Unexplained factors and their effects on second pass R-squared's and t-tests

Frank Kleibergen[1] Zhaoguo Zhan[2]

[1]Brown University Tsinghua University[2]

October 2013

Outline:

- Factor structure in observed portfolio returns
- Replacing unobserved factors by observed proxy factors
- FM two pass regression
- The R^2_{OLS} and R^2_{GLS} of the second pass regression
- Identification robust confidence sets

Factor Model for Portfolio returns

Portfolio returns exhibiting a (unobserved) factor structure with k factors result from a statistical model that is characterized by:

$$r_{it} = \mu_i + \beta_{i1} f_{1t} + \ldots + \beta_{ik} f_{kt} + \varepsilon_{it}, \qquad i = 1, \ldots, N, \ t = 1, \ldots, T;$$

with:

- r_{it} the return on the i-th portfolio in period t
- μ_i the mean return on the i-th portfolio
- f_{jt} the realization of the j-th factor in period t
- β_{ij} the factor loading of the j-th factor for the i-th portfolio
- ε_{it} the idiosyncratic disturbance for the i-th portfolio return in the t-th period

We can reflect the factor model as well using vector notation:

$$R_t = \mu + \beta F_t + \varepsilon_t,$$

with $R_t = (r_{1t} \dots r_{Nt})'$, $\mu = (\mu_1 \dots \mu_N)'$, $F_t = (f_{1t} \dots f_{kt})'$, $\varepsilon_t = (\varepsilon_{1t} \dots \varepsilon_{Nt})'$ and

$$\beta = \begin{pmatrix} \beta_{11} & \dots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \dots & \beta_{Nk} \end{pmatrix}.$$

If the factors are i.i.d. with finite variance and uncorrelated with the disturbances, the covariance matrix of the portfolio returns reads

$$V_{RR} = \beta V_{FF} \beta' + V_{\varepsilon\varepsilon},$$

with V_{RR}, V_{FF} and $V_{\varepsilon\varepsilon}$ the $N \times N$, $k \times k$ and $N \times N$ dimensional covariance matrices of the portfolio returns, factors and disturbances.

The factors affect many different portfolios simultaneously which enables identification of the number of factors using principal components analysis.

When we construct the spectral decomposition of the covariance matrix of the portfolio returns,

$$V_{RR} = P\Lambda P',$$

with

- P the $N \times N$ dimensional orthonormal matrix of characteristic vectors (or eigenvectors)
- Λ the $N \times N$ diagonal matrix of characteristic roots (or eigenvalues)

the largest k characteristic roots are distinctly larger than the remaining ones.

Factor structure in observed portfolio returns

In the paper, we use three different data-sets but here we just focus on one of them:

1. Lettau and Ludvigson (2001): quarterly portfolio returns from the third quarter of 1963 to the third quarter of 1998 of the return on twenty-five size and book-to-market sorted portfolios so $N = 25$ and $T = 141$.

	LL01
1	2720
2	113.8
3	98.6
4	18.36
5	17.61
6	13.48
7	12.11
8	9.31
9	8.42
10	7.25
largest three roots	95.5%

Table: Largest ten characteristic roots (in descending order).

The fraction of the variance of the portfolio returns that is explained by the three largest roots, which equals

$$\text{FACCHECK} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_1 + \ldots + \lambda_N}$$

with $\lambda_1 > \lambda_2 > \ldots > \lambda_N$, can be used as a test/check for three factors.

It rejects the hypothesis that the three largest characteristic roots explain less than 80% of the variation of the portfolio returns with more than 95% significance.

Factor models with observed factors

The observed factor model is identical to the unobserved factor model but F_t is observed and the number of observed proxy factors, say m, is known:

$$R_t = \mu + BG_t + \varepsilon_t,$$

with

- G_t the m-dimensional vector of observed proxy factors
- B the $n \times m$ dimensional matrix that contains the β's of the portfolio returns with the observed factors.

We compute an F-statistic for testing $\mathrm{H}_0 : \delta = 0$ in the linear model:

$$F_t = \mu_F + \delta G_t + v_t,$$

with F_t a vector that contains either the three FF (Fama-French) factors or just the HML (High-Minus-Low) and SMB (Small-Minus-Big) factors. G_t is a vector that contains the observed proxy factors.

We also report the (pseudo-) R^2 which equals one minus the total variation of the residuals over the total variation of the portfolio returns

$$\text{pseudo-}R^2 = 1 - \frac{\sum_{i=1}^N \lambda_{i,res}}{\sum_{i=1}^N \lambda_{i,port}}.$$

Lettau and Ludvigson (2001) use a combination of the factors:

- the value weighted return (R_{vw})
- the consumption-wealth ratio (cay)
- consumption growth (Δc)
- labor income growth (Δy)
- the Fama-French factors = (value weighted return, HML, SMB)
- interactions between the consumption wealth ratio and consumption growth ($cay\Delta c$) or the value weighted return ($cayR_{vw}$) or labor income growth ($cay\Delta y$).

	LL01					
	R_{vw}	Δc	FF	$cay, R_{vw},$ $cayR_{vw}$	$cay, \Delta c,$ $cay\Delta c$	$cay, R_{vw}, \Delta y,$ $cayR_{vw}, cay\Delta y$
1	435	2676	26.5	433	2414	412
2	99.5	111	22.3	98.0	105	97.2
3	26.2	98.6	14.3	26.0	96.0	25.6
4	18.36	18.1	13.9	17.9	17.9	17.8
5	13.8	16.8	11.2	12.9	16.7	12.8
FACCHECK	82.1%	95.5%	38.2%	82.5%	95.2%	82.1%
FF factors	80.1 0.000	3.73 0.292		81.8 0.000	28.9 0.001	90.3 0.000
HML-SMB				1.91 0.928		10.7 0.381
pseudo-R^2	0.78	0.016	0.95	0.78	0.10	0.79

Table: The largest five characteristic roots of the covariance matrix of the portfolio returns and residuals that result using different specifications from Lettau and Ludvigson (2001). The F-statistics at the bottom of the table result from testing the significance of the indicated factors in a regression of either the FF factors or the HML-SMB factors on them.

FM two pass procedure

Relationship between the expected return on a portfolio and the covariance between the portfolio returns and the observed factors:

$$E(R_t) = \iota_n \lambda_0 + \beta \lambda_F,$$

with

- ι_n the n-dimensional vector of ones
- λ_0 the zero-β return
- λ_F the k-dimensional vector of factor risk premia.

To estimate the risk premia, Fama and MacBeth (1973) propose a two pass procedure:

1. Estimate the observed factor model using regression to obtain:

$$\hat{B} = \sum_{t=1}^{T} \bar{R}_t \bar{G}_t' \left(\sum_{t=1}^{T} \bar{G}_t \bar{G}_t' \right)^{-1},$$

   with $\bar{G}_t = G_t - \bar{G}$, $G = \frac{1}{T} \sum_{t=1}^{T} G_t$, $\bar{R}_t = R_t - \bar{R}$ and $\bar{R} = \frac{1}{T} \sum_{t=1}^{T} R_t$.

2. Regress the average returns, $\bar{R}$, on the vector of constants ι_n and the estimated β's, to obtain estimates of the zero-β return and the risk premia:

$$\begin{pmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_F \end{pmatrix} = \left[(\iota_n \vdots \hat{B})'(\iota_n \vdots \hat{B}) \right]^{-1} (\iota_n \vdots \hat{B})' \bar{R}.$$

The observed factor model has to capture the factor structure of the portfolio returns for the FM procedure to be reliable.

Consider the (infeasible) linear regression model for the unobserved factors F_t that uses the observed factors G_t as explanatory variables:

$$F_t = \mu_F + \delta G_t + V_t$$
$$\delta = V_{FG} V_{GG}^{-1}.$$

We substitute it in the observed factor model to obtain:

$$R_t = \mu_R + \beta\mu_F + \beta\delta G_t + \beta V_t + \varepsilon_t = \mu + \beta\delta G_t + U_t,$$

with $\mu = \mu_R + \beta\mu_F$, $U_t = \beta V_t + \varepsilon_t$.

When δ is small or zero, V_t is large and implies an unexplained factor structure in the residuals U_t.

A small value of δ also implies that the estimand of $\hat{B}$, i.e. $\beta\delta$, is small.

The traditional results for the FM two pass procedure assume that the estimand of is a full rank matrix so both β and δ need to be of full rank.

For many of the observed (macro-) economic factors used in the literature δ is such that we cannot reject that it is close to zero.

This is shown by the F-statistics in the previous table.

When δ has a full rank value, these F-statistics are all proportional to the sample size.

The assumption of a full rank value of δ is therefore not supported by the data for factors other then the FF factors.

A more appropriate assumption is to assume a value of δ that leads to the smallish values of the F-statistics.

Assumption 1. The parameter δ is drifting to zero:

$$\delta = \frac{d}{\sqrt{T}}$$

with d a fixed full rank matrix.

Assumption 1 implies that the F-statistics in the Table all have non-central χ^2 distributions with finite non-centrality parameters which is in line with their realized values.

The OLS R^2 equals the explained sum of squares over the total sum of squares when we only use a constant term so its expression reads

$$\begin{aligned} R^2_{OLS} &= \frac{\bar{R}' P_{M_{\iota_N}\hat{B}} \bar{R}}{\bar{R}' M_{\iota_N} \bar{R}} \\ &= \frac{\bar{R}' M_{\iota_N} \hat{B} (\hat{B}' M_{\iota_N} \hat{B})^{-1} \hat{B}' M_{\iota_N} \bar{R}}{\bar{R}' M_{\iota_N} \bar{R}} \end{aligned}$$

with

- $P_A = A(A'A)^{-1}A'$
- $M_A = I_N - P_A$.

Theorem 3. Under Assumption 1, the behavior of R^2_{OLS} is in large samples characterized by:

$$R^2_{OLS} \approx \frac{[\beta\lambda_F + \frac{1}{\sqrt{T}}(\beta\psi_{\iota F} + \psi_{\iota\varepsilon})]' P_{M_{\iota_N}(\beta(d+\psi_{VG})+\psi_{\varepsilon G})} [\beta\lambda_F + \frac{1}{\sqrt{T}}(\beta\psi_{\iota F} + \psi_{\iota\varepsilon})]}{[\beta\lambda_F + \frac{1}{\sqrt{T}}(\beta\psi_{\iota F} + \psi_{\iota\varepsilon})]' M_{\iota_N} [\beta\lambda_F + \frac{1}{\sqrt{T}}(\beta\psi_{\iota F} + \psi_{\iota\varepsilon})]}$$

where

- $\psi_{\iota F} = V_{FF}^{\frac{1}{2}} \psi^*_{\iota F}$
- $\psi_{\iota\varepsilon} = V_{\varepsilon\varepsilon}^{\frac{1}{2}} \psi^*_{\iota\varepsilon}$
- $\psi_{VG} = V_{VV}^{\frac{1}{2}} \psi^*_{VG} V_{GG}^{-\frac{1}{2}}$
- $\psi_{\varepsilon G} = V_{\varepsilon\varepsilon}^{\frac{1}{2}} \psi^*_{\varepsilon G} V_{GG}^{-\frac{1}{2}}$

and $\psi^*_{\iota F}$, $\psi^*_{\iota\varepsilon}$, ψ^*_{VG} and $\psi^*_{\varepsilon G}$ are $k \times 1$, $N \times 1$, $k \times m$ and $N \times m$ dimensional random matrices whose elements are independently standard normally distributed.

Corollary 1. When the number of observed and unobserved factors is the same and they are highly correlated, R^2_{OLS} converges to one when the sample size increases.

Corollary 2. When the number of observed factors is less than the number of unobserved factors but the observed factors explain the unobserved factors well, so d is a large full rank rectangular $k \times m$ dimensional matrix with $m < k$, R^2_{OLS} converges to

$$\frac{\lambda_F' \beta' P_{M_{\iota_N}\beta d} \beta \lambda_F}{\lambda_F' \beta' M_{\iota_N} \beta \lambda_F}.$$

Corollaries 1 and 2 are also discussed in Lewellen *et. al.* (2010).

The cases for which Lewellen *et. al.* (2010) do not provide any analytical results are those where:

1. the observed factors are only minorly correlated with the unobserved factors and
2. when only a few of the observed factors are correlated with the unobserved factors and the number of correlated observed factors is less than the number of unobserved factors.

These cases are important since they are reminiscent of the results shown in the previous table.

The expression in the second case is similart to the one in the first case which is stated in Theorem 3.

The limiting behavior of R^2_{OLS} for cases 1 and 2 is such that only the numerator is random since the denominator of R^2_{OLS} converges to its population value.

The numerator consists of the projection of

$$M_{\iota_N}[\beta\lambda_F + \frac{1}{\sqrt{T}}(\beta\psi_{\iota F} + \psi_{\iota\varepsilon})] \qquad \text{on} \qquad M_{\iota_N}(\beta(d + \psi_{VG}) + \psi_{\varepsilon G}).$$

The first element of the part where you project on, *i.e.* $M_{\iota_N}\beta(d + \psi_{VG})$, is tangent to $M_{\iota_N}\beta(\lambda_F + \frac{1}{\sqrt{T}}\psi_{\iota F})$ since both are linear combinations of $M_{\iota_N}\beta$.

This implies that the numerator of R^2_{OLS} is big whenever $M_{\iota_N}\beta(d + \psi_{VG})$ is relatively large compared to $M_{\iota_N}\psi_{\varepsilon G}$ regardless of whether this results from a large value of d or not.

When the observed proxy factors G_t explain the unobserved factors well, d is large and V_t is small so there is no unexplained factor structure in the residuals of U_t.

When we use factors other than the FF factors, d is small and V_t often explains more than ten times as much of the variation in F_t than the observed proxy factors G_t.

This implies that d is small relative to ψ_{VG}. The resulting unexplained factor structure is then such that $\beta\psi_{VG}$ is large relative to $\psi_{\varepsilon G}$.

Taken together, large values of R^2_{OLS} result from the projection of $M_{\iota_N}\beta(\lambda + \frac{1}{\sqrt{T}}\psi_{\iota F})$ on $M_{\iota_N}\beta\psi_{VG}$ since $M_{\iota_N}\beta\psi_{VG}$ is large compared to both $M_{\iota_N}\beta d$ and $M_{\iota_N}\psi_{\varepsilon G}$.

Hence, it is the estimation error of $\hat{B}$ that leads to the large values of R^2_{OLS} when d is small.

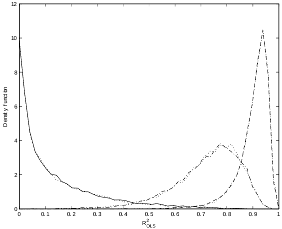
These large values of R^2_{OLS} are then not indicative of the strength of the relationship between expected portfolio returns and observed proxy factors.

Simulation experiment

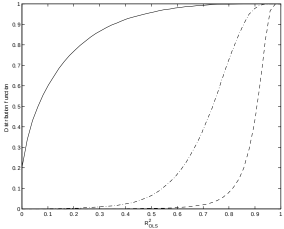
We conduct a simulation experiment calibrated to the data from Lettau and Ludvigson (2001).

We use the FM two pass procedure to estimate the risk premia on the three FF factors using returns on twenty-five size and book to market sorted portfolios from 1963 to 1998.

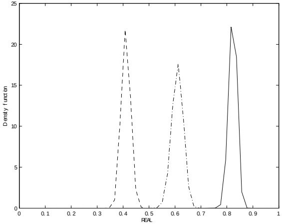
We then generate portfolio returns from the factor model with the estimated values of β, λ_0 and λ_F as the true values and factors F_t and disturbances ε_t that are generated as i.i.d. normal with mean zero and covariance matrices $\hat{V}_{FF}$ and $\hat{V}_{\varepsilon\varepsilon}$ with $\hat{V}_{FF}$ the covariance matrix of the three FF factors and $\hat{V}_{\varepsilon\varepsilon}$ the residual covariance matrix that results from regressing the portfolio returns on the three FF factors.



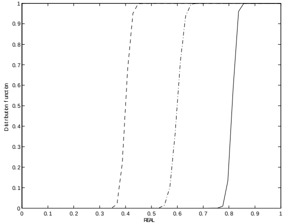
Density functions of R^2_{OLS}



Distribution functions of R^2_{OLS}

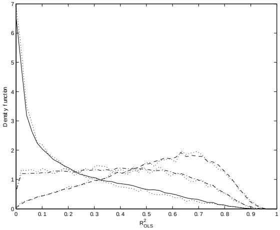


Density functions FACCHECK

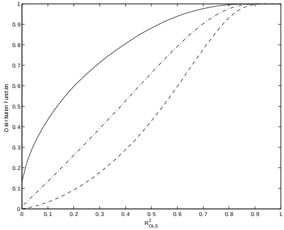


Distribution functions FACCHECK

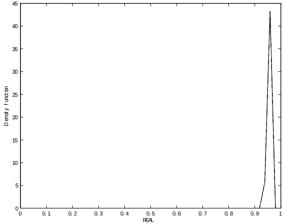
Panel 1. Using one, two or three of the true factors.



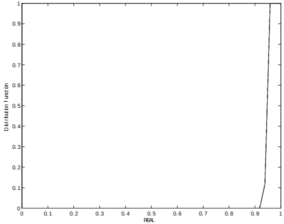
Density functions of R^2_{OLS}



Distribution functions of R^2_{OLS}

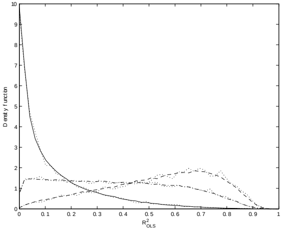


Density functions FACCHECK

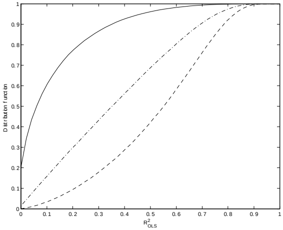


Distribution functions FACCHECK

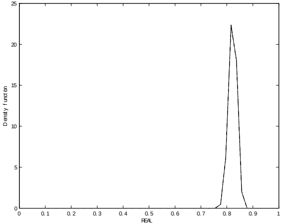
Panel 2. Using one, two or three irrelevant factors.



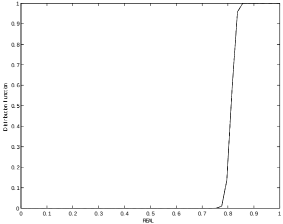
Density functions of R^2_{OLS}



Distribution functions of R^2_{OLS}

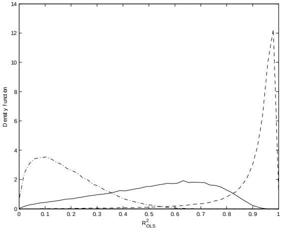


Density functions FACCHECK

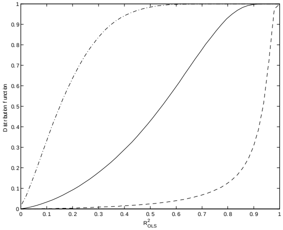


Distribution functions FACCHECK

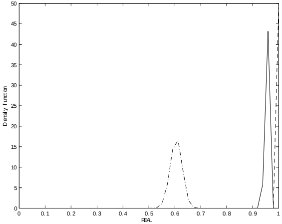
Panel 3. One true factors and zero, one or two irrelevant factors.



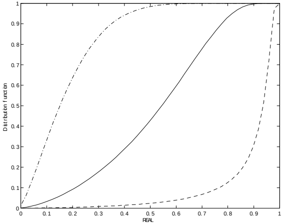
Density functions of the R^2



Distribution functions of the R^2



Density functions FACCHECK



Distribution functions FACCHECK

Panel 4. Three irrelevant factors and weak, strong and very strong factor str

	LL01					
	R_{vw}	Δc	FF factors	$cay, R_{vw},$ $cayR_{vw}$	$cay, \Delta c,$ $cay\Delta c$	$cay, R_{vw}, \Delta y$ $cayR_{vw}, cay\Delta$
R^2_{OLS}	0.01	0.16	0.80	0.31	0.70	0.77
FACCHECK	82.1%	95.5%	38.2%	82.5%	95.2%	82.1%
pseudo-R^2	0.78	0.016	0.95	0.78	0.10	0.79

For many of the specifications stated that have high values of R^2_{OLS} also the factor structure check is large.

The pseudo-R^2 of these regressions on top of the one that results from only using the value weighted return are also small.

The large values of R^2_{OLS} then result from the estimation error in the estimated β's of the observed proxy factors and are not indicative of a relationship between expected portfolio returns and observed factors

GLS R^2. The GLS R^2 equals the explained sum of squares over the total sum of squares in a GLS regression where we weight by the inverse of the covariance matrix of $\bar{R}$:

$$R^2_{GLS} = \frac{\bar{R}'\bar{M}\hat{B}(\hat{B}'\bar{M}\hat{B})^{-1}\hat{B}'\bar{M}\bar{R}}{\bar{R}'\bar{M}\bar{R}} = \frac{(V_{RR}^{-\frac{1}{2}}\bar{R})'P_{M_{V_{RR}^{-\frac{1}{2}}\iota_N}V_{RR}^{-\frac{1}{2}}\hat{B}}(V_{RR}^{-\frac{1}{2}}\bar{R})}{(V_{RR}^{-\frac{1}{2}}\bar{R})'M_{V_{RR}^{-\frac{1}{2}}\iota_N}(V_{RR}^{-\frac{1}{2}}\bar{R})}$$

with $\bar{M} = V_{RR}^{-1} - V_{RR}^{-1}\iota_N(\iota_N' V_{RR}^{-1}\iota_N)^{-1}\iota_N' V_{RR}^{-1}$.

Alongside the explanatory power of the observed proxy factors, the large sample distribution of R^2_{GLS} crucially depends on the scaled risk premia of the unobserved true factors:

$$\left(\frac{V_{FF}}{T}\right)^{-\frac{1}{2}}\lambda_F.$$

When we use the FF factors, the relative size of these size premia is small and proportional to the realization of a standard normal random variable.

	LL01		
$\lambda_F'\left(\frac{V_{FF}}{T}\right)^{-1}\lambda_F$	24.1		
	λ_F	$V_{FF}^{-\frac{1}{2}}\lambda_F$	$\left(\frac{V_{FF}}{T}\right)^{-\frac{1}{2}}\lambda_F$
R_{VW}	1.32	0.22	2.62
SMB	0.47	0.024	0.28
HML	1.46	0.35	4.14

Assumption 2. The scaled risk premia $\left(\frac{V_{FF}}{T}\right)^{-\frac{1}{2}} \lambda_F$ remain constant when the sample size increases so

$$\left(\frac{V_{FF}}{T}\right)^{-\frac{1}{2}} \lambda_F = l,$$

with l a k dimensional fixed vector, for different values of T.

Theorem 4. Under Assumptions 1, 2, the behavior of R^2_{GLS} is in large samples characterized by:

$$R^2_{GLS} \approx \frac{\left\{\begin{pmatrix} W'I \\ 0 \end{pmatrix} + \psi^*\right\}' P_{M_{V_{RR}^{-\frac{1}{2}}\iota_N}\left\{\begin{pmatrix} W'V_{FF_0}^{-\frac{1}{2}} dV_{GG}^{\frac{1}{2}} \end{pmatrix} + \varphi^*\right\}} \left\{\begin{pmatrix} W'I \\ 0 \end{pmatrix} + \psi^*\right\}}{\left\{\begin{pmatrix} W'I \\ 0 \end{pmatrix} + \psi^*\right\}' M_{V_{RR}^{-\frac{1}{2}}\iota_N} \left\{\begin{pmatrix} W'I \\ 0 \end{pmatrix} + \psi^*\right\}},$$

where the elements of ψ^* and φ^* have independent standard normal distributions, W is an orthonormal $k \times k$ dimensional matrix which contains the eigenvectors of

$$\left[(\beta'\beta)^{\frac{1}{2}\prime} V_{FF} (\beta'\beta)^{\frac{1}{2}} + (\beta'\beta)^{-\frac{1}{2}\prime} \beta' V_{\varepsilon\varepsilon} \beta (\beta'\beta)^{-\frac{1}{2}}\right]$$

and $\beta_\perp$ the $N \times (N-k)$ orthogonal complement of β, $\beta_\perp'\beta \equiv 0$, $\beta_\perp'\beta_\perp \equiv I_{N-k}$, $M_A = I_N - A(A'A)^{-1}A'$.

The large sample behavior of R^2_{GLS} in Theorem 4 differs considerably from that of R^2_{OLS}.

Corollary 1 states that R^2_{OLS} converges to one when the observed factors explain the unobserved factors well and their numbers are the same.

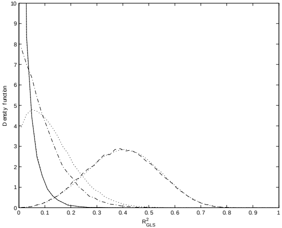
Because $W'I$ is of the same order of magnitude as the standard normal random variables in ψ^*, this is not the case for R^2_{GLS}.

Only when the scaled risk premia are very large, R^2_{GLS} is approximately equal to one.

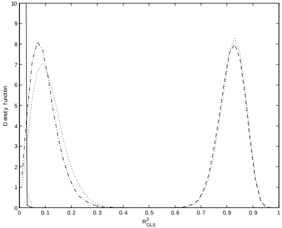
Simulation experiment We use our previous simulation experiment, calibrated to data from Lettau and Ludvigson (2001), to further illustrate the properties of R^2_{GLS} and the accuracy of the expression of the large sample distribution.

We use the data generating process that corresponds with the estimated factor model which uses the three FF factors and their risk premia. We use the true factors or irrelevant ones.

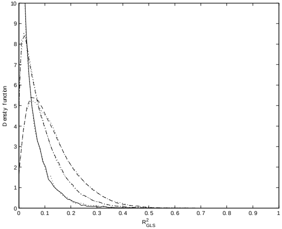
We use the estimated risk premia or ten times the estimated risk premia to show the sensitivity with respect to the estimated risk premia.



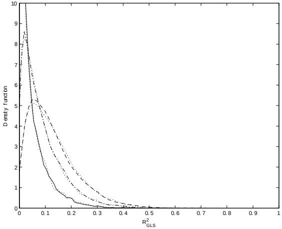
True factors, standard premia



True factors, large premia



Irrelevant factors, standard premia



Irrelevant factors, large premia

Panel 5. One factor (solid), two factors (dash-dot), three factors (dashed).

Tests on the risk premia

The FM two pass t-statistic has a standard normal distribution in large samples when the null hypothesis holds and the true values of β and δ are both full rank matrices.

If the full rank assumptions do not hold, for example, since δ is close to zero, the large sample distribution of the FM two pass t-statistic is not normal and we cannot use the FM two pass t-statistic to conduct inference, see Kleibergen (2009).

The problem of small values of δ for inference on the risk premia is analogous to the weak instrument problem for the linear instrumental variables regression model in econometrics.

So we can use the statistics developed in that literature whose large sample distributions are not affected by the weak instrument problem/small δ problem, see *e.g.* Anderson and Rubin (1949), Kleibergen (2002), Moreira (2003) and Kleibergen and Mavroeidis (2009).

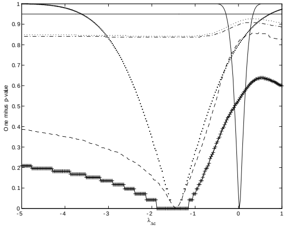
Identification robust confidence sets

We compute confidence sets using the identification robust factor statistics which are discussed in Kleibergen (2009) which are:

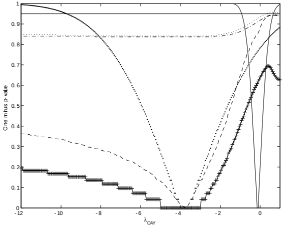
1. the factor Anderson-Rubin (FAR) statistic
2. the factor extension of Kleibergen's (2002, 2005) Lagrange multiplier statistic (FKLM)
3. the factor extension of Kleibergen's (2005) J-statistic (FJKLM)
4. the factor extension of Moreira's (2003) conditional likelihood ratio statistic (FCLR).

If we want to test a hypothesis on one element of λ_F, say $\text{H}_0 : \lambda_1 = \lambda_{1,0}$, these identification robust statistics construct a 95% confidence set for λ_1 by specifying a grid of s different values for $\lambda_{1,0}$, $(\lambda^1_{1,0} \dots \lambda^s_{1,0})$.

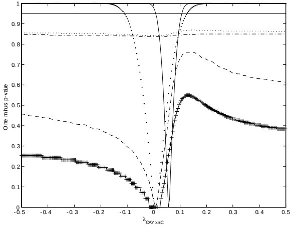
We then compute the statistics for each different value of $\lambda_{1,0}$ in the grid. The 95% confidence set consists of all values of $\lambda_{1,0}$ for which the statistic is below its 95% critical value.



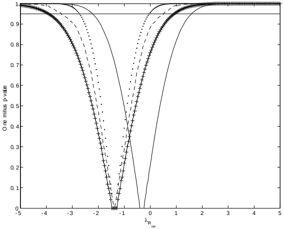
4.1 p-value plot $\lambda_{\Delta c}$



4.2 p-value plot λ_{cay}



4.3 p-value plot $\lambda_{\Delta c \times cay}$



4.4 p-value plot $\lambda_{r_{vw}}$

Lettau and Ludvigson (2001). 4.1-4.3: three factors. 4.4 single factor. FM t
MLE t (points), FKLM (solid-plusses), FCLR (dashed), FJKLM (dash-dotted

For the specification where the value weighted return is the only factor, the FM two pass and ML estimates are close to each other.

This results since the parameters of the value weighted return in the first pass regression are highly significant

Hence, the risk premium on the value weighted return is well identified so all the p-value plots are rather similar.

Conclusions

- In many empirical studies the reported R^2_{OLS} is spurious since it results from the estimation error in the estimated β's
- The R^2_{OLS} in these studies is therefore not indicative of a relationship between expected portfolio returns and the observed proxy factors
- An easy diagnostic to assess the validity of the R^2_{OLS} as a measure of the strength of the relationship between expected portfolio returns and the observed proxy factors is based on the factor structure in the first pass residuals
- R^2_{GLS} is in general small both in case of strong and weakly correlated observed factors and therefore also not a reliable measure for the strength of the relationship between expected portfolio returns and the observed proxy factors