EMPIRICAL CRYPTO ASSET PRICING USING FACTOR MODELS WITH HIGH-DIMENSIONAL CHARACTERISTICS

Adam Baybutt

Automa Capital & UCLA

May 20, 2024

http://www.adambaybutt.org/research.html

PREVIEW: SETUP

Consider a dynamic latent factor model with linear loadings

$$r_{i,t+1} = \underbrace{z_{i,t}^\top \Gamma_\beta}_{\beta_{i,t}^\top} f_{t+1} + \epsilon_{i,t+1}, \quad \mathbb{E}[\epsilon_{i,t+1} | z_{i,t}] = 0,$$

where we observe, for assets *i* and time periods *t*,

- asset excess returns $r_{i,t+1} \in \mathbb{R}$ and
- asset characteristics $z_{i,t} \in \mathbb{R}^p$.

PREVIEW: MAIN THEORY CONTRIBUTIONS

In this setup, under the novel asymptotics of $p, T, N \to \infty$, contribute a new estimation procedure for

- latent loadings $\Gamma_\beta \in \mathbb{R}^{p \times k}$ and
- latent factors $f_{t+1} \in \mathbb{R}^k$, for all t;

and, prove the consistency of these estimators.

Also, I extend to this setting a classic asset pricing test and provide an asymptotically valid inference procedure.

PREVIEW: EMPIRICAL RESEARCH QUESTIONS

Broadly: Study the dynamics of crypto asset returns.

Specifically:

- Measuring expected returns through lens of factor models.
- What characteristics are the drivers of returns?
- What is the inflation risk premium in the crypto asset class?
- If we relax interpretability, what is the maximum out-of-sample predictability that can we achieve?

MOTIVATION

Static observable factor model:

$$r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$$

$$(NT + Tk) \text{ data} \gtrsim (Nk) \text{ params.}$$

MOTIVATION

Static observable factor model:

$$r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$$

$$(NT + Tk) \text{ data} \gtrsim (Nk) \text{ params.}$$

Static latent factor model:

$$r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$$

$$NT \gtrsim Nk + Tk$$

MOTIVATION

Static observable factor model:

$$r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$$

$$(NT + Tk) \text{ data} \gtrsim (Nk) \text{ params.}$$

Static latent factor model:

$$r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$$

$$NT \gtrsim Nk + Tk$$

Dynamic latent factor model:

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

$$NT(1 + p) \gtrsim pk + Tk$$

MOTIVATION

Static observable factor model: $r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$.

Static latent factor model: $r_{i,t+1} = \beta_i^\top f_{t+1} + \epsilon_{i,t+1}$.

Dynamic latent factor model: $r_{i,t+1} = z_{i,t}^\top \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}$.

Nonparametric dynamic latent factor model:

$$r_{i,t+1} = f(z_{i,t})^\top g(r_{i,t}) + \epsilon_{i,t+1}$$

SETUP

Assume for time periods $t = 1, \ldots, T$ and assets $i = 1, \ldots, N$,

we observe realizations of random variables

- asset excess returns $r_{i,t+1} \in \mathbb{R}$ and
- asset characteristics $z_{i,t} \in \mathbb{R}^p$, often high-dim. in practice.

SETUP

Assume for time periods $t = 1, \dots, T$ and assets $i = 1, \dots, N$, we observe realizations of random variables

- asset excess returns $r_{i,t+1} \in \mathbb{R}$ and
- asset characteristics $z_{i,t} \in \mathbb{R}^p$, often high-dim. in practice.

Assume the model:

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

SETUP

Assume for time periods $t = 1, \ldots, T$ and assets $i = 1, \ldots, N$, we observe

- asset excess returns $r_{i,t+1} \in \mathbb{R}$ and asset characteristics $z_{i,t} \in \mathbb{R}^p$.

Assume the model:

$$r_{i,t+1} = \underbrace{z_{i,t}^\top \Gamma_\beta}_{\beta_{i,t}^\top} f_{t+1} + \epsilon_{i,t+1}, \quad \mathbb{E}[\epsilon_{i,t+1} | z_{i,t}] = 0,$$

where

- $f_{t+1} \in \mathbb{R}^k$ are low-dimensional latent factors;
- $\Gamma_\beta \in \mathbb{R}^{p \times k}$ are unknown factor loading parameters; and,
- $\epsilon_{i,t+1}^r, \epsilon_{i,t}^\beta \in \mathbb{R}$ are unobserved scalar idiosyncratic errors.

SETUP

Assume for time periods $t = 1, \ldots, T$ and assets $i = 1, \ldots, N$, we observe

- asset excess returns $r_{i,t+1} \in \mathbb{R}$ and asset characteristics $z_{i,t} \in \mathbb{R}^p$.

Assume the model:

$$r_{i,t+1} = \underbrace{z_{i,t}^\top \Gamma_\beta}_{\beta_{i,t}^\top} f_{t+1} + \epsilon_{i,t+1}, \quad \mathbb{E}[\epsilon_{i,t+1} | z_{i,t}] = 0,$$

where

- $f_{t+1} \in \mathbb{R}^k$ are low-dimensional latent factors and
- $\Gamma_\beta \in \mathbb{R}^{p \times k}$ are unknown factor loading parameters.
  - Key assump.: Γ_β is exactly row sparse, i.e. most rows exactly zero.

EXTENDED SETUP (1/2)

In this framework, we address a common asset pricing question.

EXTENDED SETUP (1/2)

In this framework, we address a common asset pricing question.

What is the risk premium of an observable nontradable factor $g_{t+1} \in \mathbb{R}$?

EXTENDED SETUP (1/2)

Within this framework, we address an asset pricing research question.

What is the risk premium of an observable nontradable factor $g_{t+1} \in \mathbb{R}$?

Asset pricing context:

- Risk premium: return for exposure to the factor, ceteris paribus.
- If tradable, the risk premium is the time-series average of the factor.
- If nontradable, form factor mimicking-portfolio.
- Following Giglio, Xiu, and Zhang (2021),
  - assume latent factor model recovers true factor model and
  - project observable nontradable factor onto latent factors.

EXTENDED SETUP (2/2)

What is the risk premium of an observable nontradable factor $g_{t+1} \in \mathbb{R}$?

Assume for true latent factors f_{t+1} :

$$
\begin{aligned}
f_{t+1} &:= \gamma + v_{t+1}, & \mathbb{E}[v_{t+1}] = 0 \\
g_{t+1} &= \delta + \eta^\top v_{t+1} + \epsilon^g_{t+1}, & \mathbb{E}[v_{t+1}\epsilon^g_{t+1}] = 0.
\end{aligned}
$$

EXTENDED SETUP (2/2)

What is the risk premium of an observable nontradable factor $g_{t+1} \in \mathbb{R}$?

Assume for true factors f_{t+1} :

$$
\begin{aligned}
f_{t+1} &:= \gamma + v_{t+1}, & \mathbb{E}[v_{t+1}] = 0 \\
g_{t+1} &= \delta + \eta^\top v_{t+1} + \epsilon^g_{t+1}, & \mathbb{E}[v_{t+1}\epsilon^g_{t+1}] = 0.
\end{aligned}
$$

where

- $\eta \in \mathbb{R}^k$ is an unknown parameter mapping and
- ϵ^g_{t+1} is measurement error in g_{t+1}.

EXTENDED SETUP (2/2)

What is the risk premium of an observable nontradable factor $g_{t+1} \in \mathbb{R}$?

Assume for true factors f_{t+1} :

$$f_{t+1} := \gamma + v_{t+1}, \qquad \mathbb{E}[v_{t+1}] = 0$$
$$g_{t+1} = \delta + \eta^\top v_{t+1} + \epsilon^g_{t+1}, \quad \mathbb{E}[v_{t+1}\epsilon^g_{t+1}] = 0.$$

where

- $\eta \in \mathbb{R}^k$ is an unknown parameter mapping and
- ϵ^g_{t+1} is measurement error in g_{t+1}.

Our target parameter is $\gamma_g = \eta^\top \gamma$.

THEORETICAL CONTRIBUTIONS

The model:

$$r_{i,t+1} = z_{i,t}^\top \Gamma_\beta(\gamma + v_{t+1}) + \epsilon_{i,t+1}, \quad \mathbb{E}[\epsilon_{i,t+1}|z_{i,t}] = 0, \ \mathbb{E}[v_{t+1}\epsilon_{i,t+1}] = 0,$$
$$g_{t+1} = \delta + \eta^\top v_{t+1} + \epsilon_{t+1}^g, \quad \mathbb{E}[v_{t+1}\epsilon_{t+1}^g] = 0.$$

Two contributions, under novel asymptotics of $p, T, N \to \infty$:

1. consistently estimate latent loadings Γ_β and factors f_{t+1} and
2. conduct inference on $\gamma_g = \eta^\top \gamma$
   - under novel use of a <u>dynamic</u> latent factor model.

OUTLINE

1. ~~Preview~~
2. ~~Motivation~~
3. ~~Setup~~
4. ~~Theoretical Contributions~~
5. Theory Literature Review
6. Estimation
7. Key Assumptions
8. Asymptotic Results
9. Proof Outlines
10. Monte Carlo Evidence
11. Empirical Research Questions
12. Empirical Literature Review
13. Empirical Setting
14. Motivating Empirical Facts
15. Empirical Results:
    Low Dim. Factor Models
16. Empirical Results:
    High Dimensional Models

THEORY LITERATURE REVIEW

The scope of the relevant literature is enormous. To name a few:

- *Dynamic latent factor models*: Connor and Linton (2007), Fan, Liao, and Wang (2016), Kelly, Pruitt, and Su (2019) IPCA, Kelly, Pruitt, and Su (2020), etc.

- *Tests of observable factors*: Fama and MacBeth (1973) Fama-MacBeth, Feng, Giglio, and Xiu (2020) Factor Zoo, Giglio and Xiu (2021), etc.

- *DML*: Belloni, Chernozhukov, and Hansen (2014), Chernozhukov et al. (2018), Semenova and Chernozhukov (2021), etc.

ESTIMATION (1/4)

$$r_{i,t+1} = z_{i,t}^{\top}\Gamma_{\beta} f_{t+1} + \epsilon_{i,t+1}, \qquad E[\epsilon_{i,t+1}|z_{i,t}] = 0$$

ESTIMATION (1/4)

Rewrite the model:

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t}^\top \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}, & E[\epsilon_{i,t+1} | z_{i,t}] = 0, \\
&= z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^\top c_{t+1,-j} + \epsilon_{i,t+1}, \\
c_{t+1,j} &:= \Gamma_{\beta,j}^\top f_{t+1}.
\end{aligned}
$$

ESTIMATION (1/4)

Rewrite the model:

$$\begin{aligned} r_{i,t+1} &= z_{i,t}^\top \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}, \\ &= z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^\top c_{t+1,-j} + \epsilon_{i,t+1}, \qquad E[\epsilon_{i,t+1} | z_{i,t}] = 0, \\ c_{t+1,j} &:= \Gamma_{\beta,j}^\top f_{t+1}. \end{aligned}$$

To estimate $c_{t+1,j}$ $\forall t, j$

- run Lasso to account for $p \sim N$,

ESTIMATION (1/4)

Rewrite the model:

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t}^{\top}\Gamma_{\beta} f_{t+1} + \epsilon_{i,t+1}, \\
&= z_{i,t,j}c_{t+1,j} + z_{i,t,-j}^{\top}c_{t+1,-j} + \epsilon_{i,t+1}, \qquad E[\epsilon_{i,t+1}|z_{i,t}] = 0, \\
c_{t+1,j} &:= \Gamma_{\beta,j}^{\top} f_{t+1}.
\end{aligned}
$$

To estimate $c_{t+1,j}\ \forall t, j$

- run Lasso to account for $p \sim N$, but then biased inference for γ_g;

ESTIMATION (1/4)

Rewrite the model:

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t}^\top \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}, \\
&= z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^\top c_{t+1,-j} + \epsilon_{i,t+1}, \qquad E[\epsilon_{i,t+1} | z_{i,t}] = 0, \\
c_{t+1,j} &:= \Gamma_{\beta,j}^\top f_{t+1}.
\end{aligned}
$$

To estimate $c_{t+1,j}$ $\forall t, j$

- ~~run Lasso to account for~~ $\cancel{p \sim N}$~~, but then biased inference for~~ $\cancel{\gamma_g}$~~;~~
- instead run Double Selection Lasso (DSL).

ESTIMATION (2/4)

Model:

$$r_{i,t+1} = z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^{\top} c_{t+1,-j} + \epsilon_{i,t+1}, \qquad E[\epsilon_{i,t+1} | z_{i,t}] = 0,$$
$$c_{t+1,j} := \Gamma_{\beta,j}^{\top} f_{t+1}. \tag{1}$$

Procedure:

1. To estimate $\widehat{c}_{t+1,j}$, run $T \times p$ cross sectional DSL regressions. DSL
2. To estimate $\widehat{\Gamma}_{\beta} \in \mathbb{R}^{p \times k}$ and $\widehat{F} \in \mathbb{R}^{T \times k}$, run PCA on $\widehat{C} := \widehat{F}\widehat{\Gamma}_{\beta}^{\top} \in \mathbb{R}^{T \times p}$.
3. Given exact row sparsity, soft-threshold $\widehat{\Gamma}_{\beta}$ to set most rows to zero for $\check{\Gamma}_{\beta}$.

ESTIMATION (3/4)

Model for risk premium of nontradable observable factors:

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

ESTIMATION (3/4)

Model for risk premia of nontradable observable factors:

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

Identification:

- Cannot jointly estimate η and v_{t+1} (Γ_β and f_{t+1}) without further restrictions. E.g., three classic approaches of Bai and Ng (2013).
- So parameters are identified up to rotation matrix $H \in \mathbb{R}^{k\times k}$. That is, $\eta = H^{-1}\eta_0$ and $\gamma = H\gamma_0$ ($\Gamma_\beta = \Gamma_b^0 H^{-1}$ and $f_{t+1} = H f^0_{t+1}$).
- Utilize rotation invariant result of Giglio and Xiu (2021):

$$
\gamma_g = \eta_0^\top H^{-1\top} H\gamma_0 = \eta^\top \gamma
$$

ESTIMATION (4/4)

Model for risk premia of nontradable observable factors:

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t}^{\top}\Gamma_{\beta}(\gamma + v_{t+1}) + \epsilon_{i,t+1}, & \mathbb{E}[\epsilon_{i,t+1}] = 0,\ \mathbb{E}[v_{t+1}\epsilon_{i,t+1}] = 0, \\
g_{t+1} &= \eta^{\top} v_{t+1} + \epsilon_{t+1}^{g}, & \mathbb{E}[\epsilon_{t+1}^{g}] = 0,\ \mathbb{E}[v_{t+1}\epsilon_{t+1}^{g}] = 0.
\end{aligned} \tag{2}
$$

Procedure: $\widehat{\gamma}_g = \widehat{\eta}^{\top}\widehat{\gamma}$

- Estimate factor innovations $\widehat{v}_{t+1}$ and loadings $\check{\Gamma}_{\beta}$ as before but with demeaned returns.
- Estimate latent factor risk premia $\widehat{\gamma}$ via CS OLS of average returns $\bar{r} \in \mathbb{R}^N$ on estimated latent factor loadings $\widehat{\bar{\beta}} := T^{-1}\sum_t Z_t \widehat{\Gamma}_{\beta} \in \mathbb{R}^N$.
- Estimate latent to observable factor mapping $\widehat{\eta}$ via TS OLS of demeaned g_{t+1} on estimated latent factor innovations $\widehat{v}_{t+1}$.

OUTLINE

1. ~~Preview~~
2. ~~Motivation~~
3. ~~Setup~~
4. ~~Theoretical Contributions~~
5. ~~Theory Literature Review~~
6. ~~Estimation~~
7. Key Assumptions
8. Asymptotic Results
9. Proof Outlines
10. Monte Carlo Evidence
11. Empirical Research Questions
12. Empirical Literature Review
13. Empirical Setting
14. Motivating Empirical Facts
15. Empirical Results:
    Low Dim. Factor Models
16. Empirical Results:
    High Dimensional Models

KEY ASSUMPTIONS (1/2)

Assumption (Consistency of DSL)

1. *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:* $\sum_{j=1}^{p} \mathbb{1}\left\{ \left\| \Gamma_{\beta,j} \right\|_1 > 0 \right\} \leq s$. *Additional DSL Assumptions*

KEY ASSUMPTIONS (2/2)

Assumption (Consistency of Latent Factor Model)

2. *Nonzero and distinct eigenvalues: from the infeasible eigendecomposition of* $(Tp)^{-1}CC^\top$*, the k largest eigenvalues* λ_i *for* $i \in \{1, \ldots, k\}$ *are bounded away from zero and distinct,*

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

Additional Latent Factor Model Consistency Assumptions

ASYMPTOTIC RESULTS (1/3)

Proposition (Consistency of Latent Factors)

Under the DSLFM model (1) *and aforementioned Assumptions 1 and 2, with additional Appendix Assumptions 1-6, where* $T, N, p \to \infty$, *then for all* t *the latent factor estimator has the property that*

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

PROOF OUTLINE: CONSISTENT LATENT FACTORS

Recall $C = F\Gamma_\beta^\top$, thus $(Tp)^{-1}CC^\top = (Tp)^{-1}F\Gamma_\beta^\top\Gamma_\beta F^\top$.

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

Gives control over the distance between feasible and infeasible matrix:

$$\left\|(Tp)^{-1}\widehat{C}\widehat{C}^\top - (Tp)^{-1}CC^\top\right\| = O_p\left(\frac{\log Tp}{N}\right).$$

Davis Kahan Theorem bounds distance between eigenvectors by distance between matrices.

Finally, use Wely inequality to bound distance between eigenvalues.

ASYMPTOTIC RESULTS (2/3)

Proposition (Consistency of Latent Factor Loadings)

Under the DSLFM model (1) *and aforementioned Assumptions 1 and 2, with additional Appendix Assumptions 1-6, where* $T, N, p \to \infty$, *then the latent loading estimator has the property that*

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

PROOF OUTLINE: CONSISTENT LOADINGS

Aforementioned results yield:

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

Utilizing Theorem 2.10 from Belloni et al. (2018) under exact sparsity of Γ^0_β, s.t.

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

then given $\alpha \to 0$ and $\lambda \lesssim \sqrt{\log(Tp)/N}$, we have for all $q \geq 1$

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

ASYMPTOTIC RESULTS (3/3)

Theorem (Normality of Observable Factor Risk Premium)

Under the models (1) *and* (2)*; Assumptions 1 and 2; Appendix Assumptions 1-10, and, if* $Ts^2 \log(Tp)/N \to 0$*, then as* $T, N, p \to \infty$ *the estimator* $\hat{\gamma}_g$ *obeys*

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

MONTE CARLO EVIDENCE (1/2)

Goal: study the finite-sample estimation error of our latent loading and factor estimators and the coverage properties of our risk premium estimator compared to relevant benchmarks.

DGP: for $S = 200, T = 100, N = 500, k = 3, \; p \in \{10, 50\}, s = \; p/10$

- Latent loadings: fit IPCA to empirical panel; set $p - s$ rows to zero.
- Latent factors: fit IPCA to empirical panel; fit VAR(1) to fitted latent factors; simulate from fitted VAR(1) with normal innovations.
- Characteristics: fit panel VAR(1) to demeaned empirical panel of $\{Z_t\}_{t=1}^T$ and simulate from VAR(1) with normal innovations. Set means to bs.
- Returns and observable factor are generated according to the model where errors are calibrated to empirical R^2.

MONTE CARLO EVIDENCE (2/2)

Low-Dimensional: $p = 10$ Simulation Results Low-Dim.

- Factor of ~ 3 superior estimation error for Γ_β.
- Order of magnitude inferior estimation error for f_{t+1}.
- DSLFM under-covers (6-9%) while Giglio over-covers (2-4%) γ_g.

High-Dimensional: $p = 50$ Simulation Results High-Dim.

- Factor of >3 superior estimation error for Γ_β.
- Inferior ($\times 4$) estimation error for f_{t+1}.
- DSLFM degrades 1% while Giglio degrades > 3% γ_g.

OUTLINE

1. ~~Preview~~
2. ~~Motivation~~
3. ~~Setup~~
4. ~~Theoretical Contributions~~
5. ~~Theory Literature Review~~
6. ~~Estimation~~
7. ~~Key Assumptions~~
8. ~~Asymptotic Results~~
9. ~~Proof Outlines~~
10. ~~Monte Carlo Evidence~~
11. Empirical Research Questions
12. Empirical Literature Review
13. Empirical Setting
14. Motivating Empirical Facts
15. Empirical Results: Low Dim. Factor Models
16. Empirical Results: High Dimensional Models

EMPIRICAL RESEARCH QUESTIONS

Broadly: Study the dynamics of crypto asset excess returns.

Specifically:

- Measuring expected returns through lens of factor models.
- What characteristics are the drivers of returns?
- What is the inflation risk premium in the crypto asset class?
- If we relax interpretability, what is the maximum out-of-sample predictability that can we achieve?

EMPIRICAL LITERATURE REVIEW

Empirical crypto asset pricing is a nascent literature. To name a few:

- *Asset pricing ability of factors models*: Liu et al. (2019), Shams (2020), Bianchi and Babiak (2021), Liu, Tsyvinski, and Wu (2022), etc.
- *Crypto empirical facts*: Makarov and Schoar (2020), Hu, Parlour, and Rajan (2019), Borri (2019), Bianchi (2020), Liu and Tsyvinski (2021), Bianchi, Guidolin, and Pedio (2022), Cheah et al. (2022), Zhang and Li (2020), Zhang et al. (2021), Zhang and Li (2023), etc.
- *Crypto panel*: Liebi (2022), Borri et al. (2022), Cong et al. (2022).
- *ML Factor Models*: Chen, Pelger, and Zhu (2020), Gu, Kelly, and Xiu (2020), Gu, Kelly, and Xiu (2021), Giglio, Kelly, and Xiu (2022).

EMPIRICAL SETTING

- Weekly panel of crypto asset excess returns from 2018-2022, inclusive, with 63 time-varying asset characteristics. Summary Stats. Most go to zero.
- Inclusion criteria: month by month look back over trailing 3 months
  - tradable on US CEX;
  - remove stablecoins and synthetic assets;
  - asset mean mcap above 1 bps of total crypto mcap; and,
  - asset median total weekly trade volume on US exchanges above $500k.
- When fitting models sequentially, have to reform panel monthly.
- Form price from volume-weighted average hourly candle mid price.

OUTLINE

1. ~~Preview~~
2. ~~Motivation~~
3. ~~Setup~~
4. ~~Theoretical Contributions~~
5. ~~Theory Literature Review~~
6. ~~Estimation~~
7. ~~Key Assumptions~~
8. ~~Asymptotic Results~~
9. ~~Proof Outlines~~
10. ~~Monte Carlo Evidence~~
11. ~~Empirical Research Questions~~
12. ~~Empirical Literature Review~~
13. ~~Empirical Setting~~
14. Motivating Empirical Facts
15. Empirical Results:
    Low Dim. Factor Models
16. Empirical Results:
    High Dimensional Models

MOTIVATING EMPIRICAL FACTS (1/3)

Crypto Signals:

1. There are several characteristics with significant signal for the cross-section of one-week ahead expected returns.
2. The asset characteristics contain redundant information; however, the variation cannot be captured by just a few principal components.

Figures: Char. Correlations and Signal.

MOTIVATING EMPIRICAL FACTS (2/3)

A New and Rising Asset Class:

3. From zero in 2009, Bitcoin and hundreds of other crypto assets have become a trillion dollar asset class in 2022, with several multi-billion dollar sub-industries.
4. Bitcoin achieved superior risk-adjusted returns for nearly the entire study time period as compared to traditional asset classes.
5. Bitcoin has lower correlations to the Nasdaq and S&P500 (at 0.23 and 0.21) than that of gold's correlation to these indices (at 0.26 and 0.28).
6. 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 in 2022.
7. From diversifying a risk portfolio of holding 100% Nasdaq to 60% Nasdaq and 40% CMKT, one would obtain a Sharpe Ratio gain of 0.53 (from 0.43 to 0.96).
8. The crypto market offers a positive inflation risk premium of 31 bps.

Figures: MCaps, Sharpe Ratios, Correlations, Rolling Correlations, Risk and Return, Inf. Risk Premium.

MOTIVATING EMPIRICAL FACTS (3/3)

Bitcoin Onchain Facts:

10. Bitcoin is primarily used as a store value, not speculatively trading.
11. Bitcoin is a payment network settling hundreds of billions of dollars annually where the large majority of transactions cost less then one USD.
12. Efforts to fork, that is copy, the Bitcoin blockchain have had immaterial adverse effects on it; an event study of forks observes, on the contrary, significant positive effects on price, trading volume, active addresses, and social activity.

Figures: BTC Hodling, BTC Tx, BTC Forks.

LOW DIM. MODELS: UNIVARIATE FACTORS (1/2)

- Begin our empirical study of return dynamics by forming univariate factors from each asset characteristic.
- Form zero-net investment long-short quintile strategies for $t + 1$ sorted on each characteristic at time t.
- Statistically and economically significant strategies are financial (i.e. two week momentum, 30 and 60 day industry momentum, beta, idiosyncratic skewness, and 5% shortfall). Significant Univariate Factors
- 35 of 57 remaining strategies have weekly excess return above 30 bps.
- Predictors of average excess returns appear to be characteristic-based factors formed as functions of previous returns.

LOW DIM. MODELS: MULTI, PCA, IPCA (2/2)

- Multifactor models: static observable, static latent, & dynamic latent.
- Estimation procedures are Fama-MacBeth, PCA, and IPCA.
- All three have long-short strategies with economically significant Sharpe ratios of 1-4 (IPCA has stat. sig.) although optimal number of factors is inconsistent. Low Dim. Factor Models.
- Replicates IPCA to new asset class, and suggestive of signal in characteristics.

OUTLINE

1. ~~Preview~~
2. ~~Motivation~~
3. ~~Setup~~
4. ~~Theoretical Contributions~~
5. ~~Theory Literature Review~~
6. ~~Estimation~~
7. ~~Key Assumptions~~
8. ~~Asymptotic Results~~
9. ~~Proof Outlines~~
10. ~~Monte Carlo Evidence~~
11. ~~Empirical Research Questions~~
12. ~~Empirical Literature Review~~
13. ~~Empirical Setting~~
14. ~~Motivating Empirical Facts~~
15. ~~Empirical Results:~~ ~~Low Dim. Factor Models~~
16. Empirical Results: High Dimensional Models

HIGH DIM. MODELS: DSLFM (1/4)

Study three questions:

1. out of sample predictability;
2. characteristic importance; and,
3. inflation risk premium.

HIGH DIM. MODELS: DSLFM OOS (2/4)

In comparing predictability of DSLFM to previous models:

- CV penalty hyperparameters in Q3 2021 - Q2 2022.
- Positive R^2_{pred} for only 1 factor model.
- Sharpe ratios of $\sim$ 1 for mcap weighted and $\sim$ 2.5 for equal weighted.
- IPCA outperforms on Sharpe and pricing ability, but not materially.

  DSLFM OOS

- Limited N for asymptotics to kick in; no feature selection for feat. imp.

HIGH DIM. MODELS: DSLFM CHAR. IMP. (3/4)

- Implement bootstrap procedure for characteristic importance:

$$\widehat{W}_j = \widehat{\Gamma}_{\beta,j}^{\top} \widehat{\Gamma}_{\beta,j}.$$

- According to the DSLFM, the drivers of returns are exchange inflows and outflows.

HIGH DIM. MODELS: DSLFM INFLATION (4/4)

- Outstanding question on the relationship of crypto's returns to inflation.
- Perform inference procedure for inflation risk premium γ_g.
- Inflation risk premium of statistically significant 1.4 bps per week.
- Corroborates with a dynamic latent factor model with superior pricing ability the result using the static observable factor model fit with the Fama-MacBeth procedure.
- The asset class provides investors positive compensation for holding an inflation-hedged crypto portfolio, ceteris paribus.

HIGH DIM. MODELS: DEEP LEARNING (1/6)

We ask:

- what is the maximum out of sample Sharpe ratio achievable
- in our out of sample period
- by relaxing interpretability of the factor model to specify nonparametric factors and factor loadings
- estimated with deep learning?

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\implies$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\Longrightarrow$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\implies$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\implies$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\implies$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (2/6)

Why DL to learn $f(\cdot)$ in $r_{i,t+1} = f(I_t) + \epsilon_{i,t+1}$:

- I_t is high dimensional
- I_t redundant $\implies$ regularization!
- $f(\cdot)$ is likely non-linear.
- DL's universal approximation theorems, e.g. Hornik, Stinchcombe, and White (1990).
- Minuscule improvement is the name of the game.

HIGH DIM. MODELS: DEEP LEARNING (3/6)

Now, why not?

$$r_{i,t+1} = f(I_t) + \epsilon_{t+1}, \qquad E[\epsilon_{t+1}|I_t] = 0.$$

- EMH $\implies$ $f(\cdot) = 0$, such that $r_{i,t+1} = 0 + \epsilon_{i,t+1}$.
- It's $f_t(\cdot)$, not $f(\cdot)$.
- Makes the curse of dimensionality much bigger problem.
- Perhaps SOTA DL with lots of regularization?

HIGH DIM. MODELS: DEEP LEARNING (3/6)

Now, why not?

$$r_{i,t+1} = f(I_t) + \epsilon_{t+1}, \qquad E[\epsilon_{t+1}|I_t] = 0.$$

- EMH $\implies$ $f(\cdot) = 0$, such that $r_{i,t+1} = 0 + \epsilon_{i,t+1}$.
- It's $f_t(\cdot)$, not $f(\cdot)$.
- Makes the curse of dimensionality much bigger problem.
- Perhaps SOTA DL with lots of regularization?

HIGH DIM. MODELS: DEEP LEARNING (3/6)

Now, why not?

$$r_{i,t+1} = f(I_t) + \epsilon_{t+1}, \qquad E[\epsilon_{t+1} | I_t] = 0.$$

- EMH $\implies$ $f(\cdot) = 0$, such that $r_{i,t+1} = 0 + \epsilon_{i,t+1}$.
- It's $f_t(\cdot)$, not $f(\cdot)$.
- Makes the curse of dimensionality much bigger problem.
- Perhaps SOTA DL with lots of regularization?

HIGH DIM. MODELS: DEEP LEARNING (3/6)

Now, why not?

$$r_{i,t+1} = f(I_t) + \epsilon_{t+1}, \qquad E[\epsilon_{t+1}|I_t] = 0.$$

- EMH $\implies$ $f(\cdot) = 0$, such that $r_{i,t+1} = 0 + \epsilon_{i,t+1}$.
- It's $f_t(\cdot)$, not $f(\cdot)$.
- Makes the curse of dimensionality much bigger problem.
- Perhaps SOTA DL with lots of regularization?

HIGH DIM. MODELS: DEEP LEARNING (3/6)

Now, why not?

$$r_{i,t+1} = f(I_t) + \epsilon_{t+1}, \qquad E[\epsilon_{t+1}|I_t] = 0.$$

- EMH $\implies$ $f(\cdot) = 0$, such that $r_{i,t+1} = 0 + \epsilon_{i,t+1}$.
- It's $f_t(\cdot)$, not $f(\cdot)$.
- Makes the curse of dimensionality much bigger problem.
- Perhaps SOTA DL with lots of regularization?

HIGH DIM. MODELS: DEEP LEARNING (4/6)

Utilize recent adaption of feed-forward neural networks within a factor model structure: stay true to equilibrium asset pricing theory. Autoencoder.

Several modifications:

- Nonlinear factor autoencoder.
- Sequential CV with more hp points: learning rate decay, Adam parameters, weights and bias initializers, etc.
- Weight loss by hourly trade volume over previous hour.
- Feature selection down to 50 characteristics.

Also, embed Transformers into factor model structure. Transformer.

HIGH DIM. MODELS: DEEP LEARNING (5/6)

Iterative step-forward CV: for each hyperparameter point in the grid,

- Set training data to 2018-2020.
  - For transformer, dropped 2018-2019 due to too many missing assets.
  - 2020 has 532,368 non-missing asset-hours.
- Set validation data to Q1 2021.
- For each week in the validation period,
  - Fit in training and predict in val week.
- Add current validation data to training, set next quarter to validation.

Select best model from validation period. OOS predict, **once**.

HIGH DIM. MODELS: DEEP LEARNING (6/6)

- Autoencoder dominates on out of sample Sharpe at 10.
  - It failed to achieve positive out-of-sample predictive R^2.

DL Results.

- Transformer had a stat. sig. R^2_{pred} of 3.6%,
  - But, returns were negative in the Q3-Q4 2022 data to transaction costs,
  - which motivates further research with 2023 data.
  - Scaling laws present in validation period up to 50-100k parameters, but did not replicate out of sample.

I am focused on exploring these DL scaling laws in practice.

REFERENCES

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. 2020. "Cryptocurrencies as an asset class? An empirical assessment." *The Journal of Alternative Investments* 23 (2): 162–179.

Bianchi, Daniele, and Mykola Babiak. 2021. "A risk-based explanation of cryptocurrency returns." *Available at SSRN 3935934*.

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

Borri, Nicola. 2019. "Conditional tail-risk in cryptocurrency markets." *Journal of Empirical Finance* 50: 1–19.

Borri, Nicola, Daniele Massacci, Mirco Rubin, and Dario Ruzzi. 2022. "Crypto risk premia." *Available at SSRN*.

Cheah, Jeremy Eng-Tuck, Di Luo, Zhuang Zhang, and Ming-Chien Sung. 2022. "Predictability of bitcoin returns." *The European Journal of Finance* 28 (1): 66–85.

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*.

Cong, Lin William, George Andrew Karolyi, Ke Tang, and Weiyi Zhao. 2022. "Value premium, network adoption, and factor pricing of crypto assets."

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 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.

Thank You!

APPENDIX: IPCA

The model is

$$r_{i,t} = z_{i,t-1}^{\top} \Gamma_{\delta} f_t + \epsilon_{i,t}.$$

The objective function is to minimize the sum of the squared errors:

$$\min_{\Gamma_{\delta}, f_t} \sum_{t=1}^{T} \left(r_t - Z_{t-1} \Gamma_{\delta} f_t\right)^{\top} \left(r_t - Z_{t-1} \Gamma_{\delta} f_t\right).$$

APPENDIX: IPCA

The first-order conditions are

$$\hat{f}_t = \left(\hat{\Gamma}_\delta' Z_{t-1}' Z_{t-1} \hat{\Gamma}_\delta\right)^{-1} \hat{\Gamma}_\delta' Z_{t-1}^\top r_t,$$

$$\text{vec}\left(\hat{\Gamma}_\delta'\right) = \left(\sum_{t=1}^{T-1} Z_{t-1}' Z_{t-1} \otimes \hat{f}_t \hat{f}_t'\right)^{-1} \left(\sum_{t=1}^{T-1} \left[Z_{t-1} \otimes \hat{f}_t'\right]' r_t\right).$$

Factor realizations are period-by-period cross section regression coefficients of r_t on the latent loading matrix δ_{t-1}.

Γ_δ is the coefficient of returns regressed on the factors interacted with firm-specific characteristics.

APPENDIX: IPCA

Similarities:

(Second-stage) factor model relationship and joint fitting.

Cross-sectional and time-series two step procedures a la Fama MacBeth.

Efficiency gains from using asset covariates.

Accommodate unbalanced panels.

Pro Double Lasso:

Sparse estimation

Convex objective functions

Model high dimensional p

Closed-form inference for target question

Pro IPCA:

Conceptually simpler optimization

Fewer assumptions for asymptotic theory

Rapid estimation

Back to Lit Review Back to Est

APPENDIX: FAMA-MACBETH REGRESSIONS

The classic observable factor model estimation is the Fama and MacBeth (1973) procedure.

We first run N TS regressions for each asset followed by T CS regressions for each time period.

That is, we first estimate $\hat{\beta}_i$ for each asset i by running TS OLS of $\{r_{i,t+1}\}_{t=1}^{T}$ on $\{f_{t+1}\}_{t=1}^{T}$.

Next, we run $\forall t$ the CS OLS of asset excess returns $\{r_{i,t+1}\}_{i=1}^{N}$ on estimated factor loadings $\{\hat{\beta}_i\}_{i=1}^{N}$.

We recover estimates $\hat{\lambda}_t$ for the risk premium $\lambda_t = \mathbb{E}_t[f_{t+1}]$ as well as

the pricing errors from the cross-sectional residuals, $\hat{\alpha}_{i,t+1}$.

Finally, we estimate the parameters of interest: the static risk premium $\hat{\lambda}$ and the static average

pricing error $\hat{\alpha}_i$ as the time-series averages of the relevant estimator, $\hat{\lambda}_t$ and $\hat{\alpha}_{i,t+1}$, respectively.

Back

APPENDIX: DSL ESTIMATION PROCEDURE

$$
\begin{aligned}
r_{i,t+1} &= z_{i,t,j} c_{t+1,j} + z_{i,t,-j}^{\top} c_{t+1,-j} + \epsilon_{i,t+1}, && E[\epsilon_{i,t+1} | z_{i,t}] = 0, \\
z_{i,t,j} &= z_{i,t,-j}^{\top} \delta_{t,j} + \epsilon_{i,t,j}^{z}, && E[\epsilon_{i,t,j}^{z} | z_{i,t,-j}] = 0, \\
c_{t+1,j} &:= \Gamma_{\beta,j}^{\top} f_{t+1}.
\end{aligned}
$$

For $\widehat{c}_{t+1,j}$, run $T \times p$ Double Selection Lasso CS regressions $\forall t, j$.

Lasso $\{r_{i,t+1}\}_{i=1}^{N} \rightarrow \{z_{i,t}\}_{i=1}^{N}$ for $\widehat{I}_1$ = nonzero elements of $\widehat{c}_t$.

Lasso $\{z_{i,t,j}\}_{i=1}^{N} \rightarrow \{z_{i,t,-j}\}_{i=1}^{N}$ for $\widehat{I}_2$ = nonzero elemnts of $\widehat{\delta}_{t,j}$.

Define $\widehat{I} := \widehat{I}_1 \cup \widehat{I}_2 \cup \widehat{I}_3$ where $\widehat{I}_3$ is manually chosen.

OLS $\{r_{i,t}\}_{i=1}^{N}$ on elements of $\{z_{i,t-1}\}_{i=1}^{N}$ in $\widehat{I}$. Back

APPENDIX: ASSUMPTIONS

Assumption (DSL Uniform Consistency)

1. *Bounded Characteristic Portfolios: For a finite absolute constant M and $\forall t, j$,* $\left|c_{t+1,j}\right| = \left|\Gamma_{\beta,j}^{\top} f_{t+1}\right| < M$.
2. *Sparsity rate: The sparsity index obeys* $s^2 \log^2 (p \vee N) / \left(\sqrt{N \log(Tp)}\right) \leq \delta_{N,T}$. *Additionally,* $\log^3 p/N \leq \delta_{N,T}$.
3. *Weak dependence between the first- and second-stage errors: There exists a positive constant M such that $\forall\, p, T, N$:*

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

4. *Additional standard DSL assumptions in Appendix C.2 of the paper.*

Back

APPENDIX: ASSUMPTIONS

Assumption (Consistency of Latent Factor Model)

5. *Factors:* $\mathbb{E}\left\| f_{t+1}^{0} \right\|^{4} \leq M < \infty$ *and* $T^{-1}\sum_t f_{t+1}^{0} f_{t+1}^{0\top} \to_p \Sigma_f$ *for some* $k \times k$ *positive definite matrix* Σ_f.
6. *Factor Loadings:* $\forall j$, $\left\|\Gamma_{\beta,j}\right\| \leq M < \infty$ *and* $\left\|\Gamma_\beta^\top \Gamma_\beta / p - \Sigma_\Gamma\right\| \to 0$ *for some* $k \times k$ *positive definite matrix* Σ_Γ.

Back

APPENDIX: ASSUMPTIONS

Assumption (Inference)

$\exists$ a generic absolute constant $M < \infty$ such that for all p, T, N :

7. *Bounded idiosyncratic errors:* $\mathbb{E}[(\sum_t \epsilon_{i,t+1})^2] \leq TM$.
8. *Bounded scaled factor innovations:* $\mathbb{E}[(\sum_t z_{i,t}^\top \Gamma_\beta^0 v_{t+1}^0)^2] \leq sTM$.
9. *Bounded measurement errors:* $\mathbb{E}[(\epsilon_{t+1}^g)^2] \leq M$.

APPENDIX: ASSUMPTIONS

Assumption (Inference)

9. *Convergence of characteristics:* $\frac{1}{NT}\sum_i \sum_{t'} \mathbb{E}[z_{i,t,j}] z_{i,t',j'} \to_p \mathcal{Z}_{t,j,j'}$ *uniformly over* t, j, j' *for* $j, j' \in \{1, 2, \ldots, p\}$ *and a nonstochastic finite constant* $\mathcal{Z}_{t,j,j'} \in \mathbb{R}$.
10. *CLT: As* $T \to \infty$,

$$\frac{1}{\sqrt{T}} \sum_t \begin{pmatrix} v^0_{t+1} \epsilon^g_{t+1} \\ \Pi_t v^0_{t+1} \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, \Phi)$$

for random matrix $\Pi_t \in \mathbb{R}^{k \times k}$ *and nonstochastic matrix* $\Phi \in \mathbb{R}^{2k \times 2k}$.

APPENDIX: SIMULATION LOW-DIMENSIONAL

p	Parameter	Metric	(1) IPCA	(2) Three-Pass Est.	(3) DSLFM
		MSE	0.112526		0.040480
	Γ	Bias^2	0.020931		0.029007
		Var	0.091596		0.011473
		MSE	0.046446	1.023278	1.008919
	F	Bias^2	0.000538	0.006095	0.007407
		Var	0.041890	1.006150	0.992703
		MSE	1.736775	0.348060	0.336661
	β	Bias^2	0.051617	0.027838	0.027619
10		Var	1.551492	0.008405	0.000433
		MSE	0.007724		0.034307
	C	Bias^2	0.000066		0.000184
		Var	0.012636		0.033998
		MSE		0.000086	0.000125
		Bias^2		0.000003	0.000019
	γ_g	Var		0.000028	0.000015
		Cov90		0.971000	0.835000
		Cov95		0.990000	0.855000

Back to Simulation Result Summary.

APPENDIX: SIMULATION HIGH-DIMENSIONAL

p	Parameter	Metric	(1) IPCA	(2) Three-Pass Est.	(3) DSLFM
		MSE	0.024564		0.009921
	Γ	Bias^2	0.008984		0.008385
		Var	0.015580		0.001536
		MSE	0.223446	1.034021	1.011574
	F	Bias^2	0.009573	0.033910	0.033418
		Var	0.228714	0.989699	0.967504
		MSE	4.171191	0.430072	0.396931
	β	Bias^2	0.606915	0.161588	0.155526
50		Var	4.084398	0.013159	0.000983
		MSE	0.013972		0.007161
	C	Bias^2	0.000751		0.000212
		Var	0.013849		0.007001
		MSE		0.015229	0.014656
		Bias^2		0.015084	0.014495
	γ_g	Var		0.000058	0.000069
		Cov90		1.000000	0.828571
		Cov95		1.000000	0.842857

Back to Simulation Result Summary.

APPENDIX: 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.

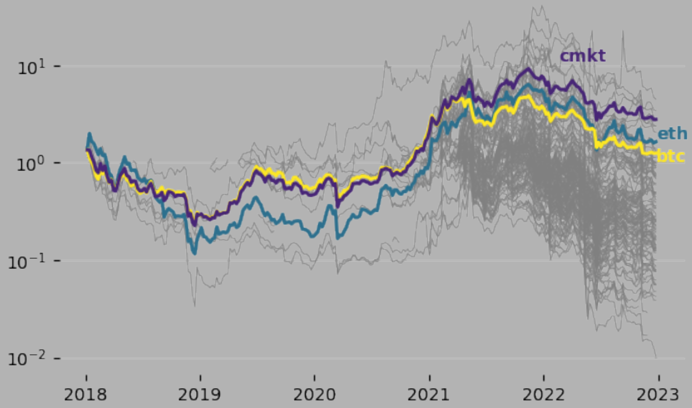
	Mean	SD	Sharpe	Skewness	Kurtosis	% > 0
CMKT	53.84%	80.61%	0.67	-0.02	0.02	0.53
Bitcoin	27.09%	75.07%	0.36	-0.02	0.02	0.52
Ethereum	52.97%	100.11%	0.53	0.03	0.04	0.52
Nasdaq	9.85%	22.98%	0.43	-0.03	0.03	0.55

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%

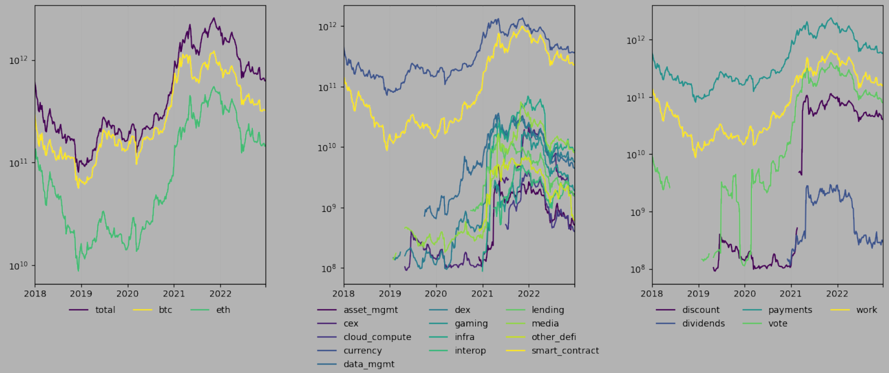
Back to Empirical Setting

APPENDIX: MOST ASSETS HAVE NEGATIVE RETURN



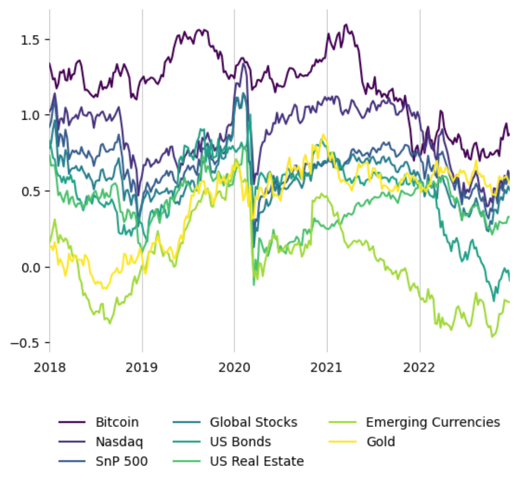
Back to Empirical Setting

APPENDIX: MARKET CAPS



Back to Motivating Empirical Facts

APPENDIX: SHARPE RATIOS



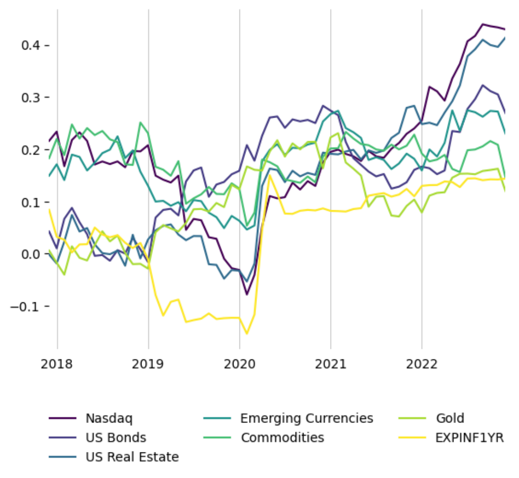
Back to Motivating Empirical Facts

APPENDIX: 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													

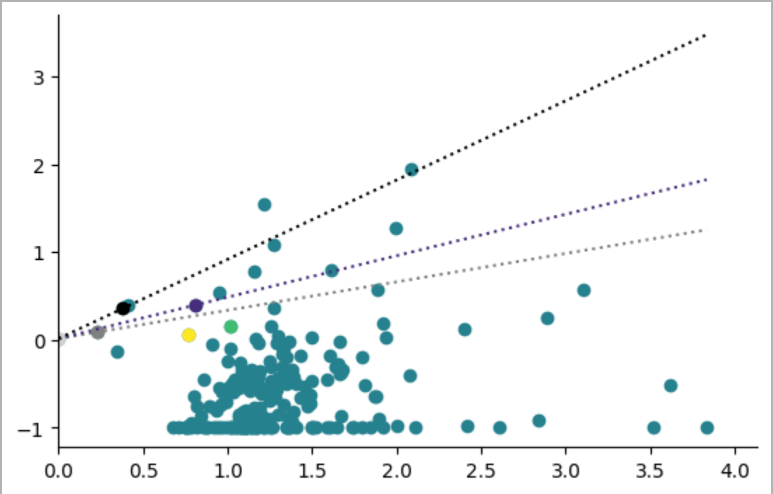
Back to Motivating Empirical Facts

APPENDIX: ROLLING CORRELATIONS



Back to Motivating Empirical Facts

APPENDIX: RISK AND RETURN



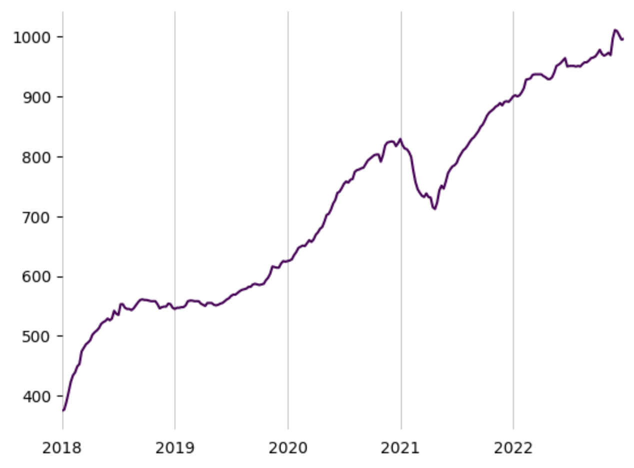
Back to Motivating Empirical Facts

APPENDIX: 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

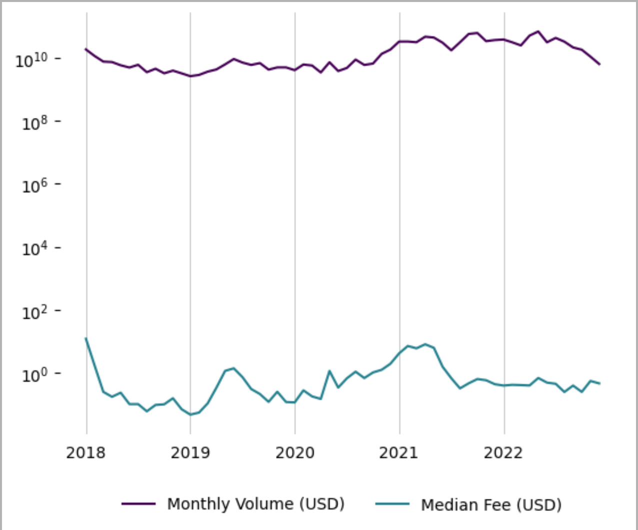
Back to Motivating Empirical Facts

APPENDIX: BITCOIN HODLING: UTXO MEDIAN AGE



Back to Motivating Empirical Facts

APPENDIX: BITCOIN ONCHAIN TRANSACTIONS



Back to Motivating Empirical Facts

APPENDIX: BITCOIN FORK: 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

Back to Motivating Empirical Facts

APPENDIX: CHARACTERISTIC CORRELATIONS AND SIGNAL

See paper appendix figures A15 through A24.

Back to Motivating Empirical Facts

APPENDIX: UNIVARIATE FACTOR RETURNS

	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)

Back to Univariate Factor Models.

APPENDIX: LOW DIMENSIONAL FACTOR MODELS

See paper appendix table A31.

Back to Low Dimensional Factor Models.

APPENDIX: DSLFM OOS PREDICTABILITY

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)

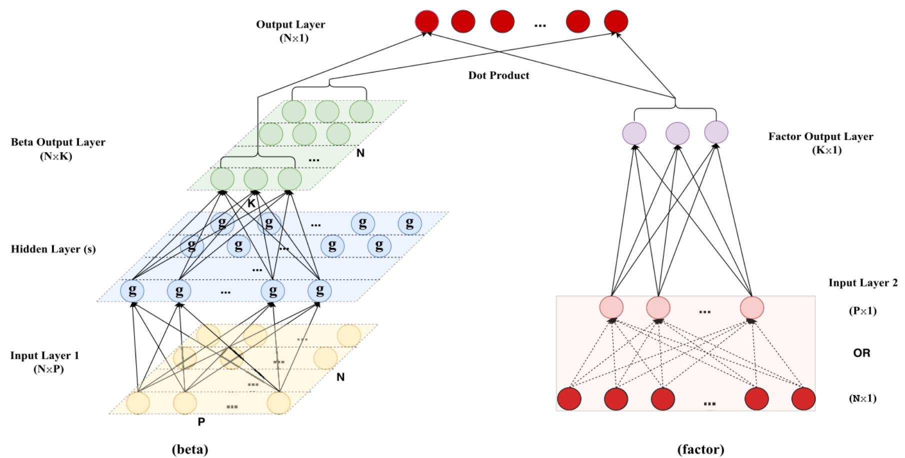
Back to DSLFM.

APPENDIX: DSLFM CHARACTERISTIC IMPORTANCE

	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

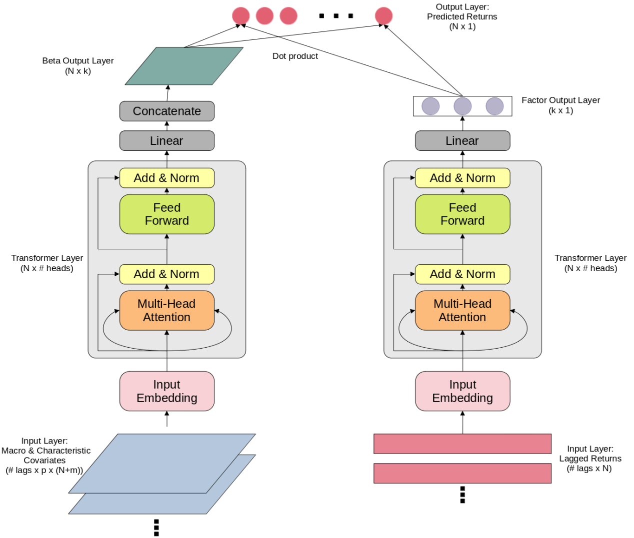
Back to DSLFM Char Imp.

APPENDIX: DL AUTOENCODER



Back to DL Models.

APPENDIX: TRANSFORMER FACTOR MODEL



Back to DL Models.

APPENDIX: DL FM OOS PORTFOLIO STATISTICS

Weighting		5-1						
	Pred. R2	TS Avg	Sharpe	Sortino	Turnover	MDD	Alpha	Beta
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)

Back to DL FM Results.