beta,j}^{\top} f_{t+1}\right| < M$.*

(ii) *Sparse Loading: Loading matrix Γ_β admits an exactly sparse form. That is, for $\exists s \in \mathbb{N}_+$, i.e. $p > s \geq 1$, Γ_β has at most s nonzero rows: $\sum_{j=1}^{p} \mathbb{1}\left\{\left\|\Gamma_{\beta,j}\right\|_1 > 0\right\} \leq s$.*

These are two critical assumptions for DSL consistency with the additional standard and technical DSL assumptions in Appendix C. Assumption 1(i) converts the bounded target parameter, in the traditional DSL context, to the DSLFM context where we require realizations of $c_{t+1,j}$ to be finite-sample bounded by a constant that does not depend on p, T, N. This imposes a bound on the return of characteristic portfolios, that is, the return of a portfolio with characteristic j set to 1 and all other characteristics set to 0. We could instead assume returns are bounded random variable to impose Assumption 1(i).

Assumption 1(ii) is the key LASSO assumption that the parameter on the control regressors admits an exactly sparse form, which follows from our assumption such that $\forall t, j$, $\left\|c_{t+1,-j}\right\|_0 = \left\|\Gamma_{\beta,-j} f_{t+1}\right\|_0 \leq s$. This sparsity of the loading matrix is supported empirically in asset pricing given the relevance of only a small number of asset characteristics, which we corroborate in our empirical setting. We have thus adapted the classic LASSO sparsity assumption to the empirical reality of cross-sectional asset pricing using high-dimensional asset characteristics. Exact sparsity could be relaxed to approximate sparsity with a similar but alternative high-dimensional econometrics toolkit.

We next turn to assumptions for consistently estimating the latent factors and loadings. The focus in our work is controlling the estimation error between the infeasible eigendecomposition of $(Tp)^{-1}CC^{\top}$ and the feasible eigendecomposition of $(Tp)^{-1}\widehat{C}\widehat{C}^{\top}$, given we do not observe $C = F\Gamma_\beta^{\top}$ but instead estimate each element via DSL and then eigendecompose using standard PCA estimators as discussed in Section 5.2.

Assumption 2 (Consistency of Latent-Factor Model). (i) *Factors: $\mathbb{E}\left\|f_{t+1}^{0}\right\|^4 \leq M < \infty$ and $T^{-1}\sum_t f_{t+1}^{0} f_{t+1}^{0\top} \rightarrow_p \Sigma_f$ for some $k \times k$ positive definite matrix Σ_f.*

(ii) *Factor Loadings: $\forall j$, $\left\|\Gamma_{\beta,j}\right\| \leq M < \infty$ and $\left\|\Gamma_\beta^{\top}\Gamma_\beta/p - \Sigma_\Gamma\right\| \rightarrow 0$ for some $k \times k$ positive definite matrix Σ_Γ.*

(iii) *Nonzero and distinct eigenvalues: from the infeasible eigendecomposition, the k largest eigenvalues λ_i for $i \in \{1, \ldots, k\}$ are bounded away from zero. Moreover, the k largest* *infeasible eigenvalues are distinct, that is,*

$$\min_{i:i\neq\kappa} |\lambda_\kappa - \lambda_i| > 0.$$

Assumptions 2(i)-(ii) are standard for factor models where the literature is styled after Assumptions A, B, and C of Bai (2003). Assumption 2(i) does not impose i.i.d. factors, as in the classical factor analysis literature, but instead imposes the factors are stationary, strong mixing, and satisfy moment conditions. Assumption 2(ii) ensures each latent factor contributes to the second moment of c_{t+1}; that is, it imposes all factors are pervasive and excludes weak factors. See Giglio, Xiu, and Zhang (2021) for adjustments for weak factors. The PCA estimation herein does not require the Assumption C of Bai (2003) given our target matrix $C = F\Gamma_\beta^\top$ is without an error term; we instead are controlling cross-sectional and temporal dependence using the moment conditions of DSL given in model (2) and more technical assumptions in Appendix C.

Assumption 2(iii) assumes the k–largest eigenvalues from the infeasible and feasible eigendecompositions remain nonzero asymptotically. In finite sample, these are real and nonzero eigenvalues given we are taking the eigendecomposition of a rank k symmetric matrix. It is reasonable to assume we have distinct eigenvalues given, for this not to hold, there would have to be two or more dimensions in the k–largest of the $T \times T$ matrix $CC^\top$ that have precisely the same variability.

Estimating Number of Factors. Given the focus of this work is on the consistency of the main estimators and the asymptotic distribution of the risk premium estimator, we assume $k = k^0$ is known.[10]

Assumption 3 (Consistent Estimator for Number of True Factors). *For $\bar{k} > k^0$, let*

[10] The asymptotic distribution of the risk premium estimator is unaffected when the number of factors is estimated because

$$\begin{aligned}\Pr\left(\widehat{\gamma}_g \leq x\right) &= \Pr\left(\widehat{\gamma}_g \leq x, \widehat{k} = k^0\right) + \Pr\left(\widehat{\gamma}_g \leq x, \widehat{k} \neq k^0\right) = \Pr\left(\widehat{\gamma}_g \leq x, \widehat{k} = k^0\right) + o(1)\\ &= \Pr\left(\widehat{\gamma}_g \leq x | \widehat{k} = k^0\right) \Pr\left(\widehat{k} = k^0\right) + o(1) = \Pr\left(\widehat{\gamma}_g \leq x | \widehat{k} = k^0\right) + o(1).\end{aligned}$$

$\widehat{k} := \arg\min_{0 \leq k \leq \bar{k}} IC(k)$ *where*

$$
\begin{aligned}
IC(k) &:= \log(V(k)) + k\left(\frac{p+T}{pT}\right)\log\left(\frac{pT}{p+T}\right) \\
V(k) &:= \min_{\Gamma_\beta, F}(pT)^{-1}\sum_{j,T}\left(c_{j,t+t} - \Gamma_{\beta,j}^{\top} f_{t+1}\right)^2 .
\end{aligned}
$$

Assume $\widehat{k} \rightarrow_p k^0$ *without further restriction on the growth rates among* p, T, N *and* $k = k^0$ *is known.*

Assumptions 1 and 2 can be shown to be sufficient for consistently estimating, with the above Information Criterion, the number of true factors k^0 using the results of Appendix C as in Bai and Ng (2002) and Bai (2003). Although providing this assumption to show the estimator to be studied in simulation, we are instead choosing to impose Assumption 3 in the asymptotic to focus on the main results of this work. Note that commonly used model selection criteria (e.g., AIC or BIC) will not yield consistent estimators, hence the specification above using the contribution of Bai and Ng (2002).

Inference on Nontradable Observable Factor Risk Premia. The final assumptions are needed to derive the limiting distribution of the risk premium estimator.

Assumption 4 (Inference). *There exists a generic absolute constant* $M < \infty$ *such that for all* p, T, N :

(i) *Bounded idiosyncratic errors:* $\mathbb{E}[(\sum_t \epsilon_{i,t+1})^2] \leq TM$.

(ii) *Bounded scaled factor innovations:* $\mathbb{E}[(\sum_t z_{i,t}^{\top}\Gamma_\beta^0 v_{t+1}^0)^2] \leq sTM$.

(iii) *Bounded measurement errors:* $\mathbb{E}[(\epsilon_{t+1}^g)^2] \leq M$.

(iv) *Convergence of characteristics:* $\frac{1}{NT}\sum_i \sum_{t'} \mathbb{E}[z_{i,t,j}]z_{i,t',j'} \rightarrow_p \mathcal{Z}_{t,j,j'}$ *uniformly over* t, j, j' *for* $j, j' \in \{1, 2, \ldots, p\}$ *and a nonstochastic finite constant* $\mathcal{Z}_{t,j,j'} \in \mathbb{R}$.

(v) *CLT: As* $T \rightarrow \infty$, *the following joint central limit theorem holds:*

$$
\frac{\sqrt{T}}{T}\sum_t \begin{pmatrix} v_{t+1}^0 \epsilon_{t+1}^g \\ \Pi_t v_{t+1}^0 \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, \Phi)
$$

where random matrix $\Pi_t \in \mathbb{R}^{k \times k}$ *and nonstochastic matrix* $\Phi \in \mathbb{R}^{2k \times 2k}$ *are defined in Appendix C.*

Assumption 4(i) bounds the second contemporaneous and cross-moments of the idiosyncratic errors, aligning with the time and cross-section dependence assumptions of Bai (2003) Assumption C. The assumption would hold if we assumed $\epsilon_{i,t+1}$ are uncorrelated across t, which is a simplified yet plausible assumption given the low signal-to-noise environment of asset pricing. We have thus relaxed the temporal dependence to the specified rate T.

Assumption 4(ii) bounds the squared time series average of the factor innovations scaled by the factor loadings. In the static factor model context of Giglio and Xiu (2021), this holds in large sample by a simple LLN argument given the static loadings are not a function of t and the factor innovations are mean zero random variables. Thus, we are ensuring the Γ_β^0 selected columns of Z_t keep the scaled ν_{t+1}^0 sufficiently small.

Assumption 4(iii) bounds the second moment of the observable factor measurement errors for use in proving $\|\epsilon^g\| = O_p(\sqrt{T})$. It is not a stringent assumption because we are bounding a zero mean scalar random variable. This is nearly an identical assumption and usage to Giglio and Xiu (2021) Assumption A8.

Assumption 4(iv) provides a convergence result such that the squared first moment for two different characteristics averaged over time and across assets is a nonstochastic finite constant. This is a weaker assumption on the distribution of characteristics than the DSL moment conditions discussed in Appendix C.

Assumption 4(v) is the assumed central limit theorem for the $2k$ (low) dimensional mean zero random variable given the models' 2 and 4 moment assumptions, which is satisfied by various mixing processes. The second moments of the later $2k$ random variables are bound already in Assumptions 4(i)-(ii). We nevertheless directly assume the needed CLT. This extends for our inference result the assumed CLT at the same rate in Assumption F4 of Bai (2003) and the assumed CLT at the same rate in Assumption A11 of Giglio and Xiu (2021). Note that although we have the same two mean zero random vectors, our factor innovations are scaled by Π_t instead of a constant 1 given the dynamic factor loadings of our model.

5.3.2. Theory Results

This section presents the three main theoretical results.

Consistent Estimators for the Latent-Factor Model. We present the first two results showing the consistency of the latent-factor model estimators.

PROPOSITION 1 (Consistency of Latent Factors). *Under the model* (2)*, Assumptions 1, 2, 3, and DSL Assumptions in Appendix C.2 where $T, N, p \to \infty$, then for all t the latent-factor estimator described above has the property that*

$$\widehat{f}_{t+1} - H^{\top} f^0_{t+1} = O_p\left(\sqrt{\frac{s \log(Tp)}{N}}\right).$$

The proof is in Appendix C. This result establishes the convergence rate of the latent factor estimator in a dynamic latent-factor model with high-dimensional characteristics. If the factor loadings were static and known, β^0_i for all i, then, f^0_{t+1} would be estimated via a cross-sectional least squares with a convergence rate of $\sqrt{N}$. Bai (2003) establishes in Theorem 1(ii), under $N, T \to \infty$ for a static latent-factor model, the foundational result of a convergence rate of $\min\left(\sqrt{N}, T\right)$ for the consistency of the latent-factor estimator for the rotated true factors $H^{\top} f^0_{t+1}$. Incorporating dynamic loadings parameterized by high dimensional characteristics comes at the cost of slowing the rate to $\sqrt{pN/s \log(Tp)}$, which is nevertheless still reasonable for typical values of p, T, N. Additional standard DSL rates, which are less restrictive, are in Appendix C.

Our rate is primarily driven by the $\sqrt{N/\log(Tp)}$ rate uniform consistency over t and j of the DSL estimation error $|\widehat{c}_{t+1,j} - c_{t+1}, j|$ as shown in Lemma A1 in Appendix C. Given the model for $C = F^0 \Gamma^{0\top}_{\beta}$ contains no error, the eigendecomposition of the unobserved C is exact for $F^0 H$ as shown in Lemma A6; and, thus, the estimation error from using $\widehat{C}$ instead of C drives this first main result. The assumed sparsity in Γ^0_{β} does improve the rate with the p/s ratio.

It is worth reiterating that under our setting F^0 and Γ^0_{β} are not separately identifiable, hence the $k \times k$ invertible matrix transformation H appears in each asymptotic result. Similarly, $\widehat{F}\widehat{\Gamma}^{\top}_{\beta}$ is an estimator of the identifiable, rotation invariant common component C, which is corroborated by simulation results. Moreover, in many cases knowing $F^0 H$ is equivalent to knowing F^0; for example, the regressor F^0 will give the same predicted values as using $F^0 H$ as a regressor given they have the same column space.

PROPOSITION 2 (Consistency of Latent-Factor Loadings). *Under the model* (2)*, Assumptions 1, 2, 3, and DSL Assumptions in Appendix C.2 where $T, N, p \to \infty$, then the latent loading estimator described above has the property that*

$$\left(\check{\Gamma}_{\beta} - \Gamma^0_{\beta} H^{-1}\right) = O_p\left(\sqrt{\frac{s \log(Tp)}{N}}\right).$$

The proof is in Appendix C. This result establishes the convergence rate of the latent loading estimator in a dynamic latent-factor model with high-dimensional characteristics. When the factors, f_{t+1}^0, for all t, are observable, static loadings β_i^0 can be estimated by a time series regression with a convergence rate of $\sqrt{T}$. Bai (2003) establishes in Theorem 2(ii), under $N, T \to \infty$ for a static latent-factor model, the foundational result of a convergence rate of $\min\left(N, \sqrt{T}\right)$ for the consistency of the latent-factor loading estimator $\widehat{\beta}_i$ for the rotated true factor loadings $H^{-1}\beta_i^0$. Incorporating dynamic loadings parameterized by high dimensional characteristics comes at the cost of slowing the rate to $\sqrt{N/s\log(Tp)}$, which is nevertheless still reasonable for typical values of p, T, N.

The rate follows similar reasoning to that of the latent-factor estimator in Proposition 1. However, here we are presenting the rate of the final soft-thresholded estimator—derived using recent results in high dimensional econometrics Belloni et al. (2018)—wherein we use the uniform estimation error between the eigendecomposition of $\widehat{C}$ for $\widehat{\beta}_{\beta,j}$ and the infeasible loading $\widetilde{\beta}_{\beta,j}$ from decomposing the unobserved C, which eliminates p in our rate from Proposition 1. The $\sqrt{N/log(Tp)}$ is similarly driven by the uniform consistency over t and j of the DSL estimation error $|\widehat{c}_{t+1,j} - c_{t+1}, j|$, which is the key result used to establish these consistency propositions along with typical high dimensional random matrix theory (e.g., Davis Kahan Theorem, Weyl Inequality, and recent tools in high dimensional econometric theory found in Belloni et al. (2018)).

Inference on Nontradable Observable Factor Risk Premium. Finally, we present the asymptotic normality of the nontradable observable factor risk premium estimator.

THEOREM 1 (Normality of Observable Factor Risk Premium). *Under the models* (2) *and* (4)*; Assumptions 1, 2, 3, 4; DSL Assumptions in Appendix C.2; and if* $Ts^2\log(Tp)/N \to 0$, *then as* $T, N, p \to \infty$ *the estimator* $\hat{\gamma}_g$ *obeys*

$$\sqrt{T}\frac{(\hat{\gamma}_g - \gamma_g)}{\sigma_g} \xrightarrow{d} \mathcal{N}(0,1),$$

where σ_g *is defined in Appendix C.*

The proof is in Appendix C. This result establishes $\sqrt{T}$ asymptotic normality of the nontradable observable factor risk premium estimator from Giglio and Xiu (2021) extended to the setting of dynamic factor loadings with high dimensionality characteristics. At a high level, our proof follows a similar approach yielding, as seen in the proof of Theorem 1, the same two asymptotically nonnegligible terms as in Giglio and Xiu (2021).

The first term arises from the time-series regression of the observed factor on the latent loadings where again the latent loading estimation error is higher order. The latter term in the $2k$ random vector in Assumption 4(v) arises from the cross-sectional regression of averaged asset excess returns on averaged factor loadings, where the factor loading estimation error and idiosyncratic error term are higher order. Although Giglio and Xiu (2021) have this same second term, ours is more complicated given the dynamic loadings, which necessitates the convergence Assumption 4(iv). A direct application of the delta method on the sum of these two terms yields the result in Theorem 1.

The crucial rate assumption is $Ts^2 \log(Tp)/N \to 0$, which controls the estimation error for the unobserved averaged latent-factor loadings $T^{-1} \sum_t \beta_{i,t}$. This is similar to Bai (2003) and Giglio and Xiu (2021), which require $T/N \to 0$ to use the estimated factors or loadings as generated regressors. However, we have slowed the rate again due to the high-dimensionality in p. This is our slowest required rate.

Given the rotation invariance of the target parameter γ_g^0, the unobserved rotation matrix H does not appear in the asymptotic distribution, in contrast to the consistency results. In finite sample, we corroborate this result in the to-be-discussed simulations. In a similar vein, it warrants noting the asymptotic efficiency loss due to not observing the factors and loadings could be large when N is relatively small. Also, we have assumed we directly observe the number of tree factors k^0, which would require estimation in practice and thus likely contribute estimation error to affect finite sample performance.

In simulation, we use the plug-in estimator for σ_g, which has satisfactory finite-sample coverage properties. However, one can establish a consistent variance estimator using a Newey and West (1987) style plug-in estimator of the asymptotic variance σ_g with lag corrections to account for temporal dependence as in Giglio and Xiu (2021) Section IV Part E.

5.4. Asset Pricing Tests

In this section we develop three tests central to our empirical analysis. The first uses the asymptotic normality of the observable factor risk premium for a statistical test of nonzero risk compensation. The second statistic informs the incremental significance of any specific asset characteristic. The third and final discusses how we empirically measure whether the DSLFM contributes predictive signal above and beyond a random walk.

Testing Nontradable Observable Factor Risk Premium. An empirical application of the DSLFM model will address whether a nontradable observable factor, namely, inflation, carries a nonzero risk premium in the crypto asset class. The target parameter γ_g captures the risk premium of the (inflation) factor-mimicking portfolio within the crypto asset class as recovered by the estimated dynamic latent-factor model. We are interested not only in the sign of the parameter, but also, in practical settings, in whether a confidence interval suggests a risk premium of economic significance.

We test the hypothesis $H_0 : \gamma_g = 0$ vs. $H_1 : \gamma_g \neq 0$ using the risk premium estimation procedure described in Section 5.2 with a plug-in variance estimator $\widehat{\sigma}_g$ for σ_g. Given the asymptotic normality of Theorem 1, we form a confidence interval

$$\gamma_g \in \left[\widehat{\gamma}_g - c(1-\alpha/2)\widehat{\sigma}_g, \widehat{\gamma}_g + c(1-\alpha/2)\widehat{\sigma}_g\right]$$

where the critical value $c(1-\alpha/2)$ is the $1-\alpha/2$ quantile of a $N(0,1)$ distribution for the researcher-specified level of the test α. We find in the coming simulation acceptable finite sample coverage for this confidence interval.

Testing Characteristic Significance. The large-sample distribution of the latent loading is unknown given the DSL regularization. Even inference in simple cross-sectional LASSO is complicated (Lee et al. 2016). Instead, we develop a simple bootstrap procedure to infer whether a specific characteristic significantly contributes to loading Γ_β. We leave for subsequent research developing the supporting theory of this bootstrap procedure or develop the asymptotic distribution of a consistent latent loading estimator in this setting. We test the hypotheses $\forall j$

$$H_0 : \Gamma_\beta^\top = \left[\Gamma_{\beta,1}, \ldots, \Gamma_{\beta,j-1}, 0, \Gamma_{\beta,j+1}, \ldots, \Gamma_{\beta,p}\right] \quad \text{vs.} \quad H_1 : \Gamma_\beta^\top = \left[\Gamma_{\beta,1}, \ldots, \Gamma_{\beta,p}\right].$$

That is, we ask whether characteristic j contributes to the factor loading through $k \times 1$ mapping vector $\Gamma_{\beta,j}$. This allows the researcher, using a large number of characteristics, to systematically ask what characteristics contribute to the latent-factor model, instead of an ad hoc selection. We thus set the entire $k \times 1$ vector to zero so the characteristic contributes to predicting the variation in returns through none of the k factors.

Our procedure is to test the alternative hypothesis model, with the unconstrained characteristic j, and then form the test statistic

$$W_{\Gamma,j} = \Gamma_{\beta,j}^\top \Gamma_{\beta,j}.$$

Using bootstrapped standard errors, we assess whether this test $W_{\Gamma,j}$ statistic is statistically distinguishable from zero.

Testing Out-of-Sample Performance. To study the out-of-sample pricing ability of the DSLFM, we use the "predictive R^2" defined as

$$\text{Predictive } R^2 = \frac{\sum_{i,t}\left(r_{i,t+1} - z_{i,t}^{\top}\check{\Gamma}_{\beta}\widehat{\lambda}_t\right)^2}{\sum_{i,t} r_{i,t+1}^2}$$

where $\widehat{\lambda}$ is the moving average of the estimated factors in previous time periods over a cross-validated window size. This measure captures whether the model forecasts realized returns better than a random walk; or, said differently, it represents the fraction of realized return variation explained by the model's description of expected returns through exposure to systematic risk. This specification allows the model's estimated conditional expected returns to be driven not just by the dynamic factor loadings, estimated using high-dimensional asset characteristics, but also by time-varying risk prices λ_t.

5.5. Simulations

This section presents a brief study of the finite-sample performance of the dynamic latent-factor model estimators and coverage properties of the inference procedure using Monte Carlo simulations. To summarize, we find the estimation errors for factors and loadings are comparable to IPCA and the Three Pass estimator of Giglio and Xiu (2021) in low-dimensional settings while superior in high-dimensional settings. This holds even in rather small samples with low signal to noise ratios, reflecting the empirical reality of cross-sectional asset pricing. Moreover, we find estimation errors and coverage properties for the observable-factor risk premium to be comparable to Giglio and Xiu (2021) in low-dimensional settings while superior in high-dimensional settings. We now present the design, followed by the results.

5.5.1. Simulation Design

First, we describe the data-generating process for given N, T, k where we follow the finite sample simulation study of IPCA (Kelly, Pruitt, and Su 2020). That is, the DGP is favorable to IPCA. We calibrated the simulated data to parameter estimates from IPCA fit to our weekly panel of crypto asset data using all 63 asset characteristics.

Latent factors f_{t+1} are simulated from a $VAR(1)$ model employing normal innovations that was fit to the estimated IPCA factors. Asset characteristics are simulated from a p variable panel $VAR(1)$ model with normal innovations, which was fit to the demeaned empirical weekly panel of randomly selected, without replacement, p asset characteristics. For each asset, we set the means of the characteristics to a bootstrap sample from the empirical distribution of time series asset characteristic means. The idiosyncratic error $\epsilon_{i,t+1}$ is simulated from an i.i.d. normal distribution whose variance is calibrated such that the population R^2 of the model is approximately 20%, matching the empirically estimated value from fitting IPCA. The measurement error ϵ^{g}_{t+1} is simulated in a simple fashion but the $R^2 = 1 - \mathbb{E}[e^g]/\mathbb{E}[g]$ is calibrated to approximately 40%. $\eta = (1, 0, \dots, 0)$ and the loadings Γ_β are set to the empirically estimated values where $p - s$ rows are set to zero where $s = p/10$. Finally, observable factors and returns are generated according to models (2) and (4).

The simulation studies results across $S = 200$ Monte Carlo draws. Hyperparameters are fixed at $N = 500$, $T = 100$, and $k = 3$. To compare the performance of estimators under low-dimensional and high-dimensional characteristics, results are generated for $p = 10$ and $p = 50$. We report results for a variety of estimators, including latent loadings Γ_β, latent factors F, average factor loadings $\bar{\beta}$, latent matrix $C = F\Gamma_\beta^\top$, and observable factor risk premium γ_g.

The benchmark estimation and inference procedures are IPCA and the three-pass estimator of Giglio and Xiu (2021), given DSLFM's basis on these foundational models.[11] We focus on two comparisons: first, the estimation error of theoretically consistent latent loading Γ_β and latent factor $\{f_{t+1}\}_{t=1}^{T}$ IPCA and DSLFM estimators; and, second, coverage properties of the observable factor risk premium estimator. We do study estimation errors for additional estimands as relevant (e.g., the three-pass estimator does not estimate latent loadings Γ_β nor latent matrix C, while IPCA does not have an observable factor risk premium estimation nor inference procedure).

5.5.2. Simulation Results

Table A25 reports results. In the low-dimensional setting of $N = 500$, $T = 100$, $p = 10$, $s = 1$, we find DSLFM to obtain smaller estimation errors for Γ_β as compared to IPCA; however, IPCA has an order of magnitude lower estimation errors for F. DSLFM's outperformance for the latent loading is driven by taking on higher bias yet substantially lower variance

[11] Many thanks to Matthias Buechner and Leland Bybee for the IPCA implementation https://github.com/bkelly-lab/ipca.

of the estimator; this is obtained from soft-thresholding many of the rows. In comparing other auxiliary estimands, DSLFM obtains lower estimation error for the time-series averaged factor loadings $\bar{\beta}$ although higher error for the latent matrix $C = F\Gamma_{\beta}^{\top}$. We attribute these results to IPCA fitting data simulated to match the fits of an empirically estimated IPCA model, yet we employ an exact row-sparsity structure in the true latent loadings Γ_{β}.

DSLFM slightly under-covers the 90% and 95% confidence intervals in the low-dimensional setting. The three-pass estimator obtains similar estimation error for the target parameter of the observable factor risk premium, but has finite-sample intervals that slightly over-cover.

Moving to the high-dimensional setting of $p = 50$, we find DSLFM to again obtain smaller estimation errors for latent loadings, yet now the estimation errors for the latent factors are of the same order as IPCA. In both cases, DSLFM takes on bias from its regularization methods, although DSLFM's latent factor estimator's variance is still higher than IPCA. DSLFM is now an order of magnitude improvement for the average factor loadings and a factor of two improvement for the latent matrix C. Finally, coverage proprieties of the risk premium estimand for both the three-pass estimator and DSLFM estimators are degraded under high-dimensionality.

We hope to add even higher dimensional results with hyperparameters closer to the empirical values. Do note, given the DSLFM's large sample theory, it is constrained by $N > T$, which although is the case for the panel of crypto asset returns, this is not the case for all cross-sectional asset pricing settings. Moreover, the DSLFM's performance was boosted by the assumed exact sparsity in the latent loadings; we hope to add results for approximate sparsity, which is likely closer to the empirical reality. Nevertheless, the DSLFM performs well as compared to state-of-the-art benchmark methods, especially under the setting for which it was developed: high-dimensional asset characteristics.

6. Empirical Applications

See the Appendix D figures A26-A39 for the final tables and figures of the empirical applications including: univariate observable factor models, multivariate observable factor models, static latent-factor models, dynamic latent-factor models, DSLFM, and a deep-learning based factor model. This drafted section will be completed when I have a few day break from interviews...

7. Conclusion

“Competition would provide better money than would government. I believe we can do much better than gold ever made possible. Free enterprise, i.e. the institutions that would emerge from a process of competition in providing good money, no doubt would.

Two hundred years ago in *The Wealth of Nations* Adam Smith wrote that

‘to expect, indeed, that the freedom of trade should ever be entirely restored in Great Britain, is as absurd as to expect that an Oceana or Utopia should ever be established in it.’

It took nearly 90 years from the publication of his work in 1776 until Great Britain became the first country to establish complete free trade in 1860...I fear that since ‘Keynesian’ propaganda has filtered through to the masses, has made inflation respectable and provided agitators with arguments which the professional politicians are unable to refute, the only way to avoid being driven by continuing inflation into a controlled and directed economy, and therefore ultimately in order to save civilisation, will be to deprive governments of their power over the supply of money. What we need now is a Free Money Movement...

I wish I could advise that we proceed slowly. But the time may be short. What is now urgently required is not the construction of a new system but the prompt removal of all legal obstacles which have for two thousand years blocked the way for an evolution which is bound to throw up beneficial results which we cannot now forsee.”

—Friedrich A. Hayek (1976) *The Denationalization of Money*.

References

Ang, Andrew, Jun Liu, and Krista Schwarz. 2009. “Using individual stocks or portfolios in tests of factor models.” *AFA 2009 San Francisco Meetings Paper*.

Babiak, Mykola, and Daniele Bianchi. 2021. “A risk-based explanation of cryptocurrency returns.” *Available at SSRN*.

Bai, Jushan. 2003. “Inferential theory for factor models of large dimensions.” *Econometrica* 71 (1): 135–171.

Bai, Jushan, and Serena Ng. 2002. “Determining the number of factors in approximate factor models.” *Econometrica* 70 (1): 191–221.

Bai, Jushan, and Serena Ng. 2013. “Principal components estimation and identification of static factors.” *Journal of econometrics* 176 (1): 18–29.

Belloni, Alexandre, Victor Chernozhukov, Denis Chetverikov, Christian Hansen, and Kengo Kato. 2018. “High-dimensional econometrics and regularized GMM.” *arXiv preprint arXiv:1806.01888*.

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

Bianchi, Daniele, Matthias Büchner, and Andrea Tamoni. 2021. “Bond risk premiums with machine learning.” *The Review of Financial Studies* 34 (2): 1046–1089.

Bianchi, Daniele, Massimo Guidolin, and Manuela Pedio. 2022. “The dynamics of returns predictability in cryptocurrency markets.” *The European Journal of Finance*: 1–29.

Chamberlain, Gary. 1983. “Funds, factors, and diversification in arbitrage pricing models.” *Econometrica: Journal of the Econometric Society*: 1305–1323.

Chen, Luyang, Markus Pelger, and Jason Zhu. 2020. “Deep learning in asset pricing.” *Available at SSRN 3350138*.

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.” *The Econometrics Journal*.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. “Valid post-selection and post-regularization inference: An elementary, general approach.” *Annu. Rev. Econ.* 7 (1): 649–688.

Chetverikov, Denis, and Jesper Riis-Vestergaard Sorensen. 2021. “Analytic and bootstrap-after-cross-validation methods for selecting penalty parameters of high-dimensional M-estimators.” *arXiv preprint arXiv:2104.04716*.

Cochrane, John H. 2009. *Asset Pricing*.: Princeton university press.

Cochrane, John H. 2011. “Presidential address: Discount rates.” *The Journal of finance* 66 (4): 1047–1108.

Connor, Gregory, Matthias Hagmann, and Oliver Linton. 2012. “Efficient semiparametric estimation of the Fama–French model and extensions.” *Econometrica* 80 (2): 713–754.

Connor, Gregory, and Oliver Linton. 2007. “Semiparametric estimation of a characteristic-based factor model of common stock returns.” *Journal of Empirical Finance* 14 (5): 694–717.

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

Fama, Eugene F, and James D MacBeth. 1973. “Risk, return, and equilibrium: Empirical tests.” *Journal of political economy* 81 (3): 607–636.

Fan, Jianqing, Yuan Liao, and Weichen Wang. 2016. “Projected principal component analysis in factor models.” *Annals of statistics* 44 (1): 219.

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

Feng, Guanhao, Jingyu He, Nicholas G Polson, and Jianeng Xu. 2018. “Deep learning in characteristics-sorted factor models.” *arXiv preprint arXiv:1805.01104*.

Giglio, Stefano, Bryan Kelly, and Dacheng Xiu. 2022. “Factor models, machine learning, and asset pricing.” *Annual Review of Financial Economics* 14.

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

Giglio, Stefano, Dacheng Xiu, and Dake Zhang. 2021. “Test assets and weak factors.”Technical report, National Bureau of Economic Research.

Gómez, David, and Alfonso Rojas. 2016. “An empirical overview of the no free lunch theorem and its effect on real-world machine learning classification.” *Neural computation* 28 (1): 216–228.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. “Empirical asset pricing via machine learning.” *The 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.

Guijarro-Ordonez, Jorge, Markus Pelger, and Greg Zanotti. 2021. “Deep Learning Statistical Arbitrage.” *Available at SSRN 3862004*.

Hayek, Friedrich A. 1976. *The Denationalization of Money: An Analysis of the Theory and Practice of Concurrent Currencies.*: The Institute of Economic Affairs.

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.

Kelly, Bryan T, Seth Pruitt, and Yinan Su. 2020. “Instrumented principal component analysis.” *Available at SSRN 2983919*.

Lamport, Leslie, Robert Shostak, and Marshall Pease. 1982. “The Byzantine generals problem.” *ACM Transactions on Programming Languages and Systems* 4 (3): 382–401.

Lee, Jason D, Dennis L Sun, Yuekai Sun, and Jonathan E Taylor. 2016. “Exact post-selection inference, with application to the lasso.”

Nakamoto, Satoshi. 2008. “Bitcoin: A peer-to-peer electronic cash system.”

Newey, Whitney K, and Kenneth D West. 1987. “Hypothesis testing with efficient method of moments estimation.” *International Economic Review*: 777–787.

Lopez de Prado, Marcos. 2022. “Causal Factor Investing: Can Factor Investing Become Scientific?” *Available at SSRN 4205613*.

Ross, Stephen A. 1976. “The arbitrage theory of capital asset pricing.” *Journal of economic theory* 13 (3): 341–360.

Semenova, Vira, and Victor Chernozhukov. 2021. “Debiased machine learning of conditional average treatment effects and other causal functions." *The Econometrics Journal* 24 (2): 264–289.

Shanken, Jay. 1992. "On the estimation of beta-pricing models." *The review of financial studies* 5 (1): 1–33.

Sharpe, William F. 1964. "The Capital Asset Pricing Model: A Theory of Market Equilibrium Under Conditions of Risk." *The Journal of Finance* 19 (3): 425–442.

Vershynin, Roman. 2018. *High-dimensional probability: An introduction with applications in data science*. vol. 47: Cambridge university press.

Yu, Yi, Tengyao Wang, and Richard J Samworth. 2015. "A useful variant of the Davis–Kahan theorem for statisticians." *Biometrika* 102 (2): 315–323.

Appendix A. Details on Crypto Asset Characteristics

The tables below present details for each of six categories of digital asset characteristics. For additional details, see Panel Construction Notebook and Feature Engineering Notebook.

TABLE A1. Digital Asset Onchain Characteristics

Acronym	Asset characteristic
Tx Volume Tm7	The total transaction volume in native units over the trailing seven days.
Active Addresses Tm7	The number of active address over the trailing seven days.
Δ Log New Addresses Tm14-Tm7	The first difference of the logarithm of new addresses from 14 to 7 days ago.
New Addresses Tm7	The total number of new addresses over the trailing seven days.
Total Addresses	The total number of unique addresses.
Circulation Tm7	The number of unique native units transferred over the trailing seven days.
Age Destroyed	The sum over the trailing week of all native units transferred times the number of days since they were previously transferred.
Δ Flow Distribution Tm7	The ratio between the total native units transferred between various entities identified by Santiment (e.g. cex, dexes, defi platforms, whales, etc.) over the trailing week and the total first absolute differences across all the flow variables over the trailing week.
Δ Holders Distribution Tm7	The same functional form as the change in flow but for the total supply held by wallets with various magnitudes, as identified by Santiment, of the total supply.
% Supply in Profit	The percentage of the total native units which last transferred at a market value below the current market value.

TABLE A2. Digital Asset Exchange Characteristics

Acronym	Asset characteristic
% Circ. Supply CEX	The percentage of circulating supply in native units in wallets associated with CEXs as identified by Santiment.
% Circ. Supply DEX	The percentage of circulating supply in native units in wallets associated with DEXs as identified by Santiment.
% Circ. Supply Defi	The percentage of circulating supply in native units in wallets associated with DeFi platforms as identified by Santiment.
% Circ. Supply Traders	The percentage of circulating supply in native units in wallets associated with active traders as identified by Santiment.
Exchange Inflow	The total number of native units transferred from wallets not associated with exchanges to wallets that are, over the trailing week, as identified by Santiment.
Exchange Outflow	The total number of native units transferred from wallets associated with exchanges to wallets that are not associated with exchanges, over the trailing week, as identified by Santiment.
Number of Trading Pairs	The number of trading pairs identified by CMC on CEXs.

Table A3. Digital Asset Social Characteristics

Acronym	Asset characteristic
Social Volume	The total number of text documents containing the asset name across Reddit, Twitter, Telegram, and BitcoinTalk over the trailing seven days.
Social Volume Reddit	The total number of text documents containing the asset name on Reddit over the trailing seven days.
Social Volume Twitter	The total number of text documents containing the asset name on Twitter over the trailing seven days.
Sentiment Pos. Reddit	The total sentiment score across all text documents with a positive sentiment on Reddit over the trailing seven days.
Sentiment Pos. Twitter	The total sentiment score across all text documents with a positive sentiment on Twitter over the trailing seven days.
Sentiment Neg. Reddit	The total sentiment score across all text documents with a negative sentiment on Reddit over the trailing seven days.
Sentiment Neg. Twitter	The total sentiment score across all text documents with a negative sentiment on Twitter over the trailing seven days.
Developer Activity	The aggregate number of GitHub actions (e.g. commits, forks, comments, etc.), as identified by CoinGecko, over the trailing seven days.
VC Owned	Whether the asset has been funded by a set of prominent venture capitalists as identified by CoinMarketCap.

Table A4. Digital Asset Momentum Characteristics

Acronym	Asset characteristic
Return Tm7	Momentum over the trailing seven days.
Return Tm14	Momentum over the trailing fourteen days.
Return Tm30	Momentum over the trailing thirty days.
Return Tm30	Momentum over the trailing sixty days.
Return Tm90	Momentum over the trailing ninety days.
Return Tm14-Tm7	Short term reversal: difference in return between trailing fourteen and seven days.
Return Tm30-Tm14	Medium term reversal: difference in return between trailing thirty and fourteen days.
Return Tm90-Tm30	Long term reversal: difference in return between trailing ninety and thirty days.
Return from ATH	The return since the all time high price.
Return from ATL	The return since the all time low price.
Return Industry Tm30	Industry momentum over the trailing thirty days.
Return Industry Tm60	Industry momentum over the trailing sixty days.

Table A5. Digital Asset Microstructure Characteristics

Acronym	Asset characteristic
Trades Sum Tm7	The total number of CEX trades in the trailing seven days.
Volume Sum Tm7	The dollar CEX trading volume in the trailing seven days.
Spread Bps	Spread in basis points.
Ask Size	Market value of orders at best ask.
Bid Size	Market value of orders at best bid.
Illiq Tm7	The average absolute hourly return over the trailing week divided by the average hourly dollar volume over the trailing week.
Turnover Tm7	The total volume over the trailing week divided by the circulating supply in native units.

Table A6. Digital Asset Financial Characteristics

Acronym	Asset characteristic
Price	The market value of one native unit in USD.
Size	The market capitalization of all free floating native units in USD.
MVRV	The ratio of the market capitalization to the realized value, or the total number of free floating native units times the dollar value at the the time of the last transfer.
Alpha Tm7	Intercept coefficient from regressing hourly excess returns on cmkt hourly returns over the trailing seven days.
Alpha Tm30	Intercept coefficient from regressing hourly excess returns on cmkt hourly returns over the trailing thirty days.
Beta Tm7	Slope coefficient from regressing hourly excess returns on cmkt hourly returns over the trailing seven days.
Beta Tm30	Slope coefficient from regressing hourly excess returns on cmkt hourly returns over the trailing thirty days.
Beta Downside Tm30	Slope coefficient from regressing negative hourly excess returns (or zero) on negative cmkt hourly returns over the trailing thirty days.
Coskew Tm30	The slope coefficient on the squared cmkt term from regressing hourly excess returns on cmkt hourly returns and squared cmkt hourly returns over the trailing thirty days.
ISkew Tm30	The skewness of the residuals from from regressing hourly excess returns on cmkt hourly returns and squared cmkt hourly returns over the trailing thirty days.
Shortfall 5% Tm7	Average hourly return of the returns below the fifth quantile of the trailing seven day hourly returns.
VaR 5% Tm7	The fifth quantile of hourly excess returns over the trailing seven days.
Vol Tm7	The standard deviation of hourly excess returns over the trailing seven days.
Vol Tm30	The standard deviation of hourly excess returns over the trailing thirty days.
Vol Tm90	The standard deviation of hourly excess returns over the trailing ninety days.
Ivol Tm7	The standard deviation of the residuals from regressing hourly excess returns on cmkt returns over the trailing seven days.
Ivol Tm30	The standard deviation of the residuals from regressing hourly excess returns on cmkt returns over the trailing thirty days.
Ivol Tm90	The standard deviation of the residuals from regressing hourly excess returns on cmkt returns over the trailing ninety days.

Appendix B. Cross Validation Procedure

Several of the statistical models studied in this manuscript contain hyperparameters set by the statistician, which we will optimize through cross validation. For example, LASSO has the single hyperparameter to scale the size of the N-diamond over which we search for optimal betas. Neural networks have a rich set of hyperparameters, perhaps the richest.

The classic method used in the machine learning literature to select hyperparameters is cross validation (CV). But, given the rich temporal dependence in financial time series, we have to adjust the CV procedure to respect the temporal order of the data. We build on top of the approach used in Gu, Kelly, and Xiu (2020) as follows.

- We first set aside Q3 and Q4 2022 data as the **test data**.
  - Traditionally test data has been of a size not many orders away from that of the training data; however, given the rise of neural networks which benefit from the maximal amount of data, test sets are now set to the minimum necessary size.
  - We aim to balance this minimize necessary size with an additional desire to capture a broad set of market DGPs given the observed differences in relationships, e.g., between bull and bear markets in crypto.
- Set our **validation data** as the calendar year of 2021 and the first two quarters of 2022.
- Given we want to not just tune a single hyperparameter grid but inform our hyperparameter grid search based on CV results, we break the validation period into six consecutive three month segments, which we will study sequentially.
- Set an initial hyperparameter grid for a model and for each hyperparameter point:
  - Initialize **training data** as 2018 through 2020.
  - Loop over each observation in the **validation data** of Q1 2021 to:
    - Fit the model in the training data.
    - Predict returns for the validation observation.
    - Add the validation observation to the training data.
  - We thus generate out of sample returns for this portion of the validation data.

- Next, we set a new hyperparameter grid and repeat this CV process but for a validation sample of Q2 2021. We repeat four more times to finally select a model based on Q2 2022 validation period results.
    - ideally, we would only touch each validation observation once so it is an unbiased estimate of the performance in the test period data, but machine learning models require significant tuning.
    - It thus becomes a tradeoff between when we we likely overfitting on the validation data because we have experimented with too many hyperparameter points versus searching hyperparameter space for the optimal model.
- We then, once and only only, fit our best model on all data before each observation in the test period to generate out of sample predictions for the entire test period.
    - Given we are simply trusting researchers to not repeatedly touch the test data, the asset pricing community should implement verification methods using some cryptography to write to a timestamped public ledger the model to be subsequently prospectively tested.
        * Obviously, researchers could write to many sequentially updated public ledgers to have many shots on goal, but if the asset pricing community were to use mechanism design to develop an incentive for a Schelling point to arise through some consensus mechanism to agree to record to a single ledger of record, then we would not have this issue.

TODO: **Add table of hps.**

Appendix C. Technical Details and Proofs

C.1. Notation

Let $\mathbb{E}_N[x_i] := N^{-1} \sum_t x_i$ for random variables $\{x_i\}_{i=1}^{N}$.

Let $\mathbb{I}_k$ be a $k \times k$ identify matrix. Let $\|\cdot\|$ be the Frobenius norm for a matrix and the ℓ_2 norm for a vector; $\|\cdot\|_1$ be the l_1-norm; $\|\cdot\|_2$ be the spectral norm for a matrix; and, $\|\cdot\|_\infty$ be the maximum element of the matrix or vector. Let $a \vee b = max(a, b)$. We also use the notation $a \lesssim_P b$ to denote $a = O_p(b)$ for $a, b \in \mathbb{R}$.

Define the following random variables: $r_{t+1} = (r_{1,t+1}, \dots, r_{N,t+1})^\top \in \mathbb{R}^N$; $z_{t,j} = (z_{1,t,j}, \dots, z_{N,t,j})^\top \in \mathbb{R}^N$; $Z_{t,-j} = (z_{1,t,-j}, \dots, z_{N,t,-j})^\top \in \mathbb{R}^{N \times (p-1)}$; $\epsilon_{t+1} = (\epsilon_{1,t+1}, \dots, \epsilon_{N,t+1})^\top \in \mathbb{R}^N$; $\epsilon^z_{t,j} = (\epsilon^z_{1,t,j}, \dots, \epsilon^z_{N,t,j})^\top \in \mathbb{R}^N$, and so on.

For $A \subset \{1, \ldots, p\}$, let $Z_{t,-j}[A]$ denote the subset of the columns of $Z_{t,-j}$ that are elements of the set A. Let $\mathcal{P}_A := Z_{t,-j}[A] \left(Z_{t,-j}[A]^\top Z_{t,-j}[A] \right)^{-1} Z_{t,-j}[A]^\top$ be the projection operator that maps vectors in $\mathbb{R}^N$ into $\text{span}(Z_{t,-j}[A])$. Let $\mathcal{M}_A = \mathbb{I}_N - \mathcal{P}_A$ be the operator that projects vectors in $\mathbb{R}^N$ into the subspace orthogonal to $\text{span}(Z_{t,-j}[A])$.

C.2. Consistency of Double Selection Lasso

We provided two critical Double Selection Lasso (DSL) assumptions in Assumption 1, to which we add the following standard DSL assumptions, adapted to the DSLFM setting. Let there exist absolute sequences $\delta_{N,T} \searrow 0$ and $\Delta_{N,T} \searrow 0$.

Assumption A1 (ASR: Approximate Sparse Regressors).

(i) *Sparsity of Confounding: The confounding mapping $\delta_{t,-j}$ admits, $\forall t, j$ an exactly sparse form $||\delta_{t,-j}||_0 \leq s$.*

(ii) *Sparsity rate: The sparsity index obeys $s^2 \log^2 (p \vee N) / \left(\sqrt{N \log(Tp)}\right) \leq \delta_{N,T}$ and the size of the amelioration set obeys $\hat{s}_3 \leq C (1 \vee \hat{s}_1 \vee \hat{s}_2)$. Additionally, $\log^3 p/N \leq \delta_{N,T}$.*

Assumption ASR(i) extends Assumption 1(ii) to include sparsity of the DSL first stage. Assumption ASR(ii) controls the rate between sparsity and the asymptotic terms p, N, T; additionally, it constrains the amelioration set to not be substantially larger than the variables selected by the LASSO procedures.

Next, we constrain the minimum and maximum m-spare eigenvalues as whenever $p > N$ the empirical design matrix $\mathbb{E}_N[z_{i,t} z_{i,t}^\top]$ will not have full rank. Define the minimal and maximal m-sparse eigenvalue of a semi-definite matrix M as

$$\phi_{\min}(m)[M] := \min_{1 \leq ||\delta||_0 \leq m} \frac{\delta^\top M \delta}{||\delta||^2} \text{ and } \phi_{\max}(m)[M] := \max_{1 \leq ||\delta||_0 \leq m} \frac{\delta^\top M \delta}{||\delta||^2}.$$

Assumption A2 (SE: Sparse Eigenvalues). *There exists an absolute sequence $l_N \to \infty$ and such that with probability of at least $1 - \Delta_{N,T}$ the maximal and minimal $l_N s$-sparse eigenvalues are bounded from above and away from zero. That is, for absolute constants $0 < \kappa' < \kappa'' < \infty$,*

$$\kappa' \leq \phi_{\min}(l_N s)[\mathbb{E}_N[z_{i,t} z_{i,t}^\top]] \leq \phi_{\max}(l_N s)[\mathbb{E}_N[z_{i,t} z_{i,t}^\top]] \leq \kappa''$$

Similarly, for $\bar{z}_i := T^{-1} \sum_i z_{i,t}$*, we have*

$$\kappa' \leq \phi_{\min}(l_N s)[\mathbb{E}_N[\bar{z}_i \bar{z}_i^\top]] \leq \phi_{\max}(l_N s)[\mathbb{E}_N[\bar{z}_i \bar{z}_i^\top]] \leq \kappa''.$$

Requiring the minimum m–sparse eigenvalue to be bounded away from zero is equivalent to assuming all empirical design submatrices formed by any m components of $z_{i,t}$ are positive definite.

Next, we impose moment conditions on the structural errors and regressors.

ASSUMPTION A3 (SM: Structural Moments). *There are absolute constants* $0 < \omega < \Omega < \infty$ *and* $4 \leq \rho < \infty$ *such that for* $(y_i, \epsilon_i) := (r_{i,t+1}, \epsilon_{i,t+1})$ *or* $(y_i, \epsilon_i) := (z_{i,t,-j}, \epsilon^z_{i,t,j})$ *we have* $\forall i, t, j$*:*

(i) $\mathbb{E}[|z_{i,t,j}|^\rho] \leq \Omega$, $\omega \leq \mathbb{E}[\epsilon^2_{i,t+1} | z_{i,t,-j}, \epsilon^z_{i,t,j}] \leq \Omega$, *and* $\omega \leq \mathbb{E}[(\epsilon^z_{i,t,j})^2 | z_{i,t,-j}] \leq \Omega$;

(ii) $\mathbb{E}[|\epsilon_i|^\rho] + \mathbb{E}[y_i^2] + \max_{1 \leq k \leq p}\{\mathbb{E}[z^2_{i,t,-j,k} y_i^2] + \mathbb{E}[|z^3_{i,t,-j,k} \epsilon_i^3|] + 1/\mathbb{E}[z^2_{i,t-1,-j,k}]\} \leq \Omega$,

(iii) $\max_{1 \leq k \leq p}\{\mathbb{E}[z^2_{i,t,-j,k} \epsilon_i^2] + \mathbb{E}[z^2_{i,t,-j,k} y_i^2]\} + \max_{1 \leq i \leq N} ||z_{i,t,-j}||^2_\infty \frac{s \log(N \vee p)}{N} \leq \delta_{N,T}$ *w.p.* $1 - \Delta_{N,T}$.

(iv) *Weak dependence between the first- and second-stage errors: There exists a positive constant M such that* $\forall\, p, T, N$:

$$\left| \sqrt{\frac{1}{N}} \sum_{i=1}^{N} \epsilon^z_{i,t,j} \epsilon_{i,t+1} \right| \leq M \log(Tp).$$

(v) *Uniformly over t, j, we have* $\frac{1}{N} \sum_i (\epsilon^z_{t,j,i})^2 \xrightarrow{p} \mathbb{Z}^0_{t,j}$ *for non-stochastic real-valued scalar finite constant* $\mathbb{Z}^0_{t,j}$*, which is bounded away from zero.*

Assumptions (SM)(i)-(iii) are standard for DSL to bound various moments of the errors, characteristics, and returns. Assumption SM(iv) is novel and bounds the dependence between the first- and second-stage errors in the DSL model, which is the non-negligible asymptotic term in the DSL estimation error. This holds trivially for i.i.d. sampling in the cross-section, which we have relaxed to this specified sum. Assumption (v) is novel and introduces a uniform consistency for the second moment of the first-stage errors.

LEMMA A1. *Under the model* (2)*; Assumption 1; and, DSL Assumptions ASR, SE, and SM, the DSL estimator has the property that*

$$\max_{t,j} |\widehat{c}_{t+1,j} - c_{t+1,j}| = O_P\left(\sqrt{\frac{\log(Tp)}{N}}\right).$$

Proof of Lemma A1. We proceed with the decomposition of the estimation error using the definition of the DSL estimator and model (2).

$$\begin{aligned}
\hat{c}_{t+1,j} - c_{t+1,j} &= \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}(Z_{t,-j}c_{t+1,-j} + \epsilon_{t+1})\right) \\
&= \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\left(Z_{t,-j}\delta_{t,j}\right)^{\top}\mathcal{M}_{\hat{I}}Z_{t,-j}c_{t+1,-j} \\
&\quad + \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\left(Z_{t,-j}\delta_{t,j}\right)^{\top}\mathcal{M}_{\hat{I}}\epsilon_{t+1} \\
&\quad + \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\epsilon_{t,j}^{z\top}\mathcal{M}_{\hat{I}}Z_{t,-j}c_{t+1,-j} \\
&\quad - \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\epsilon_{t,j}^{z\top}\mathcal{P}_{\hat{I}}\epsilon_{t+1} \\
&\quad + \left(z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j}\right)^{-1}\epsilon_{t,j}^{z\top}\epsilon_{t+1}.
\end{aligned}$$

From Belloni, Chernozhukov, and Hansen (2014) under the aforementioned DSL assumptions, the last term in this five-term decomposition is the asymptotically relevant term while the remaining terms are asymptotically negligible. We first handle the denominator of the fifth term before dealing with the entire term.

$$\begin{aligned}
N^{-1}z_{t,j}^{\top}\mathcal{M}_{\hat{I}}z_{t,j} &= N^{-1}\left(Z_{t,-j}\delta_{t,j} + \epsilon_{t,j}^{z}\right)^{\top}\mathcal{M}_{\hat{I}}\left(Z_{t,-j}\delta_{t,j} + \epsilon_{t,j}^{z}\right) \\
&= \epsilon_{t,j}^{z\top}\epsilon_{t,j}^{z}/N + \delta_{t,j}^{\top}Z_{t,-j}^{\top}\mathcal{M}_{\hat{I}}Z_{t,-j}\delta_{t,j}/N + 2\delta_{t,j}^{\top}Z_{t,-j}^{\top}\mathcal{M}_{\hat{I}}\epsilon_{t,j}^{z}/N - \epsilon_{t,j}^{z\top}\mathcal{P}_{\hat{I}}\epsilon_{t,j}^{z}/N \\
&\lesssim_P \epsilon_{t,j}^{z\top}\epsilon_{t,j}^{z}/N + o_P(1)
\end{aligned}$$

where the first equality holds by definition of the first-stage; the second equality holds by multiplying out the terms and by definition of the projection matrices; and, the probabilistic bound holds given the latter three terms are asymptotically negligible, as in the proof of Theorem 1 in Belloni, Chernozhukov, and Hansen (2014), as compared to the sum of second moments of the first-stage errors. Thus, by Assumption SM(v), we conclude $\epsilon_{t,j}^{z\top}\epsilon_{t,j}^{z}/N$ converges in probability uniformly over t, j, to $\mathbb{Z}_{t,j}^{0}$.

We proceed with the uniform consistency result.

$$
\begin{aligned}
\max_{t,j} |\hat{c}_{t+1,j} - c_{t+1,j}| &\lesssim_P \max_{t,j} \left| \left(N^{-1} z_{t,j}^\top \mathcal{M}_{\hat{I}} z_{t,j} \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} \epsilon_{i,t,j}^{z} \epsilon_{i,t+1} \right| \\
&\lesssim_P \sqrt{\frac{1}{N}} \max_{t,j} \left| \sqrt{\frac{1}{N}} \sum_{i=1}^{N} \epsilon_{i,t,j}^{z} \epsilon_{i,t+1} \right| \\
&\lesssim_P \sqrt{\frac{\log(Tp)}{N}}
\end{aligned}
$$

which holds for the first probabilistic bound by substituting the decomposition above; for the second probabilistic bound, using the above result to replace the denominator with a constant that is bounded away from zero uniformly over t, j; and, the final bound holds by assumption SM(iv). In the case of i.i.d. sampling in the cross-section or if the dependence is sufficiently weak such that SM(iv) holds, then we can invoke Lemma A.4 in Belloni et al. (2018) to conclude the mean zero scalar random variable $\epsilon_{i,t,j}^{z} \epsilon_{i,t+1}$, given the moment conditions of the DSL model, has a maximal deviation that converges in probability to zero at the specified rate if we further constrain the moment of the mean zero random variable $\mathbb{E}\left[\max_{t,j} \left|\epsilon_{i,t,j}^{z} \epsilon_{i,t+1}\right|^{q}\right] \leq M^{q}$ for $q > 2$ and absolute constant M uniformly across t, j. □

C.3. Consistency of Latent Factors and Loadings

We first prove a bound on the distance between the infeasible and feasible symmetric matrix used in the eigendecompositions. Let $\widehat{\Lambda}_{Tp} \in \mathbb{R}^{k \times k}$ be a diagonal matrix containing the k largest eigenvalues of $(Tp)^{-1}\widehat{C}\widehat{C}^\top$ and similarly for $\Lambda_{Tp} \in \mathbb{R}^{k \times k}$, a diagonal matrix containing the k largest eigenvalues of $(Tp)^{-1}\widehat{C}\widehat{C}^\top$.

Lemma A2. *Under the assumptions of Lemma A1,* $\left\|(Tp)^{-1}\widehat{C}\widehat{C}^\top - (Tp)^{-1}CC^\top\right\| = O_p\left(\frac{\log Tp}{N}\right)$.

Proof of Lemma A2.

$$\begin{aligned}
\left\|\widehat{C}\widehat{C}^\top - CC^\top\right\| &= \left\|\widehat{C}\widehat{C}^\top - C\widehat{C}^\top + C\widehat{C}^\top - CC^\top\right\| \\
&\leq \left\|C\widehat{C}^\top - CC^\top\right\| + \left\|\widehat{C}\widehat{C}^\top - C\widehat{C}^\top\right\| \\
&\leq \|C\| \left\|\widehat{C} - C\right\| + \left\|\widehat{C} - C\right\| \left\|\widehat{C}\right\| \\
&= \|C\| \left\|\widehat{C} - C\right\| + \left\|\widehat{C} - C\right\| \left\|\widehat{C} - C + C\right\| \\
&\leq \|C\| \left\|\widehat{C} - C\right\| + \left\|\widehat{C} - C\right\| \left(\left\|\widehat{C} - C\right\| + \|C\|\right) \\
&= 2\,\|C\| \left\|\widehat{C} - C\right\| + \left\|\widehat{C} - C\right\|^2 \\
&\lesssim_P \sqrt{sT} \left\|\widehat{C} - C\right\| + \left\|\widehat{C} - C\right\|^2 \\
&\leq \sqrt{spT^2} \max_{t,j} |\widehat{c}_{t+1,j} - c_{t+1,j}| + Tp \max_{t,j} |\widehat{c}_{t+1,j} - c_{t+1,j}|^2 \\
&\lesssim_P \sqrt{\frac{spT^2 \log(Tp)}{N}} + \frac{Tp\log(Tp)}{N} \lesssim \frac{Tp\log(Tp)}{N}.
\end{aligned}$$

where the first and third inequality holds by the triangle inequality; the second inequality holds by Cauchy-Schwarz; the first probabilistic bound holds by Assumption 1(i) / A1(i); and, the last probabilistic bound holds by Lemma A1. We use the final bound for simplicity. □

We next bound the estimation error between the feasible and infeasible eigenvalues.

Lemma A3. *Under the assumptions of Lemma A2 and Assumption 3,*

$$\left\|\widehat{\Lambda}_{Tp} - \Lambda_{Tp}\right\|^2 = O_p\left(\frac{s^2\log^2(Tp)}{N^2}\right).$$

Proof of Lemma A3.

$$\begin{aligned}
\left\|\widehat{\Lambda}_{Tp} - \Lambda_{Tp}\right\|^2 &= \sum_{l=1}^{k}\sum_{i=1}^{k}(\widehat{\lambda}_l - \lambda_i)^2 \\
&\leq k^2 \max_{l\in\{1,\ldots,k\}} |\widehat{\lambda}_l - \lambda_l|^2 \\
&\leq k^2 \max_{l\in\{1,\ldots,T\}} |\widehat{\lambda}_l - \lambda_l|^2 \\
&\leq \frac{k^2}{T^2p^2} \left\|\widehat{C}\widehat{C}^\top - CC^\top\right\|^2 = O_p(\frac{s^2\log^2(Tp)}{N^2})
\end{aligned}$$

where the first equality is the definition of the Frobenius norm; the first inequality bounds the sum by the maximum element; the second inequality bounds the maximum deviation between the k largest eigenvalues of the feasible and infeasible decompositions by the deviations between all T eigenvalues; the last inequality controls the stability of the spectrum by applying Weyl's inequality from Theorem 4.5.3 of Vershynin (2018); and, the probabilistic bound holds by Lemma A2. □

We next prove a lemma for the time series average of the ℓ_2 norm of the feasible and infeasible eigenvectors.

LEMMA A4. *Under Assumption 2(iii) and those of Lemma A3, there exists an orthogonal matrix* $\widehat{O} \in \mathbb{R}^{k\times k}$ *such that* $\left\|\widehat{F} - \widetilde{F}\widehat{O}^\top\right\|^2 = O_p\left(\frac{\log^2(Tp)}{N^2}\right)$.

Proof of Lemma A4. We use a variant of the Davis-Kahan theorem shown in Yu, Wang, and Samworth (2015) where

$$\delta := \min_{i:i\neq l} |\lambda_l - \lambda_i| > 0,$$

which holds by Assumption 2(iii), to conclude for some $\widehat{O} \in \mathbb{R}^{k\times k}$ orthogonal matrix that

$$2^{3/2}\delta^{-1}(Tp)^{-2}\left\|\widehat{C}\widehat{C}^\top - CC^\top\right\|^2 \geq ||\widehat{F}\widehat{O} - \widetilde{F}||^2 = ||\widehat{F} - \widetilde{F}\widehat{O}^\top||^2$$

where the inequality is the use of the variant of the Davis-Kahan theorem, bounding the distance between the eigenvectors by the distance between the original matrices, and the equality follows given post multiplying by an orthogonal matrix does not change the Frobenius norm. The rate in the result then follows given Lemma A2. □

We next bound the ℓ_2 norm between the feasible and infeasible eigenvectors.

LEMMA A5. *Under the assumptions of Lemma A4,* $\left\|\widehat{f}_{t+1} - \widehat{O}^\top\widetilde{f}_{t+1}\right\| = O_p\left(\sqrt{\frac{s\log(Tp)}{N}}\right)$.

Proof of Lemma A5. First, we perform the following decomposition using the definition of the eigenvectors.

$$\begin{aligned}
\widehat{f}_{t+1} - \widehat{O}^\top\widetilde{f}_{t+1} &= (Tp)^{-1}\widehat{\Lambda}_{Tp}^{-1}\widehat{F}^\top\widehat{C}\widehat{C}_{t+1} - (Tp)^{-1}\widehat{O}^\top\Lambda_{Tp}^{-1}\widetilde{F}^\top CC_{t+1} \\
&= (Tp)^{-1}\left(\widehat{\Lambda}_{Tp}^{-1} - \Lambda_{Tp}^{-1}\right)\widehat{F}^\top\left(\widehat{C}\widehat{C}_{t+1} - CC_{t+1}\right) \\
&\quad + (Tp)^{-1}\widehat{O}^\top\Lambda_{Tp}^{-1}\left(\widetilde{F}^\top - \widehat{F}^\top\right)CC_{t+1} \\
&\quad + (Tp)^{-1}\left(\widehat{\Lambda}_{Tp}^{-1} - \widehat{O}^\top\Lambda_{Tp}^{-1}\right)\widehat{F}^\top CC_{t+1} \\
&\quad + (Tp)^{-1}\Lambda_{Tp}^{-1}\widehat{F}^\top\left(\widehat{C}\widehat{C}_{t+1} - CC_{t+1}\right)
\end{aligned}$$

where the equality follows by adding and subtracting terms.

Thus,

$$\begin{aligned}\left\|\widehat{f}_{t+1} - \widehat{O}^{\top}\widetilde{f}_{t+1}\right\| &\leq (Tp)^{-1}\left\|\widehat{\Lambda}_{Tp}^{-1} - \Lambda_{Tp}^{-1}\right\|\left\|\widehat{F}\right\|\left\|\widehat{C}\widehat{C}_{t+1} - CC_{t+1}\right\| \\ &\quad + (Tp)^{-1}\left\|\Lambda_{Tp}^{-1}\right\|\left\|\widehat{F} - \widetilde{F}\right\|\|CC_{t+1}\| \\ &\quad + (Tp)^{-1}\left\|\widehat{\Lambda}_{Tp}^{-1} - \Lambda_{Tp}^{-1}\right\|\left\|\widehat{F}\right\|\|CC_{t+1}\| \\ &\quad + (Tp)^{-1}\left\|\Lambda_{Tp}^{-1}\right\|\left\|\widehat{F}\right\|\left\|\widehat{C}\widehat{C}_{t+1} - CC_{t+1}\right\| \\ &= O_p\left(\sqrt{\frac{s^3\log(Tp)^3}{pN^3}}\right) + O_p\left(\frac{\sqrt{T}s\log(Tp)}{TpN}\right) \\ &\quad + O_p\left(\frac{s^2\log(Tp)}{pN}\right) + O_p\left(\sqrt{\frac{s\log(Tp)}{N}}\right) \\ &= O_p\left(\sqrt{\frac{s\log(Tp)}{N}}\right)\end{aligned}$$

where the first inequality follows from the aforementioned decomposition in this proof with the use of the triangle and Cauchy-Schwarz inequalities; the first probabilistic bound holds given the normalization that $\widehat{F}^{\top}\widehat{F}/T = I_k$ then $||\widehat{F}|| = \sqrt{Tk}$, given Λ_{Tp} contains k nonzero real-valued eigenvalues bounded away from zero by Assumption 2(iii), given the rates from Lemmas A3 and A4 (which gives same rate by CMT), given—similar to lemma A2—$||\widehat{C}\widehat{C}_{t+1} - CC_{t+1}|| = O_p(\sqrt{\frac{sTp^2\log(Tp)}{N}})$, and $||CC_{t+1}|| = O_p(\sqrt{s^2T})$ by Assumption 1(ii); and, the final probabilistic bound holds for simplicity of exposition.

□

Lemma A6. *Under the assumptions of A5, for $H^{\top} = \widehat{O}^{\top}\Lambda_{Tp}^{-1}(F^{\top}F^0/T)(\Gamma_{\beta}^{0\top}\Gamma_{\beta}^{0}/p)$ we have*

$$\widehat{O}\widetilde{f}_{t+1} - H^{\top}f_{t+1}^{0} = 0.$$

Proof of Lemma A6.

$$\begin{aligned}\widehat{O}\widetilde{f}_{t+1} &= (Tp)^{-1}\widehat{O}^{\top}\Lambda_{Tp}^{-1}F^{\top}CC_{t+1} \\ &= (Tp)^{-1}\widehat{O}^{\top}\Lambda_{Tp}^{-1}F^{\top}F\Gamma_{\beta}^{\top}\Gamma_{\beta}f_{t+1} \\ &= (Tp)^{-1}\widehat{O}^{\top}\Lambda_{Tp}^{-1}F^{\top}F^{0}\Gamma_{\beta}^{0\top}\Gamma_{\beta}^{0\top}f_{t+1}^{0} \\ &= H^{\top}f_{t+1}^{0}.\end{aligned}$$

where the first equality holds by the definition of the infeasible eigenvectors; the second equality holds given the definition of C; the third equality holds given the definitions of the true loadings and factors as rotations of the observed ones; and, the final equality holds by definition of the H matrix. □

Finally, using the above lemmas, we prove Propositions 1 and 2.

Proof of Proposition 1.

$$\begin{aligned}\left\|\widehat{f}_{t+1} - H^\top f^0_{t+1}\right\| &\le \left\|\widehat{f}_{t+1} - \widehat{O}^\top \widetilde{f}_{t+1}\right\| + \left\|\widehat{O}^\top \widetilde{f}_{t+1} - H^\top f^0_{t+1}\right\| \\ &= O_p\left(\sqrt{\frac{s\log(Tp)}{N}}\right)\end{aligned}$$

by Lemmas A5 and A6. □

Next, we provide a lemma for the ℓ_∞ norm of the PCA estimation error for the loadings, which will allow us to obtain norms on the soft-thresholded estimation error for the loadings by use of a tool from the high-dimensional econometrics handbook (Belloni et al. 2018).

Lemma A7. *Under the assumptions of Lemma A1 and Assumption 3,*

$$\left\|\widehat{\Gamma}_\beta - \Gamma^0_\beta (H^\top)^{-1}\right\|_\infty = O_p\left(\sqrt{\frac{\log(Tp)}{N}}\right).$$

Proof of Lemma A7.

$$\begin{aligned}\left\|\widehat{\Gamma}_\beta - \Gamma^0_\beta (H^\top)^{-1}\right\|_\infty &= \max_j \left(\sum_{l=1}^{k} \left|\widehat{\Gamma}_{\beta,j,l} - (\Gamma^0_{\beta,j})^\top (H^\top)^{-1}_l\right|\right) \\ &= \max_j \left\|\widehat{\Gamma}_{\beta,j} - H^{-1}\Gamma^0_{\beta,j}\right\|_1 \\ &= \max_j \left\|\widehat{\Gamma}_{\beta,j} \pm \widetilde{\Gamma}_{\beta,j} - H^{-1}\Gamma^0_{\beta,j}\right\|_1 \\ &\le \max_j \left\|\widehat{\Gamma}_{\beta,j} - \widetilde{\Gamma}_{\beta,j}\right\|_1 + \max_j \left\|\widetilde{\Gamma}_{\beta,j} - H^{-1}\Gamma^0_{\beta,j}\right\|_1 \\ &= \max_j \left\|T^{-1}\widehat{F}^\top \widehat{C}_j - T^{-1}F^\top C_j\right\|_1 + \max_j \left\|T^{-1}F^\top C_j - H^{-1}\Gamma^0_{\beta,j}\right\|_1 \\ &= T^{-1}\max_j \left\|\widehat{F}^\top \widehat{C}_j \pm \widehat{F}^\top C_j - F^\top C_j\right\|_1 + \max_j \left\|T^{-1}F^\top F\Gamma_{\beta,j} - H^{-1}\Gamma^0_{\beta,j}\right\|_1\end{aligned}$$

$$\begin{aligned}
&\leq T^{-1} \max_j \left\| \widehat{F}^\top \left(\widehat{C}_j - C_j \right) \right\|_1 + T^{-1} \max_j \left\| \left(\widehat{F} - F \right)^\top C_j \right\|_1 \\
&\quad + \max_j \left\| T^{-1} F^\top F \Gamma_{\beta,j} - H^{-1} \Gamma^0_{\beta,j} \right\|_1 \\
&\leq \frac{\sqrt{k}}{T} \left\| \widehat{F} \right\|_2 \max_j \left\| \widehat{C}_j - C_j \right\|_2 + \frac{\sqrt{k}}{T} \left\| \widehat{F} - F \right\|_2 \max_j \left\| C_j \right\|_2 \\
&\quad + \sqrt{k} \left\| T^{-1} F^\top F - I_{k\times k} \right\|_2 \left\| H^{-1} \right\|_2 \max_j \left\| \Gamma^0_{\beta,j} \right\|_2 \\
&\lesssim_P \max_{t,j} \left| \widehat{c}_{t+1,j} - c_{t+1,j} \right| + T^{-1/2} \left\| \widehat{F} - F \right\| \\
&\lesssim_P O_p\left(\sqrt{\frac{\log(Tp)}{N}} \right) + O_p\left(\frac{\log(Tp)}{N} \right)
\end{aligned}$$

where the first two equalities follow from the definition of the ℓ_∞ norm; the third equality adds and subtracts; the first inequality uses the triangle inequality; the fourth equality uses the definitions of the feasible and infeasible estimator; the second inequality uses a triangle inequality; the last inequality uses $\left\| A^\top x \right\|_1 \leq \sqrt{k} \left\| A \right\|_2 \left\| x \right\|_2$; the first probabilistic bound uses $\left\| \widehat{F} \right\|_2 = O_p(\sqrt{kT})$, the sum of the estimation errors in the characteristic portfolios, $|\widehat{c}_{t+1,j} - c_{t+1,j}|$ is bounded by T times the maximum element, the elements of C_j are bounded random variables hence $\max_j \left\| C_j \right\| = O_p(\sqrt{T})$, $T^{-1} \sum_t f^0_{t+1} f^{0,\top}_{t+1} \to_p \Sigma_f$ by Assumption 2(i), $\left\| H^{-1} \right\| = O_p(1)$ by definition of the invertible H matrix, and Γ^0_β contains bounded elements uniformly over j by Assumption 2(ii); and, the last probabilistic bound follows by Lemmas A1 and A4 and assumption A1(ii). □

Proof of Proposition 2. In Lemma A7, we show $\left\| \widehat{\Gamma}_\beta - \Gamma^0_\beta (H^\top)^{-1} \right\|_\infty \lesssim_P \sqrt{\frac{\log(Tp)}{N}}$, which allows us to invoke Theorem 2.10 from Belloni et al. (2018) under exact sparsity of Γ^0_β where λ, the hyperparameter used to soft-threshold the ℓ_1 norm of the rows of $\widehat{\Gamma}_\beta$, is selected such that with probability approaching 1,

$$\lambda \geq (1-\alpha) - \text{quantile of } \left\| \widehat{\Gamma}_\beta - \Gamma^0_\beta (H^\top)^{-1} \right\|_\infty .$$

That is, λ can be set to the product of a large constant and the rate of the ℓ_∞ norm, $\sqrt{\frac{\log(Tp)}{N}}$, to ensure this inequality holds.[12] Then, by Theorem 2.10 given $\alpha \to 0$ and

[12] In practice, we cross-validate for a finite-sample optimal λ.

$\lambda \lesssim \sqrt{\log(Tp)/N}$, we have for all $q \geq 1$

$$\left\| \check{\Gamma}_{\beta,l} - \Gamma^0_\beta (H^\top)^{-1}_l \right\|_q \lesssim_P s^{1/q} \sqrt{\frac{\log(Tp)}{N}}.$$

This holds column-by-column for the matrix estimation error $\check{\Gamma}_\beta - \Gamma^0_\beta$, which we can thus square and sum together for the squared Frobenius norm of the estimation error at the same rate. □

C.4. Consistency and Normality of Observable Factor Risk Premia

In these results, we are using the eigenvectors and loadings derived from the demeaned characteristic portfolio matrix: $\widehat{C}^D := \widehat{C} - \iota_T T^{-1} \sum_t \widehat{c}^\top_{t+1}$. That is, $\widehat{V}$ are the $\sqrt{T}$ scaled eigenvectors associated with the k largest eigenvalues of $(Tp)^{-1} \widehat{C}^D \widehat{C}^{D\top}$. Further, $\widehat{\Gamma}^D_\beta = T^{-1} \widehat{C}^{D\top} \widehat{V}$. The results established in the above subsection would follow analogously for this new notation as we simply have mean zero eigenvectors. Finally, our cross sectional and time series OLS estimators are standard:

$$\widehat{\gamma} := \left(\bar{\widehat{\beta}}^\top \bar{\widehat{\beta}} \right)^{-1} \bar{\widehat{\beta}}^\top \bar{r}, \qquad \widehat{\eta} := \left(\widehat{V}^\top \widehat{V} \right)^{-1} \widehat{V}^\top G,$$

where $\bar{\widehat{\beta}} := \bar{Z} \check{\Gamma}^D_\beta$ for $\bar{Z} = T^{-1} \sum_t Z_t$ for $Z_t \in \mathbb{R}^{N \times p}$, $\forall t$. The same holds for the time series average return $\bar{r} \in \mathbb{R}^N$.

Lemma A8. *Under the models* (2) *and* (4)*; Assumptions 1, 2, and 3; and, DSL Assumptions ASR, SE, and SM; we have*

$$\left\| N^{-1} \bar{\widehat{\beta}}^\top \bar{\widehat{\beta}} - N^{-1} \bar{\beta}^\top \bar{\beta} \right\| = O_p \left(\sqrt{\frac{s^2 \log(Tp)}{N}} \right).$$

Proof of Lemma A8.

$$\begin{aligned} \left\| N^{-1} \bar{\widehat{\beta}}^\top \bar{\widehat{\beta}} - N^{-1} \bar{\beta}^\top \bar{\beta} \right\| &= N^{-1} \left\| \bar{\widehat{\beta}}^\top \bar{\widehat{\beta}} \pm \bar{\widehat{\beta}}^\top \bar{\beta} - \bar{\beta}^\top \bar{\beta} \right\| \\ &\leq N^{-1} \left\| \bar{\widehat{\beta}}^\top \bar{\widehat{\beta}} - \bar{\widehat{\beta}}^\top \bar{\beta} \right\| + N^{-1} \left\| \bar{\widehat{\beta}}^\top \bar{\beta} - \bar{\beta}^\top \bar{\beta} \right\| \\ &= N^{-1} \left\| \left(\bar{\widehat{\beta}} \pm \bar{\beta} \right)^\top \left(\bar{\widehat{\beta}} - \bar{\beta} \right) \right\| + N^{-1} \left\| \left(\bar{\widehat{\beta}} - \bar{\beta} \right)^\top \bar{\beta} \right\| \end{aligned}$$

$$
\begin{aligned}
&\leq N^{-1} \left\| \bar{\hat{\beta}} - \bar{\beta} \right\|^2 + \frac{2}{N} \left\| \left(\bar{\hat{\beta}} - \bar{\beta} \right)^\top \bar{\beta} \right\| \\
&= N^{-1} \left\| \bar{Z} \left(\check{\Gamma}_\beta^D - \Gamma_\beta^0 \right) \right\|^2 + \frac{2}{N} \left\| \left(\check{\Gamma}_\beta^D - \Gamma_\beta^0 \right)^\top \bar{Z}^\top \bar{Z} \Gamma_\beta^0 \right\| \\
&\leq \left\| \left(\check{\Gamma}_\beta^D - \Gamma_\beta^0 \right)^\top \frac{\bar{Z}^\top \bar{Z}}{N} \left(\check{\Gamma}_\beta^D - \Gamma_\beta^0 \right) \right\|^2 \\
&\qquad + 2 \left\| \left(\check{\Gamma}_\beta^D - \Gamma_\beta^0 \right)^\top \frac{\bar{Z}^\top \bar{Z}}{N} \Gamma_\beta^0 \right\| \\
&\leq \left\| \check{\Gamma}_{\beta,l}^D - \Gamma_{\beta,l}^0 \right\|^2 \left\| \check{\Gamma}_\beta^D - \Gamma_\beta^0 \right\|^2 \phi_{\max}^2(2s) \left[\frac{\bar{Z}^\top \bar{Z}}{N} \right] \\
&\qquad + 2 \left\| \check{\Gamma}_{\beta,l}^D - \Gamma_{\beta,l}^0 \right\| \left\| \Gamma_\beta^0 \right\| \phi_{\max}(2s) \left[\frac{\bar{Z}^\top \bar{Z}}{N} \right] \\
&\lesssim_P O_p \left(\frac{s^2 \log^2(Tp)}{N^2} \right) + O_p \left(\sqrt{\frac{s^2 \log(Tp)}{N}} \right)
\end{aligned}
$$

where the first equality follows by adding and subtracting; the first inequality follows by the triangle inequality; the second equality follows from adding and subtracting and rearranging; the second inequality follows from the triangle and Cauchy-Schwartz inequalities; the third equality follows from the definition of the factor loading estimator and the average factor loading; the third inequality follows from $||A|| \leq ||AA\top||$; the final inequality follows from multiplying and dividing by the norms to obtain unit vectors and then bounding with the $s + \hat{s}$-maximally sparse eigenvalue of $\frac{\bar{Z}\top\bar{Z}}{N}$ (where the $s + \hat{s} \leq 2s$ by thresholding estimator); and, the probabilistic bound holds given the maximum eigenvalue is bounded by DSL Assumption SE, $\left\| \Gamma_\beta^0 \right\| = O_p(\sqrt{s})$ by Assumption 1(ii), and $\left\| \check{\Gamma}_{\beta,l}^D - \Gamma_{\beta,l}^0 \right\| = O_p \left(\sqrt{\frac{s \log(Tp)}{N}} \right)$, which is the same rate for the entire matrix per Proposition 2. □

Lemma A9. *Under the assumptions of Lemma A8, we have*

$$
\frac{\sqrt{T}}{N} \left\| \bar{\hat{\beta}}^\top \bar{r} - \bar{\beta}^\top \bar{r} \right\| = O_p \left(\sqrt{\frac{Ts^2 \log(Tp)}{N}} \right).
$$

Proof of Lemma A9.

$$\begin{aligned}
\frac{\sqrt{T}}{N}\left\|\bar{\widehat{\beta}}^{\top}\bar{r}-\bar{\beta}^{\top}\bar{r}\right\| &= \frac{\sqrt{T}}{N}\left\|\left(\check{\Gamma}_{\beta}^{D}-\Gamma_{\beta}^{0}\right)^{\top}\bar{Z}^{\top}\left(\bar{Z}\Gamma_{\beta}^{0}\gamma^{0}+T^{-1}\sum_{t}Z_{t}\Gamma_{\beta}^{0}\nu_{t+1}^{0}+\bar{\epsilon}\right)\right\| \\
&\leq \frac{\sqrt{T}}{N}\left\|\left(\check{\Gamma}_{\beta}^{D}-\Gamma_{\beta}^{0}\right)^{\top}\bar{Z}^{\top}\bar{Z}\Gamma_{\beta}^{0}\gamma^{0}\right\| \\
&\quad+\frac{\sqrt{T}}{N}\left\|\left(\check{\Gamma}_{\beta}^{D}-\Gamma_{\beta}^{0}\right)^{\top}\bar{Z}^{\top}T^{-1}\sum_{t}Z_{t}\Gamma_{\beta}^{0}\nu_{t+1}^{0}\right\| \\
&\quad+\frac{\sqrt{T}}{N}\left\|\left(\check{\Gamma}_{\beta}^{D}-\Gamma_{\beta}^{0}\right)^{\top}\bar{Z}^{\top}\bar{\epsilon}\right\| \\
&\leq \sqrt{T}\left\|\check{\Gamma}_{\beta,l}^{D}-\Gamma_{\beta,l}^{0}\right\|\left\|\Gamma_{\beta}^{0}\gamma^{0}\right\|\phi_{\max}(2s)\left[\frac{\bar{Z}^{\top}\bar{Z}}{N}\right] \\
&\quad+\left\|\check{\Gamma}_{\beta,l}^{D}-\Gamma_{\beta,l}^{0}\right\|^{2}\phi_{\max}(2s)\left[\frac{\bar{Z}^{\top}\bar{Z}}{N}\right]\sqrt{\frac{T}{N}}\left\|T^{-1}\sum_{t}Z_{t}\Gamma_{\beta}^{0}\nu_{t+1}^{0}\right\| \\
&\quad+\left\|\check{\Gamma}_{\beta,l}^{D}-\Gamma_{\beta,l}^{0}\right\|^{2}\phi_{\max}(2s)\left[\frac{\bar{Z}^{\top}\bar{Z}}{N}\right]\sqrt{\frac{T}{N}}\left\|\bar{\epsilon}\right\| \\
&\lesssim_{p} O_{p}\left(\sqrt{\frac{Ts^{2}\log(Tp)}{N}}\right)+O_{p}\left(\frac{s^{3/2}\log(Tp)}{N}\right) \\
&\quad+O_{p}\left(\frac{s\log(Tp)}{N}\right)
\end{aligned}$$

where the first equality follows by the definitions of the factor loading estimator, the average factor loading, and the time series average return of assets; the first inequality follows by the triangle inequality; the second inequality holds by an analogous argument to Lemma A8; and, the probabilistic bound holds again by an analogous argument to Lemma A8 with the additional bounds of $\left\|\Gamma_{\beta}^{0}\gamma^{0}\right\| = O_{p}(\sqrt{s})$ by Assumptions 2(i)-(ii) and 1(ii), $\left\|T^{-1}\sum_{t}Z_{t}\Gamma_{\beta}^{0}\nu_{t+1}^{0}\right\| = O_{p}(\sqrt{\frac{sN}{T}})$ by Assumption 4(ii), and $\|\bar{\epsilon}\| = O_{p}(\sqrt{\frac{N}{T}})$ by Assumption 4(i). □

Lemma A10. *Under the assumptions of Lemma A8 and Assumption 4(iii), we have*

$$\sqrt{T}\,(\widehat{\gamma}-H\gamma_{0}) = O_{p}(1).$$

Proof of Lemma A10. We decompose the estimation error into three terms.

$$\begin{aligned}
\sqrt{T}\,(\widehat{\gamma} - H\gamma_0) &= \sqrt{T}\,(\widehat{\gamma} - \widetilde{\gamma}) + \sqrt{T}\,(\widetilde{\gamma} - H\gamma_0) \\
&= \left(\frac{\bar{\widehat{\beta}}^\top \bar{\widehat{\beta}}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\widehat{\beta}}^\top \bar{r}}{N} \pm \left(\frac{\bar{\beta}^\top \bar{\beta}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\beta}^\top \bar{r}}{N} - \sqrt{T}H\gamma_0 \\
&= \underbrace{\left(\frac{\bar{\widehat{\beta}}^\top \bar{\widehat{\beta}}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\widehat{\beta}}^\top \bar{r}}{N} - \left(\frac{\bar{\beta}^\top \bar{\beta}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\beta}^\top \bar{r}}{N}}_{\mathcal{A}_\gamma} \\
&\qquad + \underbrace{\sqrt{T}H^\top \left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1} \frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{N}\frac{1}{T}\sum_t Z_t\Gamma_\beta^0 \nu_{t+1}^0}_{\mathcal{B}_\gamma} \\
&\qquad + \underbrace{\sqrt{T}H^\top \left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1} \frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{N}\bar{\epsilon}}_{\mathcal{C}_\gamma} \\
&=: \mathcal{A}_\gamma + \mathcal{B}_\gamma + \mathcal{C}_\gamma.
\end{aligned}$$

where the first equality follows by adding and subtracting; the second equality follows from the definition of the feasible and infeasible estimator; the last equality follows from the definition of the time series average of asset returns and rearranging; and, finally, we define three terms to prove, in the rest of this proof, $\mathcal{A}_\gamma + \mathcal{C}_\gamma = o_p(1)$ and $\mathcal{B}_\gamma = O_p(1)$ to prove the lemma.

First, we prove $\mathcal{A}_\gamma = o_P(1)$. Define notation:

$$\mathcal{A}_\gamma \coloneqq \widehat{\mathcal{A}}_i^{-1}\widehat{\mathcal{A}}_{ii} - \mathcal{A}_i^{-1}\mathcal{A}_{ii} \coloneqq \left(\frac{\bar{\widehat{\beta}}^\top \bar{\widehat{\beta}}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\widehat{\beta}}^\top \bar{r}}{N} - \left(\frac{\bar{\beta}^\top \bar{\beta}}{N}\right)^{-1} \frac{\sqrt{T}\bar{\beta}^\top \bar{r}}{N}$$

and $\Delta_i \coloneqq \widehat{\mathcal{A}}_i^{-1} - \mathcal{A}_i^{-1}$ and $\Delta_{ii} \coloneqq \widehat{\mathcal{A}}_{ii} - \mathcal{A}_{ii}$. Thus,

$$\widehat{\mathcal{A}}_i^{-1}\widehat{\mathcal{A}}_{ii} - \mathcal{A}_i^{-1}\mathcal{A}_{ii} = \mathcal{A}_{ii}\Delta_i + \mathcal{A}_i^{-1}\Delta_{ii} + \Delta_i\Delta_{ii} = O_p(1)o_p(1) + O_p(1)o_p(1) + o_p(1)o_p(1) = o_p(1)$$

given: by Lemma A9, $\mathcal{A}_{ii} = O_p(1)$ and $\Delta_{ii} = o_p(1)$; and, by Lemma A8 and CMT, $\mathcal{A}_i = O_p(1)$ and $\Delta_i = o_p(1)$.

Second, $\mathcal{B}_\gamma = O_p(1)$ given

$$\begin{aligned}
\left\|\mathcal{B}_\gamma\right\| &= \sqrt{T}\left\| H^\top \left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1}\frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{N}\frac{1}{T}\sum_t Z_t\Gamma_\beta^0\nu_{t+1}^0\right\| \\
&\le \|H\|\left\|\left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1}\right\|_2 \left\|\frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{\sqrt{N}}\right\|\left\|\sqrt{\frac{1}{NT}}\sum_t Z_t\Gamma_\beta^0\nu_{t+1}^0\right\| \\
&\lesssim_P \phi_{\min}^{-1}(2s)[N^{-1}\bar{Z}^\top\bar{Z}]
\end{aligned}$$

where the first equality substitutes the notation; the first inequality follows by $\|ABx\| \le \|A\|\,\|B\|_2\,\|x\|$ for matrices A, B and vector x where $\|B\|_2$ is the spectral norm; and, the probabilistic bound follows given $\|H\| = O_p(1)$, the spectral norm of the inverse matrix is bounded by the $2s$-sparse minimum eigenvalue of $N^{-1}\bar{Z}^\top\bar{Z}$ which is bounded away from zero by DSL Assumption SE, $\left\|\bar{Z}\Gamma_\beta^0\right\| = O_p(1)$ by an analogues argument using the $2s$-sparse maximum eigenvalue, and $\left\|\sqrt{\frac{1}{TN}}\sum_t Z_t\Gamma_\beta^0\nu_{t+1}^0\right\| = O_p\,(1)$ by Assumption 4(ii).

Third, $\mathcal{C}_\gamma = o_p(1)$ given

$$\begin{aligned}
\left\|\mathcal{C}_\gamma\right\| &= \left\|\sqrt{T}H^\top\left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1}\frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{N}\bar{\epsilon}\right\| \\
&\lesssim_P \left\|\frac{\Gamma_\beta^{0\top}\left(\bar{Z}\pm\mathbb{E}[\bar{Z}]\right)^\top}{N}\sqrt{\frac{1}{T}}\sum_t \epsilon_t\right\| \\
&= \left\|\sqrt{\frac{1}{T}}\sum_t\sum_{j=1}^{s}\frac{1}{N}\sum_i \Gamma_{\beta,j}^0\left(\bar{Z}_{i,j} - \mathbb{E}[\bar{Z}_{i,j}]\right)^\top\epsilon_{i,t+1}\right\| \\
&\quad + \left\|\sqrt{\frac{1}{T}}\sum_t\sum_{j=1}^{s}\frac{1}{N}\sum_i \Gamma_{\beta,j}^0\mathbb{E}[\bar{Z}_{i,j}]^\top\epsilon_{i,t+1}\right\| \\
&= O_p(\sqrt{\frac{T}{N}}) = o_p(1)
\end{aligned}$$

where the first equality substitutes the notation; the first probabilistic bound holds by adding and subtracting $\mathbb{E}[\bar{Z}]$ and the previously established results on $\|H\|$ and the spectral norm of the inverse design matrix; the second equality holds by opening up the matrix multiplication for the two terms noting Γ_β^0 selects only s rows of $\bar{Z}^\top$; the final probabilistic bound holds given Assumption 4(i) yields $N^{-1}\sum_i \epsilon_{i,t+1} = O_p(N^{-1/2})$ and a

LLN on $T^{-1}\sum_t \bar{Z}_{i,j} - T^{-1}\sum_t \mathbb{E}[\bar{Z}_{i,j}] = o_p(1)$ using the structural moment Assumptions A3; and, the final assumption holds given $\sqrt{T/N} \to 0$.

Thus, given $\mathcal{A}_\gamma + \mathcal{B}_\gamma + \mathcal{C}_\gamma = o_p(1) + O_p(1) + o_p(1) = O_p(1)$, the lemma holds. □

Lemma A11. *Under the assumptions of Lemma A8, we have*

$$\left\| T^{-1}\hat{V}^\top \hat{V} - T^{-1}V^\top V \right\| = O_p\left(\sqrt{\frac{\log^2(Tp)}{TN^2}}\right).$$

Proof of Lemma A11.

$$\begin{aligned}
\left\| T^{-1}\hat{V}^\top \hat{V} - T^{-1}V^\top V \right\| &= T^{-1}\left\| \hat{V}^\top \hat{V} \pm \widehat{V}^\top V - V^\top V \right\| \\
&\leq T^{-1}\left\| \hat{V}^\top \hat{V} - \widehat{V}^\top V \right\| + T^{-1}\left\| \widehat{V}^\top V - V^\top V \right\| \\
&\leq T^{-1}\left\| \widehat{V} \right\| \left\| \hat{V} - V \right\| + T^{-1}\left\| \widehat{V} - V \right\| \|V\| \\
&= O_p\left(\sqrt{\frac{\log^2(Tp)}{TN^2}}\right).
\end{aligned}$$

where the first equality follows by adding and subtracting the term; the first and second inequalities follow by the triangle and Cauchy-Schwartz inequalities, respectively; and, the probabilistic bound follows given $\left\| \widehat{V} \right\| \lesssim_P \sqrt{kT}$ by the normalization, $\left\| \widehat{V} - V^0 H \right\| = O_p\left(\frac{\log(Tp)}{N}\right)$ by an analogous argument to Lemma A4; and, $\|V\| \leq \left\|V^0\right\| \|H\| = O_p(\sqrt{T})O_p(1)$ by Assumption 2(i). □

Lemma A12. *Under the assumptions of Lemma A8 and Assumption 4(ii), we have*

$$\left\| T^{-1/2}\widehat{V}^\top G - T^{-1/2}V^\top G \right\| = O_p\left(\frac{\log(Tp)}{N}\right).$$

Proof of Lemma A12.

$$\begin{aligned}
\left\| T^{-1/2}\widehat{V}^\top G - T^{-1/2}V^\top G \right\| &= T^{-1/2}\left\| \left(\widehat{V} - V\right)^\top \left(V^0 H H^{-1}\eta_0 + \epsilon^g\right) \right\| \\
&\leq T^{-1/2}\left\| \widehat{V} - V \right\| \left\| V^0 \right\| \|\eta_0\| + T^{-1/2}\left\| \widehat{V} - V \right\| \|\epsilon^g\| \\
&\lesssim_P O_p\left(\frac{\log(Tp)}{N}\right).
\end{aligned}$$

where the equality follows from definition of the observable factor model; the inequality follows from the use of the triangle and Cauchy-Schwartz inequalities; and, the probabilistic bound follows as in Lemma A11 with $\|\epsilon^g\| \lesssim_P \sqrt{T}$ by Assumption 4(ii). □

LEMMA A13. *Under the assumptions of Lemma A12, we have*

$$\sqrt{T}(\widehat{\eta} - \eta) = O_p(1)$$

Proof of Lemma A13.

$$\begin{aligned}
\sqrt{T}(\widehat{\eta} - \eta) &= \sqrt{T}\,(\widehat{\eta} \pm \widetilde{\eta} - \eta) \\
&= \left(\frac{\widehat{V}^\top \widehat{V}}{T}\right)^{-1} \sqrt{T}\frac{\widehat{V}^\top G}{T} \pm \left(\frac{V^\top V}{T}\right)^{-1} \sqrt{T}\frac{V^\top G}{T} - \sqrt{T}H^{-1}\eta_0 \\
&= \left(\frac{\widehat{V}^\top \widehat{V}}{T}\right)^{-1} \sqrt{T}\frac{\widehat{V}^\top G}{T} - \left(\frac{V^\top V}{T}\right)^{-1} \sqrt{T}\frac{V^\top G}{T} \\
&\quad + (H^\top)^{-1}\left(\frac{V^{0\top} V^0}{T}\right)^{-1} \sqrt{T}\frac{V^{0\top} \epsilon^g}{T} \\
&= (H^\top)^{-1}\left(\frac{V^{0\top} V^0}{T}\right)^{-1} \sqrt{T}\frac{V^{0\top} \epsilon^g}{T} + o_p(1) = O_p(1).
\end{aligned}$$

where the first equality follows from adding and subtracting the infeasible estimator; the second inequality follows from the definitions; the third equality follows from the definition of the model for G and rearranging; the penultimate probabilistic bound follows given the difference between the first two terms is $o_p(1)$ by the results of Lemmas A11 and A12 using an analogous argument as Lemma A10; and, the final probabilistic bound holds given $\|H\| = O_p(1)$ for invertible matrix H, the factors have positive definite second moment matrix given Assumption 2(ii), the time series mean of the mean zero random variable $v^0_{t+1}\epsilon^g_{t+1}$ is $\sqrt{T}$ by CLT Assumption 4(iv). □

Define $\Pi_t := \sum_{j=1}^{s}\sum_{j'=1}^{s}\Gamma_{\beta,j'}\mathcal{Z}_{t,j,j'}\Gamma_{\beta,j}^\top$ containing nonstochastic scalar $\mathcal{Z}_{t,j,j'}$ from Assumption 4(iv) and the asymptotic variance of the Assumption 4(v) joint CLT as

$$\Phi_{11} = \lim_{p,T,N\to\infty} \frac{1}{T}\mathbb{E}\left[V^\top \epsilon^g \epsilon^{g\top} V\right] \qquad \Phi_{22} = \lim_{p,T,N\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}\mathbb{E}\left[\Pi_t v_{t+1} v_{t'+1}^\top \Pi_{t'}^\top\right]$$

$$\Phi_{12} = \lim_{p,T,N\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}\mathbb{E}\left[v_{t+1}\epsilon^g_{t+1} v_{t'+1}^\top \Pi_{t'}^\top\right]$$

where asymptotic covariance matrix Φ is defined as

$$\Phi \coloneqq \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{pmatrix}.$$

Proof of Theorem 1.

$$\begin{aligned}
\sqrt{T}\left(\widehat{\gamma}_g - \gamma_g\right) &= \sqrt{T}\left(\widehat{\eta}^\top \widehat{\gamma} - \eta^\top \gamma\right) \\
&= \widehat{\eta}^\top \sqrt{T}\left(\widehat{\gamma} - \gamma\right) + \sqrt{T}\gamma^\top\left(\widehat{\eta} - \eta\right) \\
&= \left(\widehat{\eta} \pm \eta\right)^\top \sqrt{T}\left(\widehat{\gamma} - \gamma\right) + \sqrt{T}\gamma^\top\left(\widehat{\eta} - \eta\right) \\
&= \left(\widehat{\eta} - \eta\right)^\top \sqrt{T}\left(\widehat{\gamma} - \gamma\right) + \sqrt{T}\eta^\top\left(\widehat{\gamma} - \gamma\right) + \sqrt{T}\gamma^\top\left(\widehat{\eta} - \eta\right) \\
&= \sqrt{T}\eta^\top\left(\widehat{\gamma} - \gamma\right) + \sqrt{T}\gamma^\top\left(\widehat{\eta} - \eta\right) + o_p(1) \\
&= \gamma^{0\top}\left(\frac{V^{0\top}V^0}{T}\right)^{-1} \frac{\sqrt{T}}{T}\sum_t v_{t+1}^0 \epsilon_{t+1}^g \\
&\qquad + \eta^{0\top}\left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1} \frac{\Gamma_\beta^{0\top}\bar{Z}^\top}{N}\frac{\sqrt{T}}{T}\sum_t Z_t\Gamma_\beta^0 v_{t+1}^0 \\
&\qquad + o_p(1) \\
&= \gamma^{0\top}\left(\frac{V^{0\top}V^0}{T}\right)^{-1} \frac{\sqrt{T}}{T}\sum_t v_{t+1}^0 \epsilon_{t+1}^g + o_p(1) \\
&\quad + \eta^{0\top}\left(\frac{\Gamma_\beta^{0\top}\bar{Z}^\top\bar{Z}\Gamma_\beta^0}{N}\right)^{-1} \frac{\sqrt{T}}{T}\sum_t \underbrace{\left(\sum_{j=1}^s\sum_{j'=1}^s \Gamma_{\beta,j'}^0 \frac{1}{N}\sum_i \mathbb{E}[\bar{z}_{i,j'}]z_{i,t,j}\Gamma_{\beta,j}^\top\right)}_{\Pi_t} v_{t+1}^0 \\
&\to_d N\left(0, \sigma_g^2\right)
\end{aligned}$$

where the first equality holds by definition of the estimator and target parameter; the second and third equalities hold by adding and subtracting terms; the fourth equality holds by re-arranging; the fifth equality holds by Lemma A10 $\sqrt{T}\left(\widehat{\gamma} - \gamma\right) = O_p(1)$ and Lemma A13 $\widehat{\eta} - \eta = O_P(T^{-1/2}) = o_p(1)$; the last equality holds by Lemmas A10 and A13, which leaves the two non-asymptotically negligible terms at rate $\sqrt{T}$ scaled by the associated true parameters η^0 and γ^0; and, the convergence in distribution holds given by the joint CLT assumption 4(v) applying the delta method to the linear sum.

Define the following two invertible matrices matrices: $A \coloneqq \lim_{T\to\infty} T^{-1}\mathbb{E}\left[V^\top V\right]$

and $B := \lim_{p,T,N\to\infty} \frac{1}{N}\mathbb{E}\left[\Gamma_\beta^\top \bar{Z}^\top \bar{Z}\Gamma_\beta\right]$.

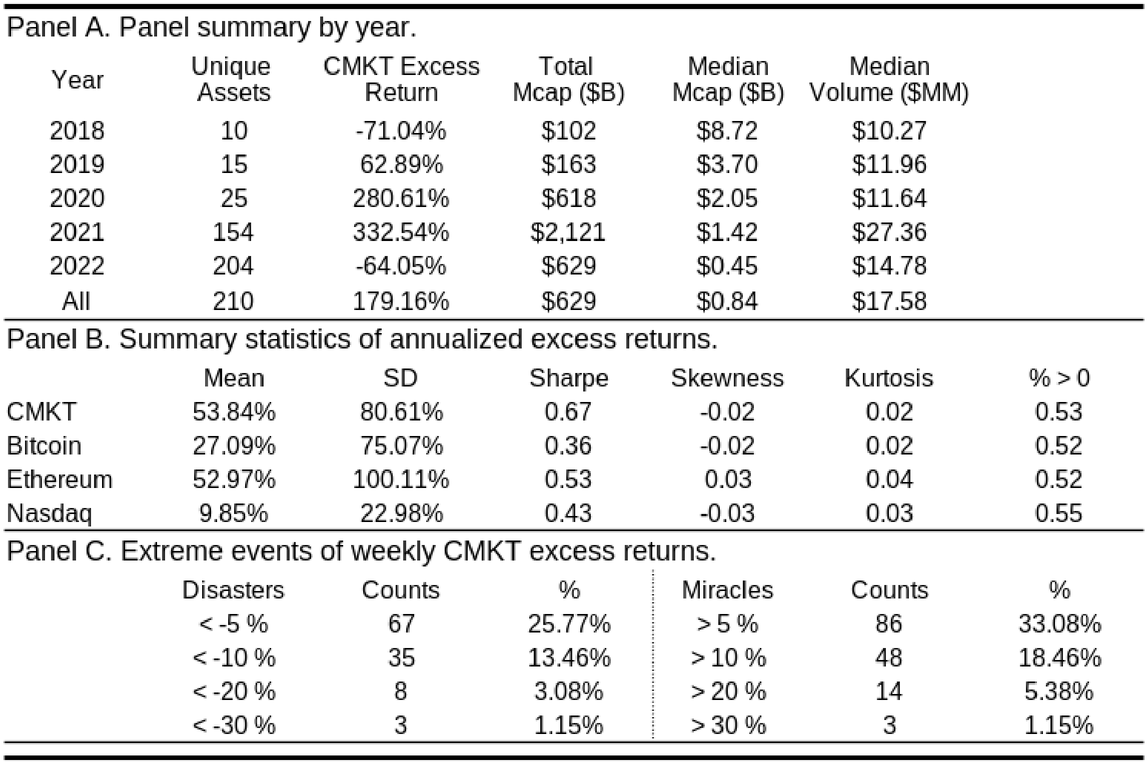
The asymptotic variance σ_g^2 is thus given by the delta method, where the rotation matrix falls out, as

$$\sigma_g^2 := \gamma^\top A^{-1}\Phi_{11}(A^\top)^{-1}\gamma + \eta^\top B^{-1}\Phi_{22}(B^\top)^{-1}\eta + \gamma^\top A^{-1}\Phi_{12}(B^\top)^{-1}\eta + \eta^\top B^{-1}\Phi_{12}^\top(A^\top)^{-1}\gamma.$$

□

Appendix D. Tables and Figures

FIGURE A1. Summary statistics.



Panel A. Panel summary by year.

Year	Unique Assets	CMKT Excess Return	Total Mcap ($B)	Median Mcap ($B)	Median Volume ($MM)
2018	10	-71.04%	$102	$8.72	$10.27
2019	15	62.89%	$163	$3.70	$11.96
2020	25	280.61%	$618	$2.05	$11.64
2021	154	332.54%	$2,121	$1.42	$27.36
2022	204	-64.05%	$629	$0.45	$14.78
All	210	179.16%	$629	$0.84	$17.58

Panel B. Summary statistics of annualized excess returns.

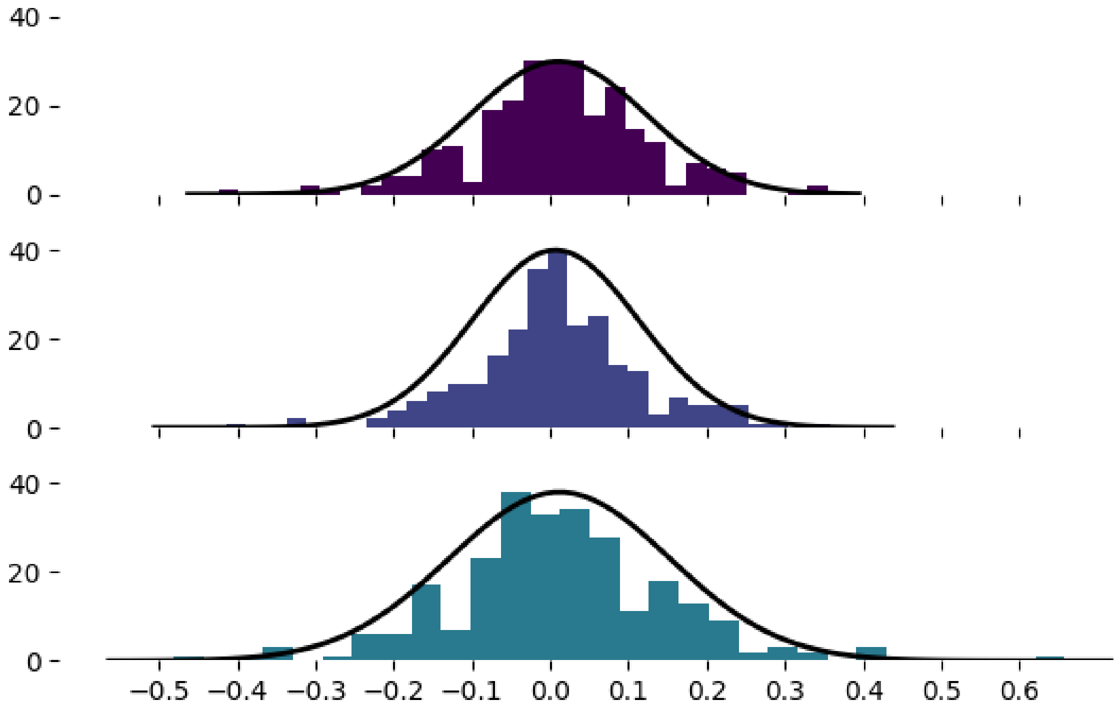
	Mean	SD	Sharpe	Skewness	Kurtosis	% > 0
CMKT	53.84%	80.61%	0.67	-0.02	0.02	0.53
Bitcoin	27.09%	75.07%	0.36	-0.02	0.02	0.52
Ethereum	52.97%	100.11%	0.53	0.03	0.04	0.52
Nasdaq	9.85%	22.98%	0.43	-0.03	0.03	0.55

Panel C. Extreme events of weekly CMKT excess returns.

Disasters	Counts	%	Miracles	Counts	%
< -5 %	67	25.77%	> 5 %	86	33.08%
< -10 %	35	13.46%	> 10 %	48	18.46%
< -20 %	8	3.08%	> 20 %	14	5.38%
< -30 %	3	1.15%	> 30 %	3	1.15%

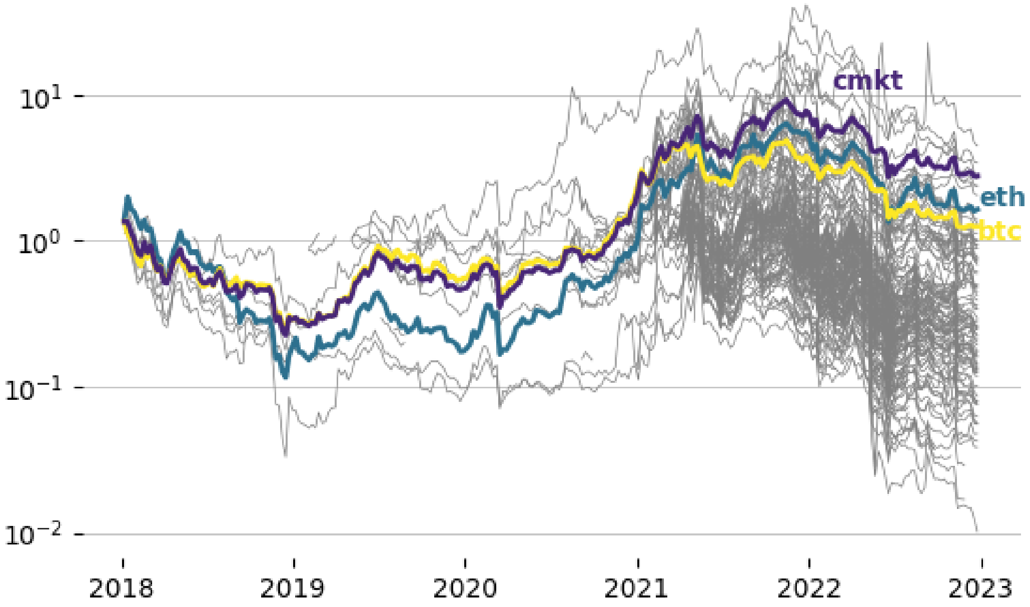
This table shows summary statistics on the weekly panel of excess returns from January 1, 2018 to December 31 2022. Panel A reports, by calendar year and for the whole panel, the number of unique assets, the cumulative excess return of the crypto market, the total market capitalization in the last week in billions of dollars, the median market capitalization in billions of dollars, and the median weekly volume in millions of dollars. Panel B reports—for the crypto market (CMKT), Bitcoin, Ethereum, and the Nasdaq—annualized excess return statistics, including the mean, standard deviation, Sharpe ratio, skewness, kurtosis, and percentage of weekly excess returns that are positive. Panel C reports the percentage of extreme events using the weekly crypto market index excess returns.

FIGURE A2. Empirical distributions of CMKT, Bitcoin, and Ethereum weekly returns.



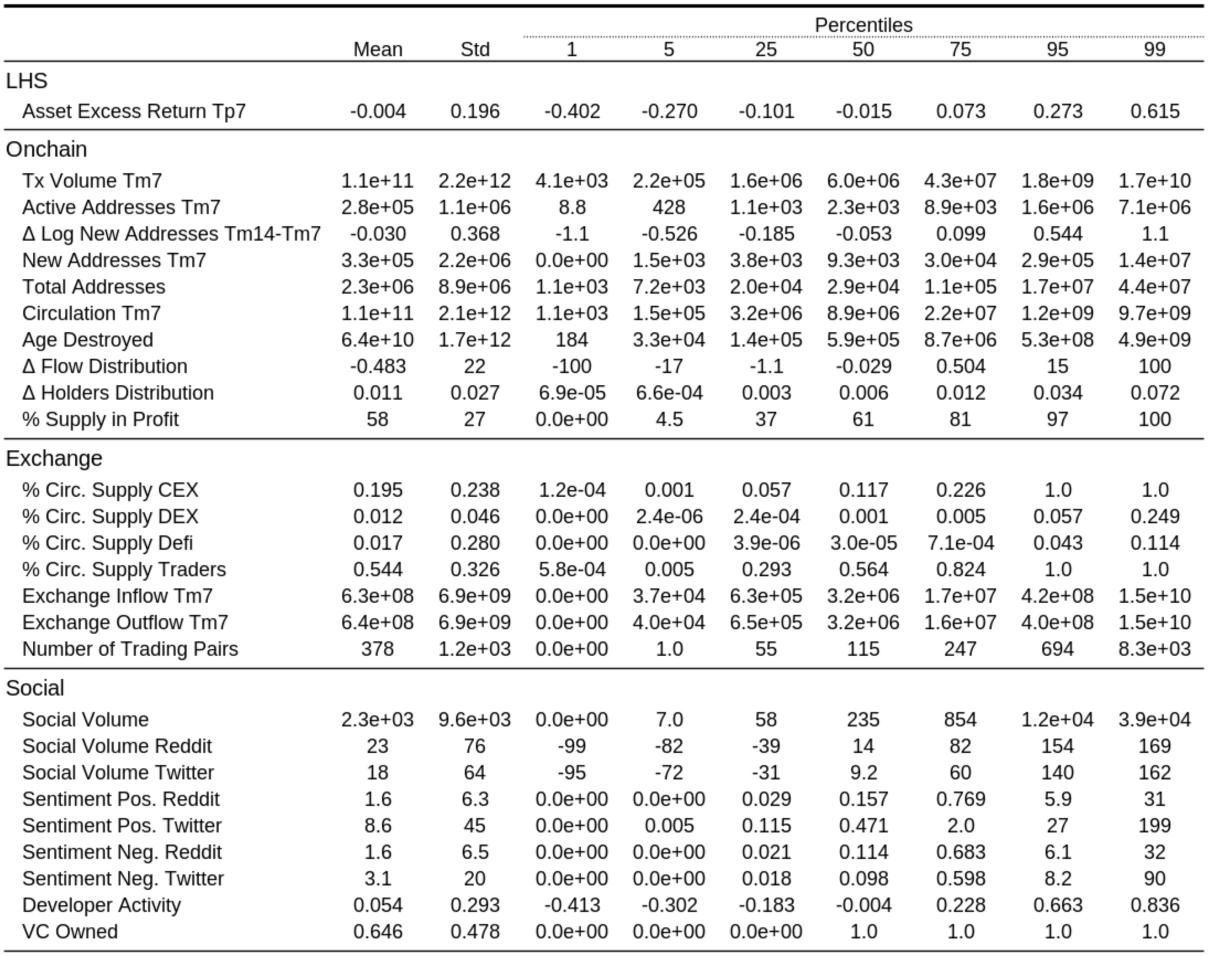
This figure shows the empirical distributions of weekly excess returns, with a normal distribution fit, for coin market (top panel), Bitcoin (middle panel), and Ethereum (bottom panel) for the January 1, 2018 to December 31, 2022 period.

FIGURE A3. Cumulative weekly return of assets in universe.



This figure shows the cumulative excess returns for each asset in the study's universe, and the crypto market, for the January 1, 2018 to December 31, 2022 time period.

FIGURE A4. Crypto Asset Characteristics: Descriptive Statistics.



	Mean	Std	Percentiles						
			1	5	25	50	75	95	99
LHS									
Asset Excess Return Tp7	-0.004	0.196	-0.402	-0.270	-0.101	-0.015	0.073	0.273	0.615
Onchain									
Tx Volume Tm7	1.1e+11	2.2e+12	4.1e+03	2.2e+05	1.6e+06	6.0e+06	4.3e+07	1.8e+09	1.7e+10
Active Addresses Tm7	2.8e+05	1.1e+06	8.8	428	1.1e+03	2.3e+03	8.9e+03	1.6e+06	7.1e+06
Δ Log New Addresses Tm14-Tm7	-0.030	0.368	-1.1	-0.526	-0.185	-0.053	0.099	0.544	1.1
New Addresses Tm7	3.3e+05	2.2e+06	0.0e+00	1.5e+03	3.8e+03	9.3e+03	3.0e+04	2.9e+05	1.4e+07
Total Addresses	2.3e+06	8.9e+06	1.1e+03	7.2e+03	2.0e+04	2.9e+04	1.1e+05	1.7e+07	4.4e+07
Circulation Tm7	1.1e+11	2.1e+12	1.1e+03	1.5e+05	3.2e+06	8.9e+06	2.2e+07	1.2e+09	9.7e+09
Age Destroyed	6.4e+10	1.7e+12	184	3.3e+04	1.4e+05	5.9e+05	8.7e+06	5.3e+08	4.9e+09
Δ Flow Distribution	-0.483	22	-100	-17	-1.1	-0.029	0.504	15	100
Δ Holders Distribution	0.011	0.027	6.9e-05	6.6e-04	0.003	0.006	0.012	0.034	0.072
% Supply in Profit	58	27	0.0e+00	4.5	37	61	81	97	100
Exchange									
% Circ. Supply CEX	0.195	0.238	1.2e-04	0.001	0.057	0.117	0.226	1.0	1.0
% Circ. Supply DEX	0.012	0.046	0.0e+00	2.4e-06	2.4e-04	0.001	0.005	0.057	0.249
% Circ. Supply Defi	0.017	0.280	0.0e+00	0.0e+00	3.9e-06	3.0e-05	7.1e-04	0.043	0.114
% Circ. Supply Traders	0.544	0.326	5.8e-04	0.005	0.293	0.564	0.824	1.0	1.0
Exchange Inflow Tm7	6.3e+08	6.9e+09	0.0e+00	3.7e+04	6.3e+05	3.2e+06	1.7e+07	4.2e+08	1.5e+10
Exchange Outflow Tm7	6.4e+08	6.9e+09	0.0e+00	4.0e+04	6.5e+05	3.2e+06	1.6e+07	4.0e+08	1.5e+10
Number of Trading Pairs	378	1.2e+03	0.0e+00	1.0	55	115	247	694	8.3e+03
Social									
Social Volume	2.3e+03	9.6e+03	0.0e+00	7.0	58	235	854	1.2e+04	3.9e+04
Social Volume Reddit	23	76	-99	-82	-39	14	82	154	169
Social Volume Twitter	18	64	-95	-72	-31	9.2	60	140	162
Sentiment Pos. Reddit	1.6	6.3	0.0e+00	0.0e+00	0.029	0.157	0.769	5.9	31
Sentiment Pos. Twitter	8.6	45	0.0e+00	0.005	0.115	0.471	2.0	27	199
Sentiment Neg. Reddit	1.6	6.5	0.0e+00	0.0e+00	0.021	0.114	0.683	6.1	32
Sentiment Neg. Twitter	3.1	20	0.0e+00	0.0e+00	0.018	0.098	0.598	8.2	90
Developer Activity	0.054	0.293	-0.413	-0.302	-0.183	-0.004	0.228	0.663	0.836
VC Owned	0.646	0.478	0.0e+00	0.0e+00	0.0e+00	1.0	1.0	1.0	1.0

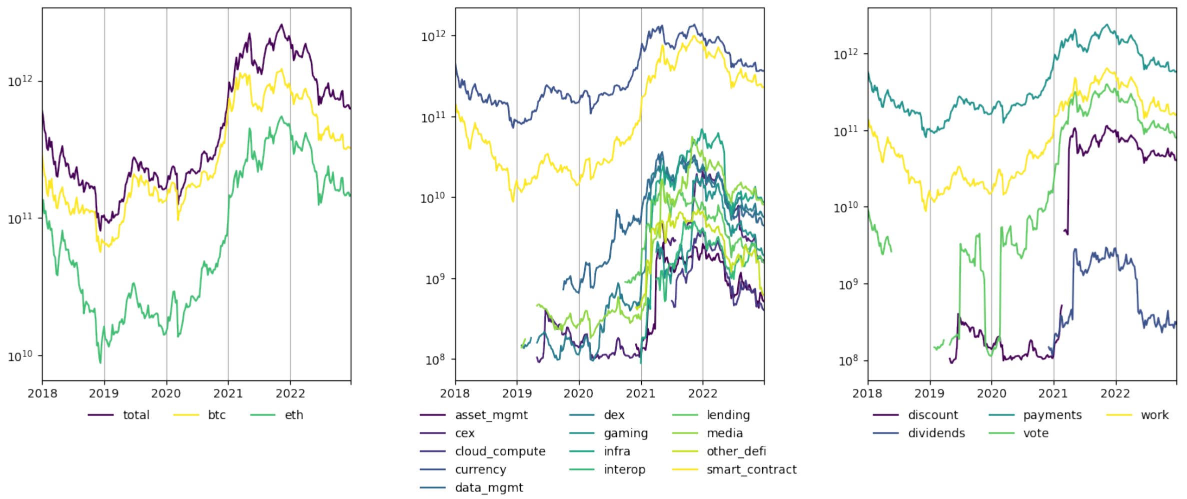
This table reports the summary statistics from the weekly asset panel for the dependent variable, asset excess returns seven days ahead, and the asset characteristics. For each variable, we report the panel mean, median, standard deviation, and selected percentiles. There are 22,678 asset-weeks from January 7, 2018 to December 15, 2022.

FIGURE A5. Crypto Asset Characteristics: Descriptive Statistics (Continued).

Momentum									
Return Tm7	0.004	0.906	-0.402	-0.270	-0.101	-0.014	0.075	0.277	0.616
Return Tm14	0.007	1.3	-0.485	-0.348	-0.166	-0.043	0.097	0.409	0.969
Return Tm30	0.010	1.3	-0.599	-0.481	-0.259	-0.095	0.132	0.698	1.7
Return Tm60	0.063	1.5	-0.740	-0.629	-0.385	-0.159	0.174	1.3	3.7
Return Tm90	0.173	2.0	-0.790	-0.676	-0.457	-0.217	0.190	2.2	6.0
Return Tm14-Tm7	0.004	0.907	-0.404	-0.272	-0.104	-0.015	0.074	0.277	0.625
Return Tm30-Tm14	0.016	1.3	-0.522	-0.377	-0.176	-0.039	0.110	0.484	1.2
Return Tm90-Tm30	0.142	1.6	-0.733	-0.620	-0.364	-0.116	0.234	1.6	4.6
Return from ATH	-0.674	0.233	-0.982	-0.948	-0.863	-0.727	-0.540	-0.195	-0.049
Return from ATL	73	1.2e+03	0.014	0.083	0.589	2.8	14	94	370
Return Industry Tm30	0.331	5.4	-0.524	-0.396	-0.187	-0.041	0.204	0.683	1.6
Return Industry Tm60	0.409	5.2	-0.645	-0.552	-0.277	-0.051	0.327	1.4	4.2
Microstructure									
Trades Sum Tm7	2.8e+05	8.5e+05	2.8e+03	7.6e+03	3.3e+04	8.2e+04	2.1e+05	1.1e+06	3.8e+06
Volume Sum Tm7	1.7e+08	7.6e+08	9.1e+05	1.9e+06	5.9e+06	1.8e+07	5.9e+07	5.5e+08	4.4e+09
Spread bps	3.5	28	3.0e-05	2.1e-04	0.001	0.003	0.015	9.3	79
Ask Size	1.6e+04	2.4e+04	11	178	1.2e+03	5.6e+03	2.0e+04	7.3e+04	9.1e+04
Bid Size	3.7e+04	9.7e+04	10	185	1.2e+03	5.6e+03	2.5e+04	2.1e+05	3.5e+05
Illiquidity Tm7	1.7e-07	3.1e-07	1.5e-10	1.9e-09	1.9e-08	6.5e-08	1.9e-07	6.4e-07	1.5e-06
Turnover Tm7	12	130	2.9e-05	3.6e-04	0.005	0.031	0.194	11	327
Financial									
Price	655	4.4e+03	0.003	0.021	0.275	1.8	23	673	2.2e+04
Size	21	2.0	17	18	19	21	22	24	27
MVRV	1.6	3.4	0.0e+00	0.278	0.557	1.1	1.8	4.1	9.2
Alpha Tm7	-2.3e-05	0.003	-0.006	-0.003	-8.4e-04	-2.1e-05	7.2e-04	0.003	0.007
Alpha Tm30	5.7e-05	0.004	-0.001	-6.3e-04	-2.7e-04	-7.7e-05	1.7e-04	7.8e-04	0.002
Beta Tm7	0.010	0.024	-0.040	-0.020	-0.002	0.007	0.020	0.048	0.081
Beta Tm30	0.007	0.017	-0.008	-0.002	0.003	0.006	0.010	0.017	0.025
Downside Beta Tm30	0.021	0.048	-0.018	-0.004	0.007	0.013	0.022	0.045	0.273
Coskew Tm30	-0.001	0.123	-0.220	-0.103	-0.038	-0.005	0.027	0.121	0.240
iSkew Tm30	0.488	1.6	-2.3	-1.1	-0.187	0.215	0.784	2.9	6.8
Shortfall 5% Tm7	-0.024	0.014	-0.077	-0.050	-0.028	-0.020	-0.015	-0.010	-0.007
VaR 5% Tm7	-0.016	0.009	-0.051	-0.035	-0.020	-0.014	-0.010	-0.006	-0.004
Vol Tm7	0.011	0.045	0.003	0.004	0.007	0.009	0.013	0.024	0.036
Vol Tm30	0.013	0.059	0.004	0.005	0.008	0.010	0.014	0.024	0.031
Vol Tm90	0.015	0.064	0.004	0.006	0.009	0.011	0.015	0.024	0.029
Ivol Tm7	0.012	0.069	0.003	0.004	0.007	0.009	0.013	0.024	0.037
Ivol Tm30	0.013	0.074	0.004	0.005	0.008	0.010	0.014	0.024	0.032
Ivol Tm90	0.013	0.064	0.0e+00	0.0e+00	0.008	0.010	0.014	0.023	0.028

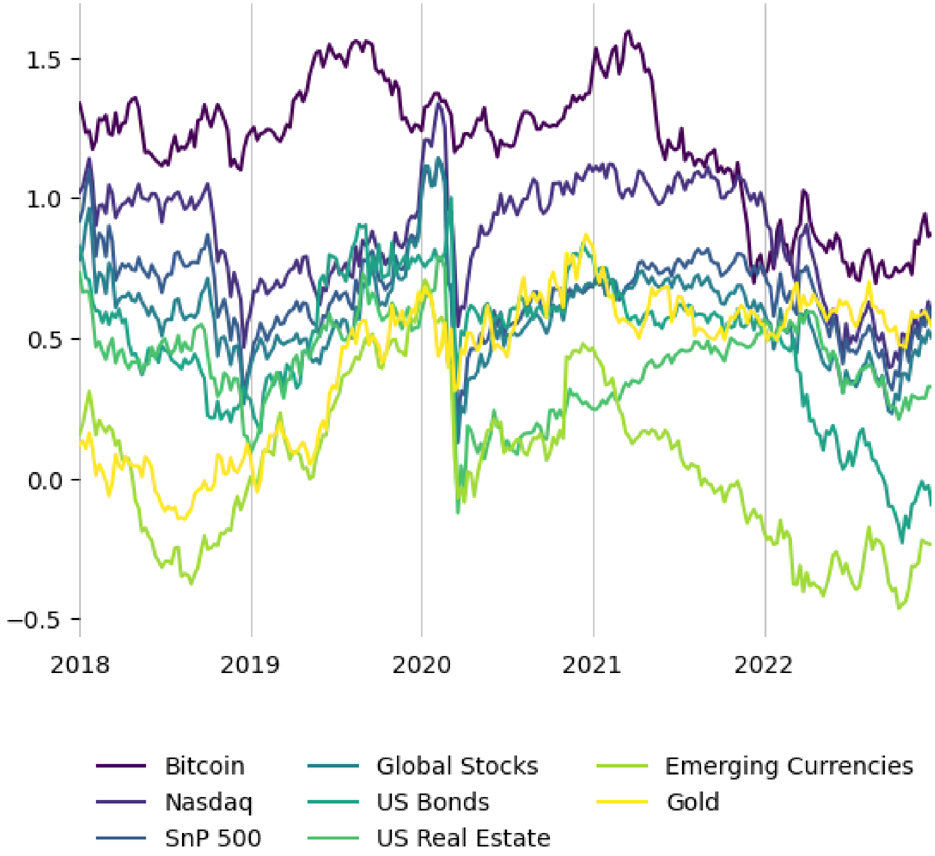
This table reports the summary statistics from the weekly asset panel for the dependent variable, asset excess returns seven days ahead, and the asset characteristics. For each variable, we report the panel mean, median, standard deviation, and selected percentiles. There are 22,678 asset-weeks from January 7, 2018 to December 15, 2022.

FIGURE A6. Market Caps (USD).



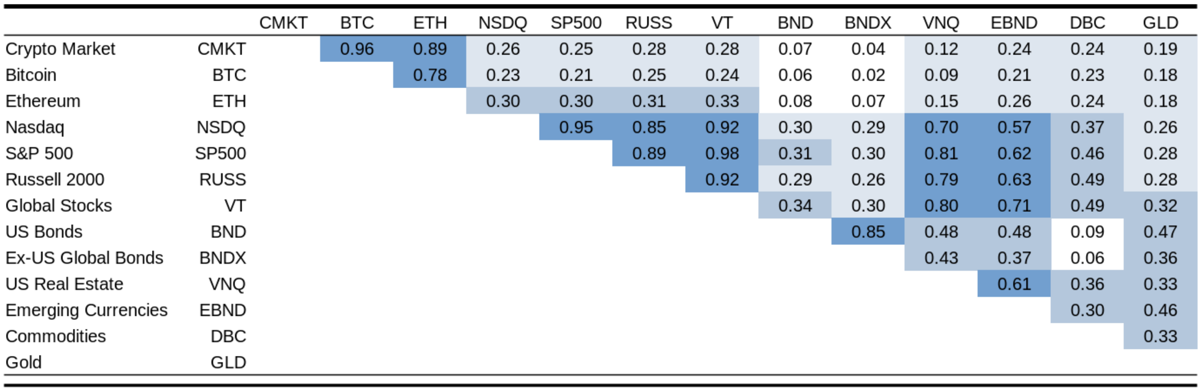
This figure shows the market capitalization of: in the first panel, the entire panel, Bitcoin, and Ethereum; in the second panel, the entire panel by asset industry classification; and, in the third panel, the entire panel by asset usage classification. The asset industry and usage classifications are from Messari.

FIGURE A7. Sharpe Ratios: Bitcoin vs Major Asset Classes.



This figure shows the rolling Sharpe Ratio over four year trailing windows using weekly excess returns for various asset classes. Bitcoin is the weekly return from Kraken's order book. Nasdaq and SnP 500 are the returns of the respective indices. The remaining series correspond to the following ETFs: Global Stocks is VT; US Bonds is BND; US Real Estate is VNQ; Emerging Currencies is EBND; and, Gold is GLD.

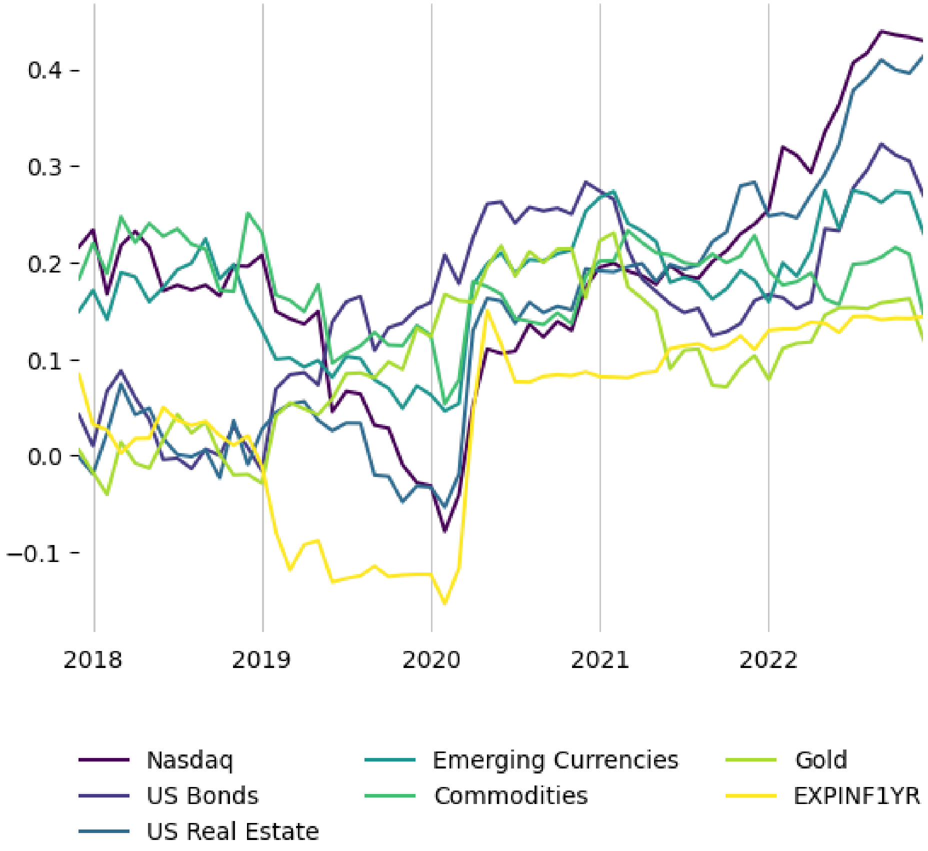
FIGURE A8. Correlations.



		CMKT	BTC	ETH	NSDQ	SP500	RUSS	VT	BND	BNDX	VNQ	EBND	DBC	GLD
Crypto Market	CMKT		0.96	0.89	0.26	0.25	0.28	0.28	0.07	0.04	0.12	0.24	0.24	0.19
Bitcoin	BTC			0.78	0.23	0.21	0.25	0.24	0.06	0.02	0.09	0.21	0.23	0.18
Ethereum	ETH				0.30	0.30	0.31	0.33	0.08	0.07	0.15	0.26	0.24	0.18
Nasdaq	NSDQ					0.95	0.85	0.92	0.30	0.29	0.70	0.57	0.37	0.26
S&P 500	SP500						0.89	0.98	0.31	0.30	0.81	0.62	0.46	0.28
Russell 2000	RUSS							0.92	0.29	0.26	0.79	0.63	0.49	0.28
Global Stocks	VT								0.34	0.30	0.80	0.71	0.49	0.32
US Bonds	BND									0.85	0.48	0.48	0.09	0.47
Ex-US Global Bonds	BNDX										0.43	0.37	0.06	0.36
US Real Estate	VNQ											0.61	0.36	0.33
Emerging Currencies	EBND												0.30	0.46
Commodities	DBC													0.33
Gold	GLD													

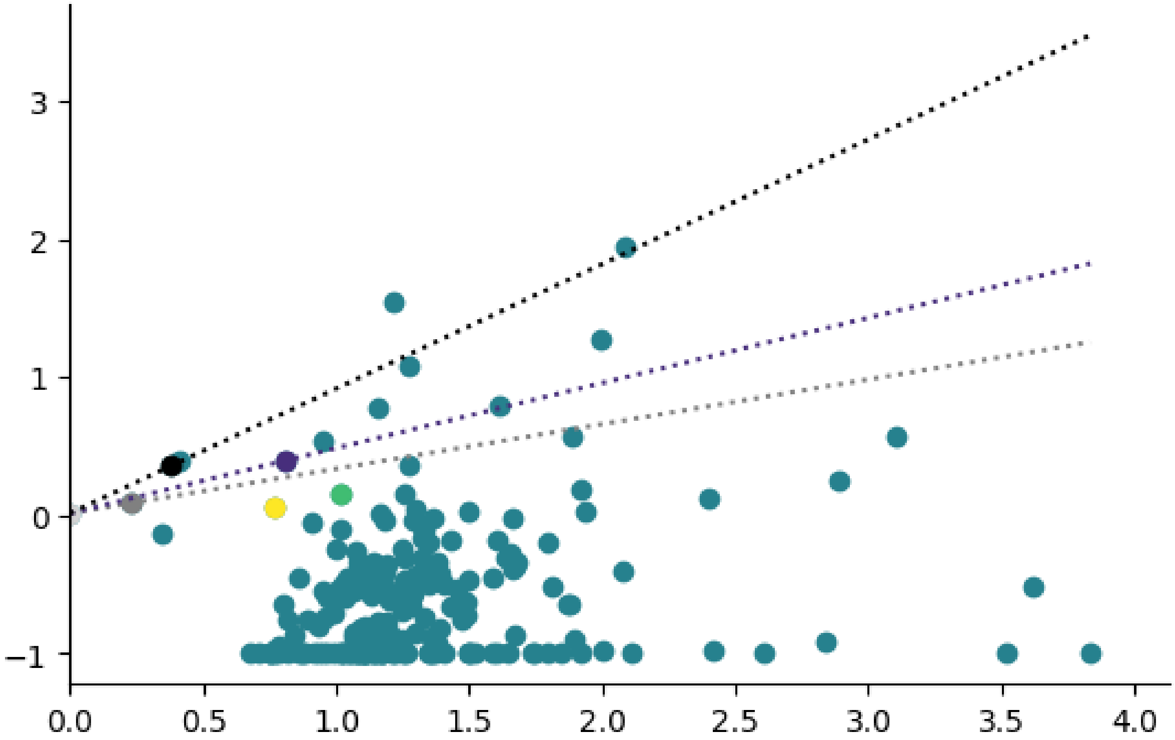
This table reports Pearson correlation coefficients between weekly excess returns of row and column assets for the January 1, 2018 to December 31, 2022 time period. CMKT, Bitcoin, and Ethereum refer to the weekly excess returns of the market cap-weighted assets in the study universe, bitcoin, and ether, respectively. Nasdaq, S&P500, and Russel 2000 refer to the weekly excess return of the IXIC, GSPC, and RUT indices, respectively. Global Stocks, US Bonds, Ex-US Global Bonds, US Real Estate, Emerging Currencies, Commodities, and Gold refer to the weekly excess returns of the following ETFs: VT, BND, BNDX, VNQ, EBND, DBC, and GLD.

FIGURE A9. Rolling Four Year Pearson Correlations: Bitcoin vs Major Asset Classes.



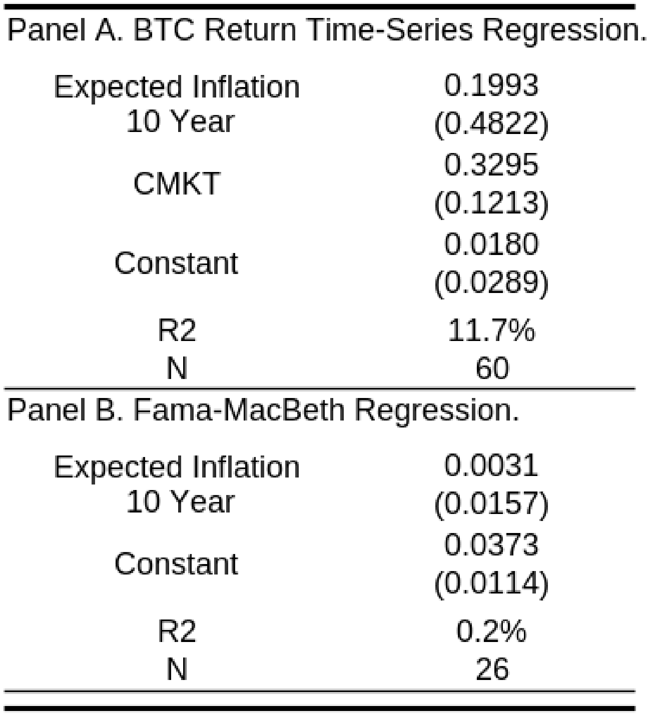
This figure shows rolling four-year Pearson Correlation coefficients between Bitcoin's weekly excess returns and those of other major asset classes for the January 1, 2018 to December 31, 2022 time period. Nasdaq refers to the weekly excess return of the IXIC index. US Bonds, US Real Estate, Emerging Currencies, Commodities, and Gold refer to the weekly excess returns of the following ETFs: BND, VNQ, EBND, DBC, and GLD. EXPINF1YR refers to the Federal Reserve's measure of expected inflation over the subsequent year.

FIGURE A10. Crypto Asset's Annualized Cumulative Returns and Volatility.



This figure shows the annualized cumulative return and annualized volatility of simple weekly excess returns of the crypto assets in the study universe (over the 2018-2022, inclusive, time period) as well as a few other portfolios. The light grey point is the risk free rate captured by the annualized cumulative return of the 1 month treasury bill during the study period. The grey point is the annualized cumulative return and annualized volatility of the Nasdaq index. For the same two measures, the yellow point corresponds to BTC, the light green point for ETH, and the purple point for CMKT. For the same two measures, the black point corresponds to a portfolio holding 60% Nasdaq and 40% CMKT. The remaining dark green points are for the rest of the assets in the study, removing three assets with outlier returns: DOGE at 9x, LUNA at 52x, and MATIC at 18x.

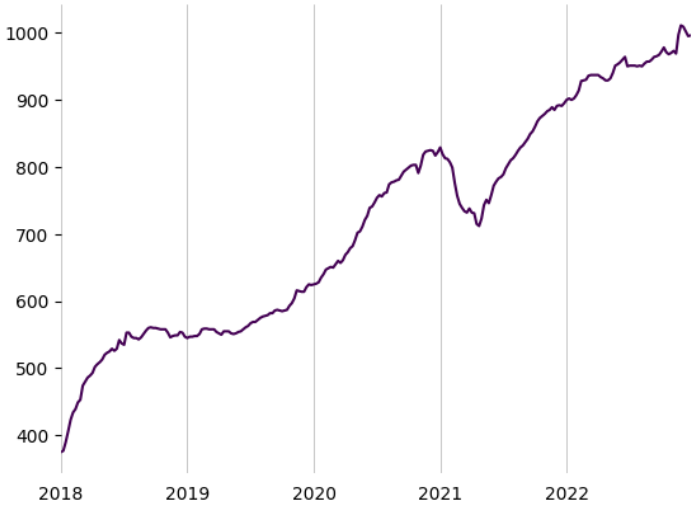
FIGURE A11. Inflation Risk Premium.



Panel A. BTC Return Time-Series Regression.	
Expected Inflation 10 Year	0.1993 (0.4822)
CMKT	0.3295 (0.1213)
Constant	0.0180 (0.0289)
R2	11.7%
N	60
Panel B. Fama-MacBeth Regression.	
Expected Inflation 10 Year	0.0031 (0.0157)
Constant	0.0373 (0.0114)
R2	0.2%
N	26

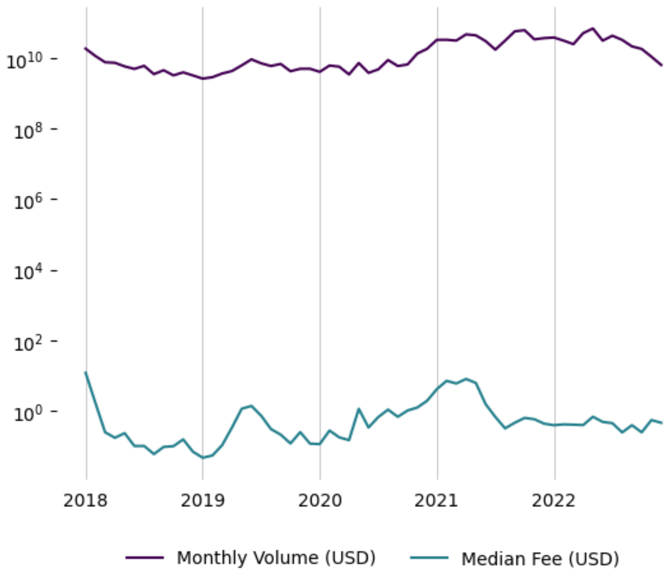
This table reports results from two regressions. Panel A reports point estimates and standard errors from the time-series regression of BTC monthly excess returns on 1 year expected inflation innovations, CMKT monthly excess returns, and a constant. Panel B reports the results from a Fama-MacBeth regression procedure to estimate the risk premium of inflation in the crypto asset class, where we use assets with at least two years of data to precisely estimate beta hats.

FIGURE A12. Hodling: Bitcoin UTXO Median Age in Days.



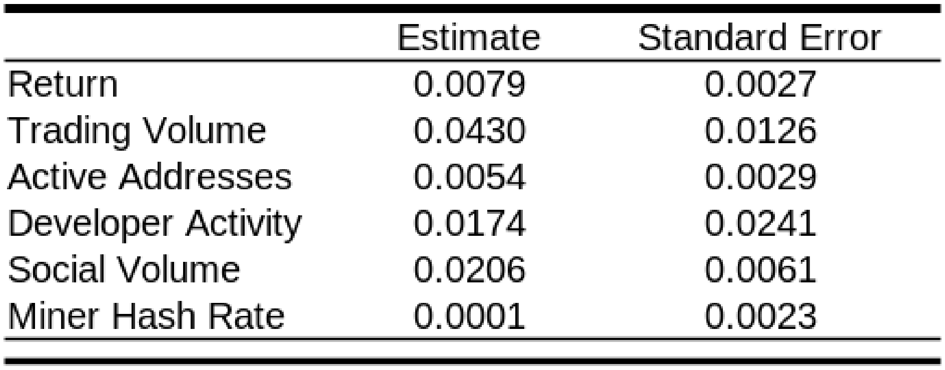
This figure shows median age in full days of all unspent transaction outputs (UTXOs), rounded down to the nearest day, on the Bitcoin ledger for each week in 2018 to 2022, inclusive. Why hodl?

Figure A13. Bitcoin Onchain Transactions.



This figure shows two time series for onchain bitcoin transactions. Monthly Volume reports, in USD, the total calendar month onchain volume transferred between distinct addresses. The Median Fee reports, in USD, the median fee paid to miners across all transactions within each calendar month.

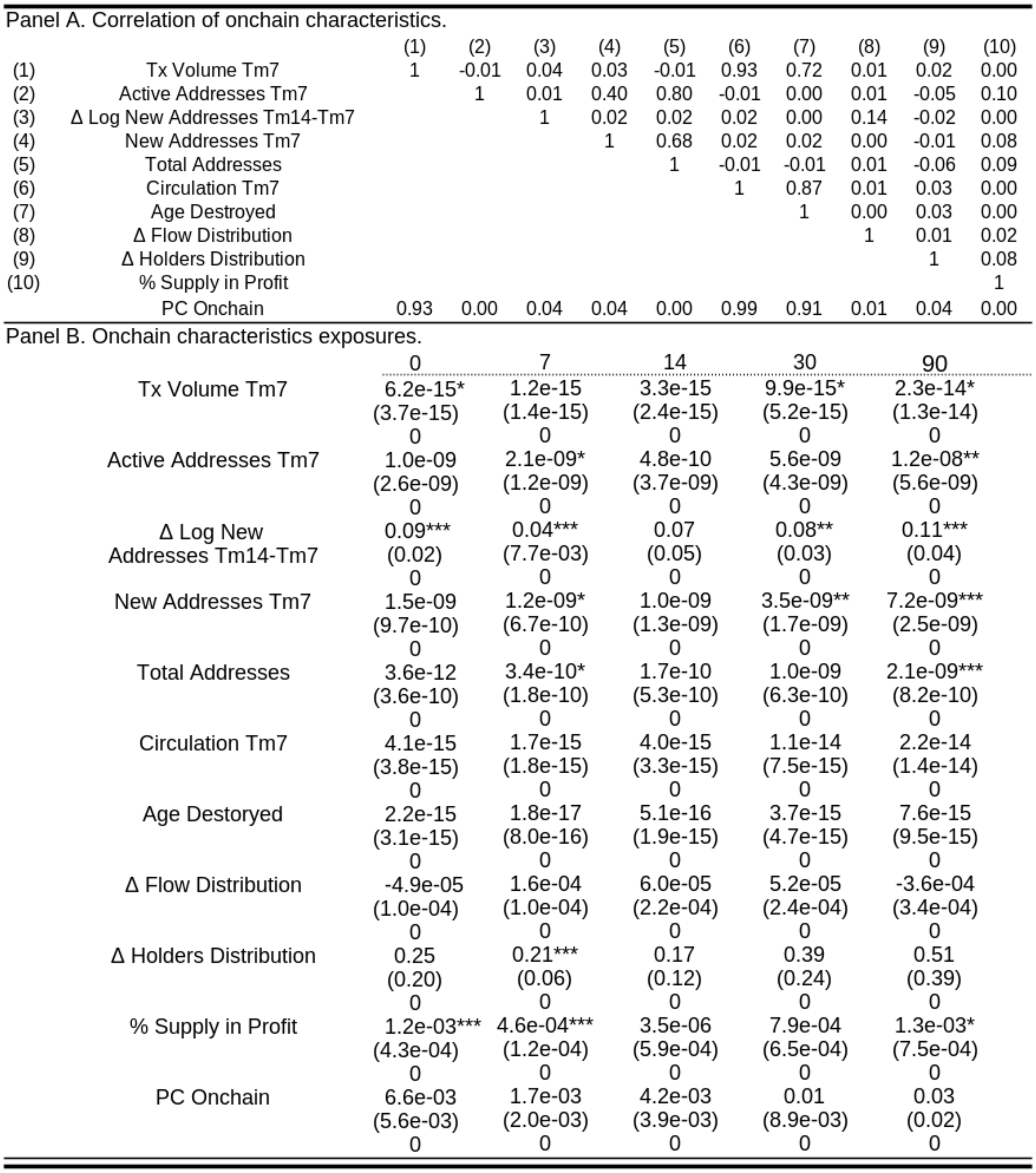
Figure A14. Bitcoin Forks: Event Study.



	Estimate	Standard Error
Return	0.0079	0.0027
Trading Volume	0.0430	0.0126
Active Addresses	0.0054	0.0029
Developer Activity	0.0174	0.0241
Social Volume	0.0206	0.0061
Miner Hash Rate	0.0001	0.0023

This table reports an event study for various Bitcoin statistics on dates on which fifteen major Bitcoin forks occurred, subsequent to January 2016. The point estimates are the difference between, in the seven days before and after the event date, the average daily change of each characteristic. Return is the daily change in bitcoin's USD price. Trading Volume is the daily change in bitcoin trading volume reported as by CoinMarketCap. Active Addresses is the daily change in the number of unique active Bitcoin addresses as reported by Santiment. Developer Activity is the daily change in the total number of GitHub events (e.g. code pushes, issue interactions, pull requests, comments on commits, etc.) as reported by Santiment. Social Volume is the daily change in the total number of text documents across Reddit, Telegram, Twitter, and BitcoinTalk containing the keyword "bitcoin" as reported by Santiment. Miner Hash Rate is the daily change in the total Bitcoin hash rate as imputed by Coinmetrics. Standard errors are bootstrapped: the standard deviation of the distribution formed by calculating each statistic for 10,000 randomly sampled, with replacement, event days.

Figure A15. Onchain Characteristics: Correlations and Signal.



Panel A. Correlation of onchain characteristics.

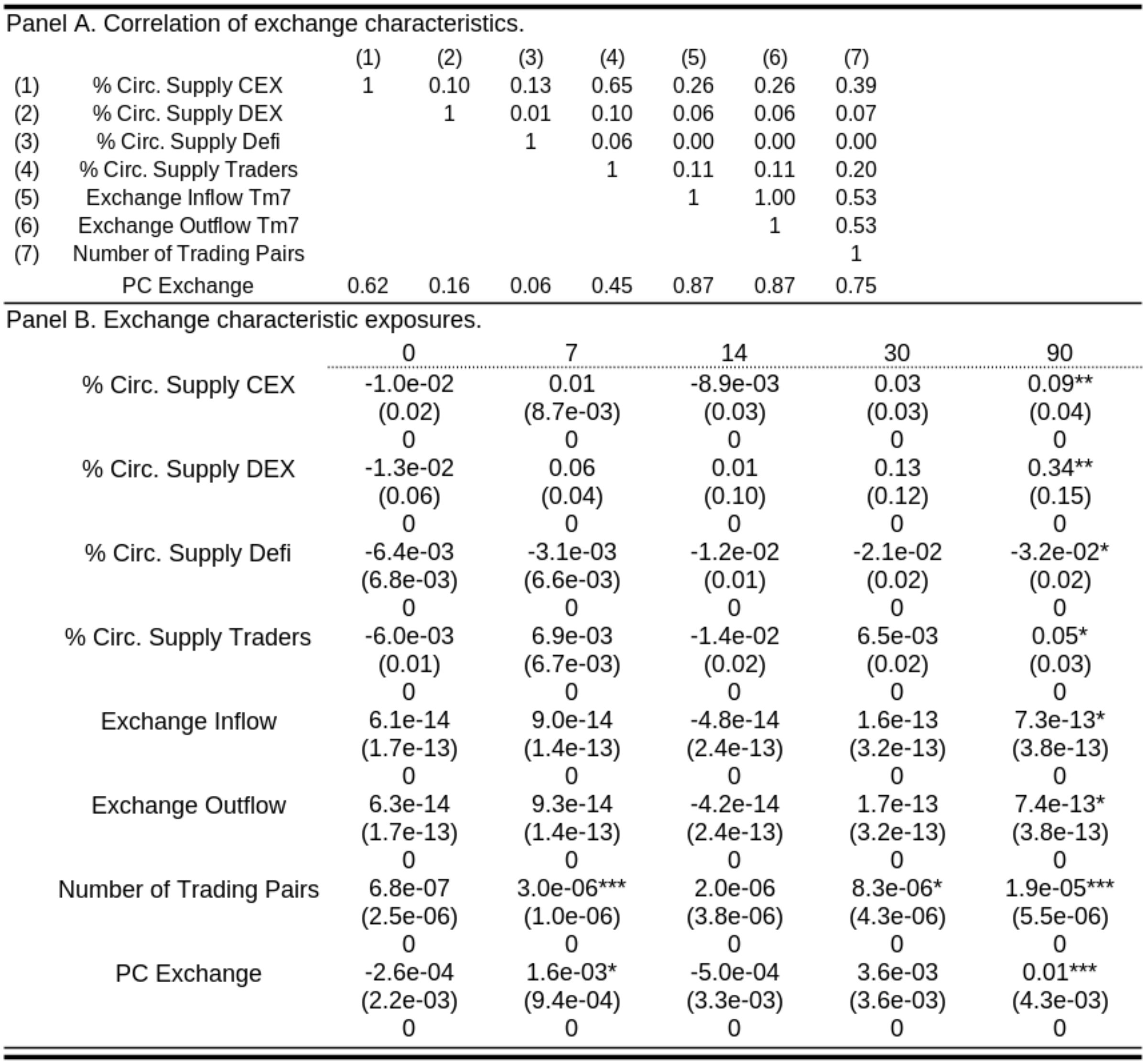
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	Tx Volume Tm7	1	-0.01	0.04	0.03	-0.01	0.93	0.72	0.01	0.02	0.00
(2)	Active Addresses Tm7		1	0.01	0.40	0.80	-0.01	0.00	0.01	-0.05	0.10
(3)	Δ Log New Addresses Tm14-Tm7			1	0.02	0.02	0.02	0.00	0.14	-0.02	0.00
(4)	New Addresses Tm7				1	0.68	0.02	0.02	0.00	-0.01	0.08
(5)	Total Addresses					1	-0.01	-0.01	0.01	-0.06	0.09
(6)	Circulation Tm7						1	0.87	0.01	0.03	0.00
(7)	Age Destroyed							1	0.00	0.03	0.00
(8)	Δ Flow Distribution								1	0.01	0.02
(9)	Δ Holders Distribution									1	0.08
(10)	% Supply in Profit										1
	PC Onchain	0.93	0.00	0.04	0.04	0.00	0.99	0.91	0.01	0.04	0.00

Panel B. Onchain characteristics exposures.

	0	7	14	30	90
Tx Volume Tm7	6.2e-15*	1.2e-15	3.3e-15	9.9e-15*	2.3e-14*
	(3.7e-15)	(1.4e-15)	(2.4e-15)	(5.2e-15)	(1.3e-14)
	0	0	0	0	0
Active Addresses Tm7	1.0e-09	2.1e-09*	4.8e-10	5.6e-09	1.2e-08**
	(2.6e-09)	(1.2e-09)	(3.7e-09)	(4.3e-09)	(5.6e-09)
	0	0	0	0	0
Δ Log New Addresses Tm14-Tm7	0.09***	0.04***	0.07	0.08**	0.11***
	(0.02)	(7.7e-03)	(0.05)	(0.03)	(0.04)
	0	0	0	0	0
New Addresses Tm7	1.5e-09	1.2e-09*	1.0e-09	3.5e-09**	7.2e-09***
	(9.7e-10)	(6.7e-10)	(1.3e-09)	(1.7e-09)	(2.5e-09)
	0	0	0	0	0
Total Addresses	3.6e-12	3.4e-10*	1.7e-10	1.0e-09	2.1e-09***
	(3.6e-10)	(1.8e-10)	(5.3e-10)	(6.3e-10)	(8.2e-10)
	0	0	0	0	0
Circulation Tm7	4.1e-15	1.7e-15	4.0e-15	1.1e-14	2.2e-14
	(3.8e-15)	(1.8e-15)	(3.3e-15)	(7.5e-15)	(1.4e-14)
	0	0	0	0	0
Age Destoryed	2.2e-15	1.8e-17	5.1e-16	3.7e-15	7.6e-15
	(3.1e-15)	(8.0e-16)	(1.9e-15)	(4.7e-15)	(9.5e-15)
	0	0	0	0	0
Δ Flow Distribution	-4.9e-05	1.6e-04	6.0e-05	5.2e-05	-3.6e-04
	(1.0e-04)	(1.0e-04)	(2.2e-04)	(2.4e-04)	(3.4e-04)
	0	0	0	0	0
Δ Holders Distribution	0.25	0.21***	0.17	0.39	0.51
	(0.20)	(0.06)	(0.12)	(0.24)	(0.39)
	0	0	0	0	0
% Supply in Profit	1.2e-03***	4.6e-04***	3.5e-06	7.9e-04	1.3e-03*
	(4.3e-04)	(1.2e-04)	(5.9e-04)	(6.5e-04)	(7.5e-04)
	0	0	0	0	0
PC Onchain	6.6e-03	1.7e-03	4.2e-03	0.01	0.03
	(5.6e-03)	(2.0e-03)	(3.9e-03)	(8.9e-03)	(0.02)
	0	0	0	0	0

This table reports the correlation matrix among Onchain Characteristics and the loadings on asset excess returns on each characteristic at various horizons. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

Figure A16. Exchange Characteristics: Correlations and Signal



Panel A. Correlation of exchange characteristics.

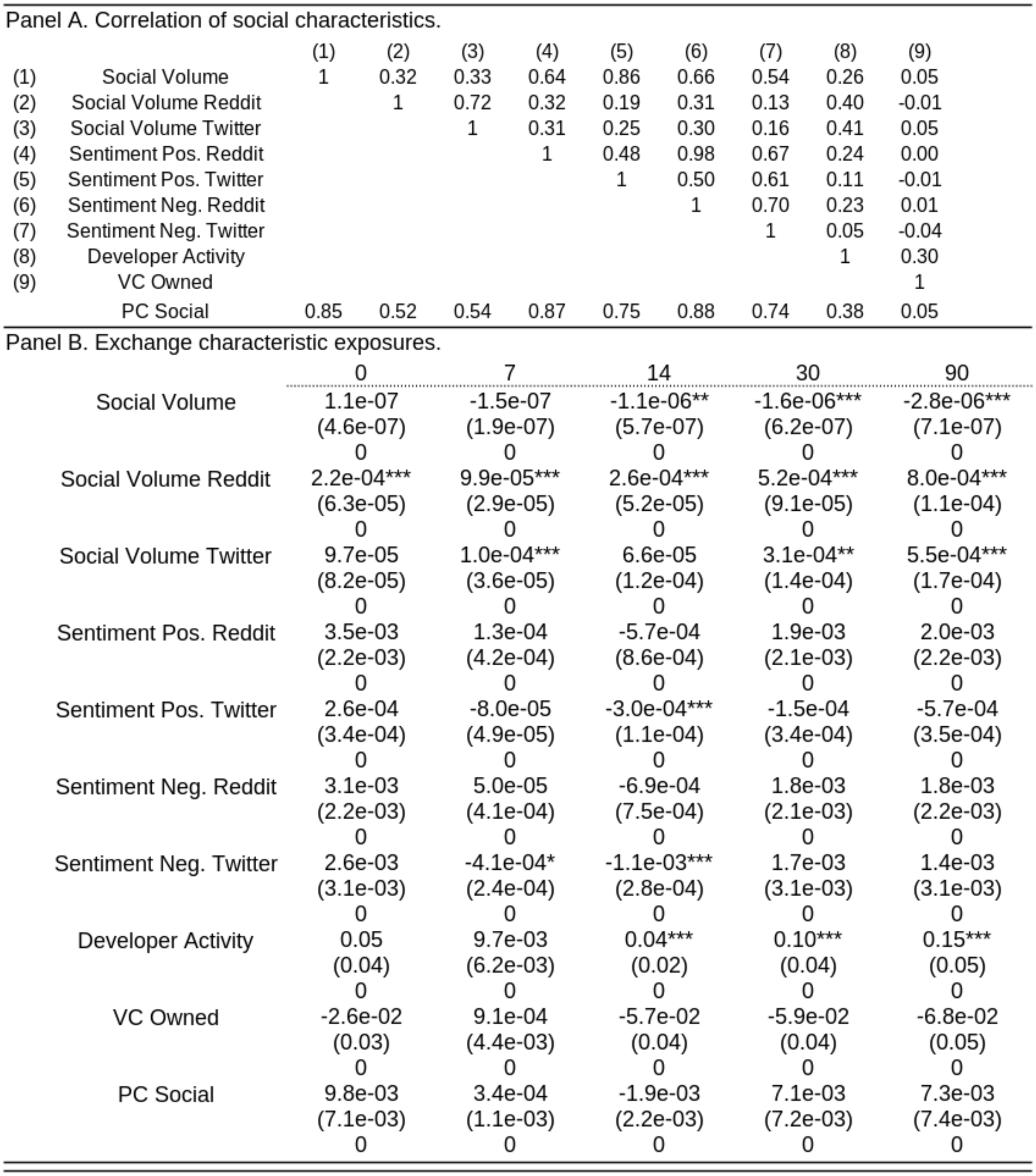
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	% Circ. Supply CEX	1	0.10	0.13	0.65	0.26	0.26	0.39
(2)	% Circ. Supply DEX		1	0.01	0.10	0.06	0.06	0.07
(3)	% Circ. Supply Defi			1	0.06	0.00	0.00	0.00
(4)	% Circ. Supply Traders				1	0.11	0.11	0.20
(5)	Exchange Inflow Tm7					1	1.00	0.53
(6)	Exchange Outflow Tm7						1	0.53
(7)	Number of Trading Pairs							1
	PC Exchange	0.62	0.16	0.06	0.45	0.87	0.87	0.75

Panel B. Exchange characteristic exposures.

	0	7	14	30	90
% Circ. Supply CEX	-1.0e-02	0.01	-8.9e-03	0.03	0.09**
	(0.02)	(8.7e-03)	(0.03)	(0.03)	(0.04)
	0	0	0	0	0
% Circ. Supply DEX	-1.3e-02	0.06	0.01	0.13	0.34**
	(0.06)	(0.04)	(0.10)	(0.12)	(0.15)
	0	0	0	0	0
% Circ. Supply Defi	-6.4e-03	-3.1e-03	-1.2e-02	-2.1e-02	-3.2e-02*
	(6.8e-03)	(6.6e-03)	(0.01)	(0.02)	(0.02)
	0	0	0	0	0
% Circ. Supply Traders	-6.0e-03	6.9e-03	-1.4e-02	6.5e-03	0.05*
	(0.01)	(6.7e-03)	(0.02)	(0.02)	(0.03)
	0	0	0	0	0
Exchange Inflow	6.1e-14	9.0e-14	-4.8e-14	1.6e-13	7.3e-13*
	(1.7e-13)	(1.4e-13)	(2.4e-13)	(3.2e-13)	(3.8e-13)
	0	0	0	0	0
Exchange Outflow	6.3e-14	9.3e-14	-4.2e-14	1.7e-13	7.4e-13*
	(1.7e-13)	(1.4e-13)	(2.4e-13)	(3.2e-13)	(3.8e-13)
	0	0	0	0	0
Number of Trading Pairs	6.8e-07	3.0e-06***	2.0e-06	8.3e-06*	1.9e-05***
	(2.5e-06)	(1.0e-06)	(3.8e-06)	(4.3e-06)	(5.5e-06)
	0	0	0	0	0
PC Exchange	-2.6e-04	1.6e-03*	-5.0e-04	3.6e-03	0.01***
	(2.2e-03)	(9.4e-04)	(3.3e-03)	(3.6e-03)	(4.3e-03)
	0	0	0	0	0

This table reports the correlation matrix among Exchange Characteristics and the loadings on asset excess returns on each characteristic at various horizons. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

FIGURE A17. Social Characteristics: Correlations and Signal.



Panel A. Correlation of social characteristics.

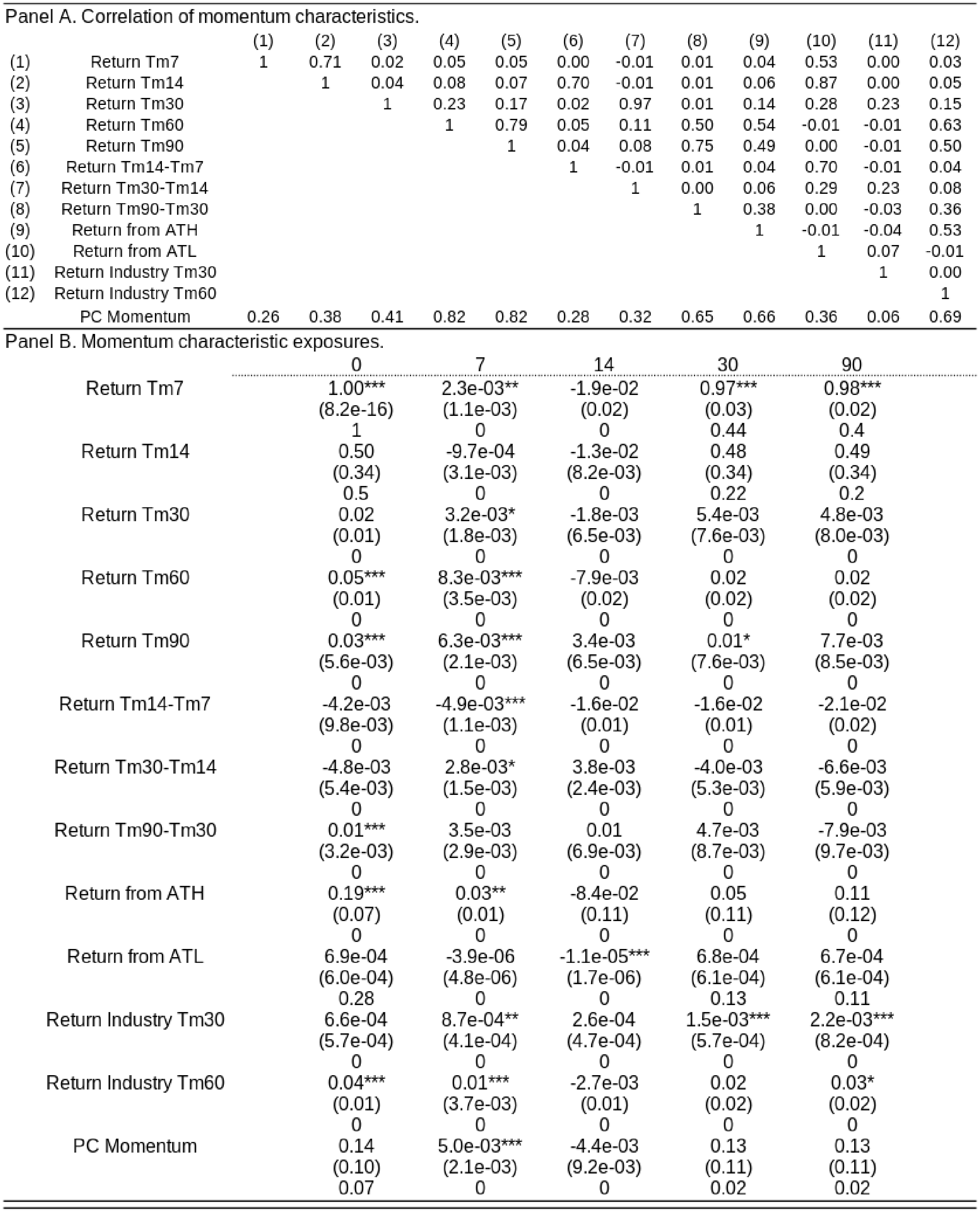
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	Social Volume	1	0.32	0.33	0.64	0.86	0.66	0.54	0.26	0.05
(2)	Social Volume Reddit		1	0.72	0.32	0.19	0.31	0.13	0.40	-0.01
(3)	Social Volume Twitter			1	0.31	0.25	0.30	0.16	0.41	0.05
(4)	Sentiment Pos. Reddit				1	0.48	0.98	0.67	0.24	0.00
(5)	Sentiment Pos. Twitter					1	0.50	0.61	0.11	-0.01
(6)	Sentiment Neg. Reddit						1	0.70	0.23	0.01
(7)	Sentiment Neg. Twitter							1	0.05	-0.04
(8)	Developer Activity								1	0.30
(9)	VC Owned									1
	PC Social	0.85	0.52	0.54	0.87	0.75	0.88	0.74	0.38	0.05

Panel B. Exchange characteristic exposures.

	0	7	14	30	90
Social Volume	1.1e-07	-1.5e-07	-1.1e-06**	-1.6e-06***	-2.8e-06***
	(4.6e-07)	(1.9e-07)	(5.7e-07)	(6.2e-07)	(7.1e-07)
	0	0	0	0	0
Social Volume Reddit	2.2e-04***	9.9e-05***	2.6e-04***	5.2e-04***	8.0e-04***
	(6.3e-05)	(2.9e-05)	(5.2e-05)	(9.1e-05)	(1.1e-04)
	0	0	0	0	0
Social Volume Twitter	9.7e-05	1.0e-04***	6.6e-05	3.1e-04**	5.5e-04***
	(8.2e-05)	(3.6e-05)	(1.2e-04)	(1.4e-04)	(1.7e-04)
	0	0	0	0	0
Sentiment Pos. Reddit	3.5e-03	1.3e-04	-5.7e-04	1.9e-03	2.0e-03
	(2.2e-03)	(4.2e-04)	(8.6e-04)	(2.1e-03)	(2.2e-03)
	0	0	0	0	0
Sentiment Pos. Twitter	2.6e-04	-8.0e-05	-3.0e-04***	-1.5e-04	-5.7e-04
	(3.4e-04)	(4.9e-05)	(1.1e-04)	(3.4e-04)	(3.5e-04)
	0	0	0	0	0
Sentiment Neg. Reddit	3.1e-03	5.0e-05	-6.9e-04	1.8e-03	1.8e-03
	(2.2e-03)	(4.1e-04)	(7.5e-04)	(2.1e-03)	(2.2e-03)
	0	0	0	0	0
Sentiment Neg. Twitter	2.6e-03	-4.1e-04*	-1.1e-03***	1.7e-03	1.4e-03
	(3.1e-03)	(2.4e-04)	(2.8e-04)	(3.1e-03)	(3.1e-03)
	0	0	0	0	0
Developer Activity	0.05	9.7e-03	0.04***	0.10***	0.15***
	(0.04)	(6.2e-03)	(0.02)	(0.04)	(0.05)
	0	0	0	0	0
VC Owned	-2.6e-02	9.1e-04	-5.7e-02	-5.9e-02	-6.8e-02
	(0.03)	(4.4e-03)	(0.04)	(0.04)	(0.05)
	0	0	0	0	0
PC Social	9.8e-03	3.4e-04	-1.9e-03	7.1e-03	7.3e-03
	(7.1e-03)	(1.1e-03)	(2.2e-03)	(7.2e-03)	(7.4e-03)
	0	0	0	0	0

This table reports the correlation matrix among Social Characteristics and the loadings on asset excess returns on each characteristic at various horizons. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

FIGURE A18. Momentum Characteristics: Correlations and Signal.



Panel A. Correlation of momentum characteristics.

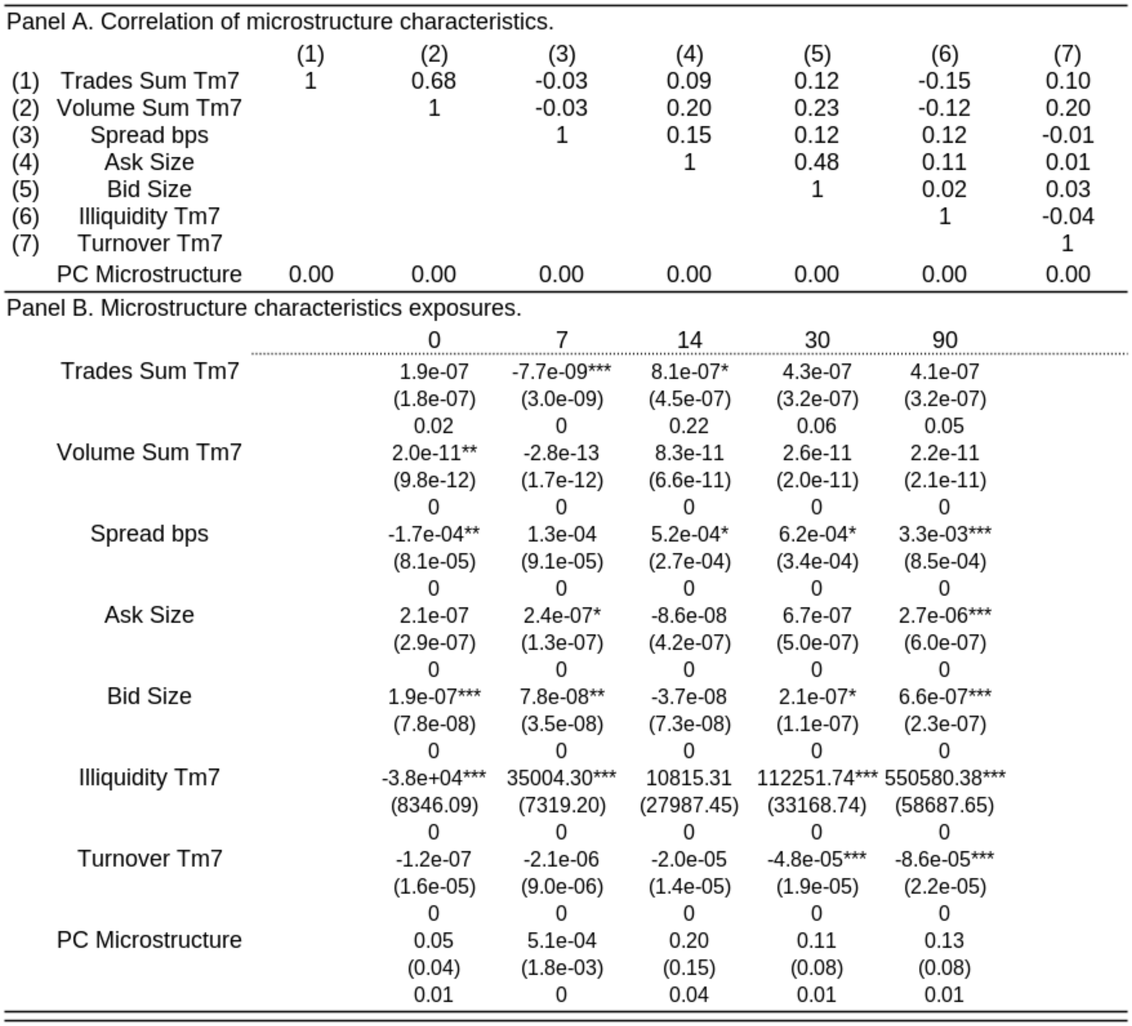
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	Return Tm7	1	0.71	0.02	0.05	0.05	0.00	-0.01	0.01	0.04	0.53	0.00	0.03
(2)	Return Tm14		1	0.04	0.08	0.07	0.70	-0.01	0.01	0.06	0.87	0.00	0.05
(3)	Return Tm30			1	0.23	0.17	0.02	0.97	0.01	0.14	0.28	0.23	0.15
(4)	Return Tm60				1	0.79	0.05	0.11	0.50	0.54	-0.01	-0.01	0.63
(5)	Return Tm90					1	0.04	0.08	0.75	0.49	0.00	-0.01	0.50
(6)	Return Tm14-Tm7						1	-0.01	0.01	0.04	0.70	-0.01	0.04
(7)	Return Tm30-Tm14							1	0.00	0.06	0.29	0.23	0.08
(8)	Return Tm90-Tm30								1	0.38	0.00	-0.03	0.36
(9)	Return from ATH									1	-0.01	-0.04	0.53
(10)	Return from ATL										1	0.07	-0.01
(11)	Return Industry Tm30											1	0.00
(12)	Return Industry Tm60												1
	PC Momentum	0.26	0.38	0.41	0.82	0.82	0.28	0.32	0.65	0.66	0.36	0.06	0.69

Panel B. Momentum characteristic exposures.

	0	7	14	30	90
Return Tm7	1.00***	2.3e-03**	-1.9e-02	0.97***	0.98***
	(8.2e-16)	(1.1e-03)	(0.02)	(0.03)	(0.02)
	1	0	0	0.44	0.4
Return Tm14	0.50	-9.7e-04	-1.3e-02	0.48	0.49
	(0.34)	(3.1e-03)	(8.2e-03)	(0.34)	(0.34)
	0.5	0	0	0.22	0.2
Return Tm30	0.02	3.2e-03*	-1.8e-03	5.4e-03	4.8e-03
	(0.01)	(1.8e-03)	(6.5e-03)	(7.6e-03)	(8.0e-03)
	0	0	0	0	0
Return Tm60	0.05***	8.3e-03***	-7.9e-03	0.02	0.02
	(0.01)	(3.5e-03)	(0.02)	(0.02)	(0.02)
	0	0	0	0	0
Return Tm90	0.03***	6.3e-03***	3.4e-03	0.01*	7.7e-03
	(5.6e-03)	(2.1e-03)	(6.5e-03)	(7.6e-03)	(8.5e-03)
	0	0	0	0	0
Return Tm14-Tm7	-4.2e-03	-4.9e-03***	-1.6e-02	-1.6e-02	-2.1e-02
	(9.8e-03)	(1.1e-03)	(0.01)	(0.01)	(0.02)
	0	0	0	0	0
Return Tm30-Tm14	-4.8e-03	2.8e-03*	3.8e-03	-4.0e-03	-6.6e-03
	(5.4e-03)	(1.5e-03)	(2.4e-03)	(5.3e-03)	(5.9e-03)
	0	0	0	0	0
Return Tm90-Tm30	0.01***	3.5e-03	0.01	4.7e-03	-7.9e-03
	(3.2e-03)	(2.9e-03)	(6.9e-03)	(8.7e-03)	(9.7e-03)
	0	0	0	0	0
Return from ATH	0.19***	0.03**	-8.4e-02	0.05	0.11
	(0.07)	(0.01)	(0.11)	(0.11)	(0.12)
	0	0	0	0	0
Return from ATL	6.9e-04	-3.9e-06	-1.1e-05***	6.8e-04	6.7e-04
	(6.0e-04)	(4.8e-06)	(1.7e-06)	(6.1e-04)	(6.1e-04)
	0.28	0	0	0.13	0.11
Return Industry Tm30	6.6e-04	8.7e-04**	2.6e-04	1.5e-03***	2.2e-03***
	(5.7e-04)	(4.1e-04)	(4.7e-04)	(5.7e-04)	(8.2e-04)
	0	0	0	0	0
Return Industry Tm60	0.04***	0.01***	-2.7e-03	0.02	0.03*
	(0.01)	(3.7e-03)	(0.01)	(0.02)	(0.02)
	0	0	0	0	0
PC Momentum	0.14	5.0e-03***	-4.4e-03	0.13	0.13
	(0.10)	(2.1e-03)	(9.2e-03)	(0.11)	(0.11)
	0.07	0	0	0.02	0.02

This table reports the correlation matrix among Momentum Characteristics and the loadings on asset excess returns on each characteristic at various horizons. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

FIGURE A19. Microstructure Characteristics: Correlations and Signal.



Panel A. Correlation of microstructure characteristics.

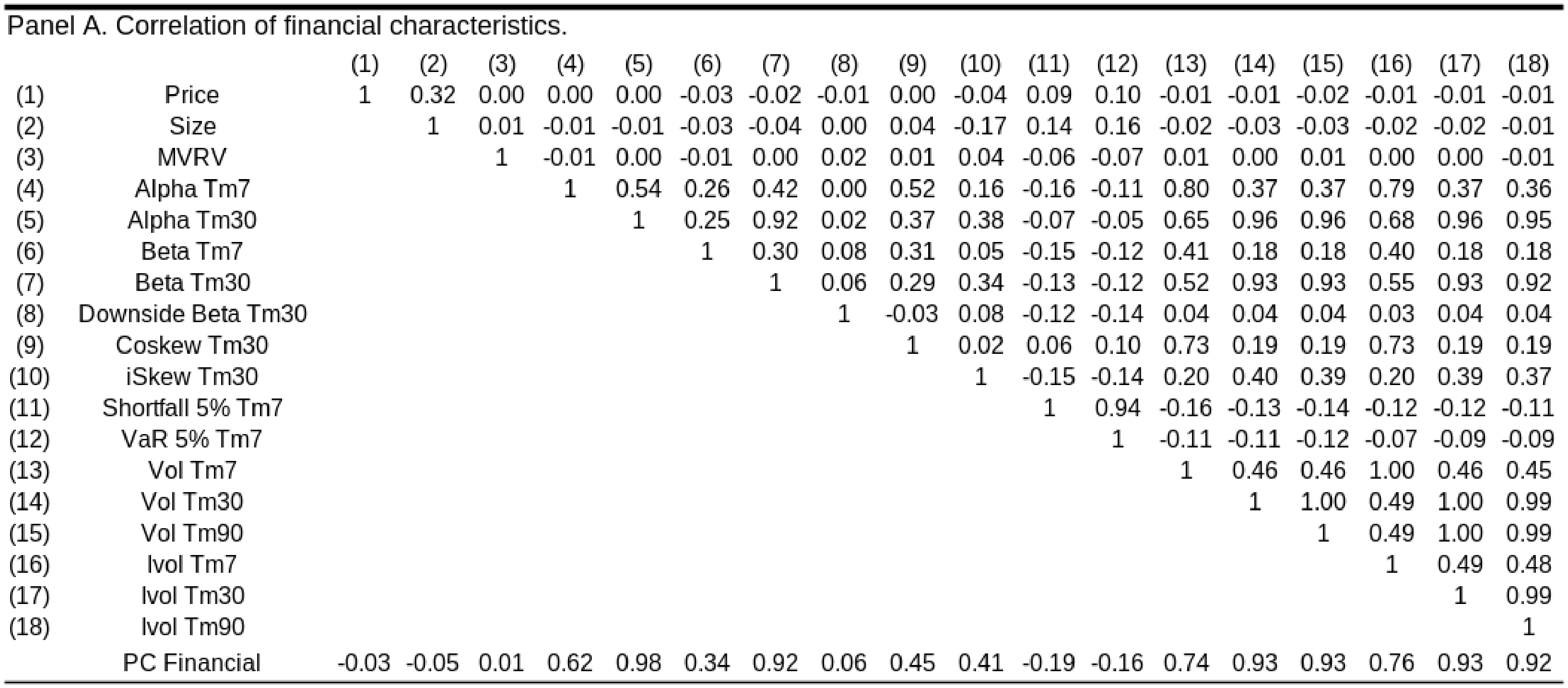
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Trades Sum Tm7	1	0.68	-0.03	0.09	0.12	-0.15	0.10
(2)	Volume Sum Tm7		1	-0.03	0.20	0.23	-0.12	0.20
(3)	Spread bps			1	0.15	0.12	0.12	-0.01
(4)	Ask Size				1	0.48	0.11	0.01
(5)	Bid Size					1	0.02	0.03
(6)	Illiquidity Tm7						1	-0.04
(7)	Turnover Tm7							1
	PC Microstructure	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B. Microstructure characteristics exposures.

	0	7	14	30	90
Trades Sum Tm7	1.9e-07	-7.7e-09***	8.1e-07*	4.3e-07	4.1e-07
	(1.8e-07)	(3.0e-09)	(4.5e-07)	(3.2e-07)	(3.2e-07)
	0.02	0	0.22	0.06	0.05
Volume Sum Tm7	2.0e-11**	-2.8e-13	8.3e-11	2.6e-11	2.2e-11
	(9.8e-12)	(1.7e-12)	(6.6e-11)	(2.0e-11)	(2.1e-11)
	0	0	0	0	0
Spread bps	-1.7e-04**	1.3e-04	5.2e-04*	6.2e-04*	3.3e-03***
	(8.1e-05)	(9.1e-05)	(2.7e-04)	(3.4e-04)	(8.5e-04)
	0	0	0	0	0
Ask Size	2.1e-07	2.4e-07*	-8.6e-08	6.7e-07	2.7e-06***
	(2.9e-07)	(1.3e-07)	(4.2e-07)	(5.0e-07)	(6.0e-07)
	0	0	0	0	0
Bid Size	1.9e-07***	7.8e-08**	-3.7e-08	2.1e-07*	6.6e-07***
	(7.8e-08)	(3.5e-08)	(7.3e-08)	(1.1e-07)	(2.3e-07)
	0	0	0	0	0
Illiquidity Tm7	-3.8e+04***	35004.30***	10815.31	112251.74***	550580.38***
	(8346.09)	(7319.20)	(27987.45)	(33168.74)	(58687.65)
	0	0	0	0	0
Turnover Tm7	-1.2e-07	-2.1e-06	-2.0e-05	-4.8e-05***	-8.6e-05***
	(1.6e-05)	(9.0e-06)	(1.4e-05)	(1.9e-05)	(2.2e-05)
	0	0	0	0	0
PC Microstructure	0.05	5.1e-04	0.20	0.11	0.13
	(0.04)	(1.8e-03)	(0.15)	(0.08)	(0.08)
	0.01	0	0.04	0.01	0.01

This table reports the correlation matrix among Microstructure Characteristics and the loadings on asset excess returns on each characteristic at various horizons. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

FIGURE A20. Financial Characteristics: Correlations.

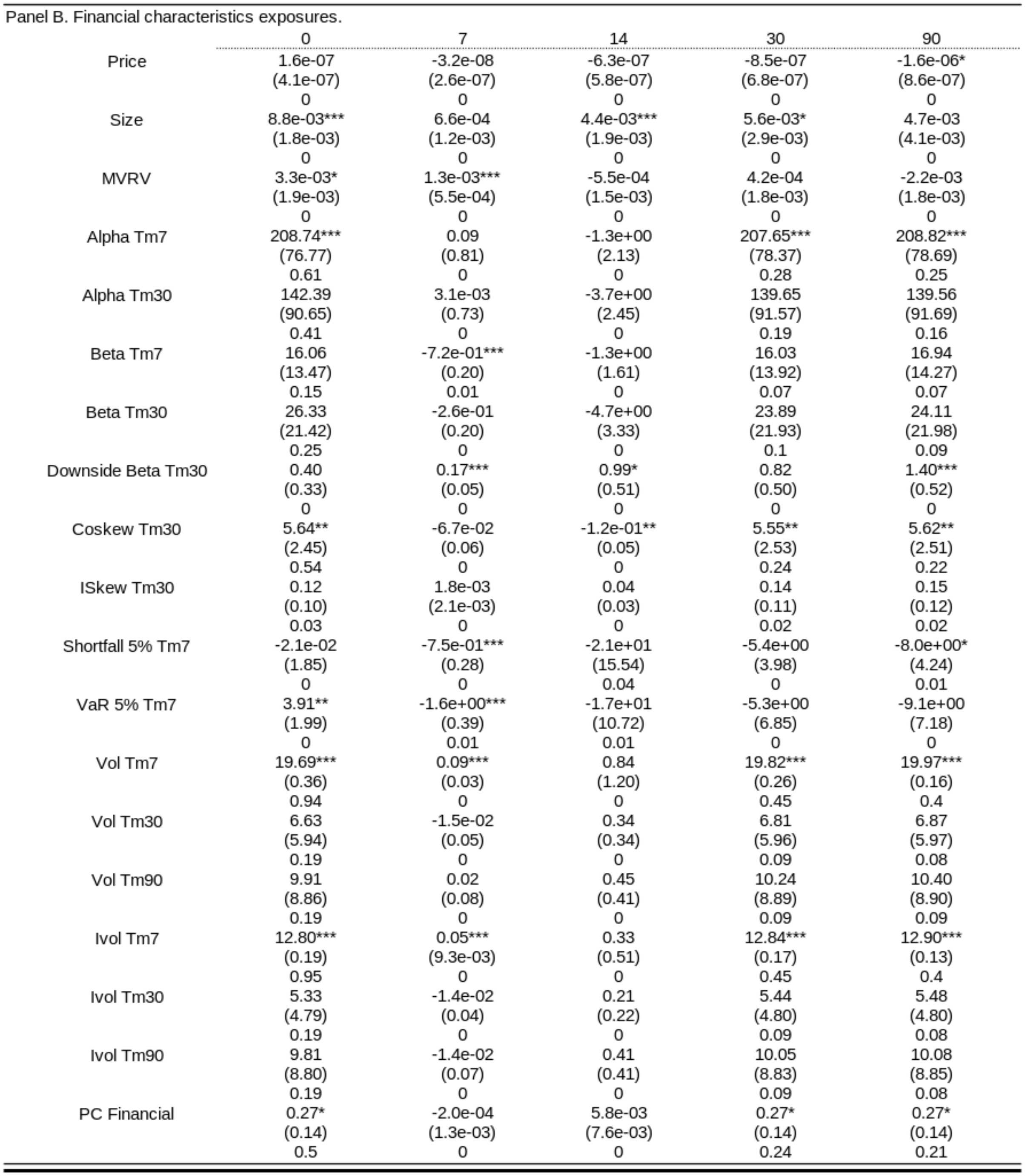


Panel A. Correlation of financial characteristics.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1)	Price	1	0.32	0.00	0.00	0.00	-0.03	-0.02	-0.01	0.00	-0.04	0.09	0.10	-0.01	-0.01	-0.02	-0.01	-0.01	-0.01
(2)	Size		1	0.01	-0.01	-0.01	-0.03	-0.04	0.00	0.04	-0.17	0.14	0.16	-0.02	-0.03	-0.03	-0.02	-0.02	-0.01
(3)	MVRV			1	-0.01	0.00	-0.01	0.00	0.02	0.01	0.04	-0.06	-0.07	0.01	0.00	0.01	0.00	0.00	-0.01
(4)	Alpha Tm7				1	0.54	0.26	0.42	0.00	0.52	0.16	-0.16	-0.11	0.80	0.37	0.37	0.79	0.37	0.36
(5)	Alpha Tm30					1	0.25	0.92	0.02	0.37	0.38	-0.07	-0.05	0.65	0.96	0.96	0.68	0.96	0.95
(6)	Beta Tm7						1	0.30	0.08	0.31	0.05	-0.15	-0.12	0.41	0.18	0.18	0.40	0.18	0.18
(7)	Beta Tm30							1	0.06	0.29	0.34	-0.13	-0.12	0.52	0.93	0.93	0.55	0.93	0.92
(8)	Downside Beta Tm30								1	-0.03	0.08	-0.12	-0.14	0.04	0.04	0.04	0.03	0.04	0.04
(9)	Coskew Tm30									1	0.02	0.06	0.10	0.73	0.19	0.19	0.73	0.19	0.19
(10)	iSkew Tm30										1	-0.15	-0.14	0.20	0.40	0.39	0.20	0.39	0.37
(11)	Shortfall 5% Tm7											1	0.94	-0.16	-0.13	-0.14	-0.12	-0.12	-0.11
(12)	VaR 5% Tm7												1	-0.11	-0.11	-0.12	-0.07	-0.09	-0.09
(13)	Vol Tm7													1	0.46	0.46	1.00	0.46	0.45
(14)	Vol Tm30														1	1.00	0.49	1.00	0.99
(15)	Vol Tm90															1	0.49	1.00	0.99
(16)	Ivol Tm7																1	0.49	0.48
(17)	Ivol Tm30																	1	0.99
(18)	Ivol Tm90																		1
	PC Financial	-0.03	-0.05	0.01	0.62	0.98	0.34	0.92	0.06	0.45	0.41	-0.19	-0.16	0.74	0.93	0.93	0.76	0.93	0.92

This table reports the correlation matrix among Financial Characteristics. Panel A reports pairwise Pearson correlation coefficients among the characteristics and the first principal component from them. The characteristics are re-scaled to be mean zero and unit variance before PCA and studying these correlations.

FIGURE A21. Financial Characteristics: Signal.

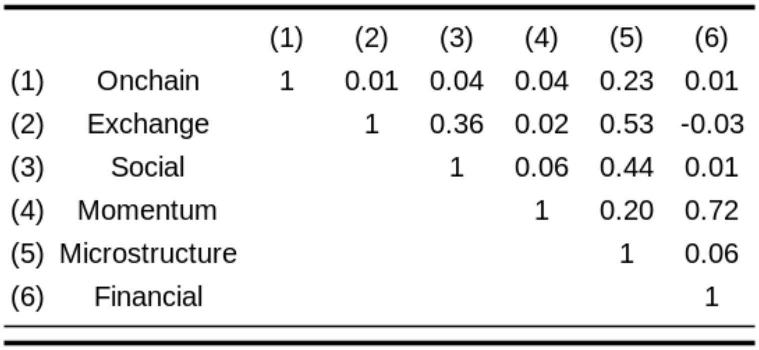


Panel B. Financial characteristics exposures.

	0	7	14	30	90
Price	1.6e-07	-3.2e-08	-6.3e-07	-8.5e-07	-1.6e-06*
	(4.1e-07)	(2.6e-07)	(5.8e-07)	(6.8e-07)	(8.6e-07)
	0	0	0	0	0
Size	8.8e-03***	6.6e-04	4.4e-03***	5.6e-03*	4.7e-03
	(1.8e-03)	(1.2e-03)	(1.9e-03)	(2.9e-03)	(4.1e-03)
	0	0	0	0	0
MVRV	3.3e-03*	1.3e-03***	-5.5e-04	4.2e-04	-2.2e-03
	(1.9e-03)	(5.5e-04)	(1.5e-03)	(1.8e-03)	(1.8e-03)
	0	0	0	0	0
Alpha Tm7	208.74***	0.09	-1.3e+00	207.65***	208.82***
	(76.77)	(0.81)	(2.13)	(78.37)	(78.69)
	0.61	0	0	0.28	0.25
Alpha Tm30	142.39	3.1e-03	-3.7e+00	139.65	139.56
	(90.65)	(0.73)	(2.45)	(91.57)	(91.69)
	0.41	0	0	0.19	0.16
Beta Tm7	16.06	-7.2e-01***	-1.3e+00	16.03	16.94
	(13.47)	(0.20)	(1.61)	(13.92)	(14.27)
	0.15	0.01	0	0.07	0.07
Beta Tm30	26.33	-2.6e-01	-4.7e+00	23.89	24.11
	(21.42)	(0.20)	(3.33)	(21.93)	(21.98)
	0.25	0	0	0.1	0.09
Downside Beta Tm30	0.40	0.17***	0.99*	0.82	1.40***
	(0.33)	(0.05)	(0.51)	(0.50)	(0.52)
	0	0	0	0	0
Coskew Tm30	5.64**	-6.7e-02	-1.2e-01**	5.55**	5.62**
	(2.45)	(0.06)	(0.05)	(2.53)	(2.51)
	0.54	0	0	0.24	0.22
ISkew Tm30	0.12	1.8e-03	0.04	0.14	0.15
	(0.10)	(2.1e-03)	(0.03)	(0.11)	(0.12)
	0.03	0	0	0.02	0.02
Shortfall 5% Tm7	-2.1e-02	-7.5e-01***	-2.1e+01	-5.4e+00	-8.0e+00*
	(1.85)	(0.28)	(15.54)	(3.98)	(4.24)
	0	0	0.04	0	0.01
VaR 5% Tm7	3.91**	-1.6e+00***	-1.7e+01	-5.3e+00	-9.1e+00
	(1.99)	(0.39)	(10.72)	(6.85)	(7.18)
	0	0.01	0.01	0	0
Vol Tm7	19.69***	0.09***	0.84	19.82***	19.97***
	(0.36)	(0.03)	(1.20)	(0.26)	(0.16)
	0.94	0	0	0.45	0.4
Vol Tm30	6.63	-1.5e-02	0.34	6.81	6.87
	(5.94)	(0.05)	(0.34)	(5.96)	(5.97)
	0.19	0	0	0.09	0.08
Vol Tm90	9.91	0.02	0.45	10.24	10.40
	(8.86)	(0.08)	(0.41)	(8.89)	(8.90)
	0.19	0	0	0.09	0.09
Ivol Tm7	12.80***	0.05***	0.33	12.84***	12.90***
	(0.19)	(9.3e-03)	(0.51)	(0.17)	(0.13)
	0.95	0	0	0.45	0.4
Ivol Tm30	5.33	-1.4e-02	0.21	5.44	5.48
	(4.79)	(0.04)	(0.22)	(4.80)	(4.80)
	0.19	0	0	0.09	0.08
Ivol Tm90	9.81	-1.4e-02	0.41	10.05	10.08
	(8.80)	(0.07)	(0.41)	(8.83)	(8.85)
	0.19	0	0	0.09	0.08
PC Financial	0.27*	-2.0e-04	5.8e-03	0.27*	0.27*
	(0.14)	(1.3e-03)	(7.6e-03)	(0.14)	(0.14)
	0.5	0	0	0.24	0.21

This table reports the loadings on asset excess returns on each characteristic at various horizons. Panel B reports the coefficient (with 1, 2, and 3 stars for significant at the 10%, 5%, and 1% levels, respectively), standard error, and R^2 for univariate panel regressions of asset excess weekly returns at 0, 7, 14, 30, and 90 days ahead on each of the characteristics and a constant. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. There are 22,678 observations.

Figure A22. Principal Components of Characteristics: Correlations.



		(1)	(2)	(3)	(4)	(5)	(6)
(1)	Onchain	1	0.01	0.04	0.04	0.23	0.01
(2)	Exchange		1	0.36	0.02	0.53	-0.03
(3)	Social			1	0.06	0.44	0.01
(4)	Momentum				1	0.20	0.72
(5)	Microstructure					1	0.06
(6)	Financial						1

This table reports the correlation matrix among the first principal components of all groupings of asset characteristics, i.e. all pairwise Pearson correlation coefficients.

FIGURE A23. Characteristic Signal by Year.

	2018	2019	2020	2021	2022	All
Onchain						
Tx Volume Tm7	0.025	0.021	0.015	0.007	0.007	0.005
Active Addresses Tm7	0.038	0.036	0.040	0.019	0.016	0.020
Δ Log New Addresses Tm14-Tm7	0.067	0.025	0.077	0.042	0.042	0.031
New Addresses Tm7	0.023	0.034	0.032	0.010	0.007	0.009
Total Addresses	0.047	0.036	0.031	0.014	0.013	0.021
Circulation Tm7	0.010	0.030	0.025	0.006	0.006	0.004
Age Destroyed	0.008	0.020	0.019	0.005	0.005	0.003
Δ Flow Distribution	0.115	0.063	0.029	0.029	0.057	0.044
Δ Holders Distribution	0.026	0.010	0.027	0.013	0.007	0.006
% Supply in Profit	0.104	0.054	0.089	0.075	0.082	0.080
Exchange						
% Circ. Supply CEX	0.028	0.026	0.059	0.037	0.038	0.034
% Circ. Supply DEX	0.050	0.011	0.049	0.019	0.019	0.018
% Circ. Supply Defi	0.050	0.068	0.017	0.002	0.002	0.001
% Circ. Supply Traders	0.031	0.030	0.056	0.044	0.048	0.041
Exchange Inflow Tm7	0.028	0.032	0.036	0.005	0.009	0.004
Exchange Outflow Tm7	0.020	0.028	0.036	0.004	0.008	0.004
Number of Trading Pairs	0.000	0.011	0.027	0.012	0.016	0.018
Social						
Social Volume	0.025	0.030	0.040	0.018	0.010	0.008
Social Volume Reddit	0.046	0.073	0.060	0.046	0.048	0.041
Social Volume Twitter	0.059	0.049	0.062	0.043	0.052	0.041
Sentiment Pos. Reddit	0.033	0.035	0.037	0.013	0.014	0.012
Sentiment Pos. Twitter	0.043	0.027	0.035	0.011	0.011	0.008
Sentiment Neg. Reddit	0.033	0.025	0.044	0.013	0.014	0.012
Sentiment Neg. Twitter	0.027	0.028	0.033	0.013	0.011	0.008
Developer Activity Tm7	0.042	0.051	0.078	0.040	0.039	0.037
VC Owned	0.008	0.006	0.010	0.003	0.003	0.002

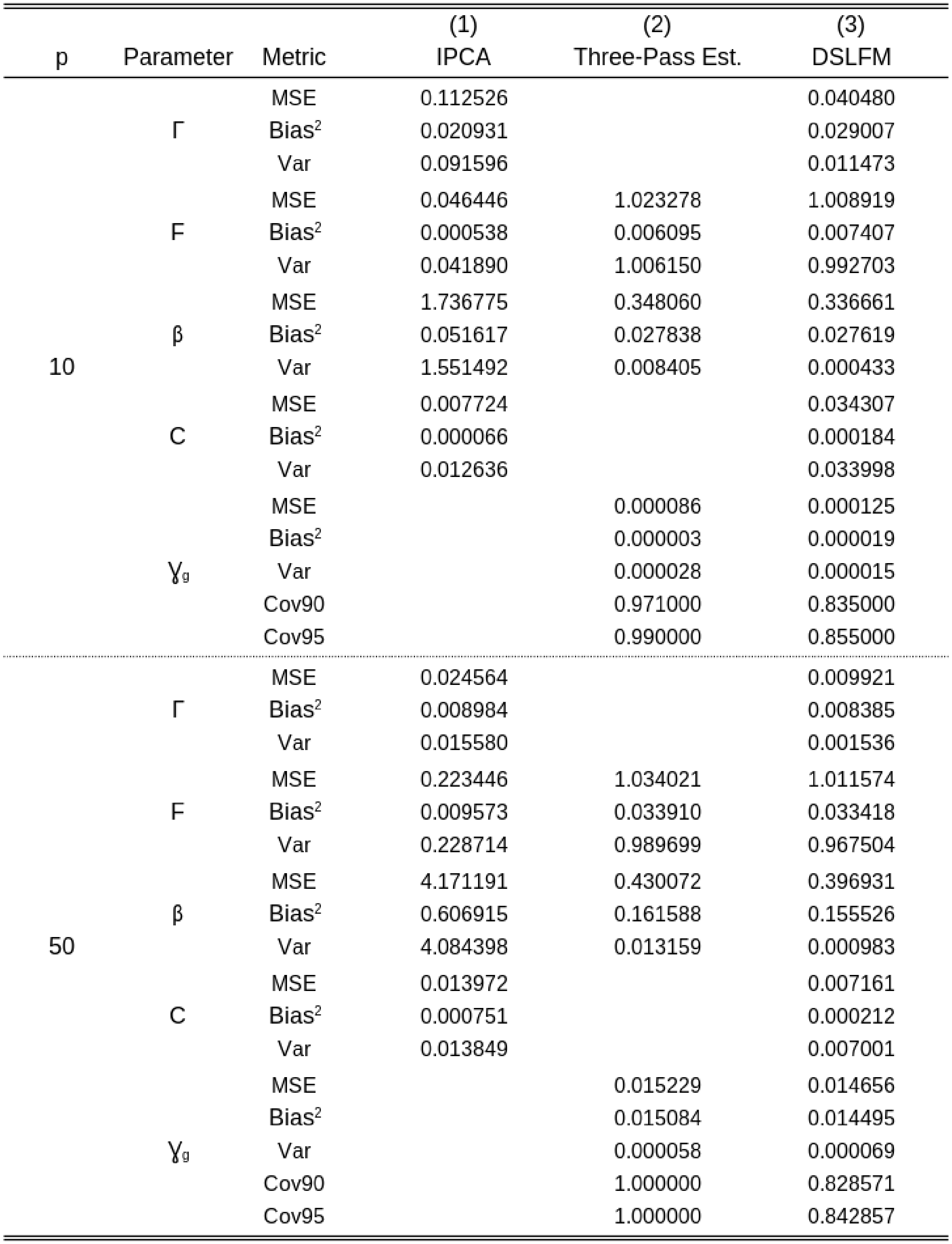
This table reports, by year and overall, the pairwise mutual information between all weekly panel characteristics and asset excess returns seven days ahead.

FIGURE A24. Characteristic Signal by Year (Continued).

Momentum						
Return Tm7	0.035	0.045	0.097	0.037	0.002	0.002
Return Tm14	0.079	0.046	0.085	0.040	0.003	0.004
Return Tm30	0.010	0.064	0.085	0.045	0.003	0.007
Return Tm60	0.030	0.050	0.082	0.038	0.003	0.013
Return Tm90	0.064	0.060	0.066	0.037	0.004	0.018
Return Tm14-Tm7	0.042	0.043	0.104	0.033	0.002	0.002
Return Tm30-Tm14	0.044	0.054	0.076	0.049	0.001	0.004
Return Tm90-Tm30	0.051	0.038	0.057	0.026	0.004	0.009
Return from ATH	0.072	0.043	0.100	0.067	0.054	0.056
Return from ATL	0.023	0.037	0.046	0.021	0.007	0.004
Return Industry Tm30	0.042	0.077	0.090	0.046	0.003	0.004
Return Industry Tm60	0.076	0.075	0.086	0.043	0.002	0.010
Microstructure						
Trades Sum Tm7	0.040	0.020	0.037	0.012	0.011	0.010
Volume Sum Tm7	0.028	0.025	0.026	0.016	0.015	0.012
Spread bps	0.034	0.031	0.018	0.019	0.001	0.011
Ask Size	0.000	0.126	0.151	0.193	0.048	0.101
Bid Size	0.000	0.074	0.072	0.036	0.014	0.018
Illiquidity Tm7	0.033	0.044	0.037	0.025	0.025	0.020
Turnover Tm7	0.023	0.025	0.033	0.007	0.008	0.005
Financial						
Price	0.035	0.033	0.026	0.011	0.014	0.012
Size	0.049	0.052	0.060	0.041	0.044	0.036
MVRV	0.059	0.043	0.059	0.006	0.008	0.007
Alpha Tm7	0.073	0.054	0.097	0.042	0.015	0.011
Alpha Tm30	0.097	0.078	0.085	0.050	0.003	0.003
Beta Tm7	0.070	0.061	0.068	0.031	0.027	0.017
Beta Tm30	0.072	0.059	0.044	0.033	0.011	0.011
Downside Beta Tm30	0.090	0.070	0.070	0.042	0.040	0.032
Coskew Tm30	0.054	0.059	0.061	0.033	0.012	0.008
ISkew Tm30	0.047	0.053	0.055	0.030	0.032	0.025
Shortfall 5% Tm7	0.047	0.049	0.060	0.071	0.034	0.035
VaR 5% Tm7	0.051	0.051	0.062	0.072	0.047	0.046
Vol Tm7	0.049	0.040	0.055	0.052	0.002	0.001
Vol Tm30	0.083	0.051	0.082	0.064	0.003	0.002
Vol Tm90	0.065	0.058	0.068	0.062	0.002	0.003
Ivol Tm7	0.052	0.050	0.053	0.053	0.002	0.002
Ivol Tm30	0.083	0.052	0.081	0.065	0.002	0.002
Ivol Tm90	0.063	0.056	0.082	0.058	0.002	0.002
T*N	410	449	671	4359	6894	

This table reports, by year and overall, the pairwise mutual information between all weekly panel characteristics and asset excess returns seven days ahead.

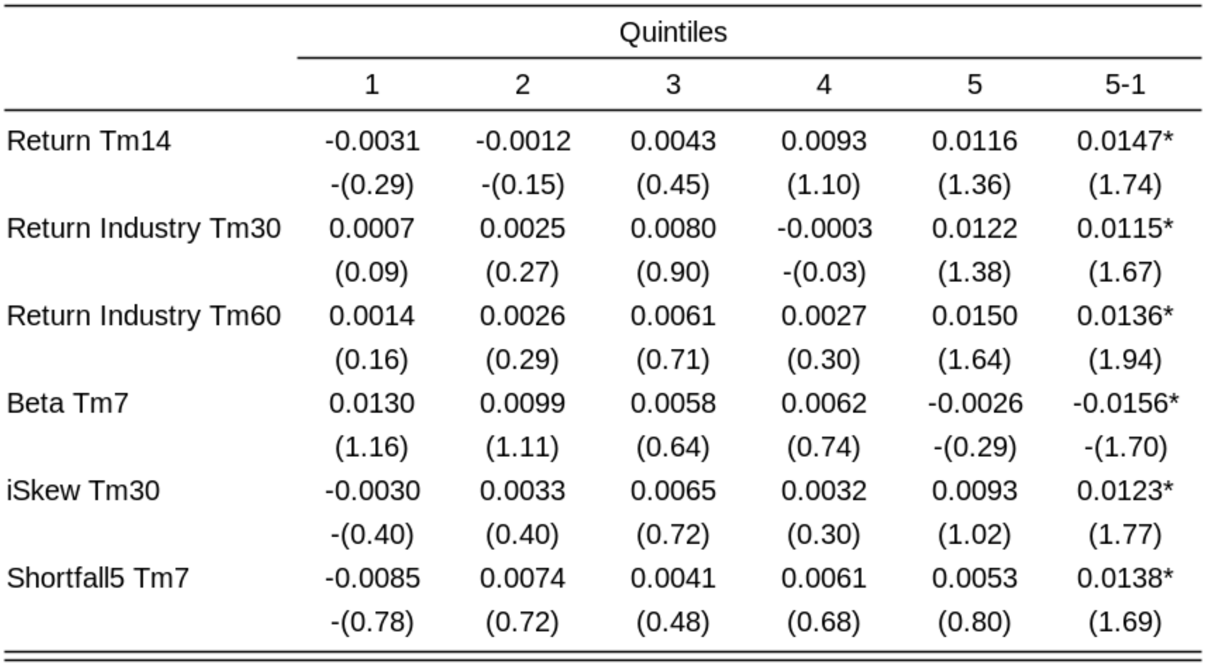
FIGURE A25. Monte Carlo Simulations.



			(1)	(2)	(3)
p	Parameter	Metric	IPCA	Three-Pass Est.	DSLFM
10	Γ	MSE	0.112526		0.040480
		Bias^2	0.020931		0.029007
		Var	0.091596		0.011473
	F	MSE	0.046446	1.023278	1.008919
		Bias^2	0.000538	0.006095	0.007407
		Var	0.041890	1.006150	0.992703
	β	MSE	1.736775	0.348060	0.336661
		Bias^2	0.051617	0.027838	0.027619
		Var	1.551492	0.008405	0.000433
	C	MSE	0.007724		0.034307
		Bias^2	0.000066		0.000184
		Var	0.012636		0.033998
	γ_g	MSE		0.000086	0.000125
		Bias^2		0.000003	0.000019
		Var		0.000028	0.000015
		Cov90		0.971000	0.835000
		Cov95		0.990000	0.855000
50	Γ	MSE	0.024564		0.009921
		Bias^2	0.008984		0.008385
		Var	0.015580		0.001536
	F	MSE	0.223446	1.034021	1.011574
		Bias^2	0.009573	0.033910	0.033418
		Var	0.228714	0.989699	0.967504
	β	MSE	4.171191	0.430072	0.396931
		Bias^2	0.606915	0.161588	0.155526
		Var	4.084398	0.013159	0.000983
	C	MSE	0.013972		0.007161
		Bias^2	0.000751		0.000212
		Var	0.013849		0.007001
	γ_g	MSE		0.015229	0.014656
		Bias^2		0.015084	0.014495
		Var		0.000058	0.000069
		Cov90		1.000000	0.828571
		Cov95		1.000000	0.842857

This table reports Monte Carlo simulations, $S = 200$, for IPCA, Three-Pass Estimators of Giglio and Xiu (2021), and the DSLFM—columns 1, 2, and 3, respectively—for target parameters: latent loadings Γ_β, latent-factors F, average factor loadings $\bar{\beta}$, latent matrix C, and observable factor risk premium γ_g. The true data-generating process has three factors, $N = 500$, $T = 100$, $p \in \{10, 50\}$, and $s = p/10$. The following metrics are reported: mean-squared error (MSE), bias squared (Bias^2), variance (Var), and 90% and 95% coverage probabilities (Cov90 and Cov95).

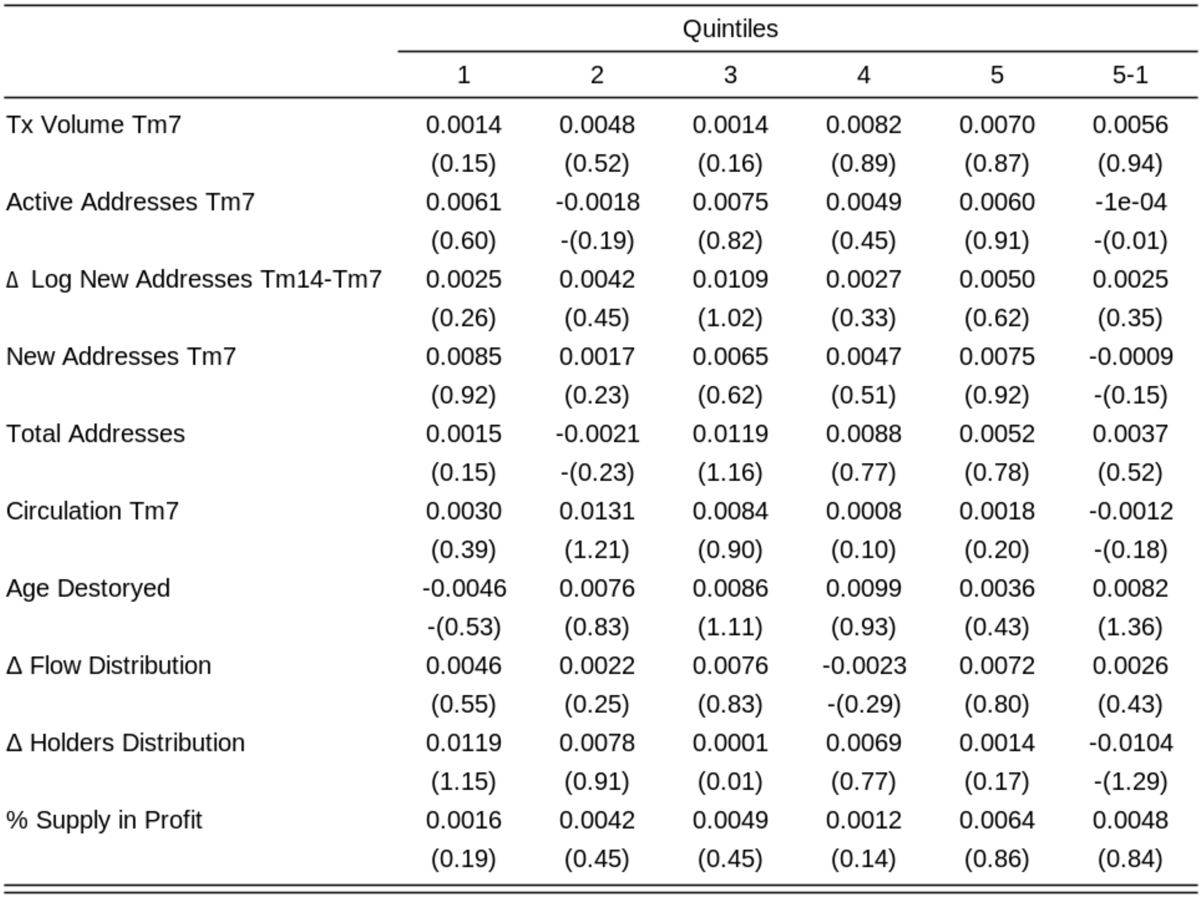
FIGURE A26. Univariate Factor Returns: Statistically Significant Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Return Tm14	-0.0031	-0.0012	0.0043	0.0093	0.0116	0.0147*
	-(0.29)	-(0.15)	(0.45)	(1.10)	(1.36)	(1.74)
Return Industry Tm30	0.0007	0.0025	0.0080	-0.0003	0.0122	0.0115*
	(0.09)	(0.27)	(0.90)	-(0.03)	(1.38)	(1.67)
Return Industry Tm60	0.0014	0.0026	0.0061	0.0027	0.0150	0.0136*
	(0.16)	(0.29)	(0.71)	(0.30)	(1.64)	(1.94)
Beta Tm7	0.0130	0.0099	0.0058	0.0062	-0.0026	-0.0156*
	(1.16)	(1.11)	(0.64)	(0.74)	-(0.29)	-(1.70)
iSkew Tm30	-0.0030	0.0033	0.0065	0.0032	0.0093	0.0123*
	-(0.40)	(0.40)	(0.72)	(0.30)	(1.02)	(1.77)
Shortfall5 Tm7	-0.0085	0.0074	0.0041	0.0061	0.0053	0.0138*
	-(0.78)	(0.72)	(0.48)	(0.68)	(0.80)	(1.69)

This table reports the mean quintile portfolio returns (and t-statistics) for characteristics with significant zero-investment strategies. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

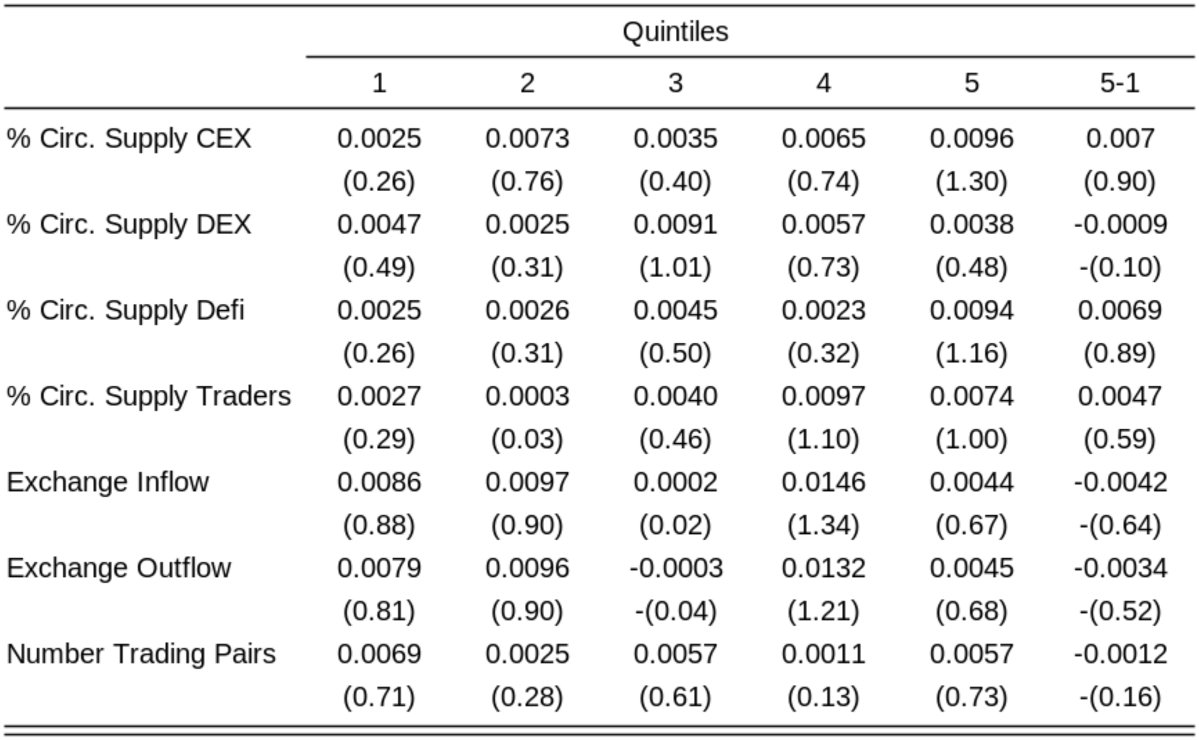
Figure A27. Univariate Factor Returns: Onchain Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Tx Volume Tm7	0.0014	0.0048	0.0014	0.0082	0.0070	0.0056
	(0.15)	(0.52)	(0.16)	(0.89)	(0.87)	(0.94)
Active Addresses Tm7	0.0061	-0.0018	0.0075	0.0049	0.0060	-1e-04
	(0.60)	-(0.19)	(0.82)	(0.45)	(0.91)	-(0.01)
Δ Log New Addresses Tm14-Tm7	0.0025	0.0042	0.0109	0.0027	0.0050	0.0025
	(0.26)	(0.45)	(1.02)	(0.33)	(0.62)	(0.35)
New Addresses Tm7	0.0085	0.0017	0.0065	0.0047	0.0075	-0.0009
	(0.92)	(0.23)	(0.62)	(0.51)	(0.92)	-(0.15)
Total Addresses	0.0015	-0.0021	0.0119	0.0088	0.0052	0.0037
	(0.15)	-(0.23)	(1.16)	(0.77)	(0.78)	(0.52)
Circulation Tm7	0.0030	0.0131	0.0084	0.0008	0.0018	-0.0012
	(0.39)	(1.21)	(0.90)	(0.10)	(0.20)	-(0.18)
Age Destoryed	-0.0046	0.0076	0.0086	0.0099	0.0036	0.0082
	-(0.53)	(0.83)	(1.11)	(0.93)	(0.43)	(1.36)
Δ Flow Distribution	0.0046	0.0022	0.0076	-0.0023	0.0072	0.0026
	(0.55)	(0.25)	(0.83)	-(0.29)	(0.80)	(0.43)
Δ Holders Distribution	0.0119	0.0078	0.0001	0.0069	0.0014	-0.0104
	(1.15)	(0.91)	(0.01)	(0.77)	(0.17)	-(1.29)
% Supply in Profit	0.0016	0.0042	0.0049	0.0012	0.0064	0.0048
	(0.19)	(0.45)	(0.45)	(0.14)	(0.86)	(0.84)

This table reports the mean quintile portfolio returns (and t-statistics) for onchain characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

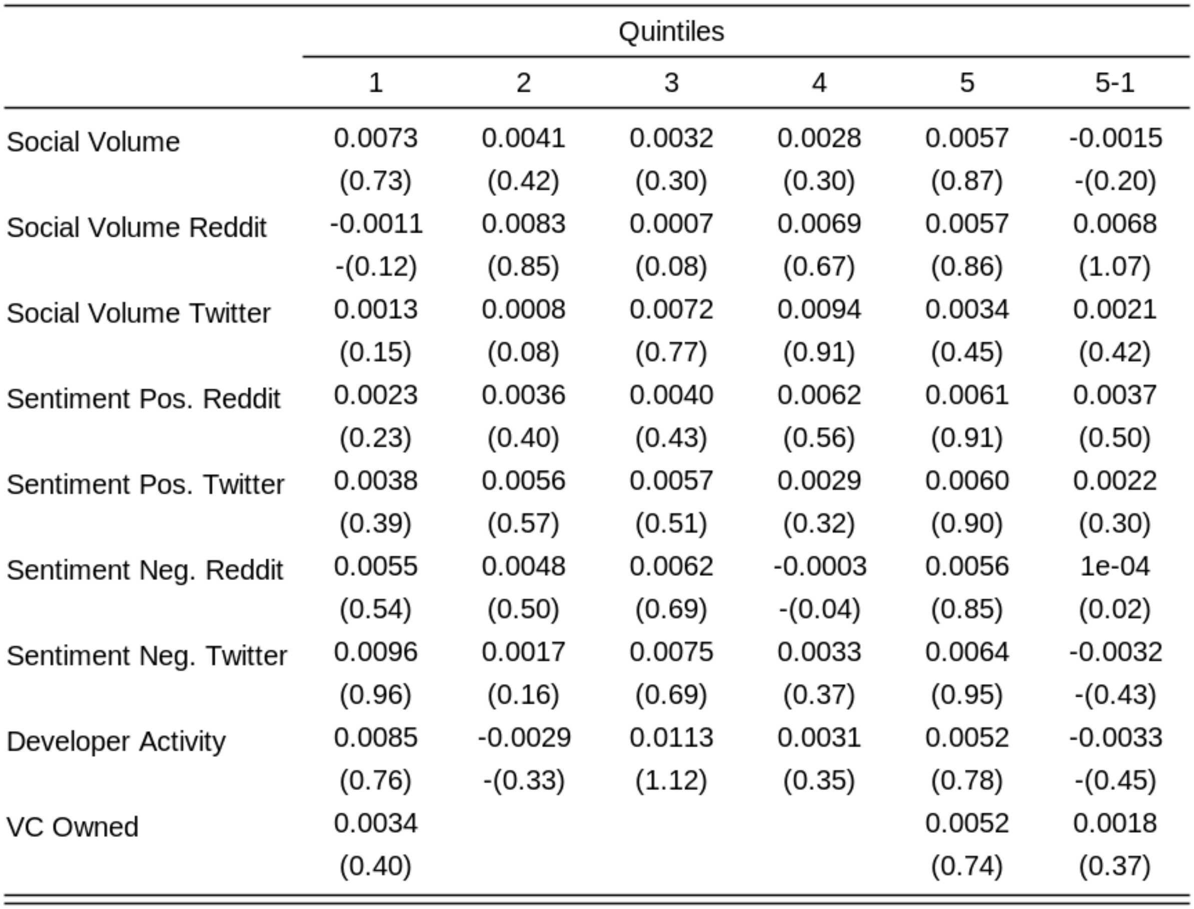
FIGURE A28. Univariate Factor Returns: Exchange Strategies.



	Quintiles					
	1	2	3	4	5	5-1
% Circ. Supply CEX	0.0025	0.0073	0.0035	0.0065	0.0096	0.007
	(0.26)	(0.76)	(0.40)	(0.74)	(1.30)	(0.90)
% Circ. Supply DEX	0.0047	0.0025	0.0091	0.0057	0.0038	-0.0009
	(0.49)	(0.31)	(1.01)	(0.73)	(0.48)	-(0.10)
% Circ. Supply Defi	0.0025	0.0026	0.0045	0.0023	0.0094	0.0069
	(0.26)	(0.31)	(0.50)	(0.32)	(1.16)	(0.89)
% Circ. Supply Traders	0.0027	0.0003	0.0040	0.0097	0.0074	0.0047
	(0.29)	(0.03)	(0.46)	(1.10)	(1.00)	(0.59)
Exchange Inflow	0.0086	0.0097	0.0002	0.0146	0.0044	-0.0042
	(0.88)	(0.90)	(0.02)	(1.34)	(0.67)	-(0.64)
Exchange Outflow	0.0079	0.0096	-0.0003	0.0132	0.0045	-0.0034
	(0.81)	(0.90)	-(0.04)	(1.21)	(0.68)	-(0.52)
Number Trading Pairs	0.0069	0.0025	0.0057	0.0011	0.0057	-0.0012
	(0.71)	(0.28)	(0.61)	(0.13)	(0.73)	-(0.16)

This table reports the mean quintile sorted portfolio returns (and t-statistics) for exchange characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

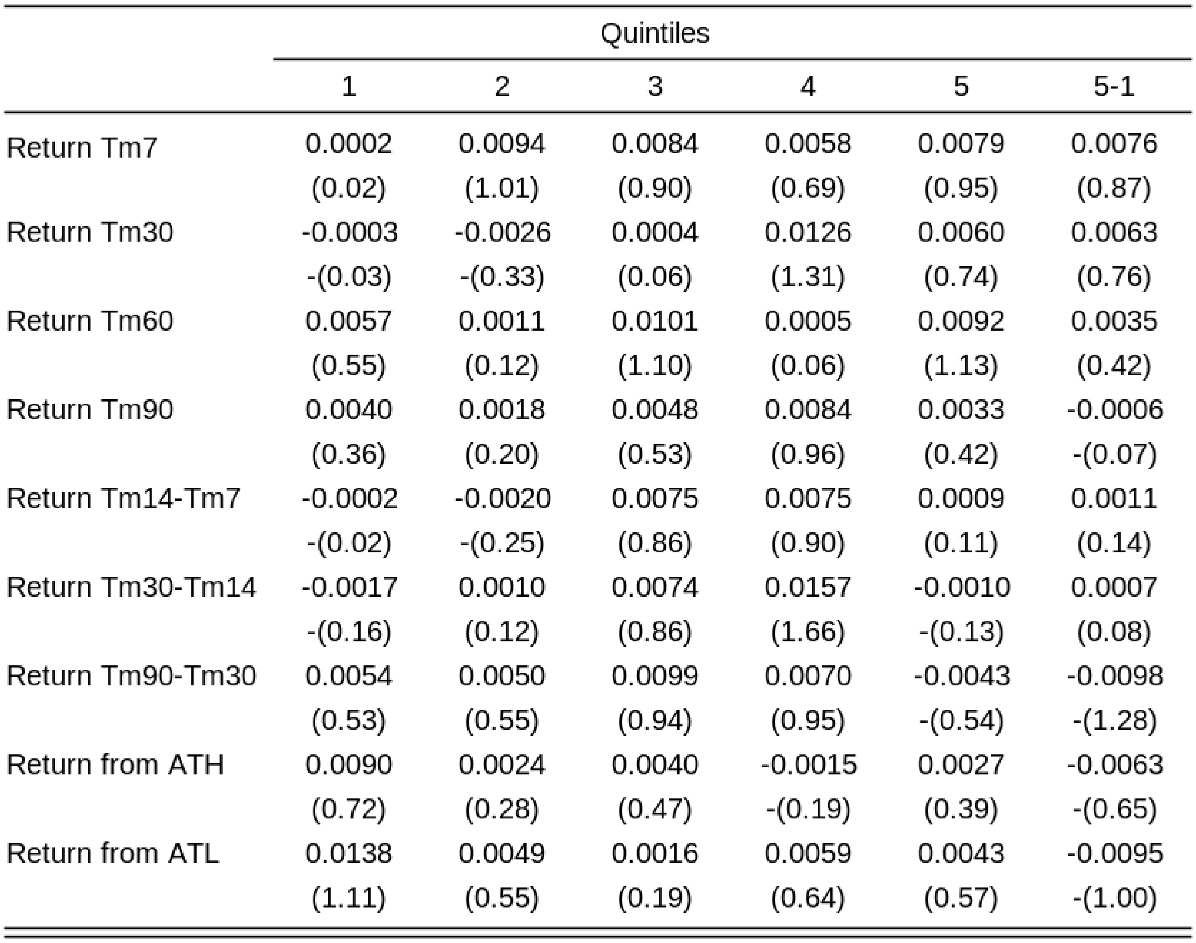
Figure A29. Univariate Factor Returns: Social Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Social Volume	0.0073	0.0041	0.0032	0.0028	0.0057	-0.0015
	(0.73)	(0.42)	(0.30)	(0.30)	(0.87)	-(0.20)
Social Volume Reddit	-0.0011	0.0083	0.0007	0.0069	0.0057	0.0068
	-(0.12)	(0.85)	(0.08)	(0.67)	(0.86)	(1.07)
Social Volume Twitter	0.0013	0.0008	0.0072	0.0094	0.0034	0.0021
	(0.15)	(0.08)	(0.77)	(0.91)	(0.45)	(0.42)
Sentiment Pos. Reddit	0.0023	0.0036	0.0040	0.0062	0.0061	0.0037
	(0.23)	(0.40)	(0.43)	(0.56)	(0.91)	(0.50)
Sentiment Pos. Twitter	0.0038	0.0056	0.0057	0.0029	0.0060	0.0022
	(0.39)	(0.57)	(0.51)	(0.32)	(0.90)	(0.30)
Sentiment Neg. Reddit	0.0055	0.0048	0.0062	-0.0003	0.0056	1e-04
	(0.54)	(0.50)	(0.69)	-(0.04)	(0.85)	(0.02)
Sentiment Neg. Twitter	0.0096	0.0017	0.0075	0.0033	0.0064	-0.0032
	(0.96)	(0.16)	(0.69)	(0.37)	(0.95)	-(0.43)
Developer Activity	0.0085	-0.0029	0.0113	0.0031	0.0052	-0.0033
	(0.76)	-(0.33)	(1.12)	(0.35)	(0.78)	-(0.45)
VC Owned	0.0034				0.0052	0.0018
	(0.40)				(0.74)	(0.37)

This table reports the mean quintile portfolio returns (and t-statistics) for social characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

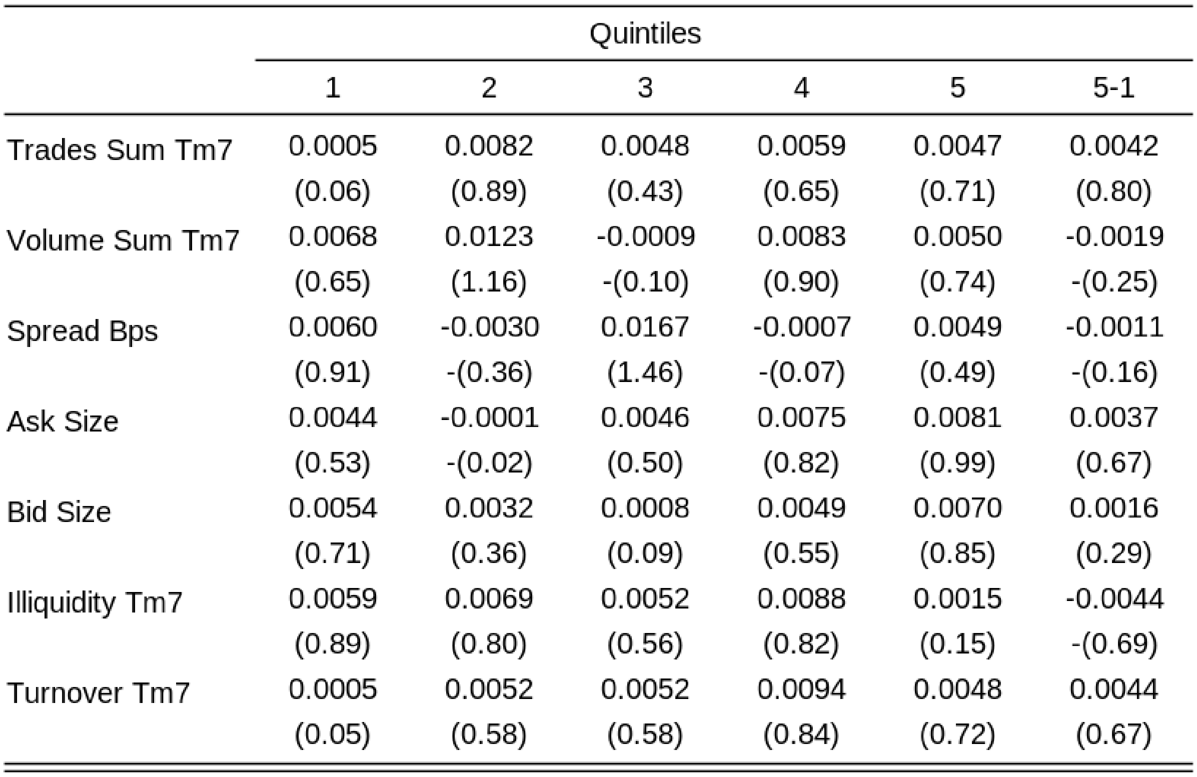
FIGURE A30. Univariate Factor Returns: Momentum Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Return Tm7	0.0002	0.0094	0.0084	0.0058	0.0079	0.0076
	(0.02)	(1.01)	(0.90)	(0.69)	(0.95)	(0.87)
Return Tm30	-0.0003	-0.0026	0.0004	0.0126	0.0060	0.0063
	-(0.03)	-(0.33)	(0.06)	(1.31)	(0.74)	(0.76)
Return Tm60	0.0057	0.0011	0.0101	0.0005	0.0092	0.0035
	(0.55)	(0.12)	(1.10)	(0.06)	(1.13)	(0.42)
Return Tm90	0.0040	0.0018	0.0048	0.0084	0.0033	-0.0006
	(0.36)	(0.20)	(0.53)	(0.96)	(0.42)	-(0.07)
Return Tm14-Tm7	-0.0002	-0.0020	0.0075	0.0075	0.0009	0.0011
	-(0.02)	-(0.25)	(0.86)	(0.90)	(0.11)	(0.14)
Return Tm30-Tm14	-0.0017	0.0010	0.0074	0.0157	-0.0010	0.0007
	-(0.16)	(0.12)	(0.86)	(1.66)	-(0.13)	(0.08)
Return Tm90-Tm30	0.0054	0.0050	0.0099	0.0070	-0.0043	-0.0098
	(0.53)	(0.55)	(0.94)	(0.95)	-(0.54)	-(1.28)
Return from ATH	0.0090	0.0024	0.0040	-0.0015	0.0027	-0.0063
	(0.72)	(0.28)	(0.47)	-(0.19)	(0.39)	-(0.65)
Return from ATL	0.0138	0.0049	0.0016	0.0059	0.0043	-0.0095
	(1.11)	(0.55)	(0.19)	(0.64)	(0.57)	-(1.00)

This table reports the mean quintile portfolio returns (and t-statistics) for momentum characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

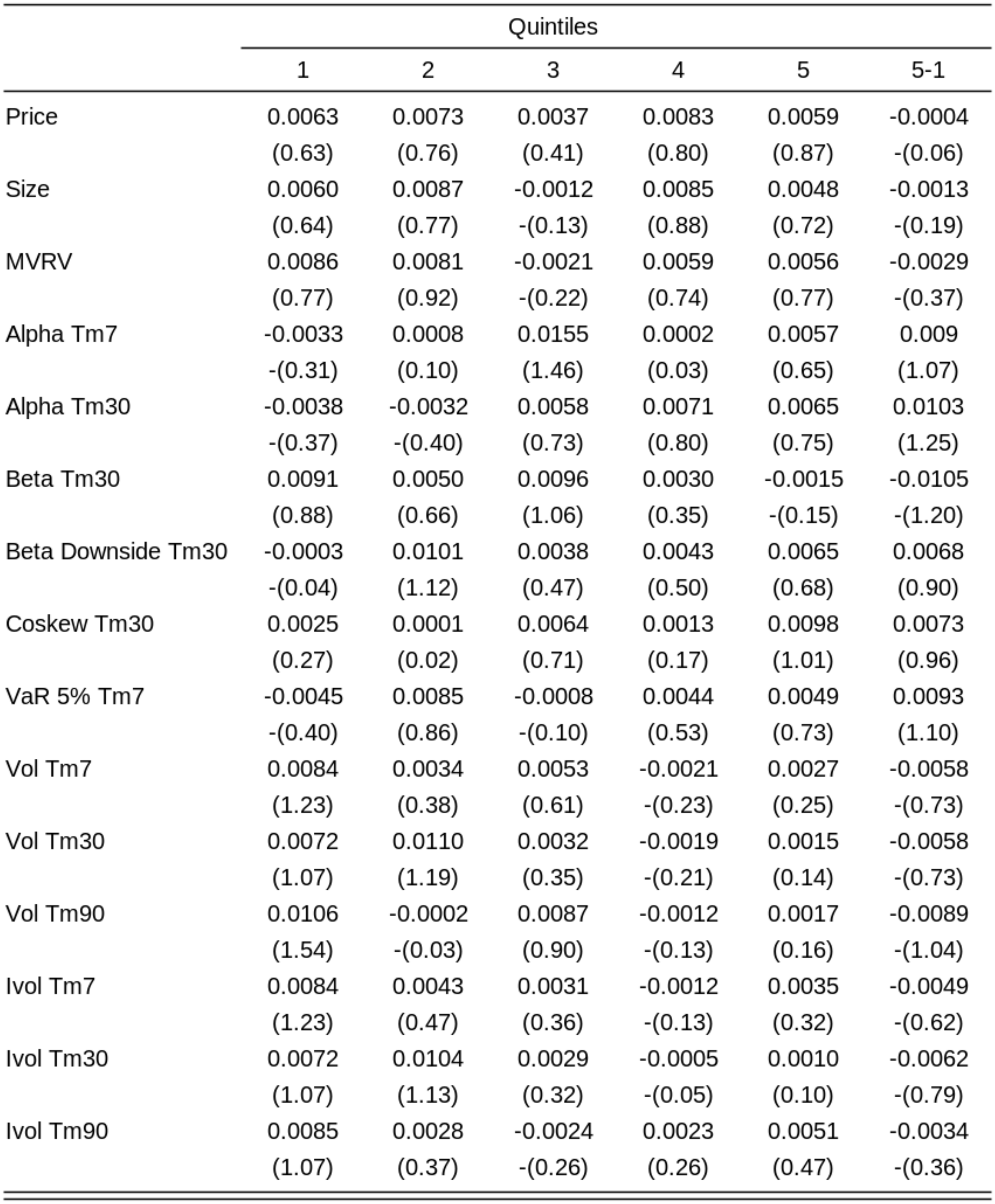
FIGURE A31. Univariate Factor Returns: Microstructure Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Trades Sum Tm7	0.0005	0.0082	0.0048	0.0059	0.0047	0.0042
	(0.06)	(0.89)	(0.43)	(0.65)	(0.71)	(0.80)
Volume Sum Tm7	0.0068	0.0123	-0.0009	0.0083	0.0050	-0.0019
	(0.65)	(1.16)	-(0.10)	(0.90)	(0.74)	-(0.25)
Spread Bps	0.0060	-0.0030	0.0167	-0.0007	0.0049	-0.0011
	(0.91)	-(0.36)	(1.46)	-(0.07)	(0.49)	-(0.16)
Ask Size	0.0044	-0.0001	0.0046	0.0075	0.0081	0.0037
	(0.53)	-(0.02)	(0.50)	(0.82)	(0.99)	(0.67)
Bid Size	0.0054	0.0032	0.0008	0.0049	0.0070	0.0016
	(0.71)	(0.36)	(0.09)	(0.55)	(0.85)	(0.29)
Illiquidity Tm7	0.0059	0.0069	0.0052	0.0088	0.0015	-0.0044
	(0.89)	(0.80)	(0.56)	(0.82)	(0.15)	-(0.69)
Turnover Tm7	0.0005	0.0052	0.0052	0.0094	0.0048	0.0044
	(0.05)	(0.58)	(0.58)	(0.84)	(0.72)	(0.67)

This table reports the mean quintile portfolio returns (and t-statistics) for microstructure characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

FIGURE A32. Univariate Factor Returns: Financial Strategies.



	Quintiles					
	1	2	3	4	5	5-1
Price	0.0063	0.0073	0.0037	0.0083	0.0059	-0.0004
	(0.63)	(0.76)	(0.41)	(0.80)	(0.87)	-(0.06)
Size	0.0060	0.0087	-0.0012	0.0085	0.0048	-0.0013
	(0.64)	(0.77)	-(0.13)	(0.88)	(0.72)	-(0.19)
MVRV	0.0086	0.0081	-0.0021	0.0059	0.0056	-0.0029
	(0.77)	(0.92)	-(0.22)	(0.74)	(0.77)	-(0.37)
Alpha Tm7	-0.0033	0.0008	0.0155	0.0002	0.0057	0.009
	-(0.31)	(0.10)	(1.46)	(0.03)	(0.65)	(1.07)
Alpha Tm30	-0.0038	-0.0032	0.0058	0.0071	0.0065	0.0103
	-(0.37)	-(0.40)	(0.73)	(0.80)	(0.75)	(1.25)
Beta Tm30	0.0091	0.0050	0.0096	0.0030	-0.0015	-0.0105
	(0.88)	(0.66)	(1.06)	(0.35)	-(0.15)	-(1.20)
Beta Downside Tm30	-0.0003	0.0101	0.0038	0.0043	0.0065	0.0068
	-(0.04)	(1.12)	(0.47)	(0.50)	(0.68)	(0.90)
Coskew Tm30	0.0025	0.0001	0.0064	0.0013	0.0098	0.0073
	(0.27)	(0.02)	(0.71)	(0.17)	(1.01)	(0.96)
VaR 5% Tm7	-0.0045	0.0085	-0.0008	0.0044	0.0049	0.0093
	-(0.40)	(0.86)	-(0.10)	(0.53)	(0.73)	(1.10)
Vol Tm7	0.0084	0.0034	0.0053	-0.0021	0.0027	-0.0058
	(1.23)	(0.38)	(0.61)	-(0.23)	(0.25)	-(0.73)
Vol Tm30	0.0072	0.0110	0.0032	-0.0019	0.0015	-0.0058
	(1.07)	(1.19)	(0.35)	-(0.21)	(0.14)	-(0.73)
Vol Tm90	0.0106	-0.0002	0.0087	-0.0012	0.0017	-0.0089
	(1.54)	-(0.03)	(0.90)	-(0.13)	(0.16)	-(1.04)
Ivol Tm7	0.0084	0.0043	0.0031	-0.0012	0.0035	-0.0049
	(1.23)	(0.47)	(0.36)	-(0.13)	(0.32)	-(0.62)
Ivol Tm30	0.0072	0.0104	0.0029	-0.0005	0.0010	-0.0062
	(1.07)	(1.13)	(0.32)	-(0.05)	(0.10)	-(0.79)
Ivol Tm90	0.0085	0.0028	-0.0024	0.0023	0.0051	-0.0034
	(1.07)	(0.37)	-(0.26)	(0.26)	(0.47)	-(0.36)

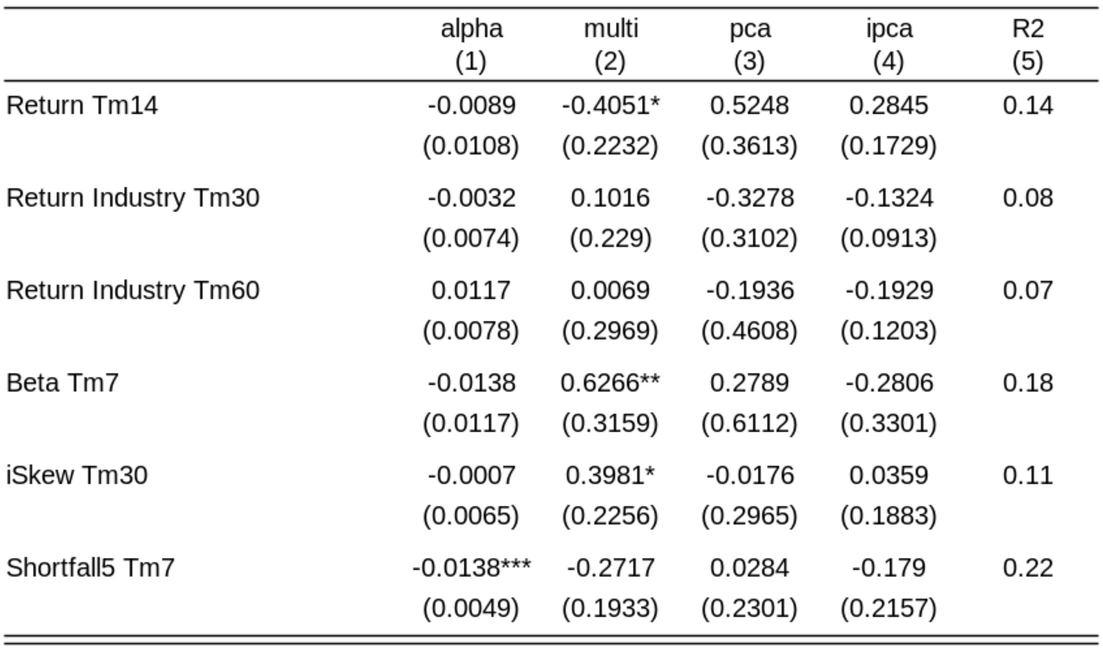
This table reports the mean quintile portfolio returns (and t-statistics) for financial characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

FIGURE A33. Low Dimensional Factor Model Out-of-Sample Returns: Multivariate, PCA, & IPCA.

Model	# Factors		Quintiles					5-1						
		Pred. R2	1	2	3	4	5	TS Avg	Sharpe	Sortino	Turnover	MDD	Alpha	Beta
Multi	1	< 0	-0.0127	-0.0121	-0.003	0.005	-0.0019	0.0107	1.31	2.74	0.06	-0.20	0.0103	-0.0441
			(-0.58)	(-0.66)	(-0.2)	(0.3)	(-0.09)	(0.93)					(0.0125)	(0.1544)
	2	< 0	-0.0101	-0.016	0.0029	-0.0073	0.0002	0.0103	1.72	3.87	0.06	-0.13	0.0106	-0.1145
			(-0.49)	(-0.87)	(0.21)	(-0.41)	(0.01)	(1.22)					(0.009)	(0.1106)
	3	< 0	-0.0129	-0.0128	-0.0027	0.0054	-0.002	0.011	1.44	2.94	0.06	-0.20	0.0099	-0.0614
			(-0.6)	(-0.7)	(-0.18)	(0.32)	(-0.1)	(1.02)					(0.0116)	(0.1424)
~ Fama	3	< 0	-0.0116	-0.0109	-0.0043	0.0032	-0.0063	0.0054	0.65	1.22	0.06	-0.20	0.0061	0.0662
			(-0.54)	(-0.62)	(-0.3)	(0.18)	(-0.28)	(0.46)					(0.0125)	(0.1544)
PCA	1	< 0	-0.0102	-0.0033	-0.0005	-0.0033	-0.0044	0.0058	0.78	1.41	0.06	-0.19	0.0053	0.0796
			(-0.47)	(-0.2)	(-0.03)	(-0.2)	(-0.2)	(0.55)					(0.0113)	(0.1394)
	2	< 0	-0.0054	-0.0147	-0.0007	-0.0076	-0.0045	0.0009	0.15	0.26	0.17	-0.13	0.0018	-0.0498
			(-0.26)	(-0.97)	(-0.05)	(-0.46)	(-0.23)	(0.11)					(0.0088)	(0.1088)
	3	< 0	-0.0108	-0.0123	-0.002	-0.0003	-0.0085	0.0023	0.45	1.01	0.21	-0.10	0.0037	0.0212
			(-0.52)	(-0.79)	(-0.13)	(-0.02)	(-0.4)	(0.32)					(0.0076)	(0.0939)
	4	< 0	-0.01	-0.0114	0.0013	-0.001	-0.0037	0.0063	1.34	3.18	0.25	-0.11	0.0059	0.0539
			(-0.54)	(-0.76)	(0.08)	(-0.06)	(-0.19)	(0.94)					(0.0072)	(0.0889)
	5	< 0	-0.017	-0.0111	-0.0031	0.0064	-0.0102	0.0068	1.24	2.45	0.29	-0.09	0.0065	0.0367
			(-0.85)	(-0.74)	(-0.21)	(0.35)	(-0.51)	(0.88)					(0.0084)	(0.1038)
IPCA	1	0.0002	-0.0188	-0.0087	-0.015	0.0003	0.0096	0.0284***	4.07	11.96	0.44	-0.08	0.0276***	-0.0234
			(-1.05)	(-0.53)	(-0.8)	(0.02)	(0.52)	(2.88)					(0.0107)	(0.1312)
	2	0.0014	-0.02	-0.0125	-0.0099	-0.0064	0.009	0.029**	3.46	11.22	0.42	-0.09	0.0274**	-0.1177
			(-1.02)	(-0.62)	(-0.62)	(-0.38)	(0.49)	(2.44)					(0.0126)	(0.1546)
	3	0.0018	-0.0206	-0.0145	-0.0086	0.0003	0.0103	0.0309**	3.59	11.13	0.41	-0.09	0.0301**	-0.3056**
			(-1.01)	(-0.73)	(-0.49)	(0.02)	(0.63)	(2.54)					(0.0121)	(0.1487)
	4	0.0004	-0.0227	-0.0139	-0.0113	0.0062	0.0013	0.024**	2.86	7.06	0.39	-0.09	0.0226*	-0.2846**
			(-1.09)	(-0.61)	(-0.76)	(0.33)	(0.09)	(2.02)					(0.0117)	(0.1446)
	5	-0.0014	-0.0139	-0.0172	-0.005	0.0002	0.0011	0.015	2.07	4.57	0.4	-0.13	0.0134	-0.1527
			(-0.72)	(-0.92)	(-0.3)	(0.01)	(0.06)	(1.47)					(0.0106)	(0.13)

This table reports—for multivariate factor models, PCA, and IPCA—the predictive R^2, the mean quintile portfolio returns, and portfolio statistics for the 5-1 strategy for July-December 2022, inclusive. For each quintile, the mean returns are the time-series averages of weekly value-weighted portfolio excess returns sorted on each model's predicted returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio from each model; for which, we report the time-series average weekly value-weighted excess return, annualized Sharpe Ratio, annualized Sortino, weekly turnover, maximum drawdown, and alpha and beta to the CMKT return. *t*-stats are reported below each strategy's point estimates where *, **, and *** denote significance at the 10%, 5%, and 1% levels. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. For the multivariate factor model with 1, 2, and 3 factors, the selected characteristics are, respectively: size; illiquidity and size; and, size, 30 day momentum, adn 90 day volatility.

FIGURE A34. Univariate Factor Returns: Alpha and Loadings on Factor Model Strategies.



	alpha (1)	multi (2)	pca (3)	ipca (4)	R2 (5)
Return Tm14	-0.0089	-0.4051*	0.5248	0.2845	0.14
	(0.0108)	(0.2232)	(0.3613)	(0.1729)	
Return Industry Tm30	-0.0032	0.1016	-0.3278	-0.1324	0.08
	(0.0074)	(0.229)	(0.3102)	(0.0913)	
Return Industry Tm60	0.0117	0.0069	-0.1936	-0.1929	0.07
	(0.0078)	(0.2969)	(0.4608)	(0.1203)	
Beta Tm7	-0.0138	0.6266**	0.2789	-0.2806	0.18
	(0.0117)	(0.3159)	(0.6112)	(0.3301)	
iSkew Tm30	-0.0007	0.3981*	-0.0176	0.0359	0.11
	(0.0065)	(0.2256)	(0.2965)	(0.1883)	
Shortfall5 Tm7	-0.0138***	-0.2717	0.0284	-0.179	0.22
	(0.0049)	(0.1933)	(0.2301)	(0.2157)	

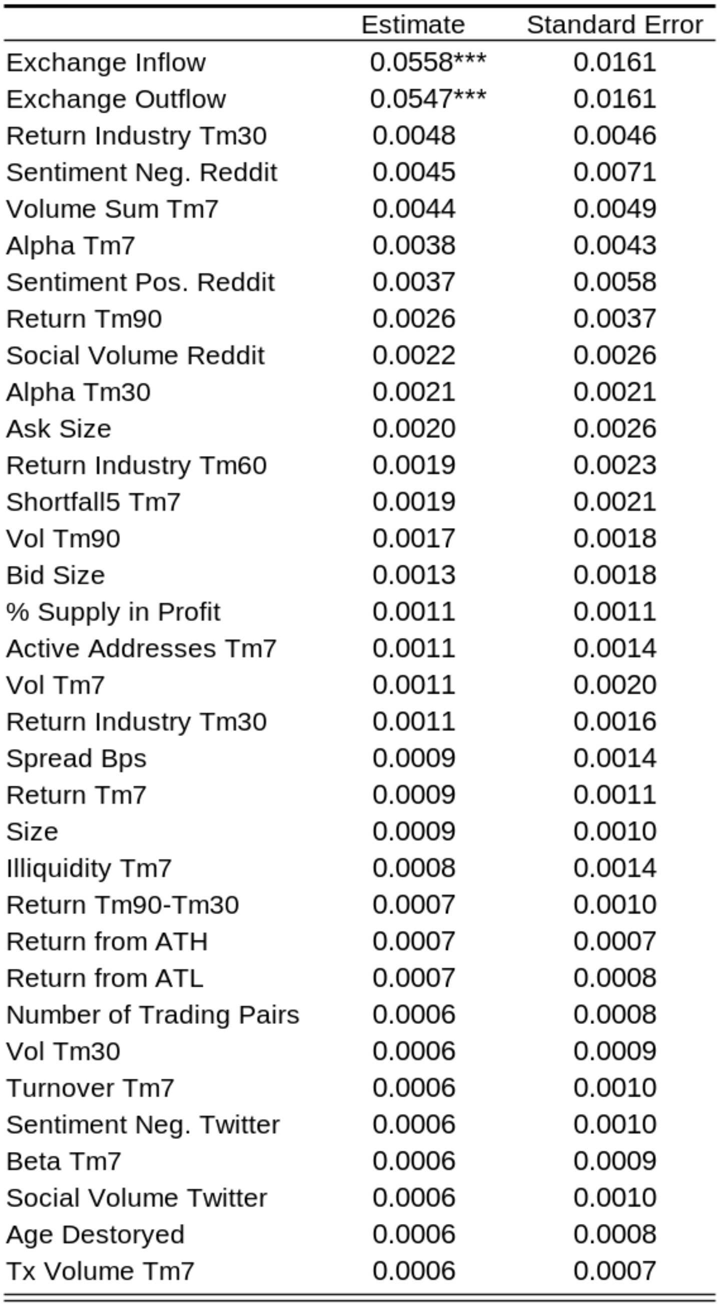
This table reports—for each univariate factor with a statistically significant 5-1 strategy—coefficients and standard errors from the contemporaneous time-series regression of the univariate factor 5-1 returns on the 5-1 returns for the best multivariate, PCA, and IPCA models. The best models were selected based on their Sharpe Ratio. The coefficients for alpha (i.e. intercept) and the three loadings are reported with standard errors in parentheses below. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The R^2 is reported in the last column.

FIGURE A35. DSLFM Out-of-Sample Portfolio Statistics.

Weighting	# Factors		Quintiles					5-1						
		Pred. R2	1	2	3	4	5	TS Avg	Sharpe	Sortino	Turnover	MDD	Alpha	Beta
Mcap	1	0.0007	-0.0206	0.0025	-0.0008	-0.0069	-0.0051	0.0156	2.32	5.7	0.41	-0.14	0.0155	0.0309
			(-1.24)	(0.13)	(-0.05)	(-0.39)	(-0.3)	(1.64)					(0.0099)	(0.1239)
	2	< 0	-0.0106	-0.0	-0.0031	0.0044	-0.0149	-0.0043	-0.7	-0.96	0.28	-0.29	-0.005	0.3425***
			(-0.73)	(-0.0)	(-0.19)	(0.18)	(-0.76)	(-0.49)					(0.0072)	(0.0899)
	3	< 0	-0.0054	-0.0052	-0.0079	0.005	-0.0053	0.0001	0.01	0.03	0.37	-0.27	-0.0009	0.4762***
			(-0.39)	(-0.3)	(-0.44)	(0.19)	(-0.25)	(0.01)					(0.009)	(0.1133)
	4	< 0	-0.0107	-0.0057	0.0016	-0.0064	0.0033	0.014	1.83	4.02	0.33	-0.26	0.0133	0.3046**
			(-0.67)	(-0.38)	(0.08)	(-0.36)	(0.16)	(1.29)					(0.0101)	(0.1271)
	5	< 0	-0.0113	-0.0013	-0.0037	-0.0	-0.0024	0.0089	1.56	2.91	0.4	-0.16	0.0086	0.1257
			(-0.69)	(-0.08)	(-0.18)	(-0.0)	(-0.14)	(1.1)					(0.0081)	(0.1019)
Equal	1	0.0007	-0.0173	-0.0058	-0.016	-0.0122	-0.0046	0.0127**	3.36	8.18	0.41	-0.06	0.0125**	0.0825
			(-0.95)	(-0.32)	(-0.91)	(-0.65)	(-0.24)	(2.37)					(0.0054)	(0.068)
	2	< 0	-0.018	-0.0103	-0.0039	-0.0125	-0.0107	0.0073	1.44	2.19	0.28	-0.13	0.0065	0.3291***
			(-1.26)	(-0.54)	(-0.2)	(-0.64)	(-0.54)	(1.02)					(0.0051)	(0.0647)
	3	< 0	-0.0216	-0.0106	-0.0127	-0.0069	-0.0039	0.0177**	2.96	8.91	0.37	-0.06	0.0171**	0.3023***
			(-1.36)	(-0.6)	(-0.71)	(-0.35)	(-0.19)	(2.09)					(0.0073)	(0.0922)
	4	< 0	-0.0185	-0.0113	-0.0103	-0.0133	-0.0024	0.0161**	3.08	8.56	0.33	-0.08	0.0156**	0.2196**
			(-1.07)	(-0.66)	(-0.57)	(-0.73)	(-0.12)	(2.18)					(0.0068)	(0.0853)
	5	< 0	-0.016	-0.0028	-0.0125	-0.0156	-0.009	0.0069	1.05	2.71	0.4	-0.16	0.0066	0.1342
			(-0.9)	(-0.16)	(-0.7)	(-0.8)	(-0.45)	(0.74)					(0.0095)	(0.1191)

This table reports—for the DSLFM with market cap and equal-weighted portfolios—the predictive R^2, the mean quintile portfolio returns, and portfolio statistics for the 5-1 strategy for July-December 2022, inclusive. For each quintile, the mean returns are the time-series averages of weekly value-weighted portfolio excess returns sorted on each model's predicted returns. 5-1 is the long-short top minus bottom quintile zero-investment portfolio for each model; for which, we report the time-series average weekly value-weighted excess return, annualized Sharpe Ratio, annualized Sortino, weekly turnover, maximum drawdown, and alpha and beta to the CMKT return. *t*-stats are reported below each strategy's point estimates where *, **, and *** denote significance at the 10%, 5%, and 1% levels. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags.

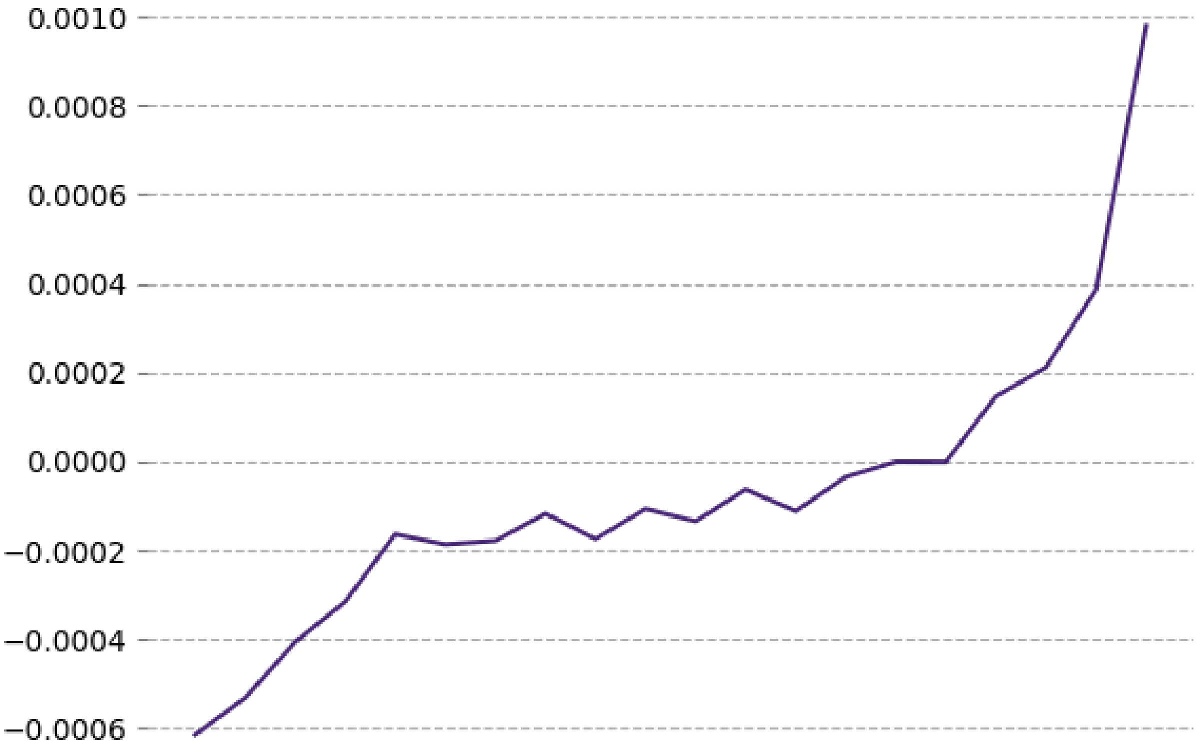
Figure A36. DSLFM: Asset Characteristic Significance.



	Estimate	Standard Error
Exchange Inflow	0.0558***	0.0161
Exchange Outflow	0.0547***	0.0161
Return Industry Tm30	0.0048	0.0046
Sentiment Neg. Reddit	0.0045	0.0071
Volume Sum Tm7	0.0044	0.0049
Alpha Tm7	0.0038	0.0043
Sentiment Pos. Reddit	0.0037	0.0058
Return Tm90	0.0026	0.0037
Social Volume Reddit	0.0022	0.0026
Alpha Tm30	0.0021	0.0021
Ask Size	0.0020	0.0026
Return Industry Tm60	0.0019	0.0023
Shortfall5 Tm7	0.0019	0.0021
Vol Tm90	0.0017	0.0018
Bid Size	0.0013	0.0018
% Supply in Profit	0.0011	0.0011
Active Addresses Tm7	0.0011	0.0014
Vol Tm7	0.0011	0.0020
Return Industry Tm30	0.0011	0.0016
Spread Bps	0.0009	0.0014
Return Tm7	0.0009	0.0011
Size	0.0009	0.0010
Illiquidity Tm7	0.0008	0.0014
Return Tm90-Tm30	0.0007	0.0010
Return from ATH	0.0007	0.0007
Return from ATL	0.0007	0.0008
Number of Trading Pairs	0.0006	0.0008
Vol Tm30	0.0006	0.0009
Turnover Tm7	0.0006	0.0010
Sentiment Neg. Twitter	0.0006	0.0010
Beta Tm7	0.0006	0.0009
Social Volume Twitter	0.0006	0.0010
Age Destoryed	0.0006	0.0008
Tx Volume Tm7	0.0006	0.0007

This table reports estimates of the importance of each asset characteristic to the fitted DSLFM using the test statistic $W_{\Gamma,j} = \Gamma_{\beta,j}^{\top}\Gamma_{\beta,j}$ discussed in the theory section. The DSLFM latent loading estimate, i.e. $\widehat{\Gamma}_{\beta}$, comes from fitting the DSLFM to the entire weekly panel with hyperparameters selected by the DSLFM CV procedure for the best k, i.e. highest mcap-weighted Sharpe. Standard errors are formed from the standard deviation of the simulated distribution of $\widehat{W}_{\Gamma,j}$ using 200 bootstrap draws, for each j. Significance is denoted with *, **, and *** for the 10%, 5%, and 1% levels, respectively. Only characteristics within two orders of magnitude of the maximum estimate are shown, i.e. 34 of the 63 characteristics.

FIGURE A37. Deep Learning Factor Model: Arithmetic Average Returns Out-of-Sample by Ventiles of Predicted Returns



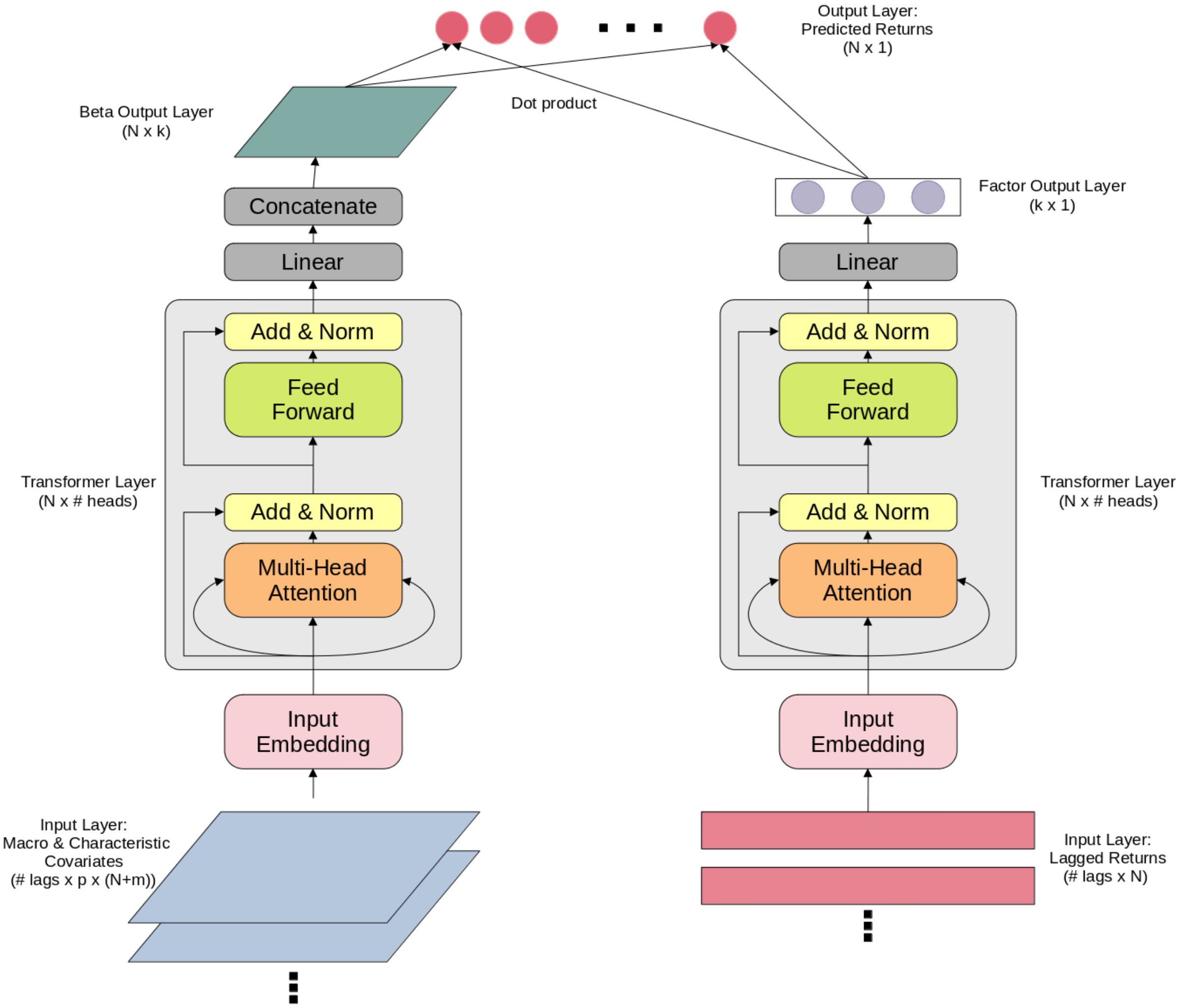
This figure shows the arithmetic average return for the out-of-sample period (Q3 and Q4 2022) of portfolios of assets in the hourly panel as sorted into ventiles by predicted returns of the deep learning factor model.

FIGURE A38. Deep Learning Factor Model: Out-of-Sample Portfolio Statistics

Restricted	Weighting		Quintiles					5-1						
		Pred. R2	1	2	3	4	5	TS Avg	Sharpe	Sortino	Turnover	MDD	Alpha	Beta
Unrestricted	Mcap	-0.0147	-0.0005***	-0.0002**	-1e-04	-0.0	0.0004***	0.0009***	37.29	81.3	0.42	-0.04	0.0009***	0.0476***
			(-6.04)	(-2.09)	(-1.53)	(-0.5)	(5.21)	(26.47)					(0.0000)	(0.0071)
	Equal	-0.0147	-0.0005***	-0.0002**	-1e-04	-0.0	0.0004***	0.0009***	37.38	81.38	0.42	-0.0397	0.0009***	0.0458***
			(-6.24)	(-2.17)	(-1.59)	(-0.48)	(5.32)	(26.53)					(0.0000)	(0.0073)
Restricted	Mcap	-0.0033	-	-	-	-	-	0.0018***	9.69	33.19	0.02	-0.45	0.0005*	0.1965***
								(6.88)					(0.0003)	(0.0567)
	Equal	-0.0033	-	-	-	-	-	0.0020***	10.32	35.02	0.02	-0.43	0.0006**	0.2037***
								(7.33)					(0.0003)	(0.0581)

This table reports—for the Deep Learning Factor Model with market cap and equal-weighted and unrestricted and restricted portfolios—the predictive R^2, the mean quintile portfolio returns, and portfolio statistics for the 5-1 strategy for July-December 2022, inclusive. The shortability and volume-limit restrictions are described in the main text. For each quintile, the mean returns are the time-series averages of weekly value-weighted portfolio excess returns sorted on each model's predicted returns. We only applied the restrictions to the long-short portfolio and do not report separate quintiles for the restricted portfolios. 5-1 is the long-short top minus bottom quintile zero-investment portfolio for each model; for which, we report the time-series average weekly value-weighted excess return, annualized Sharpe Ratio, annualized Sortino, weekly turnover, maximum drawdown, and alpha and beta to the CMKT return. *t*-stats are reported below each strategy's point estimates where *, **, and *** denote significance at the 10%, 5%, and 1% levels. Standard errors are Newey-West adjusted using Bartlett's formula for the number of lags.

FIGURE A39. Transformer Factor Model.



This figure presents a diagram of a Transformer architecture embedded into a factor model. The left-hand side describes how factor loadings $\beta_t \in \mathbb{R}^{N\times k}$ at time t (in green) are estimated using asset characteristics $Z_t \in \mathbb{R}^{N\times p}$ and macro covariates $m_t \in \mathbb{R}^{N\times m}$ and lagged values $(Z_{t-1}, m_{t-1}, \ldots)$ (in blue) mapped through the Transformer layer. The right-hand side describes how factors f_{t+1} (in purple) are estimated using lagged returns $r_t, r_{t-1}, \ldots$ (in red) through the Transformer layer. The final output layer of predicted returns $r_{t+1} \in \mathbb{R}^N$ are formed from the dot product of factor loadings and factors $\beta_t f_{t+1}$.