Advances in cross-sectional asset pricing

Empirical Asset Pricing

Mads Markvart Kjær

Department of Economics and Business Economics, Aarhus University

E-mail: mads.markvart@econ.au.dk

Spring 2023

Motivation

- Recall that one of the most fundamental research questions in empirical asset pricing is whether a given risk factor is priced in the financial markets and what its compensation is (risk premia).
- We have spend much time on improving the Fama-Macbeth methodology, which is useful in this context.
- Yet, it may still be subject to biases if we omit important *control* factors when estimating and testing a certain risk factor.
- …and/or if the risk factor is poorly measured or generally only weakly related to returns due to noise.
- How can we fix this?

$\Rightarrow$ methods from the ML toolbox, applied in a smart manner, help us!

Omitted variable bias *101*

- To recap the main point of the omitted variable bias concept, let us consider a simple OLS regressions model.

Omitted variable bias (**OVB**)

Suppose the true model (for instance a time-series or cross-sectional stage regression) is

$$y_i = \alpha + \beta x_i + \delta z_i + \varepsilon_i \tag{1}$$

for $i = 1, \ldots, N$ (where i could represent time or assets). Here, x_i and z_i are independent variables that drive the dependent variable y_i, and ε_i is a mean-zero error term. [...]

Omitted variable bias *101*

Omitted variable bias (cont'd)

[...]

- Vectorizing the model reads

$$Y = X\beta + Z\delta + E. \tag{2}$$

- Now suppose we run a regression using only X as regressor, which yields

$$\hat{\beta} = (X'X)^{-1}X'Y. \tag{3}$$

- It can then be shown that

$$\mathbb{E}[\hat{\beta}|X] = \beta + (X'X)^{-1}\mathbb{E}[X'Z|X]\delta. \tag{4}$$

Omitted variable bias *101*

- The omitted variable, z_i introduces a bias unless z_i is unrelated to the variable of interest X
- The bias is a function of the sign of the covariance between x_i and z_i, and the association between z_i and y_i measured by the coefficient δ
- In order to capture the true estimate of β, we are naturallly strongly dependent on specifying a model that properly accounts for all relevant variables...

$\Rightarrow$ this is where ML techniques can be highly useful!

Biases in standard approaches

The model

- To understand the importance of the omitted variable bias and measurement error bias in the context of asset pricing, we consider a simplistic two-factor case like Giglio and Xiu (2021).

The model

- Suppose that $v_t = (v_{1t}, v_{2t})'$ is a vector of two (nontraded) potentially correlated factors.
- They are, without loss of generality, de-meaned.
- Assuming the risk-free rate is observed, and including a constant for testing purposes, we consider the following model

$$r_t^e = \gamma_0 \iota_N + \beta(\gamma + v_t) + u_t, \tag{5}$$

where r_t^e is a vector of the entire universe of excess returns across N assets, and u_t is a mean zero error term.

The model

- Note that the model has the following implication, for the i'th asset

$$\mathbb{E}[r_{it}^e] = \gamma_0 + \beta_i \gamma, \tag{6}$$

  which is the standard β representation of expected excess returns, including a constant.
- Our main interest is in estimating the risk premia on (typically a subset of) the factors, which we denote by g_t.
- In this setting, we assume it is $g_t = v_{1t}$, and our goal is to estimate γ_1.

Biases in risk premia estimators

We have in the course examined two (actually three) different ways to test whether a nontraded factor is priced:

Recap: Two approaches to estimating risk premia

- Fama-Macbeth
- Factor-mimicking portfolio

Biases in risk premia estimators

Recap: Fama-Macbeth

- The first approach is to use the two-pass Fama-Macbeth approach.
- Here, risk premia are estimated in the second-pass cross-sectional regression step, using first-pass time series βs as regressors.

- …we are now ready to see how a omitted variable or measurement error bias causes problems in both approaches.

OVB in risk premia estimate

- Let us first consider the implication of omitted variable bias in the Fama-Macbeth setting.
- To make everything clear, note that the model can be expanded to

$$r_t^e = \beta_1(\gamma_1 + v_{1t}) + \beta_2(\gamma_2 + v_{2t}) + u_t, \tag{7}$$

where we in this example assume for simplicity that $\gamma_0 = 0$ for all i.

OVB in risk premia estimate

OVB in Fama-Macbeth

Suppose, possibly guided by a stylistic asset pricing model, that we use only $g_t = v_{1t}$ in our asset pricing analysis. The OVB arises in both the first- and second stage of the Fama-Macbeth procedure:

1. The estimate of β_1 in the first stage is biased as long as the omitted factor v_{2t} is correlated with v_{1t} (and relevant for returns) magnutide of OVB depends on the size of this correlation.
2. The estimate of γ_1 is biased due the former reason *and* that the cross-sectional stage omits $\hat{\beta}_2$ as regressor magnitude of OVB depends on the size of the correlation among β_1 and β_2.

Biases in risk premia estimators

Recap: Factor-mimicking portfolio

- When factors are traded, we can just take its sample mean as estimate of its risk premia.
- This motivates the factor-mimicking approach to estimating risk premia of a nontraded factor by transforming it into a traded portfolio.
- …We construct a traded portfolio by projecting the factor onto a set of traded base asset returns, creating a portfolio that is maximally correlated with g_t.
- The risk premium is then the time series average of this factor-mimicking portfolio return.
- Recall the lecture on portfolio sorting.

OVB in risk premia estimate

OVB in factor-mimicking portfolio

An OVB can arises in the risk premia estimate from the factor-mimicking portfolio if important base assets onto which g_t is project is omitted.

- Suppose our base assets are denoted by $\breve{r}_t^e \subseteq r_t^e$ and we project (regress) our factor $g_t = v_{1t}$ onto those asset together with a constant.
- This yields weights ω^g such that the mimicking portfolio is given by

$$r_t^g = \omega^{g\prime} \breve{r}_t^e. \tag{8}$$

- The expected excess return on this portfolio is

$$\gamma_g^{MP} = \omega^{g\prime} \mathbb{E}[\breve{r}_t^e]. \tag{9}$$

[...]

OVB in risk premia estimate

OVB in factor-mimicking portfolio

[...]

- Since $\breve{r}_t^e \subseteq r_t^e$ it implies that

$$\breve{r}_t^e = \breve{\beta}(\gamma + v_t) + \breve{u}_t. \tag{10}$$

- Using results from OLS regressions, it turns out that

$$\gamma_g^{MP} = \left((\breve{\beta}\Sigma^v \breve{\beta}' + \breve{\Sigma}^u)^{-1} (\breve{\beta}\Sigma^v e_1) \right)' \breve{\beta}\gamma \tag{11}$$

where $e_1 = (1,0)'$ is a column vector with unity in the first element and zero elsewhere, Σ^v is the covariance matrix of the factors, and $\breve{\Sigma}^u$ the covariance matrix of the idiosyncratic risks of the assets used in the projection.

OVB in risk premia estimate

OVB in factor-mimicking portfolio

[...]

- In order to get that $\gamma_g^{MP} = \gamma_1$, we need two conditions:
  1. We need that $\breve{\Sigma}^u = 0$, achieved if, for instance, the bases assets are well-diversified portfolios.
  2. $\breve{\beta}$ is invertible and $v_t = \breve{\beta}^{-1}\breve{r}_t^e$, achieved if the true factors are fully spanned (can be fully recovered by) the base assets.
- In this case, the expression in (11) reduces to the first element in γ, i.e. γ_1.
- If this is not the case, for instance by omitting some important base assets (that causes violation of either of condition 1 and 2 (or, likely, both), we have that the expression does not collapse to γ_1.
- …that is, $\gamma_g^{MP} \neq \gamma_1$, causing a bias in the risk premia estimate.

OVB in risk premia estimate

- The existing literature has typically ignored this bias…
- For instance, most people use a limited set of base assets (e.g. Fama-French type of size-value portfolios).
- Naturally, however, there are other risk sources than size and value which may be correlated with a given factor, for instance consumption growth, which are not captured by those portfolios.
- In this case, the estimator of the risk premia is affected by an OVB.

Measurement error bias in risk premia estimate

- Suppose that the econometrician can only observe the factor subject to some measurement error z_t that is orthogonal to the factors, but possibly correlated with u_t

$$g_t = v_{1t} + z_t. \tag{12}$$

- This is very common for nontraded factors that uses e.g. consumption data or tries to proxy some unobservable metric, like, liquidity.
- This adds another bias to the risk premia estimator in both the Fama-Macbeth and factor-mimicking case.

Measurement error bias in risk premia estimate

- In Fama-Macbeth it causes an attenuation bias in $\hat{\beta}_1$, which in turn leads to a bias in $\hat{\gamma}_1$.
- In the factor-mimicking case, the relevant term expands to

$$\gamma_g^{MP} = \left((\breve{\beta} \Sigma^v \breve{\beta}' + \breve{\Sigma}^u)^{-1} (\breve{\beta} \Sigma^v e_1 + \breve{\Sigma}^{z,u}) \right)' \breve{\beta} \gamma \tag{13}$$

  where $\breve{\Sigma}^{z,u} = \text{Cov}[z_t, u_t]$.
- Unless $\breve{\Sigma}^{z,u} = 0$, measurement errors cause a bias in that $\gamma_g^{MP} \neq \gamma_1$, even under the required conditions above.
- That is, if the measurement error is correlated with idiosyncratic risks of assets, this causes a bias.

A solution

Methodology

- We will now approach a methodology that tackles, jointly, both the omitted variable and measurement error issues.
- Suppose we have p many (possibly unobservable) factors such that

$$
\begin{aligned}
r_t^e &= \gamma_0 \iota_N + \beta(\gamma + v_t) + u_t, \\
\mathbb{E}[v_t] &= \mathbb{E}[u_t] = 0, \\
\text{Cov}[u_t, v_t] &= 0.
\end{aligned}
$$

- The objective is to estimate the risk premia of one or more factors g_t without necessarily observing all true factors v_t (because there is no way we can do that!).

Methodology

- To generalize the stylized setting above that assumed $g_t = v_{1t} + z_t$, we now allow for a more broad and general model of the set of d *observable* factors whose risk premia we aim to estimate

$$g_t = \delta + \eta v_t + z_t, \quad \mathbb{E}[z_t] = 0, \ \text{Cov}[z_t, v_t] = 0. \tag{14}$$

- This allows for a slope (η) and intercept (δ) coefficient in the specification, relating the observable factors to all unobservable ones.

Definition of risk premia

- The risk premium of g_t that we want to estimate is the expected excess return of a portfolio with $\beta = 1$ w.r.t g_t and $\beta = 0$ w.r.t. all other factors.
- It turns out that this implies

$$\gamma_g = \eta\gamma. \tag{15}$$

Methodology

- Thus, in order to estimate the risk premia, we need some way to estimate the entire product $\eta\gamma$ or their individual constituents.
- One of the main contributions from Giglio and Xiu (2021) is exactly this identification using an rotation invariance result.

Rotation invariance result (intuitively)

- The rotation invariance result states, intuitively, that the product $\eta\gamma$ can be identified even if one only observes an arbitrary full-rank rotation of the factors.
- That is, if one just observes

$$\hat{v}_t = H v_t, \tag{16}$$

with H any full-rank $p \times p$ matrix, but does *neither* observe v_t nor H (e.g. negate the expression).

Methodology

- To see this, note that $H^{-1}H = I_p$, such that

$$r_t^e = \gamma_0 \iota_N + \beta H^{-1} H(\gamma - v_t) + u_t \tag{17}$$

  and

$$g_t = \delta + \eta H^{-1} H v_t + z_t \tag{18}$$

  holds.
- Define now $\hat{\eta} = \eta H^{-1}$, $\hat{\gamma} = H\gamma$, and $\hat{\beta} = \beta H^{-1}$.
- We can now write the model entirely in terms of the rotated factors $\hat{v}_t$.

Methodology

Identify $\hat{\eta}$

- This implies that as long as $\hat{v}_t$ is observed (not the latent v_t) we can identify $\hat{\eta} = \eta H^{-1}$.
- How? Since g_t is observed, and so is $\hat{v}_t$, we can run a regression of g_t onto $\hat{v}_t$, including a constant.
- This estimates δ and $\hat{\eta}$ since the model reads

$$g_t = \delta + \hat{\eta}\hat{v}_t + z_t. \tag{19}$$

Methodology

Identify $\hat{\gamma}$

- This also implies that we can identify $\hat{\gamma} = H\gamma$.
- How? From standard cross-sectional regressions that uses $\hat{\beta} = \beta H^{-1}$ as regressors and average r_t^e as dependent variables a classical Fama-Macbeth with a single cross-sectional regression.
- Those $\hat{\beta}$s can be obtained in a first step from a regression of r_t^e onto $\hat{v}_t$, both being observed.

Methodology

Identify γ_g

- Now, even though we cannot identify η nor γ directly, we can identify their product via above, $\hat{\eta}\hat{\gamma}$, which in turn identifies the risk premia of interest as per

$$\hat{\eta}\hat{\gamma} = \eta H^{-1} H\gamma = \eta\gamma = \gamma_g. \tag{20}$$

- And this trick only applies to γ_g and none of the other quantities in the model.

- Big question is then: *how do we identify these rotated factors $\hat{v}_t$?*

Principal Component Analysis

- …, this is exactly what Principal Component Analysis (PCA) does!
- Recall from past lectures that PCA condenses a large data set into a smaller set of components that aim at capturing the most of the variation in data.
- Ideally, this is used to obtain proxies for the unobserved factors v_t.
- However, the factors are not completely unique in the sense that any rotation of the PCA components would explain the same amount of variance of the original data (for instance multiplying everything by -1).
- That is, PCA really identifies

$$\hat{v}_t = Hv_t \quad not \quad v_t. \tag{21}$$

- …, PCA identifies the factors up to a rotation.

Three-pass methodology

- Using the insights from above, we can formulate the following three-pass procedure in brief terms.

Three-pass procedure

To estimate the risk premia of the (observed) risk factors g_t, denoted γ_g, one conducts the following steps:

1. PCA: Estimate the rotated factors $\hat{v}_t$ via PCA on the full set of returns,
2. $\hat{\gamma}$: Estimate via a standard Fama-Macbeth two-pass cross-sectional regression $\hat{\gamma} = H\gamma$, i.e. the risk premia of the observed $\hat{v}_t$,
3. $\hat{\eta}$: Estimate $\hat{\eta} = \eta H^{-1}$ via a time-series regression of g_t onto $\hat{v}_t$.

The risk premia of g_t, $\eta\gamma$, can then be estimated by taking the product of $\hat{\eta}$ and $\hat{\gamma}$ from steps two and three above.

Three-pass methodology

- To present the methodology more formally, let us formulate the problem in matrix format and ignore the intercept γ_0 for the time being.
- R is the $N \times T$ matrix of all excess returns, V the $p \times T$ matrix of factors, G the $d \times T$ matrix of observable factors, U the $N \times T$ matrix of errors, and Z the $d \times T$ matrix of measurement errors.

$$R = \beta(\gamma \iota_T' + V) + U, \tag{22}$$

  with ι_T a conforming vector of ones.
- Writing $(\bar{R}, \bar{V}, \bar{G}, \bar{U}, \bar{Z})$ as the matrices of de-meaned variables, this rewrites to

$$\bar{R} = \beta \bar{V} + \bar{U}. \tag{23}$$

Three-pass methodology

- The de-meaned version of the observable factors are given by

$$\bar{G} = \eta\bar{V} + \bar{Z}. \tag{24}$$

- To recap, the estimation procedure does *not* require use of the unobservable V, but rather the output from PCA applied to the panel of returns in $\bar{R}$.
- Given observable returns R and factors of interest G, we can write the three steps for the estimator $\gamma_g = \eta\gamma$ more formally in the following slides.

Three-pass methodology

PCA step

- Extract the principal components of returns by conducting the PCA on the $T \times T$ matrix of data $P = (NT)^{-1}\bar{R}'\bar{R}$. This yields the etimated factors as

$$\hat{V} = T^{1/2}(\xi_1,\xi_2,\ldots,\xi_{\hat{p}})'. \tag{25}$$

  where $\xi_1,\xi_2,\ldots,\xi_{\hat{p}}$ are the eigenvectors sorted correspondingly to the largest $\hat{p}$ eigenvalues of P, and $\hat{p}$ is a consistent estimator of the number of factors.
- The estimator of $\hat{p}$ can be based on information criteria or determined by eye balling the plot of eigenvalues/explained variances, or similar arguments. Robustness is important, nevertheless.
- Estimate the factor loadings (risk exposures) as

$$\hat{\beta} = T^{-1}\bar{R}\hat{V}'. \tag{26}$$

Three-pass methodology

Cross-sectional regression step

- Run a standard cross-sectional OLS regression of average returns onto the estimated factor loadings from the prior step to obtain an estimate of the risk premia of the estimated latent factors

$$\hat{\gamma} = (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'\bar{r}. \tag{27}$$

- Note that $\bar{r}$ is not the de-meaned returns, but instead their averages, yet otherwise of same dimension.
- FYI, it turns out OLS is actually the efficient estimator here.

Three-pass methodology

Time-series regression step

- Run a standard time-series OLS regression of g_t onto the extracted factors from step 1 and obtain the slope estimator and the fitted value of the observable factor denoted by $\hat{G}$ as

$$\hat{\eta} = \bar{G}\hat{V}'(\hat{V}\hat{V}')^{-1}, \quad \text{and} \quad \hat{G} = \hat{\eta}\hat{V}. \tag{28}$$

Risk premia estimate

- Obtain the risk premia estimate as

$$\hat{\gamma}_g = \hat{\eta}\hat{\gamma}. \tag{29}$$

- This can also be expressed compactly by combining some of the components above

$$\hat{\gamma}_g = \bar{G}\hat{V}'(\hat{V}\hat{V}')^{-1}(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'\bar{r}. \tag{30}$$

Three-pass methodology

- The estimator can allow for the estimation of the intercept as well, writing excess returns as

$$r_t^e = \gamma_0 \iota_N + \beta(\gamma + v_t) + u_t. \tag{31}$$

- This simply amounts to altering step 2 (cross-sectional step) to include a constant, leading to the following compact forms

$$\begin{aligned} \hat{\gamma}_0 &= (\iota_N' M_{\hat{\beta}} \iota_N)^{-1} \iota_N M_{\hat{\beta}} \bar{r} \\ \hat{\gamma}_g &= \bar{G} \hat{V}' (\hat{V} \hat{V}')^{-1} (\hat{\beta}' M_{\iota_N} \hat{\beta})^{-1} \hat{\beta}' M_{\iota_N} \bar{r}. \end{aligned} \tag{32}$$

where

$$\begin{aligned} M_{\hat{\beta}} &= I_N - \hat{\beta}(\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}' \\ M_{\iota_N} &= I_N \iota_N (\iota_N' \iota_N)^{-1} \iota_N'. \end{aligned} \tag{33}$$

- Note that this require very little extra work - it just utilizes everything obtained in the steps without a constant.

Three-pass methodology

- Essentially the third step is the new bit compared to a standard Fama-Macbeth analysis that uses statistical factors (from step 1) in addition to the observable ones.
- This last step is critical, because it translates the uninterpretable risk premia of the latent factors to those the economic theory predicts.
- Also, it removes measurement error effects since we use only the effect in $\hat{\eta}$ and not the entire G in our procedures.
- $\ldots$, in fact $\hat{G} = \hat{\eta}\hat{V}$ is the factors cleaned from measurement error.
- Why? Note that

$$\hat{G} = \hat{\eta}\hat{V} = \bar{G} - \hat{Z}, \tag{34}$$

so that $\hat{G}$ is $\bar{G}$ when the (estimated) measurement error has been removed.

Three-pass methodology

Separation principle

- Even though g_t can be multi-dimensional ($d > 1$), the estimation of risk premia for each observable factor is separate (simply because we control for everything via $\hat{V}$).
- This is in contrast to standard Fama-Macbeth that ***has*** to include all relevant factors in the second-stage cross-sectional regression.

Interpretation

Interpretations

- This three-pass procedure is an extension of both the two-pass Fama-Macbeth procedure and the factor-mimicking approach.
- These two typically give very different results - now they do *not*.
- How to interprete $\hat{\eta}$ differs, however, a bit depending on whether we focus on the Fama-Macbeth or the factor-mimicking interpretation...

FM interpretation

- The invariance result supports directly a Fama-Macbeth interpretation
- $\hat{\eta}$ tells us how to rotate the estimated model such that $\hat{g}_t$ appears as the first factor
- $\hat{g}_t$ is used as the observables factor(s), together with $p - d$ many PCs as controls.
- This ensures no OVB in both the first and second stage of the Fama-Macbeth procedure.

Factor-mimicking portfolio interpretation

- Is also lends support to a factor-mimicking interpretation where the factor(s) is projected onto the PCs which constitute the base assets.
- Why? Because they exactly average out noise (remember the motivation from past lectures of PCA) such that $\Sigma^u \approx 0$ and they span the returns by construction of the factors.
- …, i.e. it satisfies the conditions for no OVB in the factor-mimicking portfolio.
- Using this interpretation, $\hat{\eta}$ denotes the factor weights $\rightarrow$ with risk premium of the weights times the RP of the portfolios $\hat{\gamma}$

Asymptotic theory

Theorem: Consistency

Under appropriate assumptions, and if $\hat{p} \xrightarrow{p} p$, then as $N, T \to \infty$,

$$\hat{\eta} \xrightarrow{p} \eta H^{-1}, \; \hat{\gamma} \xrightarrow{p} H\gamma, \text{ and } \hat{\eta}\hat{\gamma} = \hat{\gamma}_g \xrightarrow{p} \gamma_g. \tag{35}$$

for some invertible (with probability one) matrix H.

- This implies that the three-pass procedure indeed produces the right estimates, as long as we consistently estimate the correct number of latent factors.

Asymptotic theory

Theorem: Asymptotic normality

Under appropriate assumptions, and if $\hat{p} \xrightarrow{p} p$, then as $N, T \to \infty$, together with $T^{1/2}N^{-1} \to 0$, we have that

$$T^{1/2}\hat{\gamma}_g \xrightarrow{d} N(\gamma_g, \Phi), \tag{36}$$

where the covariance matrix is given by

$$\begin{aligned}\Phi = {} & (\gamma'(\Sigma^v)^{-1} \otimes I_d)\Pi_{11}((\Sigma^v)^{-1}\gamma \otimes I_d) \\ & + (\gamma'(\Sigma^v)^{-1} \otimes I_d)\Pi_{12}\eta' + \eta\Pi_{21}((\Sigma^v)^{-1}\gamma \otimes I_d) + \eta\Pi_{22}\eta',\end{aligned}$$

where Π is the covariance matrix of $z_t\hat{v}_t$.

Asymptotic theory

- Giglio and Xiu (2021) show that putting “hats" on everything in Φ to get $\hat{\Phi}$ yields $\hat{\Phi} \xrightarrow{p} \Phi$, provided that we use HAC type estimators.
- Under similar assumptions above, we can also get asymptotic normality of both $\hat{\gamma}_0$ and $\hat{\gamma}_g$ when we include a constant in the model.
- The variance expressions are much more involved (see Online Appendix to Giglio and Xiu (2021)). though already implemented in the live script provided to you.
- It is also quite interesting that the variance estimator does not require any EIV adjustment, like e.g. in Shanken (1992).
- And the variance estimator acts asymptotically as if all factors were observable - even though we have to estimate (a rotation of) them in step 1 via PCA.

Hypothesis testing

Hypothesis testing

With those results at hand, one can construct standard hypothesis tests on the significance of the risk premia as follows

$$t(\hat{\gamma}_g) = \frac{\hat{\gamma}_g}{\sqrt{\mathsf{Var}[\hat{\gamma}_g]}} \xrightarrow{d} N(0,1). \tag{37}$$

Ridge regression

- Remember from previous lectures that PCA essentially functions as a regularization technique that weights the relevant elements in a given data set according to their contribution to the signal.
- Another type of regularization is a version of the ENet known as ridge regression, which was defined by $\alpha = 1$ in the penalty term $\phi(\lambda, \alpha)$ in the lectures on ML and return predictability:

$$\tilde{\mathcal{L}} = \mathcal{L} + \frac{1}{2}\lambda \sum_{j=1}^{P} \theta_j^2 \tag{38}$$

- In the present context, one could imagine that instead of using PCA to select the relevant factors and project a given factor onto those for obtaining the factor-mimicking weights, one could instead use ridge regression.
- ..., that is, use ridge regression to conduct the (otherwise infeasible OLS) projection of the factor onto all available returns, handling parameter proliferation.

Ridge regression

- Giglio and Xiu (2021) shows the resulting risk premia estimator, discarding it in the same vein, will not be as efficient as the PCA approach.
- Why? For exactly the same reason we motivated dimension reduction techniques in previous lectures.
- …the ridge puts weight on all assets and will reflect some of the noise inherent in the assets.
- …PCA averages out the noise by forming the optimal linear combinations that extract the relevant signal.

Testing factor strength

Testing the strength of an observed factor

- A general and important concept in empirical asset pricing is so-called weak factors.

Weak factors

Weak factors are observable factors that are only weakly reflected in the cross-section of test assets.

- These typically challenge econometric techniques and the economic story.
- We can use the present framework to test the strength of a given observable factor.

Testing the strength of an observed factor

- Note that in the model

$$g_t = \delta + \eta v_t + z_t \tag{39}$$

an $\eta \approx 0$ indicates that either measurement error dominates or the factor is not pervasive, i.e. not a strong factor.

- In those cases, factor loadings/risk exposures would be weak (poorly estimated with little cross-sectional variation).

Testing strength (intuitively)

- The R^2 from the regression of g_t onto the estimated latent factors $\hat{v}_t$ measure the strength of the factor.
- The significance of this strength can be tested via $\mathbb{H}_0 : \eta = 0$ vs. $\mathbb{H}_1 : \eta \neq 0$.
- Those hypotheses are implied by similar hypotheses on the observed/estimated $\hat{\eta}$.

Testing the strength of an observed factor

Testing strength

To test $\mathbb{H}_0 : \eta = 0$ vs. $\mathbb{H}_1 : \eta \neq 0$ we may formulate a conventional Wald test as per

$$\hat{W} = T\hat{\eta}\left((\hat{\Sigma}^v)^{-1}\hat{\Pi}_{11}(\hat{\Sigma}^v)^{-1}\right)^{-1}\hat{\eta}' \xrightarrow{d} \chi^2_{\hat{p}}. \tag{40}$$

Moreover, as measure of strength we can use the (estimated) R^2 from the same regression as

$$\hat{R}^2_g = \frac{\hat{\eta}\hat{V}\hat{V}'\hat{\eta}'}{\bar{G}\bar{G}'} \xrightarrow{p} R^2_g \tag{41}$$

where R^2_g is the true version from a regression using v_t, not $\hat{v}_t$.

- This naturally hinges upon the assumption that the latent factors are pervasive/strong otherwise $\eta \neq 0$ wouldn't mean anything.

Testing the strength of an observed factor

- A few remarks:
  1. The degrees of freedom in the Wald test is $\hat{p}$ because we test this many null restrictions in η.
  2. This is also why we do a Wald test and not a standard t-test which would only apply if we use one latent factor $\hat{p} = 1$.
  3. We say a factor is strong (with strength $\hat{R}^2_g$) if we reject $\mathbb{H}_0 : \eta = 0$, i.e. if $\hat{W}$ exceeds the $\chi^2_{\hat{p}}(1-\alpha)$ percentile for α significance level.
- Giglio and Xiu (2021) show that

$$\mathbb{P}\left(\hat{W} > \chi^2_{\hat{p}}(1-\alpha)|\mathbb{H}_0\right) \to \alpha \text{ and } \mathbb{P}\left(\hat{W} > \chi^2_{\hat{p}}(1-\alpha)|\mathbb{H}_1\right) \to 1.$$

as $N,T \to \infty$ (size and consistency of the test).

Empirical illustration

The objective

- Instead of simply go through the results of Giglio and Xiu (2021), we will instead use the same portfolios but consider a different sample period
- You should note that Giglio and Xiu (2021) do not include an intercept in their main table
- A small hint: You should, furthermore, try to examine whether the results are robust towards including more portfolios than the 202...
- ... but now for our object:

The objective of the empirical study

We will estimate the risk premia on some traded and nontraded factors and compare their values and conclusions to the conventional approaches of Fama-Macbeth

- ... I leave the comparison with the factor-mimicking portfolio as an exercise for you...

Data on test assets

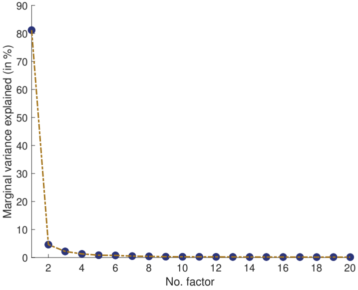
- Giglio and Xiu (2021) consider many different asset classes. We will, however, focus on the US equities from (July) 1963 to end of 2020
- We will consider 202 standard equity portfolios (pfs) from Kenneth French's website which span the most well-known dimensions of risk:
  1. 25 pfs sorted on size and book-to-market ratio
  2. 17 industry pfs
  3. 25 pfs sorted on profitability and investment
  4. 25 pfs sorted on size and variance
  5. 35 pfs sorted on size and net issuance (Note that this does not exists... We have included "Net Share Issues". These 35 portfolios have the same ticker as the one from the code of Giglio and Xiu (2021). Furthermore, the Kenneth French website states this is a 5x5 sort but the file includes 35 portfolios. I have no idea on who is right...)
  6. 25 pfs sorted on size and accruals
  7. 25 pfs sorted on size and beta
  8. 25 pfs sorted on size and momentum

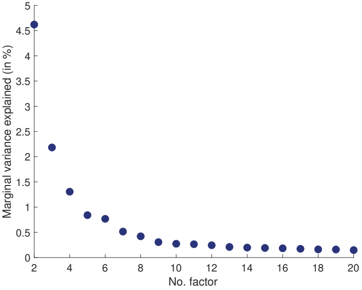
Data on observable factors g_t

- They focus on a selected few factors that have been discussed in the asset pricing literature.
- They consider single risk factors or groups of those this has no influence on the three-pass estimator whatsoever, but it influences the Fama-Macbeth estimator quite some.
- They consider the market excess return, SMB, HML, profitability (RMW), investment (CMA), momentum (MOM), betting-against-beta (BAB), quality (QMJ), and AR(1) innovations in industrial production grwoth, VAR(1) innovations in the first three PCs of 279 macro-finance variables, liquidity, two intermediary capital factors, high temperatures in Manhatten, global land surface temperatures, quasiperiodic Pacific Ocean temperatures, and the number of sunspots, and consumption growth.
- See the Online Appendix in Giglio and Xiu (2021) for data sources.
- We will, instead examine the CAPM, FF3 and innovations in Industrial product.

No. of factors from $\bar{R}$

- In order to estimate the number of factors $\hat{p}$, we may examine their individual variance contribution (or, equivalently, eigenvalues).





No. of factors from $\bar{R}$

- The first factor contributes by far the most, and the marginal variance explained generally show a hockey stick format.
- Zooming in, shows a drop after the 4th factor
- For the interested student, Giglio and Xiu (2021) also provide a concise estimator in their Online Appendix for the number of factors this also suggest 4 factors in our case

  $\rightarrow$ so let us go with 4, but note that this is our choice.
- We generally highlight the need for robustness, and a rule of thumb is to rather choose a factor too many than too few (such that we do not have an OVB issue).

No. of factors from $\bar{R}$

- We can receive a more quantitative measure to motivate our choice by looking at the R^2 of the factor model when increasing the number of included factors:

	1	2	3	4	5	6	7	8	9	10
R^2	0.00	0.17	0.28	0.64	0.64	0.64	0.66	0.67	0.68	0.68

Risk premia estimates

- Let's consider some results - We will focus on a test for the CAPM and innovations in Industrial production
- Let us start with the risk premia estimates of the two-pass and three-pass method (the estimates are measured in %)

	Avg return	two-pass	three-pass
Market	0.57	-0.10	0.37
		[-0.36]	[1.87]
ΔIP	(-)	22.56	0.64
		[1.48]	[0.55]

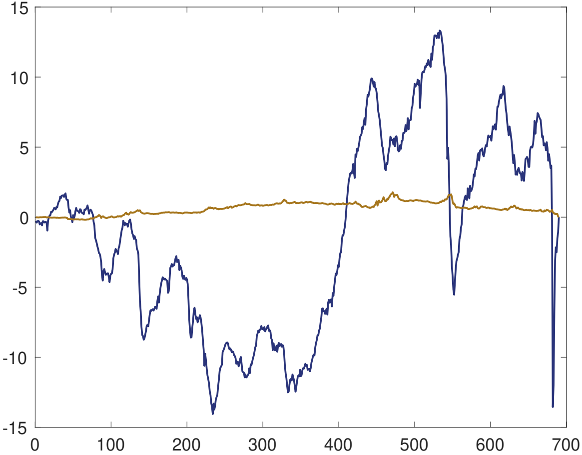
- extremely large difference in the estimates of the two-pass and three-pass!

Test for strength of the factors

- Next, we can examine whether the factors are strong:

	factor weak (p-value)	R^2_g
Market	(0.00)	0.99
ΔIP	(0.93)	0.00

A weak factor plot



Risk premia estimates

The empirical results of Giglio and Xiu (2021) is hidden in the appendix B1. Some of the conclusions are listed down below:

A small note to the results in Giglio and Xiu

1. The Fama-Macbeth estimator is subject to i) OVB as risk premia estimates vary substantially by controls (and differ from averages of traded factors) and ii) measurement error bias by finding significant risk premia on e.g. pure noise factors related to the climate, which the three-pass methodology deems as dominated by noise.
2. The three-pass methodology delivers risk premia estimates that are close economically and statistically to the averages of traded factors, and finds that several of the nontraded factors are significantly priced.
3. The strength (measured by R_g^2 and test via $\hat{W}$) indicate that traded factors are generally strong, while many nontraded ones may be deemed weak.

References

Feng, G., S. Giglio, and D. Xiu (2020): "Taming the Factor Zoo: a test of new factors," *The Journal of Finance*, Forthcoming.

Giglio, S. and D. Xiu (2021): "Asset pricing with omitted factors," *Journal of Political Economy*, Forthcoming.

Shanken, J. (1992): "On the estimation of beta-pricing models," *The Review of Financial Studies*, 5, 1–33.