Final Exam

Spring 2019

PRELIMINARY

- This is a 24-hour from 6pm June 13th, 2019, to 6pm June 14th, 2019.
- Late submission rules: The deadline is 6pm June 14th, 2018. No points will be subtracted from the final grade if you submit your exam until 6:14:59pm, i.e., 6:14pm and 59 seconds. However, your grade will be reduced by 30% if you submit after 6:14:59pm.
- No submission will be accepted after 6:29:59pm.
- Please write your full name.

By submitting your final 24-hour take-home exam, you agree to the following:

As a member of the UCLA Anderson community, I am guided in my daily actions and decisions by the principles of Honesty, Integrity, Mutual Respect, Personal Responsibility, and Professionalism.

I recognize that the integrity of the entire Anderson community—and the dignity afforded to me by my association therewith—rests with the honorable actions of every individual, both on-campus and in the community. To this end, I pledge to affirmatively uphold, in both word and deed, these principles in my dealings with all members of the UCLA Anderson community students, faculty, administration, staff, and alumni.

I acknowledge my obligations under the UCLA Anderson Honor Code and pledge to follow the ethical standards for exam taking it implies. Specifically, I pledge that I shall use only the allowed resources in taking this exam and shall neither give nor receive any type of forbidden aid.

Final Exam

This is an **individual assignment**, and you **cannot** discuss it with your classmates. We will check similarities between submissions, and, if any two submissions are too similar, then both submissions will not be considered for grading.

Please submit your code (R or Python) as well as a separate write-up. Explain the procedure clearly in the write-up (such that someone unfamiliar with the problem could solve it). Use CCLE to submit your answers. If there is something that is unclear in this assignment or in the paper, you are expected to clearly state and justify any assumptions you may make in the write-up. This is a 24-hour exam, and no changes will be made. Do not contact either the professor or the TA.

The structure of the write-up and code should follow the general guidelines used in the problem sets this quarter. You should submit two files:

- .R **FinalExam_YourStudentID** (for example, FinalExam_012345678.R), with **all** code used in answering the questions written below. Python codes are also accepted. If you have multiple files with different codes, submit a zip file with all the code used.
- .pdf **FinalExam_YourStudentID** (for example, FinalExam_012345678.pdf), with discussion on how you answered the questions written below, as well as responses to any particular questions asked.

In this exam, you will replicate the four factors developed by Hou, Xue and Zhang (2015):

Paper: Hou, K., Xue, C. and Zhang, L., 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3), pp.650-705.
Available at: https://academic.oup.com/rfs/article/28/3/650/1574802

Hou, Xue and Zhang (2015) four-factor model—or simply the q-factor model—consists of market-, size-, investment-over-assets-, and profitability-based factors. The construction of these factors is different from the Fama and French five-factor model and uses both quarterly and annual Compustat data.

Guidelines and Hints::

- The excel file 'HXZ q-Factors (monthly 1967 to 2018).xlsx' contains the four factors you will replicate in this exam.[1]

1. (15 points) Replicate Table 1 in Hou, Xue and Zhang (2015) using the q-factors available on ccle and using the Fama and French three factors from Ken French's website. Use the sample from 1972 to 2018.

Guidelines and Hints::

- The returns in the excel file 'HXZ q-Factors (monthly 1967 to 2018).xlsx' are in percent, so it is recommended to divide them 100.

[1] Thanks to Kewei Hou for sharing their data.

- In this question, you don't have to replicate the factor from CRSP/Compustat. The goal here is to replicate Table 2 for the extended sample from 1972 to 2018.
- For Panel B, you should use the market excess return from 'HXZ q-Factors (monthly 1967 to 2018).xlsx' file.
- For Panel A, you should use the market excess return from Ken French's website.
- In Panel A, both the 'Mean' and 'α' are monthly and in percent (i.e. multiplied by 100).
- As a replication reference, your outcome should be similar (or perhaps identical) to Table 2 of this exam (see last few pages of this exam).

2. (40 points) Replicate the 4 factors from Hou, Xue and Zhang (2015). Explain the replication procedure step-by-step.

(a) For each of the factors, report the following statistics for both original factor and replicated factors: annualized average returns, annualized volatility, annualized Sharpe ratio, monthly skewness, and monthly kurtosis. The outcome should be similar to Panel A of Table 3 of this exam (see last few pages of this exam).

(b) For each of the factors, compute the difference in basis points and report the following statistics: average, standard deviation, minimum and maximum difference, and quartiles. The outcome should be similar to Panel B of Table 3 of this exam (see last few pages of this exam).

(c) For each of the factors, compute the correlation between the original series and the replicated one (report 4 decimal digits). The outcome should be similar to Panel C of Table 3 of this exam (see last few pages of this exam).

(d) For each of the factors, plot three time series: the original factor, the replicated factor, and the difference. Report the factors in percent (i.e. multiplied by 100) and the difference in basis points. The outcome should be similar to Figures 1, 2, 3, 4, 5, 6, 7, and 8 of this exam (see last few pages of this exam).

Guidelines and Hints::

- In this question, you have to replicate the factors using CRSP and Compustat data.
- As a replication reference, your outcome maybe be similar (or perhaps identical) to Table 3 available in end of this exam. The replication results available in this exam were done within a reasonable amount of time. It is not a perfect replication, because the goal was to focus on implementing the method implied by the assumptions described in the paper as well as the assumptions mentioned in this exam. Feel free to further improve this replication!
- Carefully read Section 2 in the original paper (specially Section 2.1). The portoflio rebalancing is tricky because it uses both quarterly and annual data.
- Book values are quarterly (Compustat quarterly). See footnote 9 in the original paper for details on its construction.
- This replication requires you to use 4 different data:

  - CRSP Monthly Stock
    - Here is a list of variables you might need: permno, permco, date, shrcd, exchcd, siccd, dlret, dlretx, dlstdt.
    - The paper does not specific which exchange and shares codes to use (at least I did find). My recommendation is to keep it consistent with Fama and French factor construction methodology and use share codes 10 and 11 as well as exchange codes 1, 2, and 3. Feel free to implement a different selection (e.g. one could include share code 12). In either case, justify your choice.
  - CRSP-Compustat Linktable
    - Here is a list of variables you might need: gvkey, lpermco as permco, linktype, linkprim, liid, linkdt, linkenddt.
    - Follow the cleaning and gvkey-permco matching procedure recommended for the previous assignments (no need to explain in the write up though).
  - Compustat Fundamentals Quarterly
    - Here is a list of variables you might need: gvkey, datadate, fyr, fqtr, fyearq, atq, txditcq, seqq, ceqq, pstkq, ltq, pstkrq, ibq, rdq.
    - These data has a few duplicated observations. This happens because some firms change their fiscal calendar. (see detailed discussion here: http://kaichen.work/?p=387). Please use the account data of the fiscal calendar the company is switching to. The following website has a detail discussion of this issue and how to solve it: http://kaichen.work/?p=387
  - Compustat Fundamentals Annual
    - Here is a list of variables you might need: gvkey, datadate, at, pstkl, txditc, fyr, fyear, pstkrv, seq, pstk, ceq, lt, mib, itcb, txdb.

- For Panel B, you should use the replicated market factor.
- For Panel A, you should use the market excess return from Ken French's website.
- In Panel A, both the 'Mean' and 'α' are monthly and in percent (i.e. multiplied by 100).
- When constructing the breakpoints, justify which companies to be included based on which criteria.
- As a replication reference, your outcome should be similar (or perhaps identical) to Table 3 available in end of this exam. This is the outcome using my own replication of these factors.

3. (15 points) Replicate Table 1 in Hou, Xue and Zhang (2015) using your replicated q-factors and using the Fama and French three factors from Ken French's website. Use the sample from 1972 to 2018.

Guidelines and Hints::

- For Panel B, you should use the market excess return from 'HXZ q-Factors (monthly 1967 to 2018).xlsx' file.
- For Panel A, you should use the market excess return from Ken French's website.
- In Panel A, both the 'Mean' and 'α' are monthly and in percent (i.e. multiplied by 100).
- As a replication reference, your outcome should be similar (or perhaps identical) to Table 2 available in end of this exam. This is the outcome using my own replication of these factors.

4. (10 points) The paper uses breakpoints based on NYSE stocks. Explain the reason behind this choice. What are the main benefits of constructing breakpoints in this way?

5. (20 points) Table 1 in the original paper shows the factors from the q-factor model are not explained by the CAPM, Fama and French three-factor model, or the Carhart model. How can you further test the idea tested in Table 1? Is this result alone enough to argue that the q-factor model should be used instead of the usual Fama and French three-factor model? What are the main advantages and disadvantages of the q-factor model over other asset pricing models (i.e. CAPM, Fama and French three-factor model, Fama and French five-factor model, and Carhart model).

Hints:

- No need to code the robustness exercises.
- Explain in detail the exercises you suggest and justify.

Table 1: Same as Table 1 in Hou, Xue and Zhang (2015) using 1972-2018 sample (longer sample).

Panel A

0	Mean	alpha	beta(MKT)	beta(SMB)	beta(HML)	beta(UMD)	R2
ME	0.26	0.17	0.17				0.06
	1.98	1.31	5.96				
		0.04	0.02	0.98	0.18		0.93
		1.04	3.12	84.91	15.16		
		0.01	0.03	0.98	0.20	0.03	0.93
		0.21	3.89	85.99	15.85	3.94	
IA	0.37	0.45	-0.15				0.13
	4.75	6.14	-9.10				
		0.28	-0.07	-0.03	0.39		0.49
		4.98	-5.54	-1.53	19.57		
		0.24	-0.06	-0.03	0.41	0.05	0.50
		4.14	-4.72	-1.56	20.15	3.82	
ROE	0.54	0.60	-0.11				0.04
	5.04	5.66	-4.70				
		0.70	-0.09	-0.32	-0.20		0.20
		7.19	-3.89	-9.79	-5.83		
		0.46	-0.03	-0.32	-0.10	0.27	0.40
		5.31	-1.68	-11.28	-3.31	13.35	

Panel B

index	IA	ROE	MKT-RF	SMB	HML	UMD
ME	-0.11	-0.32	0.25	0.95	-0.02	-0.00
	0.01	0.00	0.00	0.00	0.60	0.98
IA		0.06	-0.36	-0.21	0.68	0.02
		0.17	0.00	0.00	0.00	0.69
ROE			-0.21	-0.38	-0.11	0.49
			0.00	0.00	0.01	0.00
MKT-RF				0.27	-0.27	-0.15
				0.00	0.00	0.00
SMB					-0.19	-0.00
					0.00	0.94
HML						-0.19
						0.00

Table 2: Replication of Table 1 in Hou, Xue and Zhang (2015) using 1972-2018 sample (replication and longer sample).

Panel A

0	Mean	alpha	beta(MKT)	beta(SMB)	beta(HML)	beta(UMD)	R2
ME_rep	0.24	0.14	0.18				0.06
	1.80	1.11	6.19				
		0.02	0.03	0.99	0.17		0.92
		0.45	3.47	79.40	13.10		
		-0.01	0.04	0.99	0.18	0.03	0.93
		-0.28	4.15	80.19	13.69	3.53	
IA_rep	0.35	0.43	-0.15				0.13
	4.41	5.76	-9.06				
		0.27	-0.08	-0.02	0.39		0.46
		4.46	-5.69	-0.77	18.39		
		0.23	-0.07	-0.02	0.40	0.04	0.47
		3.78	-5.02	-0.79	18.69	2.98	
ROE_rep	0.55	0.60	-0.10				0.03
	5.00	5.53	-4.11				
		0.70	-0.08	-0.32	-0.19		0.19
		6.94	-3.21	-9.50	-5.40		
		0.46	-0.02	-0.32	-0.09	0.26	0.37
		5.07	-1.02	-10.82	-2.92	12.65	

Panel B

index	IA_rep	ROE_rep	MKT-RF_rep	SMB	HML	UMD
ME_rep	-0.10	-0.31	0.25	0.95	-0.04	-0.00
	0.02	0.00	0.00	0.00	0.38	0.98
IA_rep		0.06	-0.36	-0.19	0.65	0.00
		0.13	0.00	0.00	0.00	0.99
ROE_rep			-0.17	-0.37	-0.11	0.47
			0.00	0.00	0.01	0.00
MKT-RF_rep				0.26	-0.27	-0.14
				0.00	0.00	0.00
SMB					-0.19	-0.00
					0.00	0.94
HML						-0.19
						0.00

Table 3: Summary Statistics of q-factor model replication using 1972-2018 sample.

Panel A: Summary Statistics of original and replicated factors

	MKT-RF		ME		IA		ROE	
	Original	Replication	Original	Replication	Original	Replication	Original	Replication
Average	6.21	6.49	3.07	2.87	4.44	4.21	6.50	6.56
Volatility	15.52	15.55	10.63	10.90	6.41	6.55	8.84	8.99
Sharpe Ratio	0.40	0.42	0.29	0.26	0.69	0.64	0.74	0.73
Skewness	-0.59	-0.57	0.63	0.79	0.16	0.19	-0.73	-0.67
Kurtosis	2.28	2.13	5.77	6.94	1.69	1.58	4.99	4.60

Panel B: Summary Statistics of difference between original and replicated factors (basis points)

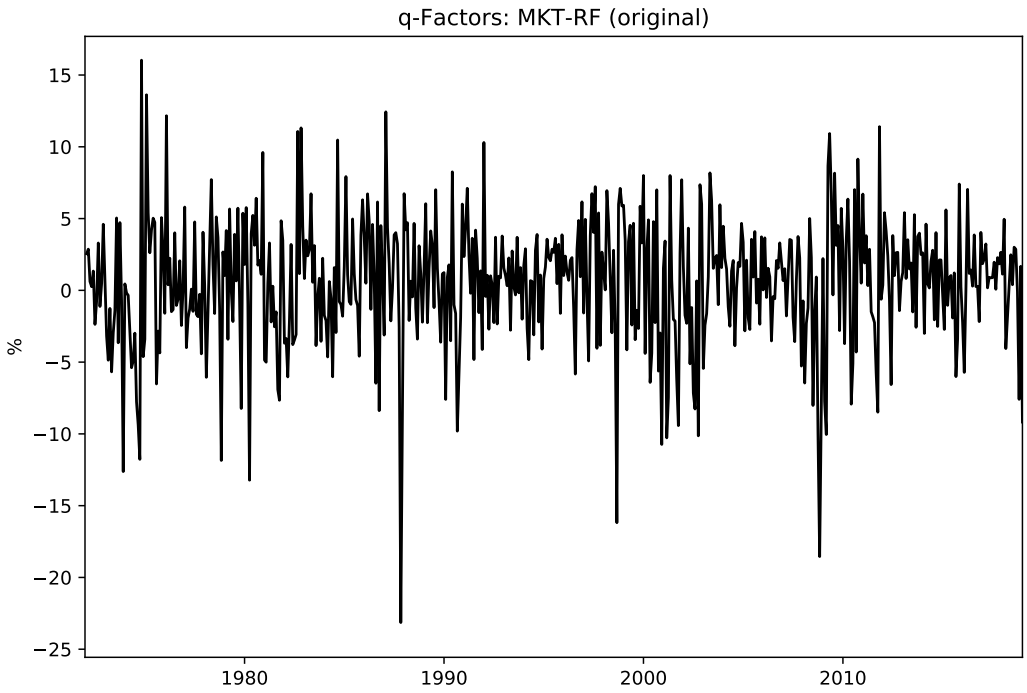
index	MKT-RF	ME	IA	ROE
mean	15.88	31.34	40.05	42.83
std	17.02	27.11	33.32	35.33
min	0.01	0.07	0.09	0.18
25%	5.04	11.77	14.60	15.41
50%	10.01	23.87	31.86	37.31
75%	20.97	45.22	56.27	58.77
max	156.81	203.84	183.73	227.16

Panel C: Correlation between original and replicated factors (basis points)

Statistic	MKT-RF	ME	IA	ROE
Correlation	0.9987	0.9914	0.9614	0.9768

Figure 1: q-factor model: MKT-RF (original and replication)

Panel A: Original Factor



Panel B: Replicated Factor

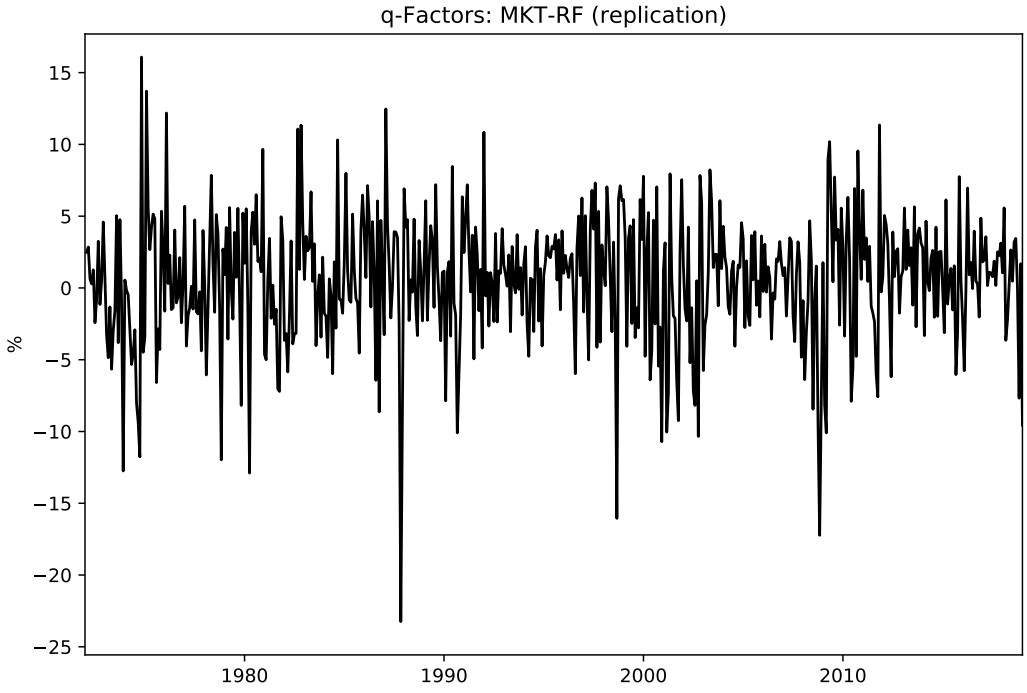


Figure 2: q-factor model: MKT-RF (difference between original and replication)

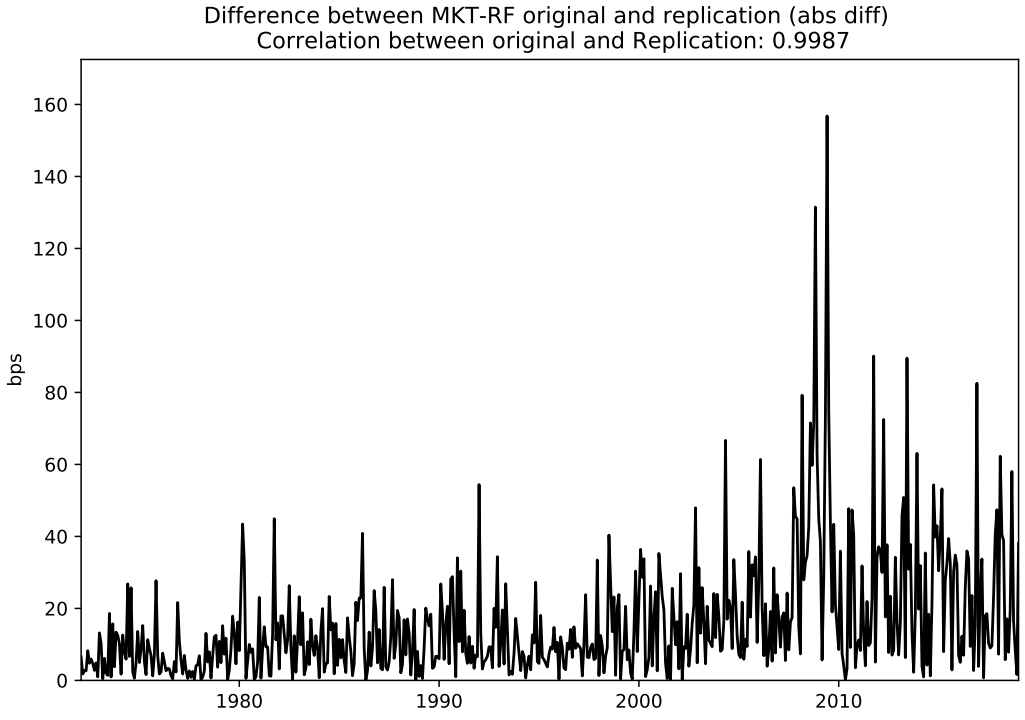
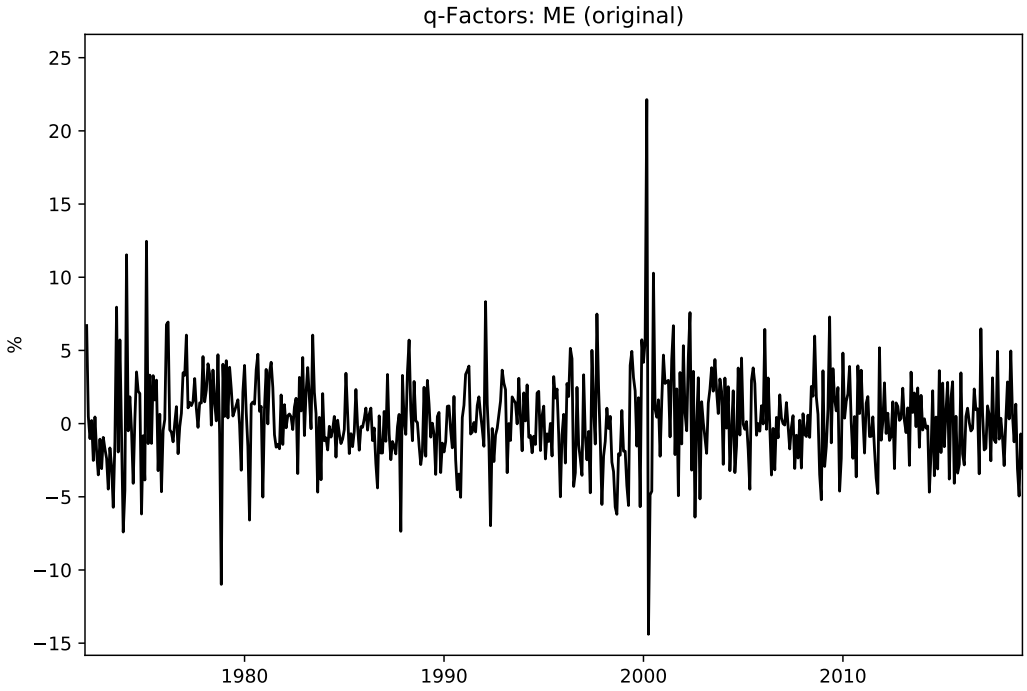


Figure 3: q-factor model: ME (original and replication)

Panel A: Original Factor



Panel B: Replicated Factor

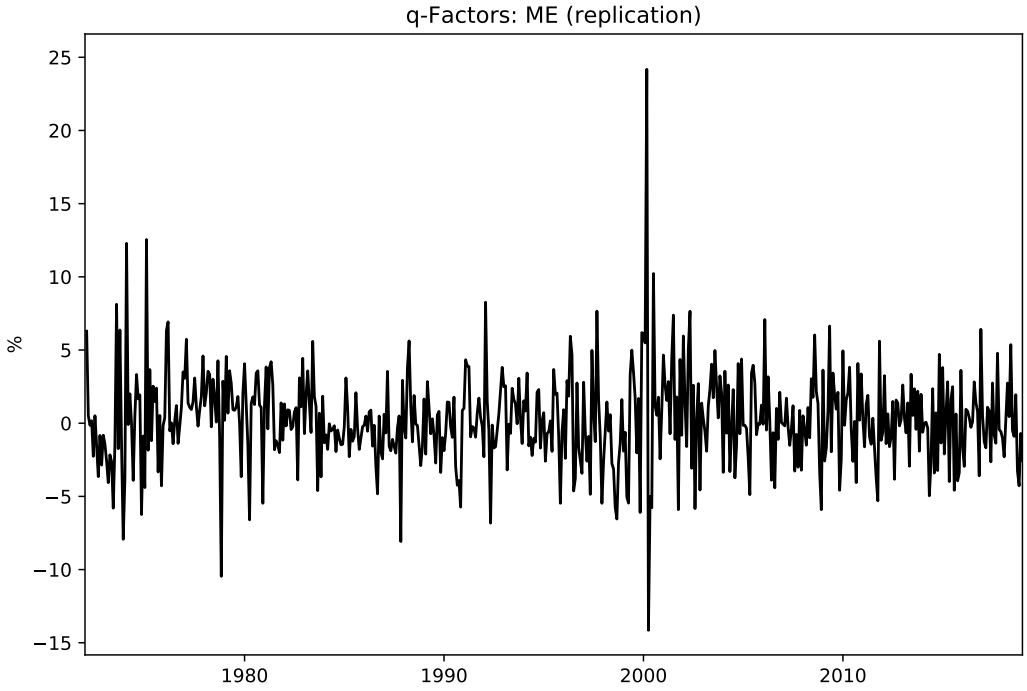


Figure 4: q-factor model: ME (difference between original and replication)

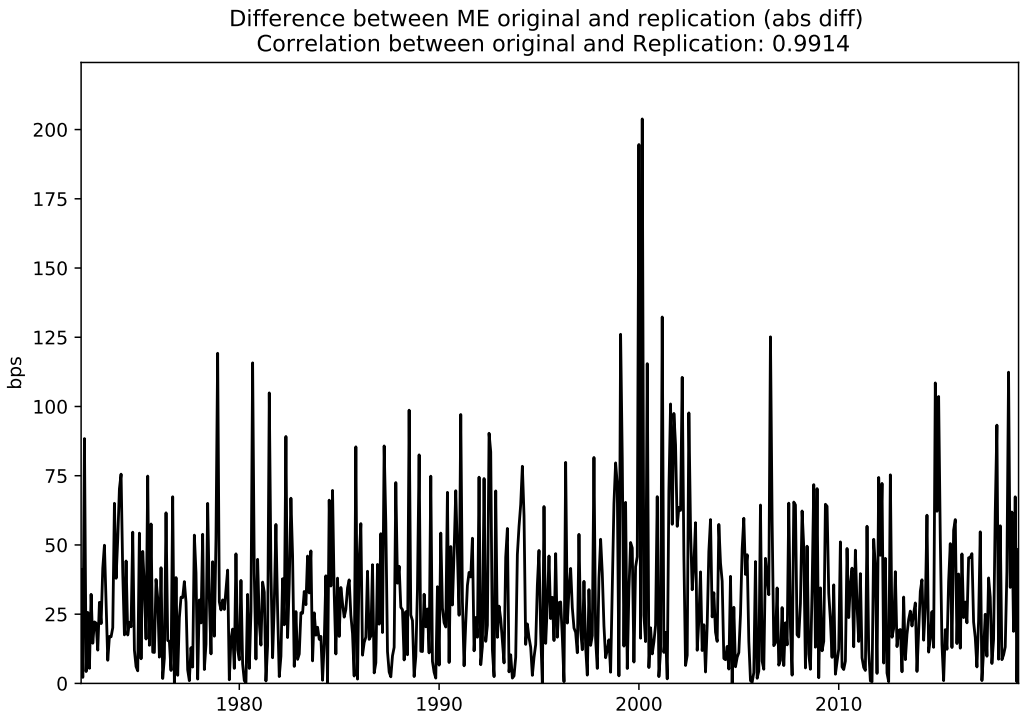
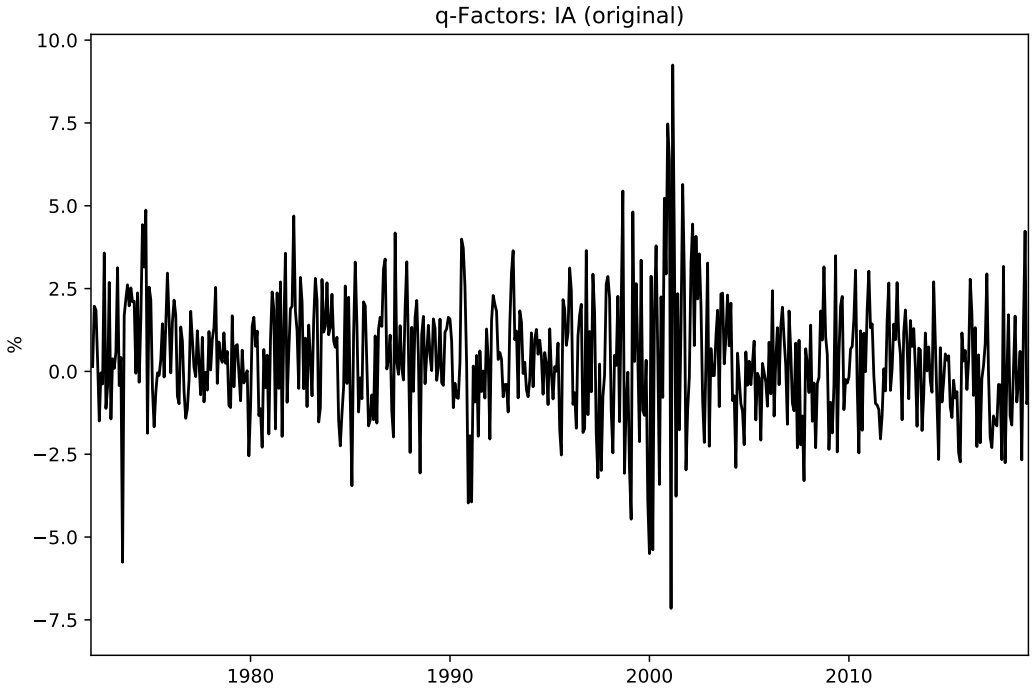


Figure 5: q-factor model: IA (original and replication)

Panel A: Original Factor



Panel B: Replicated Factor

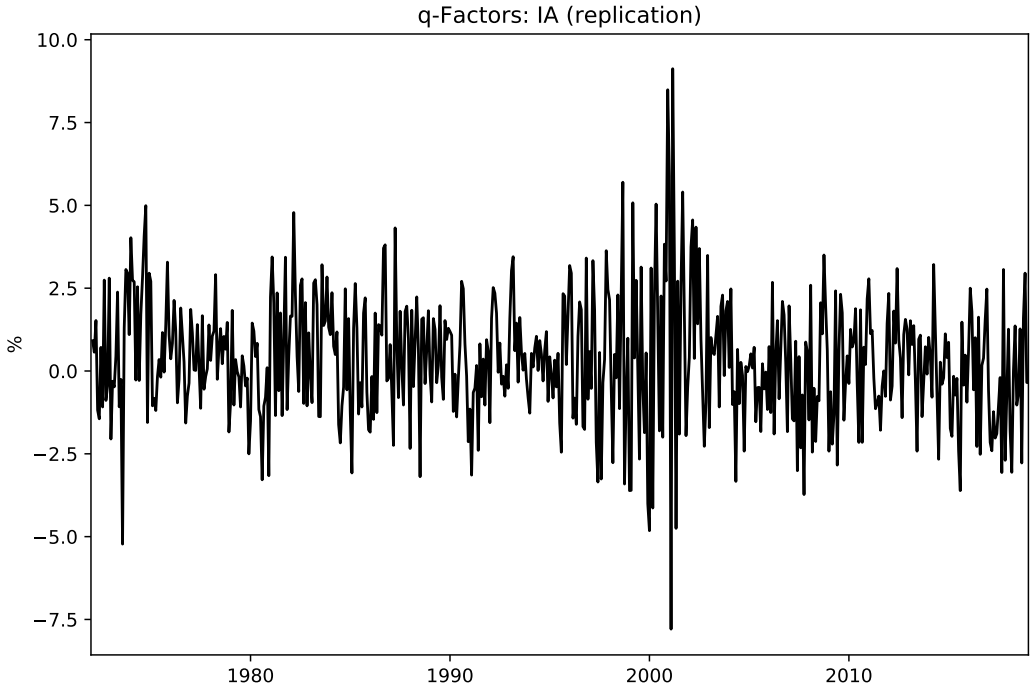


Figure 6: q-factor model: IA (difference between original and replication)

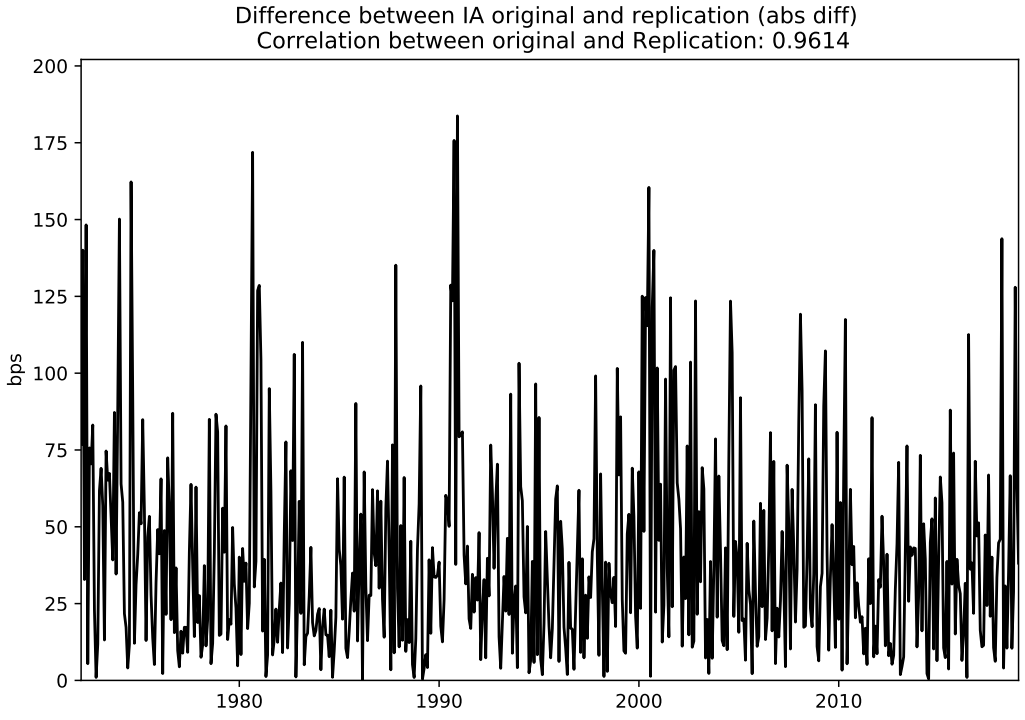
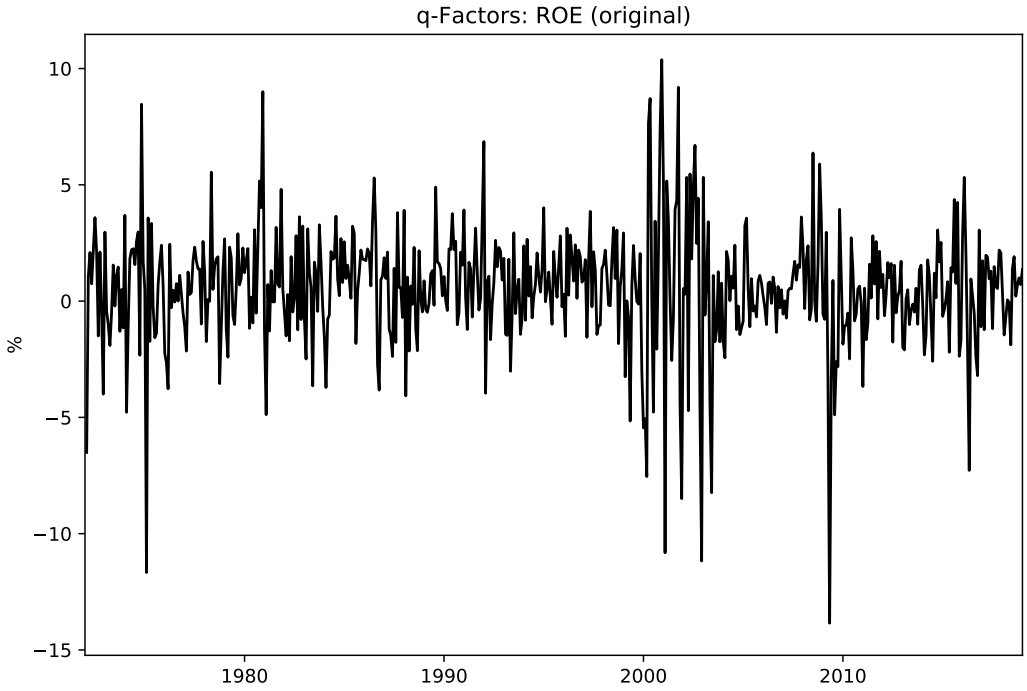


Figure 7: q-factor model: ROE (original and replication)

Panel A: Original Factor



Panel B: Replicated Factor

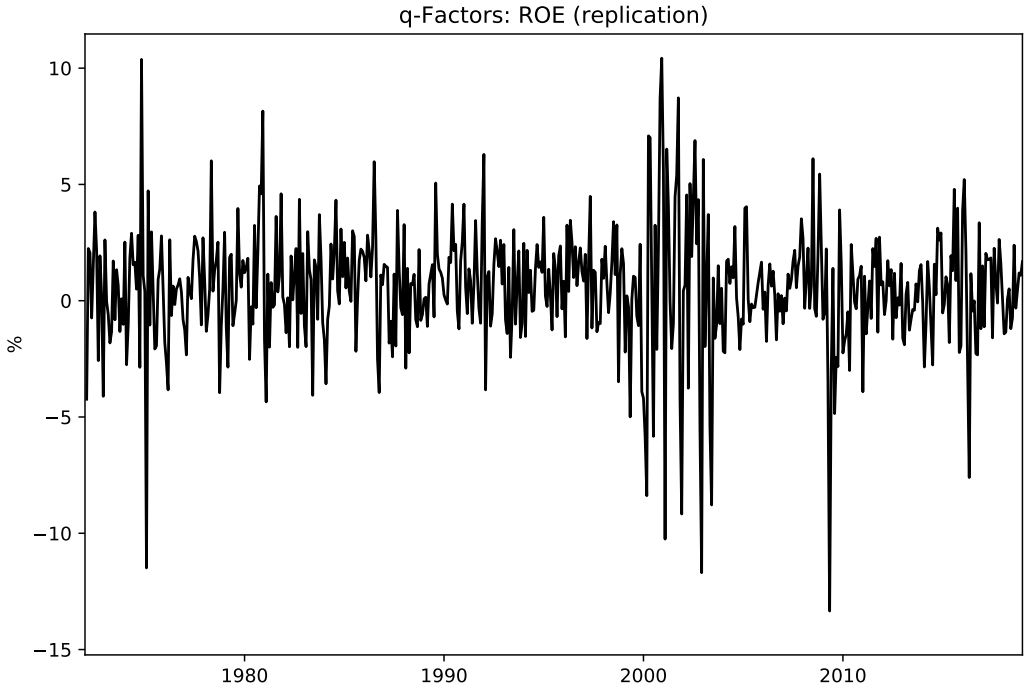


Figure 8: q-factor model: ROE (difference between original and replication)