Extracting Statistical Factors When Betas Are Time-Varying

Patrick Gagliardini, Hao Ma
Università della Svizzera italiana (USI, Lugano) & Swiss Finance Institute (SFI)

Abstract
This paper deals with identification and inference on the unobservable conditional factor space and its dimension in large unbalanced panels of asset returns. The model specification is nonparametric regarding the way the loadings vary in time as functions of common shocks and individual characteristics. The number of active factors can also be time-varying as an effect of the changing macroeconomic environment. The method deploys Instrumental Variables (IV) which have full-rank covariance with the factor beta in the cross-section. It allows for a large dimension of the vector generating the conditioning information by machine learning techniques. In an empirical application, we infer the conditional factor space in the panel of monthly returns of individual stocks in the CRSP dataset between January 1971 and December 2017.

Introduction
Motivation
Asset Pricing literature on time-varying beta specifications has mostly focused on models with observable factors. However, there is a great latitude in the choice of observable economic factors. In this paper, we aim to find a dynamic PCA method for inference in latent factor models with time-varying betas.

Contributions
Methodological aspects
1. Deploy No-Arbitrage restrictions with conditioning information
2. (Almost) "model-free" regarding the dynamics of betas
3. First paper that allows for time-varying number of conditional factors and provides a consistent selection procedure
4. Cope with large-dimensional conditioning information via machine learning techniques
5. Use large unbalanced panels of individual stock returns
6. Develop asymptotic theory (as $T \to \infty$) building on results for plug-in sieve estimation and Double Machine Learning.

Empirical findings
1. There are two dominant factors across the sample period, together explaining over 70% measured by the AEP ratio.
2. The number of factors tends to be smaller during recession periods.
3. The first conditional factor is best explained by MKT, and the second one is best spanned by SMB.

No-Arbitrage Conditional Factor Model
Consider a conditional factor model for individual asset $i$ in period $t$: $r_{i,t} = \beta_{i,t}^0 + \beta_{i,t}f_{t} + \epsilon_{i,t}, t = 1, \ldots, T; i = 1, \ldots, n$.

- $\epsilon_{i,t}$ is the idiosyncratic error term
- No-arbitrage restriction (Gagliardini, Ossola and Scaillet (2006)): $\beta_{i,t}^0 = \beta_{i,t}^0|F_{t-1}$

Insert (2) into (1) and get the no-arbitrage conditional factor model: $f_{t} = \beta_{i,t}^0 + \epsilon_{i,t}$

where $\beta_{i,t}$ is $\beta_{i,t}^0 + \epsilon_{i,t}$ and $\beta_{i,t}$ is $\beta_{i,t}^0 + \epsilon_{i,t}$ is the risk premium.

Identification Strategy
Assumption 1. There exist a $K \times k$ ($K \geq k$) vector of instrumental variables $w_{i,t}$, measurable w.r.t. $F_{t-1}$ such that:

\begin{align}
(1) & \quad \text{plim } \sum_{i=1}^{n} w_{i,t}^2 = 0, \\
(2) & \quad \text{plim } \sum_{i=1}^{n} w_{i,t}w_{i,t}' = \Gamma_{t-1} \text{ is a } K \times k \text{ full-rank matrix}, \text{ a.s.}
\end{align}

\[ f_{t} = \beta_{i,t}^0 + \epsilon_{i,t} \]

where $\beta_{i,t}$ is $\beta_{i,t}^0 + \epsilon_{i,t}$ and $\beta_{i,t}$ is $\beta_{i,t}^0 + \epsilon_{i,t}$ is the risk premium.

Identification Strategy
Assumption 1. There exist a $K \times k$ ($K \geq k$) vector of instrumental variables $w_{i,t}$, measurable w.r.t. $F_{t-1}$ such that:

\begin{align}
(1) & \quad \text{plim } \sum_{i=1}^{n} w_{i,t}^2 = 0, \\
(2) & \quad \text{plim } \sum_{i=1}^{n} w_{i,t}w_{i,t}' = \Gamma_{t-1} \text{ is a } K \times k \text{ full-rank matrix}, \text{ a.s.}
\end{align}

Instrument-Weighted Portfolio Returns $\mathcal{E}_t$
Define the large $n$ limit of cross-sectional portfolio returns: $\xi_t = \frac{1}{n} \sum_{i=1}^{n} w_{i,t}^2$, which is measurable w.r.t. $F_{t-1}$. Then, (3) and Assumption 1 imply:

$$ \xi_t = \beta_{i,t}^0 + \epsilon_{i,t} $$

Identification of Number of Factors $k$
The conditional variance of $\xi_t$ is a $K \times K$ symmetric matrix:

\[ V(\xi_t) = \Gamma_{t-1} - \gamma_t(y_{t-1} - \Gamma_{t-1} y_{t-1}) \]

The number of latent conditional factors $k$ is identifiable by the rank:

$$ \text{rank}(\xi_t) = \text{rank}(\Gamma_{t-1}) $$

Identification of Latent Factors $\xi_t$
Let $A_{t-1}$ be a $K \times k$-dimensional $\xi_t$-full-rank matrix whose columns are the standardized eigenvectors of $\text{cov}(\xi_t)$. \(
A_{t-1} \text{ is the time-varying beta, } \beta_{i,t}^0 \text{ is a } k \times 1 \text{ vector, } f_{t} \text{ is the systematic risk factor, a } \xi_t \text{-measurable vector, } \gamma_t \text{ is a } k \times 1 \text{ vector.}
\)

Estimation Methodology
Assumption 4. The information set $\mathcal{F}_t$ is generated by the observable $d$-dimensional vector Markov process $Z_t$.

\[ E(\xi_t|\mathcal{F}_t) = E(\xi_t|\mathcal{F}_{t-1}) = \psi(Z_{t-1}) \]

where $\psi$ is a function of one and any random vector $Z_t$. Since $Z_t$ could be large-dimensional, machine learning techniques are used in estimating conditional expectations and variances.

- Post-Lasso method used in e.g. Belloni et al. (2012)
- Artificial Neural Networks with different network structures

Double Machine Learning Inference on Avg. Cond’l Features
Let the finite-dimensional parameter $c = \psi(\theta)$ be defined by

\[ c = E(\xi_t|\mathcal{F}_t) = E(\xi_t|\mathcal{F}_{t-1}) \]

e.g. average conditional correlation.

Double Machine Learning (DML): use the "locally robust" moment restriction

\[ \sum_{t=1}^{n} [\phi(\theta, Z_{t-1}) - \phi(\theta, \theta_{t-1})] = 0 \]

where function $\phi(\theta, Z_{t-1})$ is the Ritz representer of the Gauaux derivative

\[ \frac{\partial}{\partial \theta} g(\theta) = \phi(\theta, Z_{t-1}) \]

Split the sample in subintervals $I_{T_{1}, T_{2}}, \ldots, I_{T_{l}, T_{l+1}}$, and the DML estimator of $c$ is given by

\[ \hat{c} = \frac{1}{T} \sum_{t=1}^{T} \phi(\theta, Z_{t}) + \hat{o}(Z_{t}) \]

where $\hat{o}$ and $\hat{o}$ are obtained using observations not in $I_{t_{i}}$.

Empirical Results

- $\beta_{i,t}$: monthly excess returns of individual stocks
- $\alpha_{i,t}$: firm characteristics from Freyberger, Neuhierl and Weber (2017)
- $g_{i,t}$: 15 characteristics-based cross-sectional averages plus Fama-French 5 factors, Momentum, and Betting Against Beta
- $\gamma_{i,t}$: 19 variables - financial indicators from Goyal and Welch (2008), risk factors and macroeconomic variables

Figure 1: Accumulative Explanatory Power Ratio (6-month average)
The number of conditional factors is rather small. There are two dominant factors together with over 70% explanatory power across the whole sample period. Moreover, there tend to be fewer factors during recession periods.

Average Conditional Correlation
Figure 2 shows the average conditional correlation between our conditional latent factors $f_{t}$ and each variable in the information set $Z_{t}$.

Figure 2: Estimates of $E(\xi_t|\mathcal{F}_t, Z_{t})$