UvA ECONOMETRICS

Discussion Paper: 2010/12

Reality checks for and of factor pricing

Frank Kleibergen

www.feb.uva.nl/ke/UvA-Econometrics

Amsterdam School of Economics
Department of Quantitative Economics
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Reality checks for and of factor pricing

Frank Kleibergen*

December 24, 2010

Abstract

We propose two reality checks to gauge factor pricing. We do so since the typically used R^2 and t-statistics from the second pass of the Fama and MacBeth (1973) two pass procedure signal factor pricing when the observed factors miss the unobserved factor structure of portfolio returns. A large number of (macro-economic) factors that are commonly used, like, for example, consumption and labor income growth, housing collateral, consumption-wealth ratio, labor income-consumption ratio, interactions of either one of the latter three with other factors, fail to capture any of the unobserved factor structure in the portfolio returns so the R^2's and t-statistics reported for (the risk premia on) them are spurious. The reality checks reveal the (un)reliability of these statistics. The first reality check shows if there is any factor structure left in the residuals from the first pass of the Fama-MacBeth two pass procedure by computing the percentage of the variation of the residuals that is explained by their three largest principal components. If this percentage is large, the R^2 of the second pass regression is spuriously large so we cannot interpret it. The second reality check computes 95% confidence sets for the risk premia on the factors using statistics whose large sample distributions remain valid when the observed factors miss the unobserved factor structure. These confidence sets show, since they are then unbounded, that we cannot determine risk premia on observed factors that do not capture any of the unobserved factor structure of the portfolio returns.

1 Introduction

An important part of the asset pricing literature is concerned with the relationship between portfolio returns and (macro-) economic factors. Support for such a relationship is often established using linear factor models, see *e.g.* Fama and French (1992,1993,1996), Jagannathan and Wang (1996,1998), Lettau and Ludvigson (2001), Lustig and Van Niewerburgh (2005), Li *et. al.* (2006), Santos and Veronesi (2006) and Yogo (2006). Linear factor models postulate a linear relationship between expected portfolio returns and the β's of the (macro-) economic factors, see *e.g.* Lintner (1965), Fama and MacBeth (1973), Gibbons (1982), Shanken (1992) and Cochrane (2001). This relationship is typically estimated using the Fama-MacBeth (FM) two pass procedure, see Fama and MacBeth (1973). The FM two pass procedure first estimates the β's by regressing the portfolio returns on the (macro-) economic factors. The second pass then obtains the estimates of the risk premia on the factors by regressing the average portfolio returns on the estimated β's from the first pass. The R^2 of the second pass

*Department of Economics, Brown University, Box B, Providence, RI 02912, United States, Email: Frank_Kleibergen@brown.edu. Homepage: http://www.econ.brown.edu/fac/Frank_Kleibergen. I thank Zhaoguo Zhan for research assistance.

regression and the t-statistics of the risk premia are used to indicate a relationship between expected asset returns and the involved factors.

Portfolio returns exhibit a (unobserved) factor structure, see *e.g.* Merton (1973), Ross (1976), Roll and Ross (1980) and Chamberlain and Rothschild (1983). The presence of a factor structure is typically established using the principal components (or characteristic roots of the covariance matrix) of the portfolio returns, see *e.g.* Anderson (1984, Chap 11). The number of factors equals the number of principal components that are distinctly larger than the other principal components. This leads to three factors for portfolio returns which explain around ninety five percent of the variation of the portfolio returns when we consider twenty-five portfolios and around eighty five percent of the variation when we consider one hundred portfolios. These factors are, however, unobserved so linear factor models replace them with observed factors. Given the pervasiveness of the factor structure, it is of the utmost importance that the linear factor model properly captures the (unobserved) factor structure. We measure the extent to which the observed factors eradicate the factor structure using the percentage of the variation of the residuals from the first pass of the FM two pass procedure that is explained by their three largest principal components. This reality check shows that the Fama-French factors, see Fama and French (1992, 1993), eradicate the factor structure but many other factors proposed in the literature do not capture any of the (unobserved) factor structure. For example, none of the factors other then the Fama-French factors that are proposed in Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Lustig and Van Niewerburgh (2005), Li *et. al.* (2006), Santos and Veronesi (2006) and Yogo (2006), capture any of the (unobserved) factor structure. When we use (some of) these factors in the first pass of the FM two pass procedure, the three largest principal components of the residuals are almost identical to the three largest principal components of the original portfolio returns so these factors do not capture any of the (unobserved) factor structure. It is therefore doubtful if they represent any risk characteristics as well.

The novel macro-economic factors, which we refer to as those not included in the Fama-French factors, are proposed largely on the basis of the R^2 of the second pass regression and the t-statistics associated with their risk premia. We reveal that, because of the (unobserved) factor structure of the portfolio returns, both of these criteria show support for factors that do not capture the (unobserved) factor structure of the portfolio returns. Since the findings from the R^2 and FM t statistics reinforce one another, usage of these criteria easily leads to inclusion of factors that are irrelevant.

We show that the R^2 of the second pass regression is spurious when there is a (unobserved) factor structure but the observed factors fail not capture it by computing the distribution of the R^2 for such instances. This distribution lies at rather high values. For instance, there is a sixty percent probability of a R^2 that exceeds fifty percent when there are three (unobserved) factors but we happen to use three unrelated ones. The distribution of our percentage of variation reality check, *i.e.* the percentage of the variation of the first pass residuals that is explained by their three largest principal components, then lies around ninety five percent so it shows that the observed factors have completely missed the (unobserved) factor structure. It thus shows that the high R^2's are spurious. This explains the high R^2's from the second pass regressions that are reported in Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Lustig and Van Niewerburgh (2005), Li *et. al.* (2006), Santos and Veronesi (2006) and Yogo (2006), using factors that differ from the Fama-French factors since for all these settings we also find large values for our percentage of variation reality check. Hence, the R^2's reported in these studies are spurious. It shows that the R^2 cannot be interpreted without a measure that reflects the remaining factor structure in the residuals of the first pass regression like, for example, our percentage of variation reality check. If this reality check shows that there is no factor structure left in the residuals, we can interpret the R^2 but not if there remains a strong factor structure.

The distribution of the R^2 of the second pass regression is also analyzed in Lewellen *et. al.* (2009). They show that if the factors of the linear factor model are correlated with the true unobserved factors that the distribution of the R^2 lies at a rather high value. This results because although the factors are not priced themselves they are correlated with the factors which are priced and can therefore replace them in the linear factor model. We obtain a rather different, albeit not contradictory, result and show that, in the presence of an unobserved factor structure for the portfolio returns, the distribution of the R^2 lies at a high value when there is no correlation between the factors in the linear factor model and the true unobserved factors. Hence, when the observed factors completely miss the factor structure in the portfolio returns, the R^2 of the second pass regression is spuriously large. We can therefore only sensibly interpret the R^2 if there is no factor structure left in the residuals from the first pass of the FM two pass procedure and we propose our reality check to verify this.

When the factors of the linear factor model fail to capture the factor structure of the portfolio returns, the β's from the first pass regression are relatively small. This makes the standard normal approximation of the large sample distribution of the FM t-statistic from the second pass invalid since it only applies for sizeable values of the β's, see Kleibergen (2009). Hence, the statistical significance of the FM t-statistic when the factors of the linear factor model fail to capture the factor structure is anomalous. In Kleibergen (2009) statistics for testing the risk premia are proposed that remain trustworthy when the β's are small or zero. These statistics are extensions of the identification robust statistics proposed for the instrumental variables regression model and the generalized method of moments, see *e.g.* Anderson and Rubin (1949), Kleibergen (2002, 2005), Kleibergen and Mavroeidis (2009) and Moreira (2003). We use these statistics to propose a second reality check of factor pricing, the 95% confidence sets of the risk premia that result from them. These 95% confidence sets accurately show the amount of information that we have on the risk premia even when the factors fail to capture the factor structure. In the latter case, these confidence sets are unbounded which shows that, as expected, we learn little about the risk premia. We find such unbounded 95% confidence sets for the risk premia on the factors proposed in Lettau and Ludvigson (2001), Lustig and Van Niewerburgh (2005), Li *et. al.* (2006), Santos and Veronesi (2006) and Yogo (2006). Except for the risk premia on the Fama-French factors, we find all 95% confidence sets to be unbounded. When we use the Fama-French factors, we find that our second reality check shows that the hypothesis of factor pricing is rejected which does not happen when we use the other factors. This occurs because of the smaller covariance matrix that results when we use the Fama-French factors, since they capture the factor structure, which increases the test statistics so they are significant. Rejection of factor pricing using the Fama-French factors is also reported in Lettau and Ludvigson (2001).

The paper is organized as follows. In the second section we discuss the (missed) factor structure in the portfolio returns and our first reality check based upon the percentage of the variation of the residuals from the first pass of the FM two pass procedure that is explained by their three largest principal components. The third section analyzes the implications of the missed factor structure for the R^2 and the FM t-statistic. It also proposes our second reality check based upon the 95% confidence sets that result from the identification robust factor statistics. The fourth section concludes.

2 Factor Model for Portfolio returns

Portfolio returns exhibiting a (unobserved) factor structure with k factors result from a statistical model that is characterized by, see *e.g.* Merton (1973), Ross (1976), Roll and Ross (1980) and Cham-

	LL01	JW96	F52-01
1	2720	3116	2434
2	113.8	180.2	140.5
3	98.6	80.6	108.9
4	18.36	28.5	26.7
5	17.61	25.4	19.9
6	13.48	16.2	14.0
7	12.11	14.8	11.6
8	9.31	14.0	10.9
9	8.42	12.6	9.92
10	7.25	12.1	8.18
11	6.02	12.1	7.19
12	5.4	11.4	6.32
13	4.9	11.3	6.17
14	4.38	11.1	5.63
15	4.26	10.8	5.21
16	3.93	10.3	5.02
17	3.5	10.2	4.4
18	3.39	9.9	3.83
19	3.02	9.6	3.43
20	2.71	9.5	2.9
21	2.5	9.2	2.79
22	2.18	9.0	2.75
23	1.74	8.9	2.47
24	1.46	8.7	2.12
25	0.93	8.4	1.77
$REAL_1$	95.5%	86%	94.3%

Table 1: Largest twenty five principal components (in descending order) of the covariance matrix of the portfolio returns (LL01 stands for Lettau and Ludvigson (2001), JW96 stands for Jagannathan and Wang (1996) and F52-01 stands for the portfolio returns from Ken French's website from 1952-2001). REAL1 equals the percentage of the variation explained by the three largest principal components.

berlain and Rothschild (1983):

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

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 and N and T the number of portfolios and time periods. We can reflect the factor model in (1) as well using vector notation:

$$R_t = \mu + \beta F_t + \varepsilon_t, \tag{2}$$

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

$$\beta = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \cdots & \beta_{Nk} \end{pmatrix}. \tag{3}$$

The vector notation of the factor model in (2) shows that, if the factors are i.i.d. with finite variance and are uncorrelated with the disturbances which are i.i.d. with finite variance as well, the covariance matrix of the portfolio returns reads

$$V_{RR} = \beta V_{FF} \beta' + V_{\varepsilon\varepsilon}, \tag{4}$$

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 respectively.

The covariance matrix of the factors, V_{FF}, is considered to be much larger (in a matrix sense) than the covariance matrix of the disturbances, $V_{\varepsilon\varepsilon}$. This assumption allows us to identify the number of factors using principal components analysis, see *e.g.* Anderson (1984, Chap 11). When we construct the spectral decomposition of the covariance matrix of the portfolio returns,

$$V_{RR} = P \Lambda P', \tag{5}$$

with $P = (p_1 \ldots p_N)$ the $N \times N$ orthonormal matrix of characteristic vectors (or eigenvectors) and Λ the $N \times N$ diagonal matrix of principal components or characteristic roots (eigenvalues) which are in descending order on the main diagonal, the number of factors can be estimated as the number of principal components that are distinctly larger than the other principal components. The literature on selecting the number of factors is vast and contains further refinements of this factor selection procedure and settings with fixed and increasing number of portfolios. We do not contribute to this literature but just use some elements of it to shed light on statistical issues involved with estimating risk premia using the FM two pass procedure.

2.1 Factor structure in observed portfolio returns

We use three different data-sets to discuss the statistical issues that arise because of the factor structure in the portfolio returns for the FM two pass procedure. The first data-set is the one used by Lettau and Ludvigson (2001). The portfolio returns used by Lettau and Ludvigson (2001) consist of quarterly observations 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 $T = 141$. The second data-set results from

Jagannathan and Wang (1996) and their portfolio return series consist of the monthly returns on one hundred size and beta sorted portfolios. Their return series begin in July 1963 and end in December 1990 so $T = 330$ and $N = 100$. The third data-set consists of quarterly returns on twenty-five size and book to market sorted portfolio's and are obtained from Ken French's website. The return series are from the first quarter of 1952 to the fourth quarter of 2001 so $T = 200$ and $N = 25$.

Table 1 lists the (largest) twenty-five principal components[1] of the three different sets of portfolio returns. It is clear from these principal components that there is a rapid decline of the value of the principal components from the largest to the third largest one and a much more gradual decline from the fourth largest one onwards. This indicates that the number of factors is (most likely) equal to three. We use the fraction of the variance of the portfolio returns that is explained by these three largest principal components as our first reality check for the presence of a factor structure. Our first reality check therefore reads

$$\text{REAL}_1 = \frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_1 + \ldots + \lambda_N} \tag{6}$$

with $\lambda_1 > \lambda_2 > \ldots > \lambda_N$ the principal components in descending order. Our reality check, REAL_1, equals 95.5% for the Lettau-Ludvigson (LL01) data, 86% for the Jagannathan and Wang (JW96) data and 94.3% for the French (F52-01) data. Using the statistic proposed in, for example, Anderson (1984, Section 11.7.2), it can be shown that the hypothesis that the three largest principal components explain less than 80% of the variation of the portfolio returns is rejected with more than 95% significance for each of these three data sets.

2.2 Factor models with observed factors

Table 1 shows that there is compelling evidence for a factor structure in portfolio returns. Alongside describing portfolio returns using "unobserved factors" as briefly discussed previously, a large literature exists which explains portfolio returns using observed factors. The observed factors that are used in this literature consist both of asset return based factors and macro-economic factors. The observed factor model is identical to the factor model in (2) but with a value of F_t that is observed and a known value of the number of factors, say m :

$$R_t = \mu + BG_t + \varepsilon_t, \tag{7}$$

with $G_t = (g_{1t} \ldots g_{mt})'$ the m-dimensional vector of observed factors and B the $n \times m$ dimensional matrix that contains the β's of the portfolio returns with the observed factors. In the sequel we discuss the observed factors used in seven different articles: Fama and French (1993), Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Li *et. al.* (2006), Lustig and van Nieuwerburgh (2005), Santos and Veronesi (2006) and Yogo (2006).

Fama and French (1993) show that the variation in portfolio returns can be explained using a factor model with three observed asset return based factors: the return on a value weighted portfolio, a "small minus big" (SMB) factor which consists of the difference in returns on a portfolio consisting of assets with a small market capitalization minus the return on a portfolio consisting of assets with a large market capitalization and a "high minus low" (HML) factor which consists of the difference in the returns on a portfolio consisting of assets with a high book to market ratio minus the return

[1] The data-set from Jagannathan and Wang (1996) consists of one hundred portfolio returns so Table 1, for reasons of brevity, only shows the largest twenty-five principal components.

on a portfolio consisting of assets with a low book to market ratio. We use the portfolio returns of the twenty-five size and book to market sorted portfolio's from Ken French's website to estimate the observed factor model. The three Fama-French factors are obtained from this website as well.

Table 2 shows the principal components of the covariance matrix of the portfolio returns and of the covariance matrix of the residuals that results from the observed factor model with the three Fama-French factors. Table 2 shows that the three largest principal components of the covariance matrix have decreased substantially after incorporating the three Fama-French factors. Our reality check shows that the factors associated with the three largest principal components of the residuals explain 47.5% of their variation while the factors associated with the three largest principal components of the covariance matrix of the portfolio returns explain 94.3% of its variation. This shows that the three Fama-French factors basically eradicate the factors associated with the three largest principal components of the portfolio returns, see also Bai and Ng (2006). After incorporating the Fama-French factors, there is no factor structure left over in the residuals since the principal components of their covariance matrix decline gradually and do not have a specific cutoff after which the decline of the principal components becomes much more gradual.

The principal components of the covariance matrices can be used to test the significance of the parameters associated with the observed factors. The likelihood ratio (LR) statistic for testing the null hypothesis that the parameters associated with the observed factors are all equal to zero, $\text{H}_0 : B = 0$, against the alternative hypothesis that they are unequal to zero, $\text{H}_1 : B \neq 0$, can be specified as

$$\text{LR} = T \left[\sum_{i=1}^{N} \log(\lambda_{i,\text{port}}) - \log(\lambda_{i,\text{res}}) \right], \tag{8}$$

with $\lambda_{i,\text{port}}$, $i = 1, \ldots, N$, the principal components of the portfolio returns and $\lambda_{i,\text{res}}$, $i = 1, \ldots, N$, the principal components of the residuals of the observed factor model. The LR statistic in (8) has a $\chi^2(3N)$ distribution in large samples. The value of the LR statistic using the Fama-French factors stated in Table 2 equals 2064 which is strongly significant since its p-value is 0.000 which is in line with our finding that the Fama-French factors eradicate the factors associated with the largest principal components of the covariance matrix of the portfolio returns.

Alongside the LR statistic that tests the significance of all the parameters associated with the Fama-French factors, Table 2 also lists the LR statistic that tests the significance of the parameters associated only with the SMB and HML factors. The expression for this LR statistic is identical to that in (8) if we replace the principal components of the covariance matrix of the raw portfolio returns, $\lambda_{i,\text{port}}$, with the principal components of the covariance matrix of the residuals of an observed factor model that has the value weighted return as the only factor. This LR statistic is equal to 1285 which is strongly significant if compared with its 95% critical value that results from a $\chi^2(2N)$ distribution so the parameters of the SMB and HML factors are significant.

Jagannathan and Wang (1996) propose a conditional version of the capital asset pricing model which they estimate using three observed factors: the return on a value-weighted portfolio, a corporate bond yield spread and a measure of per capita labor income growth. Table 2 contains the largest twenty-five principal components of the covariance matrices that result for the Jagannathan and Wang (1996) data set. Besides the principal components of the covariance matrix of the raw portfolio returns, Table 2 contains the principal components of the covariance matrix of the residuals of three observed factor models. The first one of these uses the three Fama-French factors (for which we use the SMB and HML factors from Ken French's website), the second one has the value weighted return,

	F52-01		JW96			
	raw	FF factors	raw	FF factors	R_{vw}	JW96 factors
1	2434	54.94	3116	70.5	600.6	594.0
2	140.5	38.87	180.2	50.0	81.24	78.9
3	108.9	22.77	80.6	38.6	48.4	48.0
4	26.7	18.24	28.5	27.3	28.5	28.0
5	19.9	12.47	25.4	16.0	17.1	17.0
6	14.0	10.8	16.2	14.7	14.8	14.8
7	11.6	9.17	14.8	13.9	14.4	14.3
8	10.9	8.61	14.0	13.5	13.7	13.6
9	9.92	7.84	12.6	12.2	12.2	12.2
10	8.18	6.55	12.1	11.7	12.1	12.0
11	7.19	6.43	12.1	11.6	11.8	11.7
12	6.32	5.77	11.4	11.3	11.3	11.2
13	6.17	5.16	11.3	10.7	11.1	11.1
14	5.63	5.09	11.1	10.6	10.8	10.8
15	5.21	4.64	10.8	10.4	10.5	10.5
16	5.02	4.31	10.3	10.3	10.3	10.3
17	4.4	3.88	10.2	9.81	10.2	9.94
18	3.83	3.47	9.9	9.58	9.80	9.71
19	3.43	2.95	9.6	9.38	9.54	9.51
20	2.9	2.83	9.5	9.03	9.18	9.17
21	2.79	2.74	9.2	8.90	9.11	9.11
22	2.75	2.51	9.0	8.88	8.91	8.85
23	2.47	2.19	8.9	8.64	8.70	8.51
24	2.12	1.76	8.7	8.39	8.47	8.44
25	1.77	1.64	8.4	8.15	8.36	8.30
REAL_1	94.3%	47.5%	86%	23%	57%	57%
LR against raw		2064 0.000		2994 0.000	1586 0.000	1845 0.000
LR against R_{vw}		1285 0.000		1408 0.000		259 0.004

Table 2: The largest twenty five principal components (in descending order) of the covariance matrix of the portfolio returns and residuals that result using Fama-French factors (French's website data 1952-2001) and those that result from using the Jagannathan and Wang (1996) data with different observed factors. The likelihood ratio (LR) statistic tests against the indicated specification (p-value is listed below). REAL1 equals the percentage of the variation explained by the largest three principal components.

R_{vw}, as the only factor while the third specification uses all three factors from Jagannathan and Wang (1996).

The principal components associated with the Jagannathan and Wang (1996) data show that the Fama-French factors remove the factor structure in the portfolio returns since the principal components of the covariance matrix of the residuals decline gradually so there are no principal components that are distinctly larger than all the others. This is further reflected by the value of the reality check of 23%. The LR statistics indicate that the parameters of all the Fama-French factors are strongly significant. All these findings are identical to those previously discussed for the quarterly portfolio returns.

If the value weighted return is the only factor in the observed factor model, Table 2 shows that there remains a considerable factor structure in the residuals since the principal components of the covariance matrix of the residuals indicate that there are (probably) still two factors present. This is further indicated by the reality check of 57% which indicates a factor structure in the residuals. The LR statistic in Table 2 shows that the parameters of the value weighted return are highly significant.

If we add the yield spread and labor income factors to the factor model with only the value weighted return, Table 2 shows that the principal components of the covariance matrix of the residuals hardly change. The reality check remains equal to 57% since the principal components only decrease minorly compared with the specification that has the value weighted return as the only factor. It indicates that the yield spread and labor income factors do not capture the two remaining factors present in the residuals that result from the factor model with only the value weighted return. This is further reflected by the LR statistic that tests the significance of the parameters associated with the yield spread and labor income factors. This LR statistic is equal to 259 and results from an expression identical to (8) but using the principal components from the residual covariance matrices that result from factor models with only the value weighted return and all three factors used by Jagannathan and Wang (1996) for which the largest twenty-five are stated in the sixth and seventh columns of Table 2. Although this LR statistic is significant compared with its 95% critical value from a $\chi^2(2N = 200)$ distribution, its p-value is only 0.004. This indicates that the parameters associated with the yield spread and labor income factors are all rather close to zero.

Lettau and Ludvigson (2001) use a number of specifications of an observed factor model to estimate different conditional asset pricing models. The observed factors that they consider are the value weighted return (R_{vw}), the consumption-wealth ratio (cay), consumption growth (Δc), labor income growth (Δy), the Fama-French factors and interactions between the consumption wealth ratio and consumption growth ($cay\Delta c$), the value weighted return ($cayR_{vw}$) and labor income growth ($cay\Delta y$). The principal components of the covariance matrix of the residuals that result from six different observed factor models estimated in Lettau and Ludvigson (2001) are stated in Table 3.

The principal components in Table 3 show that for all specifications of the observed factor model, except the one using the Fama-French factors, the residuals still contain a factor structure. This is reflected by our reality check which is above 82% for all these models. For the observed factor models that do not contain the value weighted return as one of the factors, the reality check equals the one which results for the portfolio returns itself. Hence, there are still three factors present in the residuals of these factor models while there are two factors present in the residuals of the factors models that include the value weighted return (except for the specification that includes all three Fama-French factors).

The likelihood ratio statistics reported in Table 3 test the significance of the parameters associated

	LL01						
	raw	R_{vw}	Δc	FF factors	cay, R_{vw}, $cayR_{vw}$	$cay, \Delta c$, $cay\Delta c$	$cay, R_{vw}, \Delta y$, $cayR_{vw}, cay\Delta y$
1	2720	435	2676	26.5	433	2414	412
2	114	99.5	111	22.3	98.0	105	97.2
3	98.6	26.2	98.6	14.3	26.0	96.0	25.6
4	18.4	18.36	18.1	13.9	17.9	17.9	17.8
5	17.6	13.8	16.8	11.2	12.9	16.7	12.8
6	13.5	12.2	13.5	8.91	12.1	13.0	11.8
7	12.1	9.33	12.0	8.0	9.20	11.9	9.19
8	9.31	8.51	9.28	7.03	8.31	9.24	8.18
9	8.42	7.26	8.37	6.27	7.06	8.24	6.91
10	7.25	6.03	7.12	5.17	5.75	7.09	5.65
11	6.02	5.53	6.01	4.65	5.02	5.61	4.95
12	5.4	4.91	5.34	4.40	4.88	5.27	4.87
13	4.9	4.38	4.86	4.25	4.36	4.86	4.25
14	4.38	4.26	4.38	3.73	3.99	4.34	3.93
15	4.26	3.97	4.18	3.61	3.85	3.89	3.85
16	3.93	3.54	3.93	3.34	3.40	3.85	3.40
17	3.5	3.39	3.44	3.06	3.38	3.35	3.32
18	3.39	3.03	3.30	2.91	2.99	3.12	2.92
19	3.02	2.77	3.01	2.52	2.77	2.98	2.73
20	2.71	2.51	2.69	2.29	2.49	2.68	2.48
21	2.5	2.33	2.46	2.07	2.30	2.42	2.29
22	2.18	1.99	2.17	1.54	1.99	2.16	1.97
23	1.74	1.47	1.73	1.44	1.42	1.71	1.41
24	1.46	1.44	1.46	0.95	1.33	1.38	1.30
25	0.93	0.92	0.92	0.68	0.90	0.91	0.90
Real_1	95.5%	82.1%	95.5%	38.2%	82.5%	95.2%	82.1%
LR against raw		765 0.000	36.4 0.064	1940 0.000	856.4 0.000	128.2 0.001	902.5 0.000
LR against R_{vw}				1175 0.000	91.4 0.000		138.5 0.007
LR against Δc						91.8 0.000	
LR against cay, R_{vw}, $cayR_{vw}$							46.1 0.630

Table 3: The twenty five principal components (in descending order) of the covariance matrix of the portfolio returns and residuals that result using different specifications from Lettau and Ludvigson (2001). The likelihood ratio (LR) statistic tests against the indicated specification (p-value is listed below). REAL1 equals the percentage of the variation explained by the largest three principal components.

with the observed factors. They result from a specification identical to the one in (8) when the principal components are chosen from the appropriate columns in Table 3. The degrees of freedom of their χ^2 large sample distribution equals the number of tested parameters. The likelihood ratio statistics show that only the parameters of the Fama-French factors and the value weighted return are strongly significant. For all the other specifications, the likelihood ratio statistic are always less than twice the number of estimated parameters. This indicates that although the LR statistic might be significant at the 95% significance level, which is not even the case for the consumption growth, the value of the parameters associated with the observed factors are all close to zero.

Li, Vassalou and Xing (2006) use investments made by households (HHOLD), nonfinancial corporate firms (NFINCO) and financial companies (FINAN) as factors in an observed factor model which they then use to estimate a stochastic discount factor model. We estimate this observed factor model using the previously discussed quarterly portfolio returns from French's website.

Table 4 contains the estimation results for the observed factor model from Li *et. al.* (2006). It contains two specifications of the observed factor model, one which uses all three factors and one which only uses the FINAN factor. For both specifications, the principal components of the covariance matix of the residuals are comparable to those of the covariance matrix of the raw portfolio returns. Hence, both specifications do not capture the factor structure of the portfolio returns. This is further reflected by the reality check which for both specifications equals the one of the raw portfolio returns which is 94.3%. The LR statistics reported in Table 2 therefore show that the parameters of the FINAN factors are not significant at the 95% significance level while the parameters of the HHOLD and NFINCO factors are only significant at the 99% significance level. Hence, the parameters of all three factors are close to zero.

Lustig and Van Nieuwerburgh (2005) employ an observed factor model that contains nondurable consumption growth (Δc_{nondur}), a housing collateral ratio (myfa) and the interaction between nondurable consumption growth and the housing collateral ratio ($\Delta c_{nondur}\times$myfa). We estimate this model using the quarterly portfolio returns from French's website. The results are reported in Table 4.

Table 4 contains two specifications of the factor model used by Lustig and Van Nieuwerburgh (2005). One of these specifications uses all three factors while the other one has nondurable consumption growth as the only factor (Δc_{nondur} column). The principal components in Table 4 show that both specifications do not capture the factor structure of portfolio returns since Table 4 shows that the residuals of both of these two models still contain three factors. This is further reflected by the reality check which equals 93.9% for the specification with nondurable consumption growth as the only factor and 93.8% for the specification that uses all three factors. The likelihood ratio statistics that test the significance of the parameters of the consumption growth and the joint significance of the housing collateral and the interaction factor are only significant at the 98% and 59% significance level. This further shows that the empirical support for the factors proposed by Lustig and Van Nieuwerburgh (2005) is rather minor.

Santos and Veronesi (2006) use adaptations of the factors from Lettau and Ludvigson (2001). Alongside the value weighted returns, Santos and Veronesi (2006) use both the consumption-wealth ratio (*cay*), previously used by Lettau and Ludvigson (2001), and a labor income to consumption ratio

	F52-01							
	raw	R_{vw}	$\Delta c_{nondurable}$	FINAN	LVX06	LN05	SV06	Y06
1	2434	465.2	2250	2422	2404	2204	439.7	465.1
2	140.5	140.3	139.5	138.9	137.7	137.0	122.7	139.1
3	108.9	36.7	108.8	105.9	103.9	108.0	36.3	36.1
4	26.7	21.6	26.6	26.6	25.1	26.5	21.5	21.6
5	19.9	16.8	19.9	19.5	19.2	19.7	16.3	16.1
6	14.0	12.3	13.9	13.9	13.8	13.8	12.2	11.9
7	11.6	10.9	11.1	11.6	11.3	11.1	10.9	10.8
8	10.9	9.98	10.8	10.8	10.8	10.8	9.90	9.86
9	9.92	8.24	9.89	9.71	9.70	9.76	8.01	8.19
10	8.18	7.22	8.13	8.14	8.12	8.12	6.91	7.22
11	7.19	6.50	7.19	7.16	7.00	7.16	6.33	6.47
12	6.32	6.19	6.24	6.31	6.31	6.23	6.16	6.15
13	6.17	5.70	6.17	6.17	6.16	6.09	5.64	5.66
14	5.63	5.59	5.61	5.62	5.59	5.55	5.57	5.56
15	5.21	5.03	5.20	5.20	5.19	5.19	4.98	5.01
16	5.02	4.62	5.02	5.02	5.00	4.99	4.57	4.51
17	4.4	4.31	4.39	4.29	4.25	4.37	4.31	4.21
18	3.83	3.62	3.81	3.82	3.77	3.79	3.48	3.53
19	3.43	3.43	3.42	3.40	3.32	3.38	3.37	3.41
20	2.9	2.89	3.88	2.84	2.77	2.84	2.87	2.88
21	2.79	2.79	2.78	2.77	2.74	2.75	2.69	2.77
22	2.75	2.71	2.73	2.73	2.68	2.53	2.60	2.67
23	2.47	2.39	2.46	2.47	2.43	2.45	2.27	2.40
24	2.12	2.12	2.10	2.10	2.00	2.09	2.10	2.10
25	1.77	1.62	1.76	1.73	1.72	1.74	1.61	16.2
largest three roots	94.3%	81.4%	93.9%	94.3%	94.3%	93.8%	80.5%	81.5%
LR against raw		854 0.000	41.9 0.019	35.2 0.085	111 0.004	93.5 0.07	972 0.000	904 0.00
LR against FINAN					75.87 0.011			
LR against $\Delta c_{nondurable}$						51.6 0.41		863 0.000
LR against R_{vw}							118 0.000	50.4 0.46

Table 4: The twenty five principal components (in descending order) of the covariance matrix of the portfolio returns and residuals that result using different specifications from Li et. al. (2006) (LVX06), Lustig and Van Nieuwerburgh (2005) (LN05), Santos and Veronesi (2006) (SV06) and Yogo (2006) (Y06). All use the quarterly portfolio returns from French's website. The likelihood ratio (LR) statistic tests against the indicated specification (p-value is listed below). REAL1 equals the percentage of the variation explained by the largest three principal components.

(s_w) interacted with the value weighted return as factors. We estimate their specification using the portfolio returns from French's website.

Table 4 contains the estimation results for two specifications of the observed factor model used by Santos and Veronesi (2006). The first specification has the value weighted return as the only factor while the other specification uses all three factors. The principal components of the covariance matrix of the residuals that results for the first specification show that the value weighted return removes one factor so the residuals contain two factors. Table 4 shows that if we consider all three observed factors, the residuals still have a factor structure with two factors. This is further reflected by the reality check which equals 81.4% for the specification with the value weighted return as the only factor and 80.5% for the specification with all three factors. Hence, the additional two factors explain little of the variation in the portfolio returns. The likelihood ratio statistic that tests the joint significance of the *cay* and labor-consumption ratios is equal to 118. Although this is a rather significant value of the LR statistic, we have to consider that it tests fifty parameters so it is only twice the number of tested parameters which shows that all these parameters are relatively close to zero. The LR statistic which tests the significance of the parameters of the value weighted return is obviously highly significant.

Yogo (2006) considers a specification of the observed factor model that alongside the value weighted return has consumption growth in durables (Δc_{dur}) and nondurables (Δc_{nondur}) as the three observed factors. We estimate his specification using the portfolio returns from French's website.

Table 4 contains two specifications of the observed factor model from Yogo (2006). The first one has the value weighted return as the only factor and has been discussed previously. The other specification has all three factors. The principal components of the factor model with all three factors show that the additional two consumption growth factors fail to absorb the two factors left over in the residuals from the model with the value weighted return as the only factor. The principal components of the covariance matrices of the residuals of these two models are almost identical which indicates that there are still two factors left in the residuals of the three factor model. This is further reflected by the reality check which equals 81.4% for the specification with only the value weighted return and 81.5% for the specification with all three factors. The likelihood ratio statistic which tests the joint significance of the parameters associated with the two consumption growth factors is also only significant at the 54% significance level so we cannot rule out that the parameters associated with the two consumption growth factors are all equal to zero.

3 Implications of missed factors for the FM two pass procedure

Stochastic discount factor models, see *e.g.* Cochrane (2001), stipulate a relationship between the expected returns on the portfolios and the β's of the portfolio returns with their (unobserved) factors:

$$E(R_t) = \iota_n \lambda_0 + \beta \lambda_F, \tag{9}$$

with ι_n the n-dimensional vector of ones, λ_0 the zero-β return and 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 in (7) by regressing the portfolio returns R_t on the observed factors G_t to obtain the least squares estimator:

$$\hat{B} = \sum_{t=1}^{T} \bar{R}_t \bar{G}_t' \left(\sum_{t=1}^{T} \bar{G}_t \bar{G}_t' \right)^{-1}, \tag{10}$$

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:

$$\begin{pmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_F \end{pmatrix} = \left[(\iota_n \vdots \hat{B})'(\iota_n \vdots \hat{B}) \right]^{-1} (\iota_n \vdots \hat{B})'\bar{R}. \tag{11}$$

The FM two pass procedure uses the least squares estimator that results from the observed factor model to estimate the risk premia. The adequacy of the results that stem from the FM two pass regression therefore crucially hinge on the ability of the observed factor model to capture the factor structure of the portfolio returns. We show this for two statistics that are commonly reported to reflect the validity of the results and the significance of the risk premia: the R^2 of the second pass regression and the t-statistics of the risk premia.

3.1 The R^2 of the second pass regression

It is common practice to measure the explanatory power of a regression using a goodness of fit measure like the R^2. The R^2 of the second pass regression of the FM two pass procedure is therefore often reported for this purpose. It equals the explained sum of squares over the total sum of squares when we only use a constant term so its expression reads

$$R^2 = \frac{\bar{R}' P_{M_{\iota_n}\hat{B}} \bar{R}}{\bar{R}' M_{\iota_n} \bar{R}} = \frac{\bar{R}' M_{\iota_n} \hat{B}(\hat{B}' M_{\iota_n} \hat{B})^{-1} \hat{B}' M_{\iota_n} \bar{R}}{\bar{R}' M_{\iota_n} \bar{R}}, \tag{12}$$

with $\bar{R}$ the average portfolio returns and $P_A = A(A'A)^{-1}A'$, $M_A = I - P_A$ for a full rank matrix A and I the identity matrix. To assess the adequacy of the R^2 to reflect the goodness of fit, we assume that the portfolio returns are generated by the unobserved factor model in (2). The least squares estimator of the observed factors $\hat{B}$ then provides an estimate of

$$\hat{B} \approx \beta V_{FG} V_{GG}^{-1}, \tag{13}$$

with V_{FG} the covariance between the unobserved and observed factors, $V_{FG} = cov(F_t, G_t)$, and V_{GG} is the covariance matrix of the observed factors, $V_{GG} = var(G_t)$. If $\beta V_{FG} V_{GG}^{-1}$ in (13) is a full rank matrix, the number of observed and unobserved factors is identical and $\hat{B}$ is a precise estimator of the expression in (13), so its approximation error is small, the R^2 in (12) is close to one which results if we substitute (13) and that $\bar{R}$ is approximately equal to (9) into the expression of the R^2 in (12):

$$R^2 \approx \frac{\lambda_F' \beta' M_{\iota_N} \beta V_{FG} V_{GG}^{-1} \left[V_{GG}^{-1\prime} V_{FG}' \beta' M_{\iota_N} \beta V_{FG} V_{GG}^{-1} \right]^{-1} V_{GG}^{-1} V_{FG}' \beta' M_{\iota_n} \beta \lambda_F}{\lambda_F' \beta' M_{\iota_n} \beta \lambda_F} = 1. \tag{14}$$

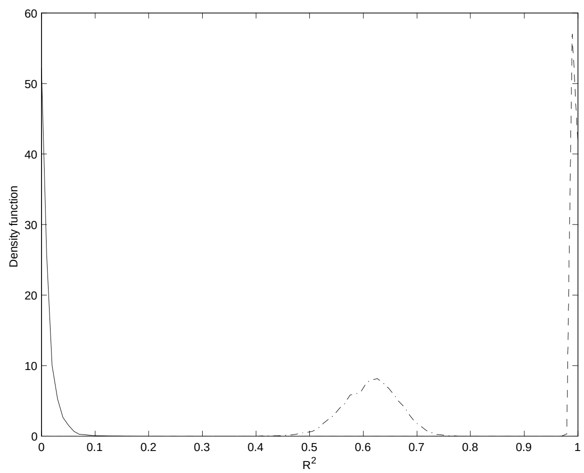

Figure 1.1. Density functions of the R^2

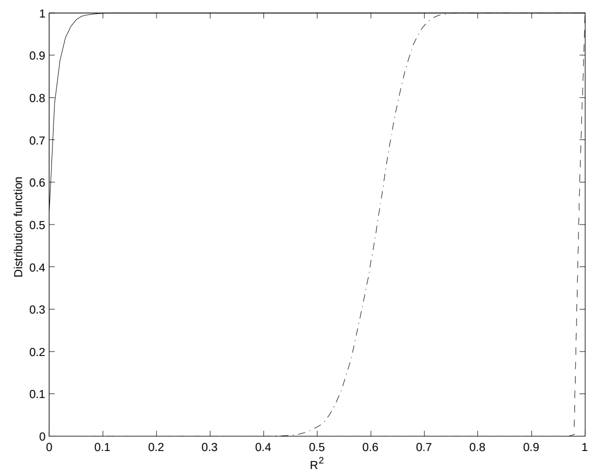

Figure 1.2. Distribution functions of the R^2

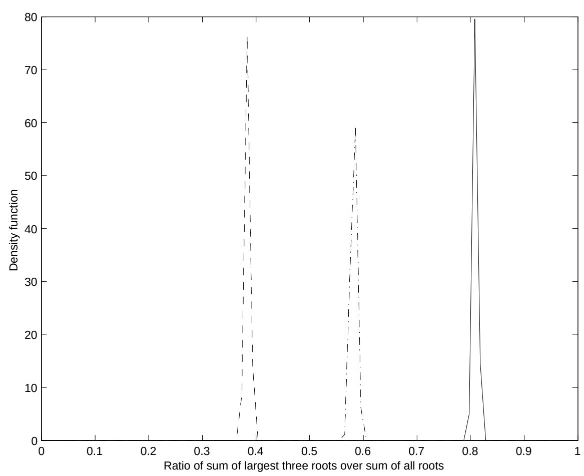

Figure 1.3. Density functions REAL_1

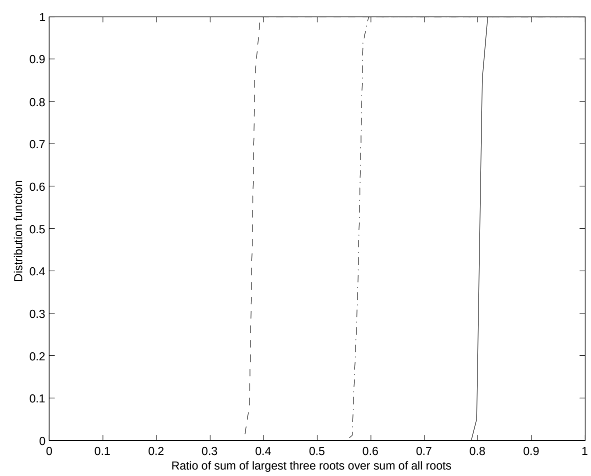

Figure 1.4. Distribution REAL_1

Panel 1. Density and distribution functions of the R^2 and REAL_1 (the ratio of the sum of the three largest principal components of the residual covariance matrix over the sum of all principal components) when we use one of the three factors (solid line), two (dashed-dotted line) and all three (dashed line).

A large value of the R^2 is therefore typically seen as an indication that the observed factors explain a large part of the variation of the average portfolio returns. We obtained the R^2 using the observed factors which are not necessarily the factors that generated the portfolio returns. When the observed factors are correlated with the unobserved true factors, they can serve as proxies for them and explain the average portfolio returns equally well as the unobserved true factors. We just showed the validity of this argument in (14) which is also put forward in Lewellen *et. al.* (2009). Lewellen *et. al.* (2009) also show that when the number of observed factors is less than the number of unobserved true factors but are still correlated with it, the R^2 goes down proportionally.

The argument goes, however, much further since the R^2 is strongly influenced by the (missed) factor structure in the portfolio returns irrespective of the correlation between the observed and unobserved factors. To illustrate this phenomenon, we conduct a simulation experiment calibrated to data from Lettau and Ludvigson (2001). We therefore estimate the risk premia using their data of returns on twenty-five size and book to market sorted portfolios from 1963 to 1998 on the three Fama-French factors. We then generate portfolio returns from the factor model in (2) with the estimated values of β, λ_1 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 Fama-French factors and $\hat{V}_{\varepsilon\varepsilon}$ the residual covariance matrix that results from regressing the portfolio returns on the three Fama-French factors.

We use the simulated portfolio returns to compute the density and distribution functions of the R^2 from (12) using an observed factor G_t that first only consists of the first (observed) factor, then of the first two factors and then of all three factors. The Figures in Panel 1 show that, as expected, the distribution of the R^2 moves to the right when we add an additional true factor. Figures 1.1 and 1.2 also show that the R^2 is close to one when we use all three factors as indicated by (14). It is somewhat surprising though that the R^2 in case we use just one factor lies around zero. To show that the observed factor model only captures part of the factor structure of the portfolio returns, Panel 1 also shows the density and distribution functions of our reality check. These Figures show that when we use only one factor, the three largest roots explain around 81% of the variation which is roughly equal to the 82% that we stated in Tables 3 and 4 when we use the value weighted return as the only factor.[2] The variation explained by the largest three roots decreases to 58% when we use two factors and 38% when we use all three factors. The last percentage is again similar to the percentage in Table 3 when we use all three Fama-French factors.

Panel 2 shows the density and distribution functions that result from another simulation experiment where we simulate from the same model as used previously but now we estimate an observed factor model with only useless factors. We begin with an observed factor model with one useless factor and then add one or two additional useless factors. The density and distribution functions of the R^2 in Figures 2.1 and 2.2 are rather surprising. They dominate the distribution of the R^2 in case we only use one of the true factors. Hence, based on the R^2, observed factors models with useless factors outperform an observed factor model which just has one of the three true factors. It is even such that the R^2 when we use three useless factors often exceeds the R^2 when we use two valid factors. This becomes even more pronounced when we add more useless factors which we do not show. To reveal that the observed factor models with the useless factors do not explain anything, we also computed the density and distribution functions of the reality check. As expected, its' density and distribution functions that result from the three specifications with the useless factors all lie on top of one another at 95% which is identical to the value of the ratio is Tables 2 and 4 when the observed factors matter very little.

[2] We note that the Jagannathan-Wang data contains one hundred portfolio returns so the explained percentage of the variation is not comparable with that which results when we use the value weigthed return as the only factor for the Jagannathan-Wang data.

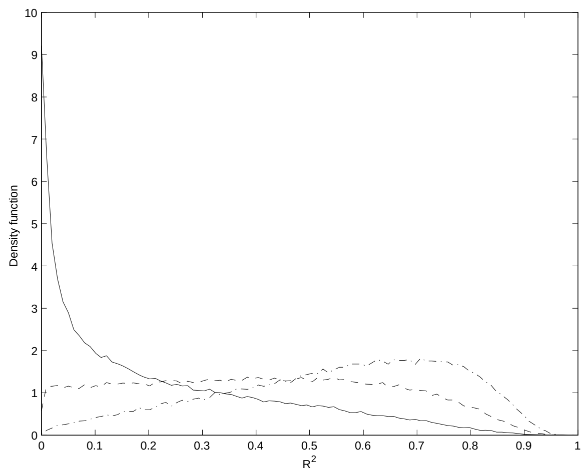

Figure 2.1. Density functions of the R^2

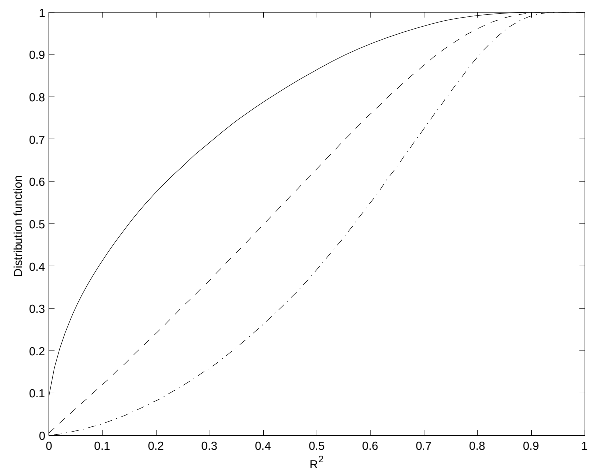

Figure 2.2. Distribution functions of the R^2

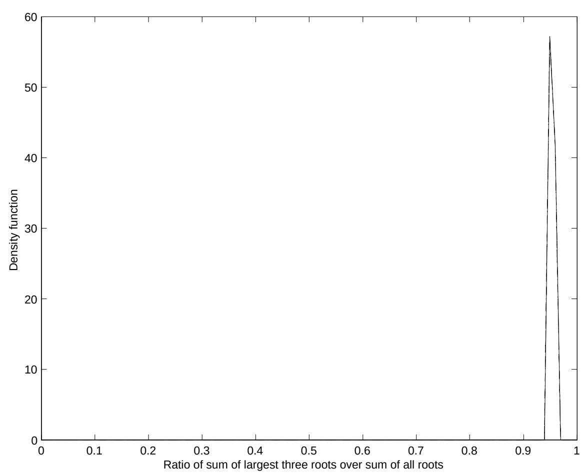

Figure 2.3. Density functions REAL_1

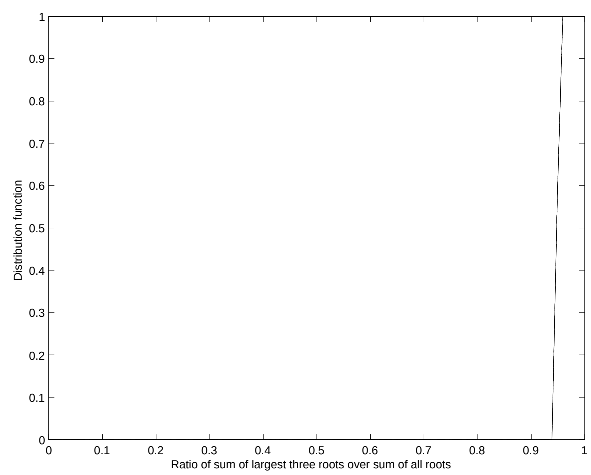

Figure 2.4. Distribution REAL_1

Panel 2. Density and distribution functions of the R^2 and REAL_1 (the ratio of the sum of the three largest principal components of the residual covariance matrix over the sum of all principal components) when we use one useless factor (solid line), two (dashed-dotted line) and three (dashed line).

Panels 1 and 2 show that the distribution of the R^2 is heavily influenced by the (missed) factor structure in the portfolio returns. To further emphasize this, we conducted another simulation experiment where we specifically analyze the influence of the factor structure. We therefore estimate an observed factor model that has three useless factors. To show the sensitivity of the R^2 to the factor structure, we simulate from the same model as used previously but we now use three different settings of the covariance matrix $V_{\varepsilon\varepsilon}$ of the disturbances in the original factor model: $V_{\varepsilon\varepsilon} = 25\hat{V}_{\varepsilon\varepsilon}$ (weak factor structure), $V_{\varepsilon\varepsilon} = \hat{V}_{\varepsilon\varepsilon}$ (factor structure) and $V_{\varepsilon\varepsilon} = 0.04\hat{V}_{\varepsilon\varepsilon}$ (strong factor structure) with $\hat{V}_{\varepsilon\varepsilon}$ the residual covariance matrix that results from regressing the portfolio returns on the three Fama-French factors. The results are reported in Panel 3.

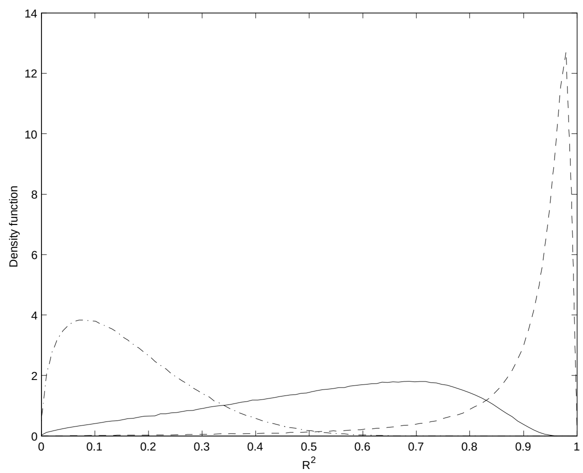

Figure 3.1. Density functions of the R^2

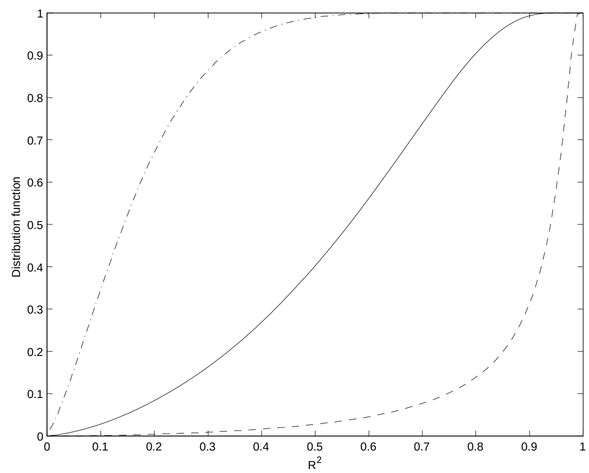

Figure 3.2. Distribution functions of the R^2

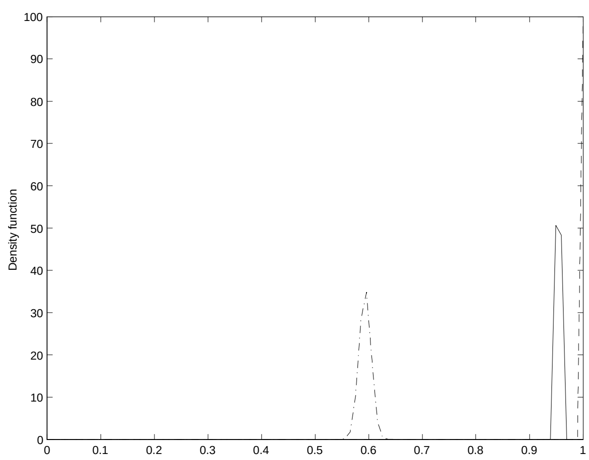

Figure 3.3. Density functions REAL_1

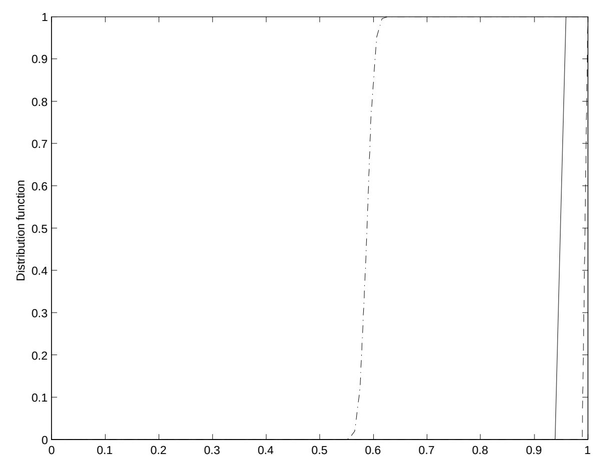

Figure 3.4. Distribution REAL_1

Panel 3. Density and distribution functions of the R^2 and REAL_1 (the ratio of the sum of the three largest principal components of the residual covariance matrix over the sum of all principal components) when we use three useless factors and there is a factor structure (solid line), strong factor structure (dashed line) and weak factor structure (dashed-dotted line).

The Figures in Panel 3 show that the distribution of the R^2 is very sensitive to the remaining factor structure in the residuals. Figures 3.1 and 3.2 show that for the same irrelevant explanatory power of the observed factor model, the R^2 can vary greatly. Figures 3.3 and 3.4 show that for the observed factor models where the R^2 is high in Figures 3.1 and 3.2 that also the factor structure in the residuals is very strong. For the observed factor model where the factor structure in the residuals is rather mild, the density of the R^2 is as expected and close to zero. Hence, for the models where there is still a strong factor structure in the residuals, the R^2 does not represent the quality of the model.

Panels 1-3 show that the R^2 of the second pass regression cannot be interpreted sensibly without

	LL01					
	R_{vw}	Δc	FF factors	cay, R_{vw}, $cayR_{vw}$	$cay, \Delta c$, $cay\Delta c$	$cay, R_{vw}, \Delta y$, $cayR_{vw}, cay\Delta y$
R^2	0.01	0.16	0.80	0.31	0.70	0.77
REAL_1	82.1%	95.5%	38.2%	82.5%	95.2%	82.1%

Table 5: R-squared of the second pass regression of the FM two pass procedure and REAL1 (the percentage of the variation explained by the largest three principal components) for different specifications from Lettau and Ludvigson (2001).

	F52-01						
	R_{vw}	Δc	FINAN	LVX06	LN05	SV06	Y06
R^2	0.07	0.04	0.51	0.58	0.74	0.65	0.54
REAL_1	81.4%	93.9%	94.3%	94.3%	93.8%	80.5%	81.5%

Table 6: R-squared of the second pass regression of the FM two pass procedure and REAL1 (the percentage of the variation explained by the largest three principal components) using the factors from Li et. al. (2006) (LVX06), Lustig and Van Nieuwerburgh (2005) (LN05), Santos and Veronesi (2006) (SV06) and Yogo (2006) (Y06). All use the quarterly portfolio returns from French's website.

some diagnostic statistics, like, for example, our reality check, that report on the remaining factor structure in the residuals from the first pass of the FM two pass procedure. If there is a strong factor structure in the residuals, the R^2 can be spuriously low as is the case with one valid factor in Panel 1 or spuriously high as shown in the cases with the useless factors. Panels 1-3 therefore show that the R^2 is only trustworthy when there is hardly any factor structure left in the residuals.

Tables 5 and 6 report the R^2 of the second pass regression and our reality check, REAL_1, for the specifications in Tables 3 and 4. Many of the specifications stated in Tables 5 and 6 have high values of the R^2. Except for the specification using the Fama-French factors, all of these specifications also have large values of the reality check which indicates that a factor structure remains present in the first pass residuals. We just showed that the R^2 of the second pass regression is spuriously high when the reality check is large. Hence, the high values of the R^2 associated with a large value of the reality check in Tables 5 and 6 are spuriously large as well. Tables 5 and 6 correspond with Lettau and Ludvigson (2001), Li *et. al.* (2006), Lustig and Van Nieuwerburgh (2006), Santos and Veronesi (2006) and Yogo (2006) so the R^2's reported in these papers are spuriously high. They do, because of the remaining factor structure in the first pass residuals, not appropriately reflect the explanatory power of the average returns by the β's of the involved observed factors.

3.2 Tests on the risk premia

We just showed that we can only interpret the R^2 of the second pass regression when there is no factor structure left in the first pass residuals. We therefore analyze the R^2 given a statistic that reflects if such a factor structure remains present. Analyzing the R^2 without such a statistic, for example to compute the confidence set of the R^2 as in Lewellen *et. al.* (2009), is not feasible since we cannot characterize the distribution of the R^2 without knowing the remaining factor structure in the first pass residuals. The R^2 is not the only statistic whose (large sample) distribution crucially depends on the value of some of the other parameters of the observed factor model. Another example of such a statistic is the t-statistic of the risk-premia estimator from the second pass of the FM two pass procedure.

The FM two pass t-statistic testing the hypothesis that the i-th risk premia is equal to $\lambda_{i,0}$, $\mathrm{H}_0 : \lambda_i = \lambda_{i,0}$, reads

$$t(\lambda_i) = \frac{\hat{\lambda}_i - \lambda_{i,0}}{\sqrt{var(\hat{\lambda}_i)}}, \tag{15}$$

with $\hat{\lambda}_F = (\hat{\lambda}_1 \ldots \hat{\lambda}_m)'$ resulting from (11) and $var(\hat{\lambda}_i)$ the variance of the i-th element of $\hat{\lambda}_F$. The variance of $(\hat{\lambda}_0,\ \hat{\lambda}_F)$ equals

$$var(\hat{\lambda}_0,\ \hat{\lambda}_F) = \ \frac{1}{T}\left[(\iota_n \vdots \hat{B})'(\iota_n \vdots \hat{B})\right]^{-1}(\iota_n \vdots \hat{B})'\hat{\Theta}(\iota_n \vdots \hat{B})\left[(\iota_n \vdots \hat{B})'\hat{\Theta}(\iota_n \vdots \hat{B})\right]^{-1}, \tag{16}$$

where $\hat{\Theta} = \hat{V}_{\varepsilon\varepsilon}(1 + \hat{\lambda}_F'(\hat{V}_{GG})^{-1}\hat{\lambda}_F)$, and the variance of the i-th element of $\hat{\lambda}_F$ equals the $(i+1)$-th diagonal element of this covariance matrix, see Shanken (1992). The FM two pass t-statistic has a standard normal distribution in large samples when the null hypothesis holds and the estimand of $\hat{B}$, $\beta V_{FG} V_{GG}^{-1}$ defined in (13), has a full rank value. This implies that both β and V_{FG} have to be full rank matrices. If the full rank assumption does not hold, for example, since V_{FG} is close to zero because the observed and unobserved factors are uncorrelated, 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) and Kan and Zhang (1999). We just showed that many observed factors used in practice fail to eradicate the factor structure in portfolio returns. For these factors, the estimand of $\hat{B}$ is close to zero so the two pass t-statistics of the, with these factors associated, risk premia do not have a standard normal distribution in large samples. Thus we cannot determine if these t-statistics are significant and usage of standard normal critical values leads to erroneous conclusions.

The statistical issues associated with small values of the β's and/or of the covariance between the true and observed factors, V_{FG}, for inference on the risk premia using the FM two pass t-statistic are analogous to the weak instrument problem in the linear instrumental variables regression model in econometrics, see *e.g.* Staiger and Stock (1997). The traditional manner to analyze the linear instrumental variables regression model is also by means of a two stage regression procedure, *i.e.* two stage least squares. There is an extensive literature on weak instruments in econometrics which has led to the development of statistics whose large sample distributions are not affected by the weak instrument problem, see *e.g.* Anderson and Rubin (1949), Kleibergen (2002), Moreira (2003) and Kleibergen and Mavroeidis (2009). Still, the traditional t-statistic is commonly used. Stock and Yogo (2001) therefore analyze the sensitivity of the large sample distribution of the t-statistic to the value of a pre-test for weak instruments that results from the first stage regression, *i.e.* the first stage F-statistic. Stock and Yogo (2001) show that the large sample distribution of the t-statistic is standard normal when the first stage F-statistic exceeds ten and differs from a normal distribution when the first stage F-statistic is less than ten. Hence, for values of the first stage F-statistic that exceed ten we can use the t-statistic for inference and for values of the first stage F-statistic that are less than ten we have to use a weak instrument robust statistic to conduct inference.

The first stage F-statistic tests if the instruments used for the instrumental variables regression are jointly significant in the first stage regression. With respect to the FM two pass regression, this is analogous to testing if the observed factors in the first pass regression are jointly significant. The LR statistics in Tables 2-4 test the significance of the observed factors in the first pass regression. Because the first stage F-statistic is defined as an F-statistic, so it results from dividing by the number of tested parameters, a value of the first stage F-statistic of around ten is comparable with a value of the LR statistic that equals ten times the number of tested parameters. This means that the cutoffs for the LR statistic that allow us to use the FM t-statistic are two hundred and fifty in case of twenty five parameters, five hundred in case of fifty parameters, seven hundred and fifty in case of seventy five parameters etc.. The only two settings in Tables 2-4 for which the LR statistics exceed these thresholds are when there is one observed factor which equals the value weighted return or when the observed factors equal the Fama-French factors. For all other specifications either the LR statistic that tests the parameters of all observed factors or the LR statistic that tests the value of the parameters of the observed factors that are incremental to the value weighted return are below the threshold value which allows the FM two pass t-statistic to be compared to a standard normal distribution. These specifications are obviously also those for which the observed factors fail to eradicate the factor structure in the first pass residuals. For all these specifications we should use the analogs of the weak instrument robust statistics for factor models which are developed in Kleibergen (2009). A brief review of them that discusses how to compute them and what their properties are is provided in Appendix B, see also Kleibergen (2009).

Usage of confidence sets as reality check Alongside the reality check for factor pricing based on the remaining factor structure in the first pass residuals, we propose a second reality check that uses the 95% confidence sets that result from the identification robust factor statistics. We use the four identification robust factor statistics discussed in Appendix B: the factor Anderson-Rubin (FAR) statistic, the factor extension of Kleibergen's (2002, 2005) Lagrange multiplier statistic (FKLM), the factor extension of Kleibergen's (2005) J-statistic (FJKLM) and 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 $\mathrm{H}_0 : \lambda_1 = \lambda_{1,0}$, these statistics test it in different manners. The FAR statistic tests the joint hypothesis of factor pricing stated in (9) and H_0. The hypothesis of factor pricing given that H_0 holds is tested using the FJKLM statistic while the FKLM and FCLR statistics both just test $\mathrm{H}_0 : \lambda_1 = \lambda_{1,0}$. We can use each of these four statistics to construct a 95% confidence set for λ_1 by specifying a grid of s different values for $\lambda_{1,0}$, $(\lambda_{1,0}^1 \dots \lambda_{1,0}^s)$. 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. In case one of the elements of λ_F is not well identified, because of the small value of β and/or V_{FG}, the 95% confidence sets are unbounded, see *e.g.* Dufour (1997). Our second reality check for factor pricing therefore uses the 95% confidence sets for the risk premia that result from the identification robust factor statistics. These confidence sets are trustworthy for all possible values of β and V_{FG} and show if the observed series are informative about the risk premia. If the latter is not the case, these confidence sets are unbounded which shows that the observed series do not contain (much) information about the pricing of the associated factors. In the sequel, we discuss these confidence sets and how we turn them into a reality check for the previously analyzed series.

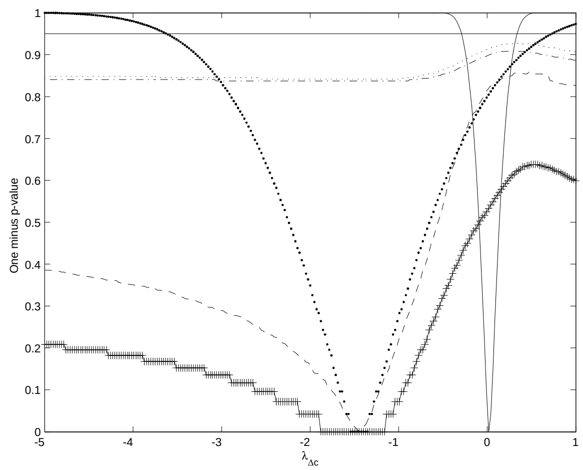

Figure 4.1: p-value plot $\lambda_{\Delta c}$

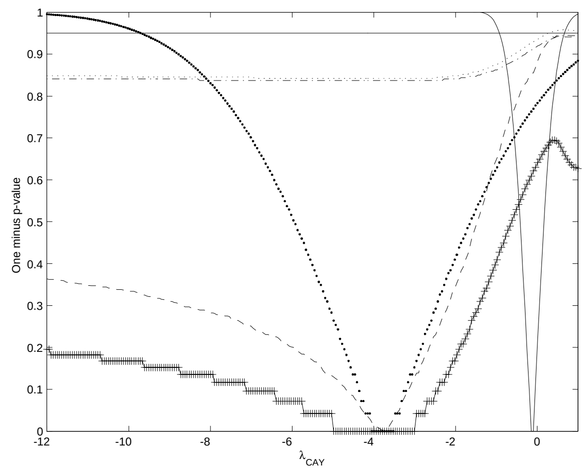

Figure 4.2: p-value plot λ_{cay}

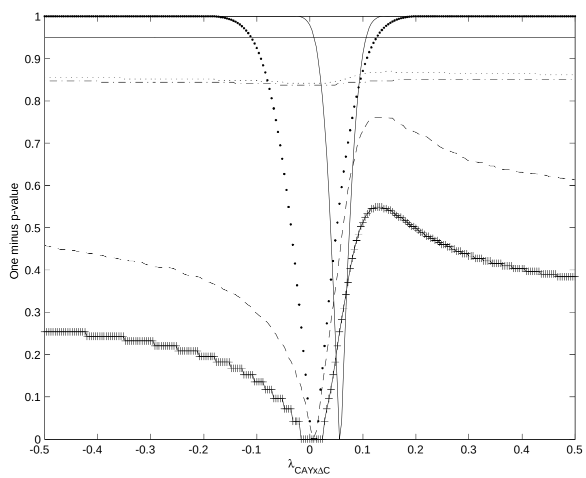

Figure 4.3: p-value plot $\lambda_{\Delta c \times cay}$

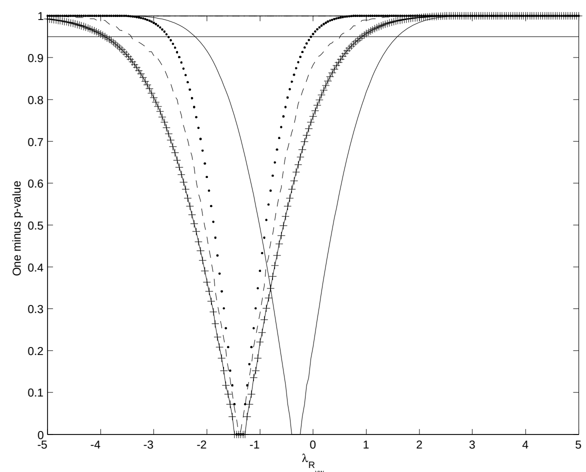

Figure 4.4: p-value plot $\lambda_{r_{vw}}$

Panel 4. p-value plots for the risk premia that result from different specifications used in Lettau and Ludvigson (2001). Figures 4.1-4.3: three factor model. Figure 4.4 single factor model. FM two pass t-statistic (solid line), MLE t-statistic (points), FKLM (solid-plusses), FCLR (dashed), FJKLM (dash-dotted), FAR (dotted).

Lettau-Ludvigson (2001): consumption based and value weighted return factors We discuss two different specifications from Lettau and Ludvigson (2001) which are representative for the results at large. After that we briefly mention the results for the other specifications in Lettau and Ludvigson (2001), Lustig and Van Niewerburgh (2005), Li *et. al.* (2006), Santos and Veronesi (2006) and Yogo (2006) that we discussed previously. The first specification has three consumption based factors: consumption growth, the consumption wealth ratio (cay) and the interaction between consumption growth and the consumption wealth ratio. The second specification has the value weighted return as its single factor. The specifications correspond with the "cay, Δc, $cay\Delta c$" and "R_{vw}" columns in

	LL01			
	R_{vw}	Δc	cay	$cay\Delta c$
FM		0.022 0.16	−0.13 0.40	0.057 0.025
MLE		−1.44 1.10	−3.78 3.00	0.0077 0.06
FM	−0.31 0.96			
MLE	−1.38 0.68			

Table 7: FM two pass and ML estimates of the risk premia for different specifications used in Lettau and Ludvigson (2001). Standard errors (with Shanken correction for FM standard errors) are listed below the estimates.

Table 3. Table 7 contains the FM two pass and maximum likelihood (ML) estimates, see Gibbons (1982).[3] We use the four identification robust factor statistics, the FM two pass t-statistic and the t-statistic based on the ML estimator to construct the confidence sets for the risk premia. Figures 4.1-4.4 in Panel 4 therefore each contain six (one minus the) p-value plots, one for every statistic. These figures also contain a straight line at 0.95 which enables us to construct the 95% confidence set using the intersection of the p-value plot with the line at 0.95.

For the first specification with the consumption based factors, Figures 4.1-4.3 show the p-value plots of the risk premia using the different statistics. The FM t-statistic is the statistic that is almost solely used in the literature. Figures 4.1-4.3 show that according to the FM t-statistic the risk premia on the consumption growth factor and consumption-wealth ratio are not significant at the 95% level while the risk premium on the interaction between consumption growth and the consumption wealth ratio is. This can be concluded from the intersection of the p-value plot of the FM t-statistic and the line at 0.95. It also results from the risk premia estimates and standard errors in Table 7. Lettau and Ludvigson (2001) use the significant risk premium on the interaction between consumption growth and the consumption-wealth ratio to advocate an extension of the consumption capital asset pricing model in which the consumption growth's risk premium changes over the business cycle. These findings, however, do not correspond with those obtained from the identification robust factor statistics. For all three risk premia, the 95% confidence sets that result from any of the identification robust factor statistics are unbounded. The p-value plots of these statistics in Figures 4.1-4.3 do namely not intersect with the line at 0.95 or lie below it at large values of the risk premia. Hence, the 95% confidence sets that result from the identification robust factor statistics are unbounded. It shows that we cannot determine the risk premia. This results since the consumption based factors are not able to eradicate the factor structure as shown in Table 3. Because the consumption based factors do not eradicate the factor structure, β and/or V_{FG} are rather small which is revealed by the LR statistics in Table 3.[4] The risk premia are then hard to identify so a large range of values is plausible as indicated by the unbounded confidence sets. We therefore propose to use these confidence sets as another reality check for factor pricing. When β and/or V_{FG} are sizeable, the confidence sets are small and similar for all the different statistics. When β and/or V_{FG} are small, we can, however, only use the identification robust factor statistics since their distributions do, unlike that of the FM t-statistic, not depend on it. Because the distribution of the FM t-statistic differs from a standard normal one

[3] As discussed in Appendix B, the four identification robust statistics all have the ML estimator as their minimizer so we do not report separate estimators for these statistics.

[4] For example, Table 3 shows that the p-value for testing the significance of consumption growth is only 6%.

when β and/or V_{FG} are small, it gives rise to misleading conclusions when we assess its significance using the standard normal distribution. This is what happens here with the FM t-statistic of the risk premium on the interaction between the consumption growth and the consumption wealth ratio. When the confidence sets that result from the identification robust factor statistics are small, the data is informative about the risk premia so factor pricing occurs. When they are, however, unbounded, the data is not informative about the risk premia so it is unlikely that factor pricing occurs. The confidence sets that result from the identification robust factors statistics therefore provide another reality check of factor pricing.

The differences in the p-value plots of the identification robust factor statistics in Figures 4.1-4.3 result since they do not test the same hypothesis. The FKLM and FCLR statistics just test if the respective risk premium is equal to a specific value which explains why they are rather similar and equal to zero at the ML estimate. The FAR and FJKLM statistics also or just, in case of the FJKLM statistic, test the hypothesis of factor pricing which is stated in (9). This explains why the p-value plots associated with them are never equal to zero because the number of moment equations in (9) exceeds the number of risk premia.

Besides the identification robust factor statistics and the FM t-statistic, Figures 4.1-4.3 also show the p-value plots of the ML t-statistic which results from the estimates and standard errors in Table 7. This t-statistic suffers from the same statistical issues as the FM t-statistic being only valid when β and/or V_{FG} are sizeable. Identical to the FM t-statistic, the confidence sets that it leads to can therefore not be trusted when β and/or V_{FG} are small. To show that the confidence sets that result from the different statistics are similar when β and/or V_{FG} are sizeable, we compute them using an observed factor that has such sizeable values, the value weighted return.

Figure 4.4 contains the (one minus the) p-value plots for the risk premium on the value weighted return when it is the only factor. They are computed using the statistics discussed before. The results in Figure 4.4 correspond with the "R_{vw}" column in Table 3. Table 3 shows that the value weighted return eradicates one of the factors in the portfolio returns. The values of its β and V_{FG} parameters are therefore sizeable as revealed by the huge value of the likelihood ratio statistic testing the significance of these parameters, 765. Because of these large values, the p-value plots that result from the identification robust factor statistics and the FM and MLE t-statistics are almost identical and all indicate that the risk premium on the value weighted return is not significant at the 95% level.

The p-value plots of the identification robust FAR and FJKLM statistics are not visible in Figure 4.4. This results because they are significant for all values of the risk premium. It shows that the hypothesis of factor pricing is rejected when we use the value weighted return as the only factor. Hence, the factor pricing moment equation in (9) does not hold so the average portfolio returns are not spanned by the β's of the value weighted return. Given the small value of the R^2 in Table 5, 0.01, this does not come as a surprise. We have, however, previously shown that the R^2 is unreliable when there is still a considerable factor structure left over in the residuals of the first pass of the FM two pass procedure. Table 3 shows that two factors remain in these residuals so the R^2 in Table 5 is unreliable but coincidentally not in disagreement with the rejection of the factor pricing hypothesis by the FAR and FJKLM statistics.

The p-value plots of the FAR and FJKLM statistics in Figures 4.1-4.3 lie below 0.95 so the factor pricing hypothesis is not rejected when we use the consumption based factors while it is when we use the value weighted return as the only factor. Similar results are documented in Table 8 in Lettau and Ludvigson (2001) where the factor pricing hypothesis is tested using the Hansen-Jagannathan distance, see Hansen and Jagannathan (1997). The rejection of factor pricing when we use the value weighted return as the single factor compared to the non-rejection for the three factor model with the consumption based factors is surprising since the three factors in the latter specification do not explain any of the factor structure in the portfolio returns while the value weighted return at least eradicates one of the factors as shown in Table 3. The non-rejection of the factor pricing hypothesis when we use the consumption based factors therefore results because these factors do not eradicate any of the factor structure. The value of the covariance matrix involved in the FAR and FJKLM statistics is therefore much larger when we use the consumption based factors compared to the value weighted return. Since we divide by the covariance matrix to obtain the statistics, a smaller covariance matrix leads to a larger and thus also more likely to be significant value of the statistics. This explains why the FAR and FJKLM statistics are significant when we use the value weighted return as the only factor but not when we use the three consumption based factors.

The ML estimates for the risk premia on the consumption based factors in Table 7 can be large and quite different from the FM two pass estimates. Table 3 shows that the β and/or V_{FG} parameters associated with the consumption based factors are all quite small. In the second pass, we kind of divide by the estimated parameters from the first pass so their proximity to zero explains the large values of the ML estimates[5]. This also explains why the FM and ML estimates of the risk premium on the value weighted return in Table 7 are quite similar because its β and/or V_{FG} parameters are sizeable.

Lettau-Ludvigson (2001): Remaining factor specifications Table 8 and Panels 5 and 6 in Appendix A contain the estimates and p-value plots for the remaining specifications from Lettau and Ludvigson (2001) which are listed in Table 3. These results can be classified along the lines of the two specifications discussed previously. First there are the specifications for which, according to Table 3, the observed factors do not fully eradicate the factor structure. These are the single factor model with consumption growth as its single factor, the three factor model with the value weighted return, consumption wealth ratio and the interaction of the consumption wealth ratio with the value weighted return as its factors and the five factor model with the value weighted return, consumption wealth ratio, income growth and the interactions of the consumption wealth ratio with the value weighted return and income growth as its factors. Second is the specification for which Table 3 shows that the observed factors do eradicate the factor structure which is the three factor model with the Fama-French factors.

For the first set of specifications, the ML estimates of the risk premia in Table 8 are all considerably different from the FM two pass estimates. Also the p-value plots in Figures 5.1 and 6.1-6.6 show that the 95% confidence sets that result from the identification robust factor statistics for all of the risk premia are unbounded. These results are identical to those for the three factor model with the consumption based factors discussed previously. The conclusion for that specification based on our second reality check, that there is no information on any of the risk premia, applies to all of these specifications therefore as well. Because of the proximity to zero of the β and V_{FG} parameters in the first pass, the behavior of the FM two pass t-statistic is anomalous. Its' indication of factor pricing through significant t-statistics, which applies here, for example, for the interaction of the value

[5]For the same reason, the distribution of the ML estimator, just like the distribution of the ML estimator in the linear instrumental variables regression model, has no moments so the mean and variance of the distribution of the ML estimator are infinite. The same holds to a lesser extent for the FM two pass risk premia estimator for which we can show that its moments exist up to the order of the number of portfolios minus the number of factors which is around 25. Its' distribution therefore has a finite mean and variance so extreme values of the FM risk premium estimator are less likely than for the ML estimator, see Kleibergen (2009).

weighted return with the consumption wealth ratio in Figure 6.3 and which we have shown previously for the interaction between consumption growth and the consumption wealth ratio as well, is therefore spurious so it provides no indication of factor pricing.

For the second kind of specification, the ML estimates of the risk premia in Table 8 are somewhat similar to that of the FM two estimates which results because the parameters in the first pass are all considerably different from zero. The p-value plots of the FAR and FJKLM factor pricing statistics are, however, hardly visible because of their proximity to one which makes them highly significant. The significance of the FAR and FJKLM factor pricing statistics again results since the Fama-French factors, as shown in Table 3, explain a lot of the variation of the portfolio returns so the residual covariance matrix is small which leads to large values of the FAR and FJKLM factor pricing statistics. This conclusion is in line with that of the specification discussed previously which has the value weighted return as its single factor.

Li et. al. (2001), Lustig and Van Nieuwerburgh (2005), Santos and Veronesi (2006) and Yogo (2006) The same distinction based on whether the observed factors can eradicate the factor structure reported in Tables 3 and 5 can be made to classify the results for the factor specifications used in Li *et. al.* (2001), Lustig and Van Nieuwerburgh (2005), Santos and Veronesi (2006) and Yogo (2006). Except for the specification that has the value weighted return as its single factor, whose p-value plots are shown in Figure 7.1 in Appendix A, Tables 4 and 6 show that none of these specifications manage to eradicate the same number of unobserved factors as the number of observed factors that they contain. Table 9 in Appendix A therefore shows that the ML estimates of the risk premia for these specifications typically differ considerably from the FM two pass estimates. It also implies that the p-value plots in Figure 7.2-7.6, 8.1-8.6 and 9.1-9.3 in Appendix A all indicate that the 95% confidence sets that result from the identification robust factor statistics for all of the risk premia are unbounded. Our second reality check thus shows that we cannot determine the risk premia. Because of the proximity to zero of the β and V_{FG} parameters in the first pass, the FM two pass t-statistic is unreliable. Its' indications of significant risk premia, as, for example, for FINAN and HHOLD in the specification of Li *et. al.* (2001) shown in Figures 7.4 and 7.5 and the interaction of the growth of the consumption of nondurables and MYFA in the specification of Lustig and Van Nieuwerburgh (2005) shown in Figure 8.3, are therefore anomalous.

Identical to the previous specification that we analyzed that had the value weighted return as its single factor, Figure 7.1 shows that all p-value plots are rather similar since the β and V_{FG} parameters from the first pass are sizeable. Again the p-value plot of the factor pricing FAR and FJKLM statistics lies at one which shows that factor pricing is strongly rejected.

We do not compute the confidence sets for the specifications used by Jagannathan and Wang (1996) because the results are rather similar to those from the other specifications.

4 Conclusions

Portfolio returns exhibit a (unobserved) factor structure. This factor structure has to be captured appropriately to obtain reliable results from stochastic discount factor models using the traditional statistical criteria. If the factor structure is not captured by the observed factors, the traditional statistical criteria to assess stochastic discount factors, which are the R^2 and t-statistics from the second pass of the FM two pass procedure, falter. Hence, they often signal relationships between observed factors and portfolio returns which are essentially non-existent. Many of the factors proposed in the literature, like, for example, consumption and labor income growth, housing collateral, consumption-wealth ratio, labor income-consumption ratio, interactions of either one of the latter three with other factors, etc., are supported by their high R^2's and significant t-statistics but fail to capture any of the factor structure of the portfolio returns. These results can therefore not be trusted. To indicate whether we can trust the support for factor pricing, we propose two reality checks. The first reality check shows if the observed factors eradicate the factor structure. The second reality check assesses whether factor pricing occurs using the 95% confidence sets that result from the identification robust factor statistics. Unlike the FM t-statistic, these statistics remain reliable when the observed factors do not capture the factor structure. The first reality check shows that many factors used in the literature fail to eradicate the factor structure in portfolio returns. The second one shows that, since the 95% confidence sets for the risk premia on these factors are then unbounded, we cannot determine the risk premia on observed factors that do not eradicate any of the unobserved factors.

Appendix A: Tables and Figures

	LL01								
	R_{vw}	Δc	SMB	HML	cay	$cayR_{vw}$	$cay\Delta c$	Δy	$cay\Delta y$
FM		0.21							
		0.18							
MLE		−2.68							
		1.95							
FM	1.32		0.47	1.46					
	1.58		0.1	0.12					
MLE	−11.2		0.68	1.52					
	3.44		0.11	0.15					
FM	−0.52				−0.05	1.13			
	3.28				1.47	0.46			
MLE	32.4				0.67	1.79			
	14.32				2.41	0.87			
FM	−1.98				−0.44	0.34		0.56	−0.17
	1.58				0.45	0.32		0.44	0.12
MLE	−10.75				−5.66	−2.93		−2.86	−1.94
	17.97				9.70	5.40		5.41	3.18

Table 8: FM two pass and ML estimates of the risk premia for different specifications used in Lettau and Ludvigson (2001). Standard errors (with Shanken correction for FM standard errors) are listed below the estimates.

	R_{vw}	Δc_{Nondur}	FIN	HHOLD	NFIN	$s_w R_{vw}$	$cayR_{vw}$	Δc_{Dur}	$myfa$	$myfa\Delta c$
FM	−0.85									
	0.94									
MLE	−1.90									
	0.69									
FM		0.13								
		0.21								
MLE		−4.65								
		4.84								
FM			0.078							
			0.034							
MLE			−54.9							
			4138							
FM			0.092	0.027	0.008					
			0.04	0.016	0.011					
MLE			0.079	0.10	−0.003					
			0.061	0.047	0.016					
FM		0.14							0.015	0.050
		0.33							0.028	0.025
MLE		0.016							−0.047	0.11
		0.43							0.047	0.067
FM	0.49					0.61	0.016			
	1.02					0.88	0.03			
MLE	5.69					3.71	−2.86			
	5.77					4.34	0.36			
FM	0.064	0.68						−0.12		
	1.70	0.38						0.30		
MLE	32.6	10.5						38.2		
	292	90.4						327		

Table 9: FM two pass and ML estimates of the risk premia for different specifications used in Li et. al. (2006), Lustig and Van Nieuwerburgh (2005), Santos and Veronesi (2006) and Yogo (2006). All use the portfolio returns from the website of Ken French. Standard errors (with Shanken correction for FM standard errors) are listed below the estimates.

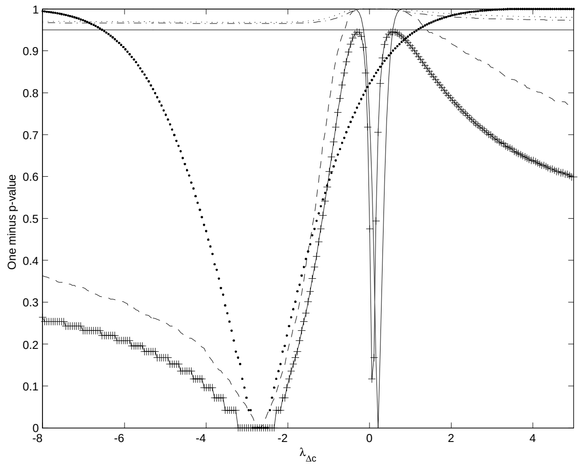

Figure 5.1. $\lambda_{\Delta c}$

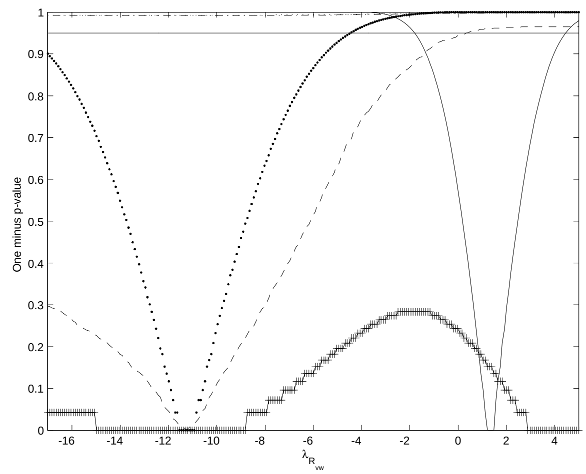

Figure 5.2. $\lambda_{r_{vw}}$

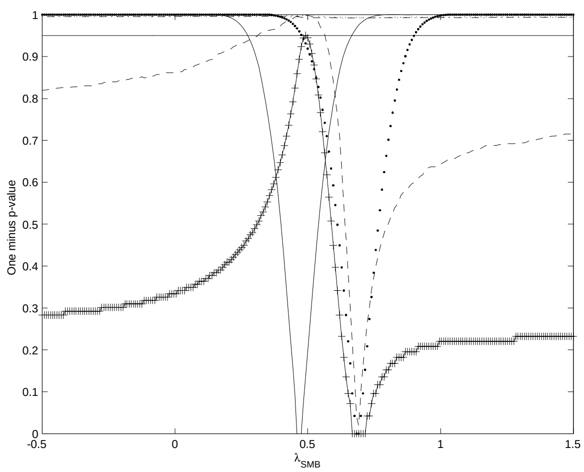

Figure 5.3. λ_{SMB}

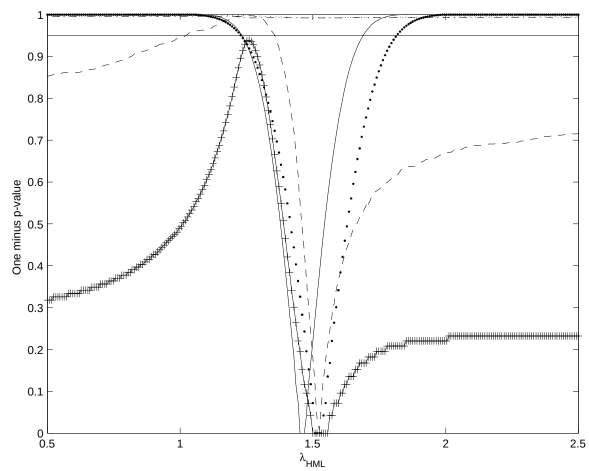

Figure 5.4. λ_{HML}

Panel 5. p-value plots for the risk premia that result from different specifications used in Lettau and Ludvigson (2001). Figure 5.1: single factor model, Figures 5.2-5.4: Three factor model. FM two pass t-statistic (solid line), MLE t-statistic (points), FKLM (solid-plusses), FCLR (dashed), FJKLM (dash-dotted), FAR (dotted).

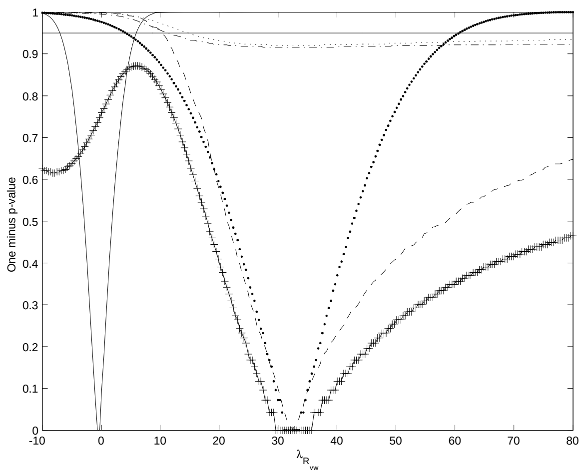

Figure 6.1. $\lambda_{r_{vw}}$

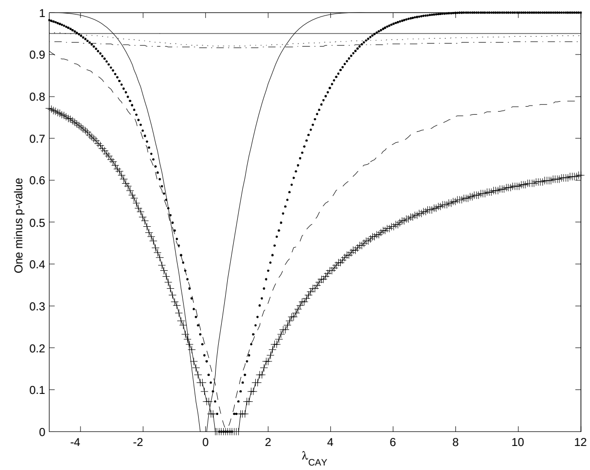

Figure 6.2. λ_{CAY}

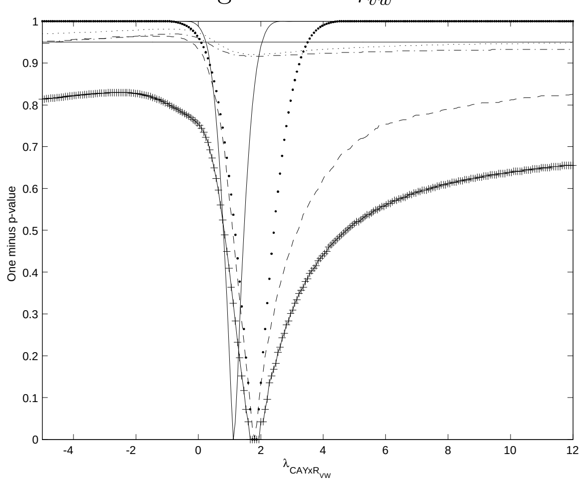

Figure 6.3. $\lambda_{r_{vw} \times \mathrm{CAY}}$

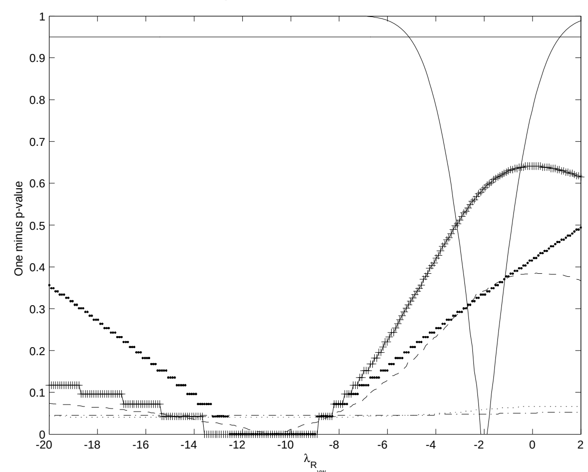

Figure 6.4. $\lambda_{r_{vw}}$

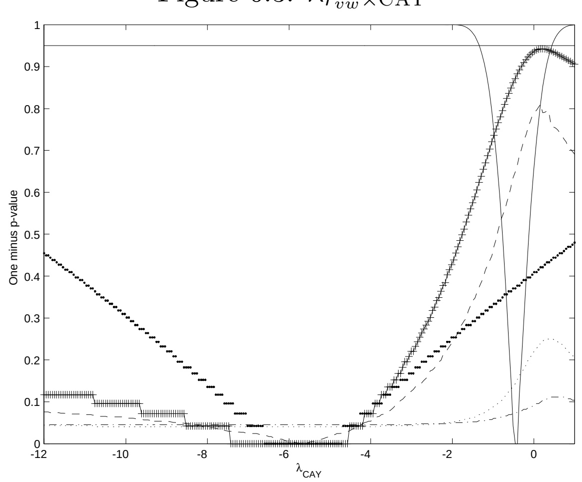

Figure 6.5. λ_{CAY}

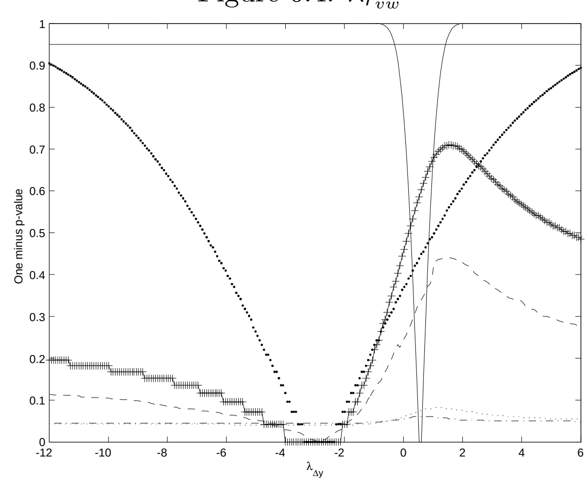

Figure 6.6. $\lambda_{\Delta y}$

Panel 6. p-value plots. Figure 6.1-6.3 three factor model. Figure 6.4-6.5: five factor model.

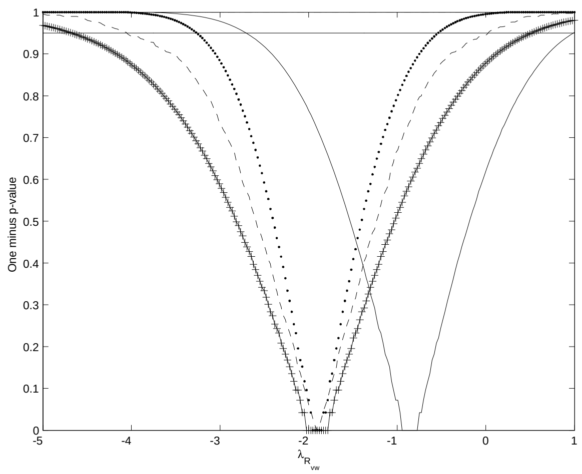

Figure 7.1. $\lambda_{r_{vw}}$

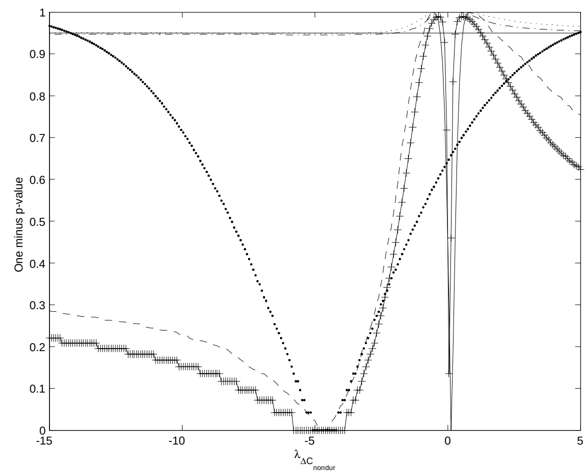

Figure 7.2. $\lambda_{\Delta c_{\text{nondur}}}$

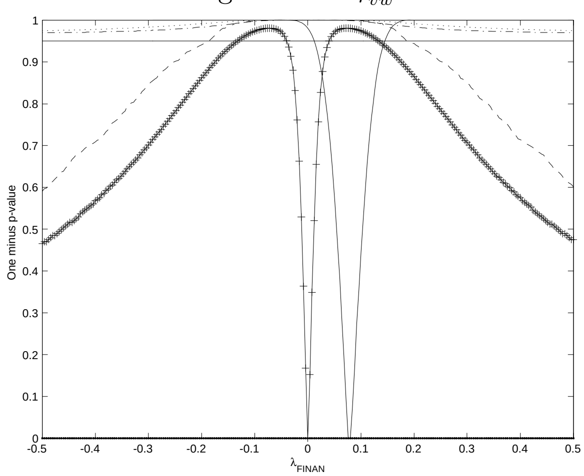

Figure 7.3. λ_{FINAN}

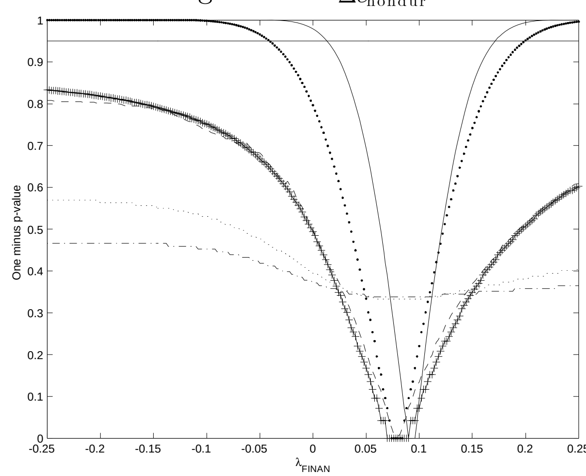

Figure 7.4. λ_{FINAN}

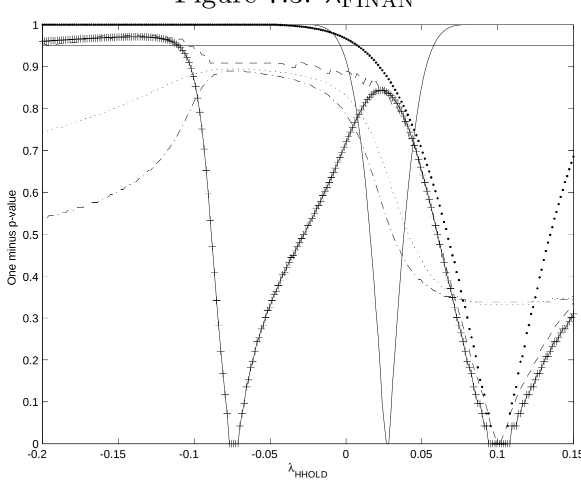

Figure 7.5. λ_{HHOLD}

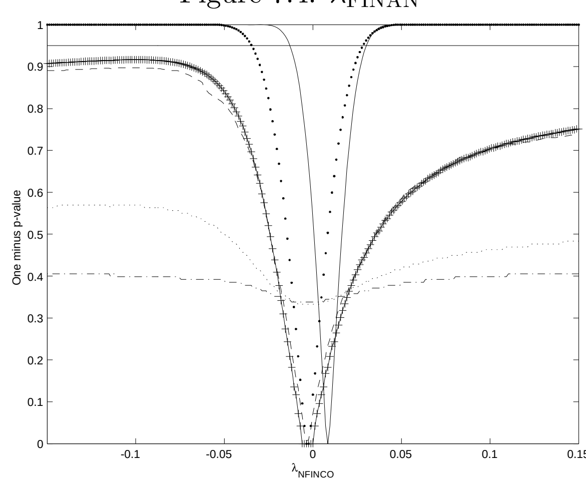

Figure 7.6. λ_{NFINCO}

Panel 7. p-value plots. Figures 7.1-7.3 single factor models. Figure 7.4-7.6, three factor model from Li *et*. *al* (2001).

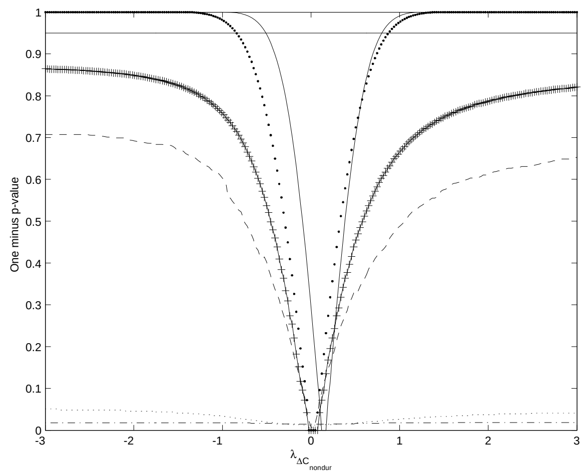

Figure 8.1. $\lambda_{\Delta c_{\text{nondur}}}$

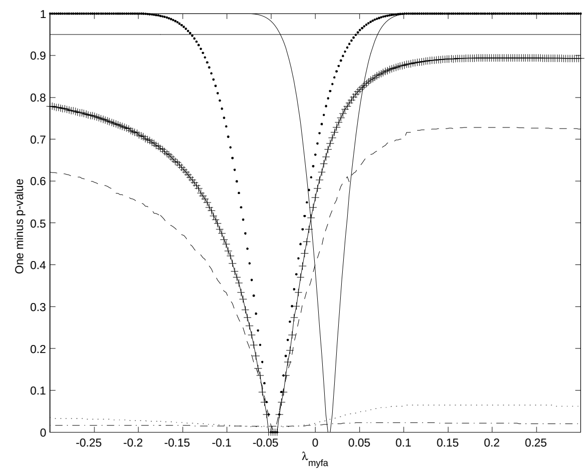

Figure 8.2. λ_{MYFA}

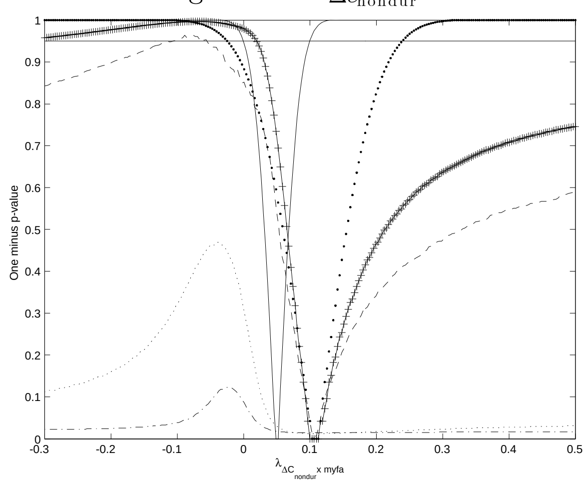

Figure 8.3. $\lambda_{\Delta c_{\text{nondur}} \times \text{MYFA}}$

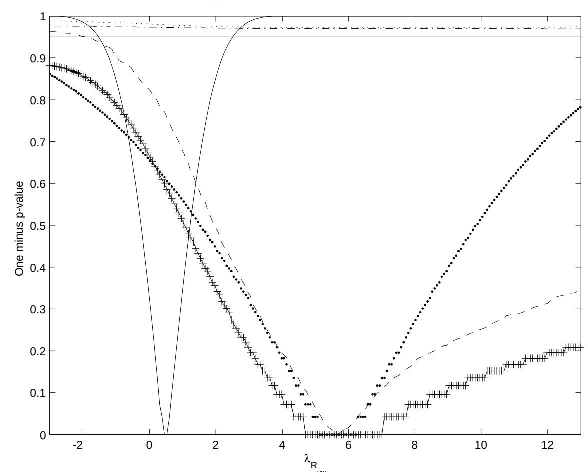

Figure 8.4. $\lambda_{r_{vw}}$

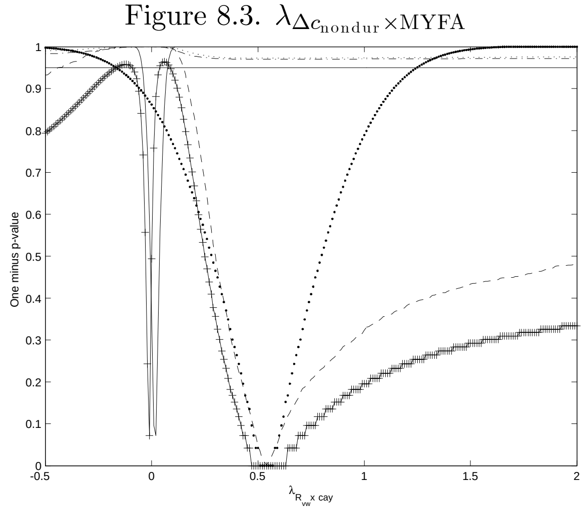

Figure 8.5. $\lambda_{r_{vw} \times \text{CAY}}$

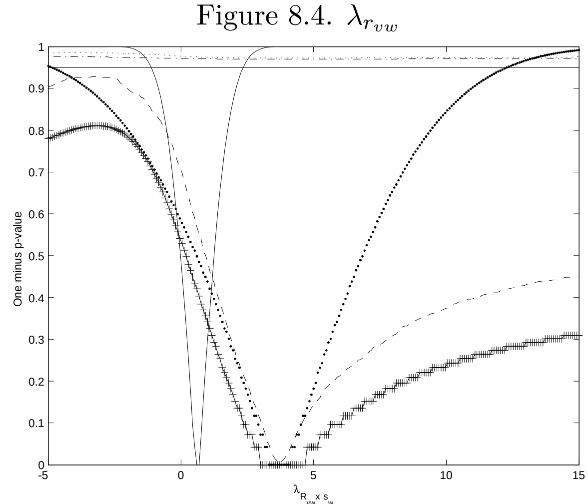

Figure 8.6. $\lambda_{r_{vw} \times s_w}$

Panel 8. p-value plots. Figure 8.1-8.3 three factor model from Lustig and Van Nieuwerburgh (2005). Figure 8.4-8.6 three factor model from Santos and Veronesi (2006).

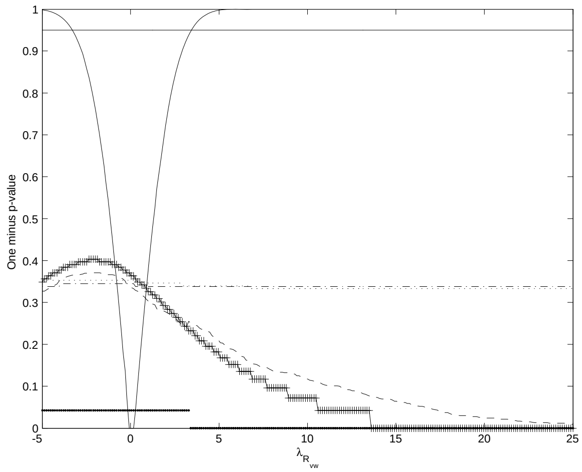

Figure 9.1. $\lambda_{r_{vw}}$

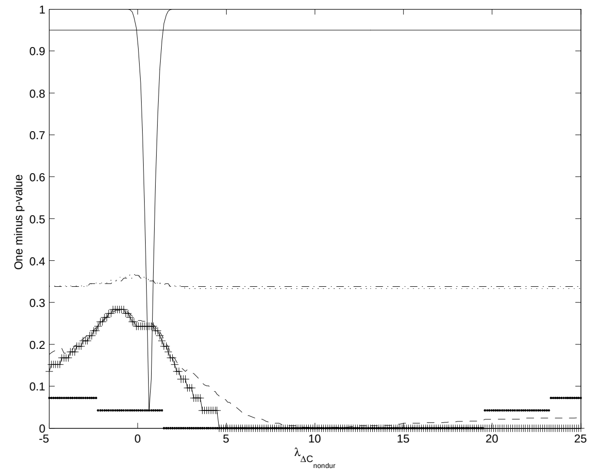

Figure 9.2. $\lambda_{\Delta c_{\mathrm{nondur}}}$

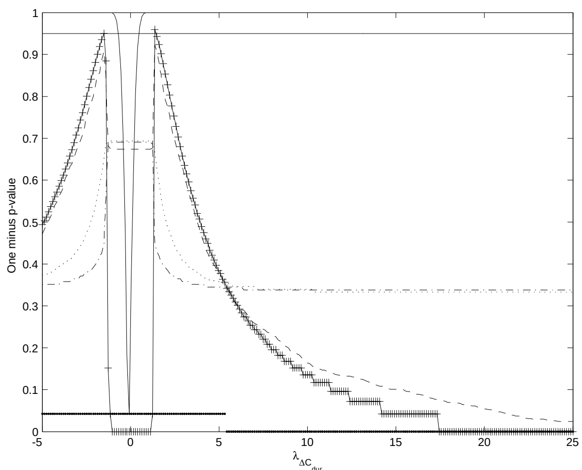

Figure 9.3. $\lambda_{\Delta c_{\mathrm{dur}}}$

Panel 9. p-value plots. Figure 9.1-9.3 three factor model from Yogo (2006).

Appendix B. Identification robust factor statistics

To test a hypothesis on one element of λ_F, say λ_1, the identification robust factor statistics first remove the zero-β return λ_0 by taking all portfolio returns in deviation from a baseline portfolio return, say $r_{t,n}$:

$$\mathcal{R}_t = R_{1t} - \iota_{n-1} r_{nt}$$

with $R_{1t} = (r_{1t} \dots r_{n-1t})'$ and ι_{n-1} is a $(n-1)$-dimensional vector of ones. The identification robust factor statistics are invariant with respect to the choice of the baseline portfolio and also with respect to other transformations.

If λ_F has more than one element and we want to test the hypothesis, $\mathrm{H}_0 : \lambda_1 = \lambda_{1,0}$, we first compute the maximum likelihood estimates under H_0 of the other elements of $\lambda_F : \lambda_2 \ldots \lambda_m$, see Gibbons (1982). These result from the characteristic vectors (eigenvectors) of a characteristic polynomial for which we use the residuals of three linear regressions:

$$\begin{aligned} \mathcal{R}_t =& \quad \alpha_1(\bar{G}_{1t} + \lambda_{1,0}) + \tilde{\mathcal{R}}_t \\ 1 =& \quad \alpha_2(\bar{G}_{1t} + \lambda_{1,0}) + \tilde{1}_t \\ \bar{G}_{2t} =& \quad \alpha_3(G_{2t} + \lambda_{1,0}) + \tilde{G}_{2t} \end{aligned} \tag{17}$$

where $G_t = \binom{G_{1t}}{G_{2t}}$ with G_{1t} a scalar and G_{2t} a $(m-1)$-dimensional vector, G_t is the vector with observed factors in (7). When we estimate α_1, α_2 and α_3 using least squares and $\tilde{\mathcal{R}}_t$, $\tilde{1}_t$ and $\tilde{G}_{2t}$ are the residuals of these least squares regressions, the characteristic polynomial from which we obtain the maximum likelihood estimators reads

$$\left|\theta\left[\tfrac{1}{T}\textstyle\sum_{t=1}^T \binom{\tilde{1}_t}{\tilde{G}_{2t}}\binom{\tilde{1}_t}{\tilde{G}_{2t}}'\right] - \left[\tfrac{1}{T}\textstyle\sum_{t=1}^T \binom{\tilde{1}_t}{\tilde{G}_{2t}}\tilde{\mathcal{R}}_t'\right]\Sigma^{-1}\left[\tfrac{1}{T}\textstyle\sum_{t=1}^T \tilde{\mathcal{R}}_t\binom{\tilde{1}_t}{\tilde{G}_{2t}}'\right]\right| = 0, \tag{18}$$

with $\Sigma = \frac{1}{T}\sum_{t=1}^T \tilde{\mathcal{R}}_t\tilde{\mathcal{R}}_t'$. The maximum likelihood estimator of $\lambda_{F,2} = (\lambda_2 \ldots \lambda_m)'$ then equals

$$\tilde{\lambda}_{F,2} = W_2^{-1\prime}w_1', \tag{19}$$

with w_1 a m dimensional row vector and W_2 a $m \times m$ dimensional matrix which are such that $\binom{w_1}{W_2}$ contains the m characteristic vectors that are associated with the m largest roots of the characteristic polynomial in (18), see Kleibergen (2009) and Shanken and Zhou (2007). We then proceed with computing the least squares estimator of B under H_0[6] :

$$\hat{B} = \textstyle\sum_{t=1}^T \mathcal{R}_t\binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}\left[\textstyle\sum_{t=1}^T \binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}\binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}'\right]^{-1}, \tag{20}$$

which provides the base for our four identification robust factor statistics, see Kleibergen (2009):

1. The factor Anderson-Rubin (FAR) statistic which is based on the Anderson-Rubin statistic, see Anderson and Rubin (1949):

$$\mathrm{FAR}(\lambda_{1,0}) = \frac{T}{1-\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}'\hat{Q}_{GG}^{-1}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}}\left[\bar{\mathcal{R}} - \hat{B}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}\right]'\tilde{\Sigma}^{-1}\left[\bar{\mathcal{R}} - \hat{B}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}\right], \tag{21}$$

with $\bar{\mathcal{R}} = \frac{1}{T}\sum_{t=1}^T \mathcal{R}_t$, $\hat{Q}_{GG} = \frac{1}{T}\sum_{t=1}^T \binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}\binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}'$ and $\tilde{\Sigma} = \frac{1}{T-k}\sum_{t=1}^T \left[\mathcal{R}_t - \hat{B}\binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}\right]\left[\mathcal{R}_t - \hat{B}\binom{\bar{G}_{1t}+\lambda_{1,0}}{\bar{G}_{2t}+\tilde{\lambda}_{F,2}}\right]'$ and whose large sample distribution is bounded by a $\chi^2(N-m)$ distributed random variable when the sample size T gets large.

2. The factor extension of Kleibergen's (2002,2005) Lagrange Multiplier statistic:

$$\mathrm{FKLM}(\lambda_{1,0}) = \frac{T}{1-\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}'\hat{Q}_{GG}^{-1}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}}\left[\bar{\mathcal{R}} - \hat{B}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}\right]'\tilde{\Sigma}^{-1}\hat{B}(\hat{B}'\tilde{\Sigma}^{-1}\hat{B})^{-1}\hat{B}'\tilde{\Sigma}^{-1}\left[\bar{\mathcal{R}} - \hat{B}\binom{\lambda_{1,0}}{\tilde{\lambda}_{F,2}}\right], \tag{22}$$

whose large sample distribution is bounded by a $\chi^2(1)$ distributed random variable when the sample size gets large.

[6] We are essentially estimating $B_1 - \iota_{n-1}b_n$ if $B = \binom{B_1}{b_n}$, with B_1 a $(n-1) \times m$ dimensional matrix and b_n a $1 \times m$ dimensional row vector.

3. The factor extension of Kleibergen's (2005) J-statistic:

$$\text{FJKLM}(\lambda_{1,0}) = \text{FAR}(\lambda_{1,0}) - \text{FKLM}(\lambda_{1,0}), \tag{23}$$

whose large sample distribution is bounded by a $\chi^2(N-m-1)$ distributed random variable when the sample size gets large. This $\chi^2(N-m-1)$ distributed random variable is independent of the $\chi^2(1)$ distributed random variable which bounds the large sample distribution of the FKLM statistic.

4. The factor extension of Moreira's (2003) conditional likelihood ratio statistic:

$$\begin{aligned}\text{FCLR}(\lambda_{1,0}) = \; & \tfrac{1}{2}\Big[\text{FKLM}(\lambda_{1,0}) + \text{FJKLM}(\lambda_{1,0}) - \text{r}(\lambda_{1,0}) + , \\ & \sqrt{(\text{FKLM}(\lambda_{1,0}) + \text{FJKLM}(\lambda_{1,0}) + \text{r}(\lambda_{1,0}))^2 - 4\text{r}(\lambda_{1,0})\text{FJKLM}(\lambda_{1,0})}\Big],\end{aligned} \tag{24}$$

with $\text{r}(\lambda_{1,0})$ the smallest root of the characteristic polynomial:

$$\left|\mu\hat{Q}_{GG}^{-1} - \hat{B}'\tilde{\Sigma}^{-1}\hat{B}\right| = 0. \tag{25}$$

In large samples the distribution of the FCLR statistic is bounded by a random variable whose distribution is conditional on the value of $\text{r}(\lambda_{1,0})$. The (bounding) critical values of the FCLR statistic are therefore a function of $\text{r}(\lambda_{1,0})$ and the independent $\chi^2(1)$ and $\chi^2(N-m-1)$ large sample distributions of the FKLM and FJKLM statistics. We obtain these critical values by fixing $\text{r}(\lambda_{1,0})$ and simulating the FKLM and FJKLM statistics from their independent large sample distributions to compute the FCLR statistic using the simulated values of the FKLM and FJKLM statistics, see Kleibergen (2009).

The critical values that result from the bounding distributions for each of the four identification robust factor statistics are such that these statistics are size correct when $(\lambda_2 \dots \lambda_m)$ are well identified so the associated values of β and V_{FG} constitute sizeable full rank matrices. When $(\lambda_2 \dots \lambda_m)$ are not well identified because of small or zero values of the associated values of β and V_{FG}, the critical values are such that the rejection frequencies of the identification robust factor statistics are smaller than the size of the test. The maximal rejection frequencies of the identification robust factor statistics are therefore equal to the size of the test which makes them size correct, see Kleibergen and Mavroeidis (2009).

The identification robust factor statistics test different hypothezes. The FAR statistic tests the joint hypothesis of factor pricing and $\lambda_1 = \lambda_{1,0}$: $\text{H}_{\text{FAR}} : E(\mathcal{R}_t) = B\binom{\lambda_{1,0}}{\lambda_{F,2}}$ where $\lambda_{F,2}$ is to be estimated. The hypothesis of factor pricing given that $\lambda_1 = \lambda_{1,0}$ is tested using the FJKLM statistic while the FKLM and FCLR statistics both test $\text{H}_0 : \lambda_1 = \lambda_{1,0}$.

We can use each of the four different statistics to construct a 95% confidence set for λ_1 by specifying a grid of s different values for $\lambda_{1,0}$, $(\lambda_{1,0}^1 \dots \lambda_{1,0}^s)$. 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% (conditional) critical value. In case that one of the elements of λ_F is not well identified, the 95% confidence sets is unbounded, see *e.g.* Dufour (1997).

The above discussion deals with testing one element of λ_F so it assumes that λ_F has more than one element, $m > 1$. When λ_F just consists of one element, $\lambda_{F,2}$ is not present so we do not conduct the computations in (17)-(19). The expressions of the identification robust factor statistics then extend to the case that there is just one observed factor when we remove $\tilde{\lambda}_{F,2}$ and $\bar{G}_{2t} + \tilde{\lambda}_{F,2}$ from (20)-(25).

The performance of the identification robust FKLM and FCLR statistics is similar to that of t-statistics based on FM two pass or maximum likelihood estimators when the β's and V_{FG} are sizeable. The latter two, however, become unreliable when the β's and/or V_{FG} are quite small. The identification robust factor statistics remain reliable in these cases. It is therefore important to use them for computing confidence sets on the risk premia when the β's and/or V_{FG} are quite small.

References

[1] Anderson, T.W. *An Introduction to Multivariate Statistical Analysis.* John Wiley (New York), second edition, 1984.

[2] Anderson, T.W. and H. Rubin. Estimation of the Parameters of a Single Equation in a Complete Set of Stochastic Equations. *The Annals of Mathematical Statistics*, **21**:570–582, (1949).

[3] Chamberlain, G. and M. Rothschild. Arbitrage, Factor Structure, and Mean-Variancce Analysis on Large Asset Markets. *Econometrica*, **51**:1281–1302, 1983.

[4] Cochrane, J.H. *Asset Pricing.* Princeton University Press, 2001.

[5] Dufour, J.-M. Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models. *Econometrica*, **65**:1365–388, 1997.

[6] Fama, E.F. and J.D. MacBeth. Risk, Return and Equillibrium: Empirical Tests. *Journal of Political Economy*, **81**:607–636, 1973.

[7] Fama, E.F. and K.R. French. The Cross-Section of Expected Stock Returns. *Journal of Finance*, **47**:427–465, 1992.

[8] Fama, E.F. and K.R. French. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, **33**:3–56, 1993.

[9] Fama, E.F. and K.R. French. Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, **51**:55–84, 1996.

[10] Gibbons, M. Multivariate Tests of Financial Models: A New Approach. *Journal of Financial Economics*, **10**:3–27.

[11] Hansen, L.P. and R. Jagannathan. Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance*, **52**:557–590, 1997.

[12] Jagannathan, R. and Z. Wang. The Conditional CAPM and the Cross-Section of Expected Returns. *Journal of Finance*, **51**:3–53, 1996.

[13] Jagannathan, R. and Z. Wang. An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression. *Journal of Finance*, **53**:1285–1309, 1998.

[14] Kan, R. and C. Zhang. Two-Pass Tests of Asset Pricing Models with Useless Factors. *Journal of Finance*, **54**:203–235, 1999.

[15] Kleibergen, F. Pivotal Statistics for testing Structural Parameters in Instrumental Variables Regression. *Econometrica*, **70**:1781–1803, 2002.

[16] Kleibergen, F. Testing Parameters in GMM without assuming that they are identified. *Econometrica*, **73**:1103–1124, 2005.

[17] Kleibergen, F. Tests of Risk Premia in Linear Factor Models. *Journal of Econometrics*, **149**:149–173, 2009.

[18] Kleibergen, F. and S. Mavroeidis. Inference on subsets of parameters in GMM without assuming identification. 2008. Working Paper, Brown University.

[19] Lettau, M. and S. Ludvigson. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying. *Journal of Political Economy*, **109**:1238–1287, 2001.

[20] Lewellen, J., S. Nagel and J. Shanken. A Skeptical Appraisal of Asset-Pricing Tests. *Journal of Financial Economics*, 2009. Forthcoming.

[21] Li, Q., M. Vassalou and Y. Xing. Sector investment growth rates and the cross section of equity returns. *The Journal of Business*, **79**:1637–1665, 2006.

[22] Lintner, J. The Valuation of Risk Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, **47**:13–37, 1965.

[23] Lustig, H. and S. Van Nieuwerburgh. Housing collateral, consumption insurance, and risk premia: An empirical perspective. *Journal of Finance*, **60**:1167–1219, 2005.

[24] Merton, R.C. An intertemporal capital asset pricing model. *Econometrica*, **41**:867–887, 1973.

[25] Moreira, M.J.,. A Conditional Likelihood Ratio Test for Structural Models. *Econometrica*, **71**:1027–1048, 2003.

[26] Roll, R. and S.A. Ross. An Empirical Investigation of the Arbitrage Pricing Theory. *Journal of Finance*, **5**:1073–1103, 1980.

[27] Ross, S. The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, **13**:341–360, 1976.

[28] Santos, T. and P. Veronesi. Labor income and predictable stock returns. *Review of Financial Studies*, **19**:1–44, 2006.

[29] Shanken, J. On the Estimation of Beta-Pricing Models. *Review of Financial Studies*, **5**:1–33, 1992.

[30] Shanken, J. and G. Zhou. Estimating and testing beta pricing models: Alternative Methods and their performance in simulations. *Journal of Financial Economics*, **84**:40–86, 2007.

[31] Staiger, D. and J.H. Stock. Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**:557–586, 1997.

[32] Stock, J.H. and M. Yogo. Testing for Weak Instruments in Linear IV Regression. In D.W.K. Andrews and J.H. Stock, editor, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge: Cambridge University Press, 2005. Prepared for Festschrift in honor of Thomas Rothenberg.

[33] M. Yogo. A consumption-based explanation of expected stock returns. *Journal of Finance*, **61**:539–580, 2006.