Moggle

Sirius and Terrarium

Let’s play, create and farm

**Factor Series for All - II**

**BAB, QMJ Factor**

**May 2022**

Kiyoon Jang

Geonsu Kim

**Contents**

2

3

4

7

13

15

Abstract

Motivation

Paper Review

Verification

Conclusion

Appendix

**Abstract**

CAPM 모형의 확립 이후로 시장에 존재하는 리스크 요소를 팩터로 정의하며 이를 활용하여 포트폴리오의 수익률을 설명하고, 궁극적으로 초과수익률을 거두기 위한 시도가 지금까지 계속됐다. 그 중에서 가장 유명한 연구는 Fama and French (1992, 1993)의 연구로, CAPM의 시장수익률 팩터에 더해, 소형주와 가치주가 갖는 리스크 팩터를 추가하여 포트폴리오의 초과수익률을 설명하는 팩터 모델링을 보편화시켰다. 이는 팩터의 대표적인 연구로 일컬어지며 현재까지도 팩터 모델링의 대명사로 자리매김하고 있다. Fama and French는 이어서 2015년 논문을 통해 투자와 수익성에 대한 리스크 팩터를 추가하여 5팩터 모델을 추가적으로 발표하였고, 이는 기존 3팩터로 설명되지 못한 초과수익률에 대해 보완한 연구였다. 🡪 수정필요

우리 모형의 결과가 들어가야함.

Moggle-Labs의 이번 리포트는 한국 주식시장 데이터를 이용해 BAB(Betting Against Beta), QMJ(Quality Minus Junk) 팩터가 한국 주식시장에서도 작동하는 팩터인지 판단하는 작업을 담고 있다. 우리의 리포트는 단순히 결과를 대중에게 알리는 것뿐만 아니라, 관심있는 사람들이 직접 데이터를 다루고 우리의 API를 이용해서 우리와 같은 결론에 도달할 수 있도록 하는 것이 목적이며, 금융시장의 대표적인 팩터들에 대해 산재되어 있는 정보들을 아티클에 모으고 많은 사람들이 손쉽게 따라해보며 팩터에 대한 이해를 넓힐 수 있다는 것에서 이번 리포트의 의의가 있다고 할 수 있다.

**Motivation**

‘쉽게 따라할 수 있는 퀀트 투자방식’이라는 동기로 설립된 Moggle-Labs는 리포트의 형식으로 많은 사람들이 퀀트 투자방식에 대해 깊이 이해하고 직접 데이터를 만지며 따라할 수 있는 환경을 지향합니다. Moggle-Labs가 처음으로 발간하는 리포트 시리즈인 ‘모두를 위한 팩터 시리즈’는 팩터 모델링에 대한 이해를 돕고 API와 데이터를 이용해 직접 따라할 수 있도록 도움을 주는 것을 목적으로 합니다. 기본적인 Fama-French 5팩터 모델에서 시작하여 궁극적으로는 AI기반 팩터 탐색 엔진을 소개하고, 이를 기반으로 저희 홈페이지에서 모든 사람들이 자유롭게 다양한 전략과 아이디어를 실험해볼 수 있도록 할 것입니다. Fama-French 5팩터 모델에 대해 설명한 아티클1에 이어 이번에는 금융시장의 대표적인 팩터인 BAB(Betting Against Beta), QMJ(Quality Minus Junk) 팩터를 소개하고 한국 주식시장에 적용해봄으로써 팩터 모델에 대한 넓고 깊은 이해를 가져다줄 것입니다.

이번 리포트의 의의는 금융시장 대표 팩터인 BAB, QMJ 팩터를 한국 주식시장에 적용시키는 과정과 결과를 모든 사람이 보면서 따라할 수 있도록 공개했다는 것에 있습니다. 지금껏 여러 기관 및 개인이 해당 논문을 한국 주식시장에 적용한 시도는 있어왔지만, 중간 과정 없이 팩터값 또는 팩터 계수만을 공개하거나 팩터를 정의하는데 필요한 세세한 기준을 자의적으로 변경하여 각기 다른 기준을 갖고 팩터를 정의하는 문제점이 있었습니다. 이런 부분 때문에 개인이 실습으로 해당 팩터들을 한국 주식시장에서 추출하고 이를 활용한 투자전략을 구현하기에도 어려움이 따랐을 거라 생각합니다. Moggle-Labs의 이번 리포트는 이런 문제점을 해결하기 위해 결과를 도출하는 코드가 담긴 API를 공개해 모두가 동일한 데이터와 기준을 갖고 팩터를 정의하고 모델링할 수 있도록 하는데 집중했습니다. 저희의 이런 시도가 팩터 모델링의 대중적인 이해를 넓히는 데 긍정적인 영향을 주기를 바랍니다.

**Paper review**

Frazzini, Pedersen (2013)와 Asness, Frazzini ,Pedersen (2018) 연구는 Fama and French model 이후로 시장의 초과수익률을 설명하는 팩터를 찾는 연구였습니다. 해당 연구들에서는 기존의 연구들로 설명될 수 없던, 시장의 이상현상으로 받아들여지는 현상들을 설명하는 팩터를 찾는 데에 집중했고 각각 BAB팩터와 QMJ팩터를 발견했습니다. BAB팩터는 시장위험 노출도가 낮은 저베타 종목일수록 기대수익률이 높다는 의미를 갖으며, QMJ팩터는 투자자들이 더 높은 가격을 지불하고자 하는 성격들을 정의한 후 해당 성격들이 종합적으로 높은 종목일수록 기대수익률이 높다는 의미를 갖고 있습니다. 데이터는 Asness, Frazzini, ,Pedersen (2018)는 1957년 6월 ~ 2016년 12월까지의 24개국에 대한 데이터를 활용했으며 Frazzini, Pedersen (2013)는 1926년 1월 ~ 2012년 3월까지의 20개국에 대한 데이터를 연구 대상으로 삼았습니다. 두 연구 모두 팩터스코어를 기준으로 오름차순 정렬하여 10개의 포트폴리오로 나눈 후 초과수익률의 단조성을 확인하였고, Three-Factor model, Four-Factor model(모멘텀팩터[[1]](#footnote-1)가 포함됨), Five-Factor model에서의 알파를 확인하여 해당 팩터의 유의성을 확인했습니다. 팩터는 Article 1의 Fama and French model가 동일하게 팩터스코어에 대해 상위 포트폴리오와 하위 포트폴리오를 구성한 후, 그 둘의 수익률 차이로 나타냈으며. Rm-Rf 팩터는 CAPM에서와 마찬가지로 시장 초과수익률로 이루어진 팩터입니다.

BAB팩터를 구성할 때 사용하는 팩터스코어는 아래의 식을 통해 산출되었으며, 여기서 와 는 각각 주식 종목과 시장의 수익률에 대한 변동성 추정량이고 는 주식 종목과 시장 수익률에 대한 상관계수입니다.

여기서 주의해야할 점은, BAB팩터의 팩터스코어가 단순히 CAPM모델로부터 도출되는 Beta값이 아니라, 변동성 추정량과 상관계수를 계산할 때 사용한 데이터의 기간을 달리하여 따로 구한다는 점입니다. 변동성 추정량은 1년씩 rolling하여 계산한 표준편차인 반면, 상관계수는 5년씩 rolling하며 구하는 수치입니다. 이렇게 차이를 두는 이유는 둘이 변화하는 속도가 다르기 때문입니다. 또한 이에 필요한 수익률을 계산할 때는 수익률에 log값을 취한 후 3일씩 중첩하여 계산합니다.

뿐만 아니라, 아웃라이어 효과를 막기 위해 시계열 Beta를 횡단면 Beta에 맞춰 축소시키는 작업도 필요로 합니다.

이 아티클에서는 원 논문을 따라 으로 설정하고, 로 설정한 후 작업을 진행했습니다. 이렇게 구한 Beta의 중앙값을 기준으로 고베타, 저베타로 포트폴리오를 나눈 후 저베타에 속하는 포트폴리오에서는 베타가 낮을수록 높은 가중치를 주었으며, 고베타에 속하는 포트폴리오에서는 베타가 높을수록 높은 가중치를 주어 BAB팩터를 완성했습니다.

QMJ팩터에 대한 퀄리티 팩터스코어는 수익성, 성장성, 안정성이라는 세 가지 하위기준의 합으로 구성되며, 각각은 항목에 대한 z-score로 계산되어 합산합니다. 주의해야할 부분은 z-score로 변환하기 전에 rank()를 통해 항목들을 정렬한 후 해당 순위값을 z-score로 변환시키는 부분입니다. 하위기준을 구성하는 항목에 대한 설명은 Appendix를 참고해주시기 바랍니다.

이후, Factor Series for All – I에서 소개한 것과 같이 시가총액(Size)를 기준으로 먼저 두 포트폴리오로 나눈 후, 그 안에서 퀄리티 팩터스코어를 기준으로 나누어 2x3 포트폴리오를 만든 후, 상위 점수의 포트폴리오는 Quality포트폴리오, 하위 점수의 포트폴리오는 Junk포트폴리오로 부르며 두 수익률의 차이를 QMJ팩터라고 계산했습니다. 이 때 breakpoint는 이전 아티클에서와 같이 Size는 NYSE의 Median값, 퀄리티 팩터스코어는 NYSE기준 상하위 30%를 택했습니다.[[2]](#footnote-2)

**Verification**

저희는 위에서 살펴본 Frazzini, Pedersen (2013)와 Asness, Frazzini, ,Pedersen (2018)의 방법론이 한국 주식시장에서도 적용되는지 알아보기 위해 한국 데이터로 재구성하였습니다. BAB팩터의 경우 2005년 1월부터 2021년 6월까지의 198개월 데이터를 활용하였고, Quality팩터는 2011년 1월부터 2020년 12월까지의 132개월 데이터를 활용하였으며, KOSPI, KOSDAQ 기업 재무 및 주식시장 데이터입니다. 기업의 가격과 시가총액데이터는 KRX데이터를 이용했으며, 재무데이터는 Sirius API[[3]](#footnote-3)를 활용하여 가져왔습니다. 팩터를 계산하기 위한 포트폴리오 masking은 Paper Review의 정의식과 동일한 형태를 취했으며, NYSE대신 KOSPI로 대체하여 breakpoint를 설정했습니다. 또한 시장수익률은 KOSPI와 KOSDAQ의 시가총액 가중 수익률을 택하였고, 무위험이자율은 CD91일물 수익률로 계산하였습니다. 팩터 및 회귀의 설명변수에 필요한 포트폴리오를 구성할 때는 이상치 문제를 해결하기 위해 5% winsorization을 시행한 후 진행했습니다.[[4]](#footnote-4)

팩터의 유의성을 검증하기 위해 팩터스코어별로 종목을 정렬시킨 후 오름차순 10분위로 나누어 포트폴리오를 구성하였고, 각 포트폴리오의 팩터스코어와 초과수익률, 샤프지수 등의 단조성을 확인하였습니다. 그 이후, 포트폴리오의 초과수익률을 종속변수로 하고 기존에 알려진 팩터들[[5]](#footnote-5)의 초과수익률을 설명변수로 하여 그 Alpha값의 통계적 유의성과 분포를 확인하는 방식으로 진행했습니다. 또한 마지막으로 정렬시킨 포트폴리오가 아닌, 팩터의 초과수익률로 변환하여 종속변수로 삼은 회귀분석을 추가적으로 진행하여, 각 팩터의 팩터 계수 값을 확인하여, 해당 팩터가 다른 팩터들과 어떤 관계에 있는지 확인하는 작업을 진행했습니다.

표1은 포트폴리오에 따른 팩터스코어의 분포를 기간을 나눠 보여주고 있는 표입니다. Quality 스코어의 경우 데이터의 부족으로 2016년 12월부터의 팩터스코어만 산출할 수 있었기에 해당

표 1. 기간에 따른 팩터스코어의 분포

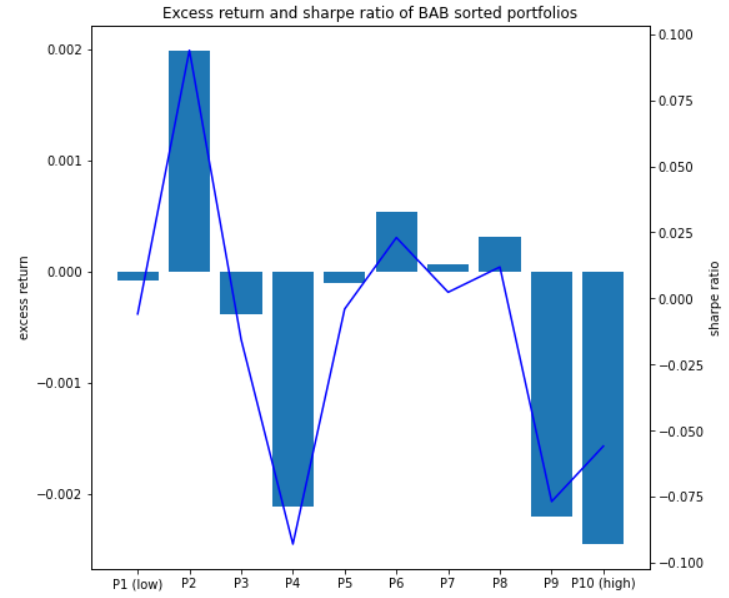
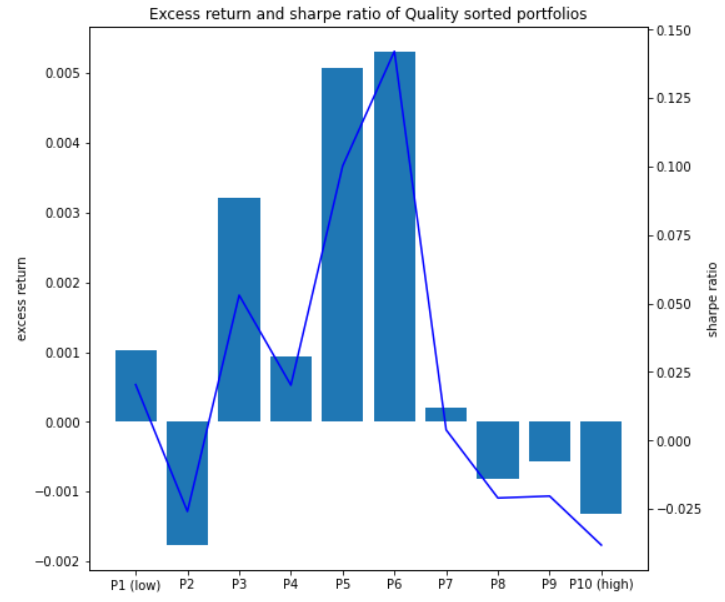
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | P1  (Low) | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 (High) |
| Panel A: BAB | | | | | | | | | | |
| 2005. 01 ~ 2021. 06. | 0.21 | 0.44 | 0.58 | 0.70 | 0.81 | 0.92 | 1.03 | 1.16 | 1.33 | 1.73 |
| 2008. 12 ~ 2021. 06. | 0.21 | 0.46 | 0.60 | 0.72 | 0.83 | 0.94 | 1.06 | 1.19 | 1.36 | 1.76 |
| 2013. 12 ~ 2021. 06. | 0.19 | 0.44 | 0.57 | 0.68 | 0.78 | 0.89 | 1.00 | 1.12 | 1.29 | 1.69 |
| 2016. 12. ~ 2021. 06. | 0.14 | 0.36 | 0.49 | 0.58 | 0.67 | 0.76 | 0.86 | 0.96 | 1.10 | 1.47 |
| Panel B: Quality |  |  |  |  |  |  |  |  |  |  |
| 2016. 12. ~ 2021. 06. | -1.56 | -1.22 | -0.87 | -0.52 | -0.18 | 0.17 | 0.51 | 0.86 | 1.21 | 1.55 |

기간에 대한 결과만 실었습니다.[[6]](#footnote-6) 해당 표는 기간별로 포트폴리오에 따른 팩터스코어의 분포가 강건하게 나오는지 확인할 수 있는 결과인데, 기간별로 비교가 가능한 BAB팩터스코어의 경우, 가장 기간이 짧은 2016년 12월부터의 결과에서 팩터스코어가 다른 기간들에 비해 낮은 경향을 볼 수 있습니다. 특히 상위 포트폴리오로 갈수록 다른 기간들과의 스코어 차이가 많이 나는 것을 확인할 수 있는데, 이는 한국 주식시장에서 BAB팩터가 최근으로 올수록 강건성이 떨어지고 있다고 해석할 수 있습니다. 그 이전까지의 세 기간에 대해서는 비교적 강건하게 팩터스코어가 분포하고 있습니다.

그래프1은 팩터스코어를 기준으로 종목을 정렬한 후, 10분위의 포트폴리오로 구성하여 포트폴리오별 초과수익률과 샤프지수를 확인한 그래프입니다. 왼쪽 그래프가 Beta를 기준으로 정렬한 포트폴리오에 대한 결과치이며, 오른쪽 그래프가 Quality 스코어를 기준으로 정렬한 포트폴리오입니다. Beta를 기준으로 정렬한 포트폴리오의 경우 두 번째 포트폴리오에서 가장 높은 초과수익률과 샤프지수를 기록하였고, 아홉 번째와 열 번째 포트폴리오에서 가장 낮은 초과수익률과 샤프지수를 기록하며, 저Beta일수록 높은 초과수익률을 보인다는 원논문과 비슷한 결과가 나왔습니다. Quality 스코어를 기준으로 정렬한 포트폴리오의 경우, 첫 번째와 열 번째 포트폴리오와 같은 양 극단의 포트폴리오는 낮은 수익률을 보이고 있으며, 오히려 중간단계의 포트폴리오들이 높은 수익률을 기록하고 있음을 알 수 있고 샤프지수도 거의 동일한 양상을 띄고 있습니다.

그래프2는 위에서 구한 포트폴리오의 수익률을 Factor Series for All – I에서 다룬 FF5팩터 모델의 종속변수로 넣어 회귀분석을 돌렸을 때의 alpha값에 대한 결과치입니다. Beta기준과 Quality 스코어 기준에서 모두 그래프1과 상당히 유사한 값이 나왔으며 이는 alpha값과 초과수익률의 변동이 거의 비슷하다고 볼 수 있습니다. FF5 모델의 5가지 팩터들이 두 기준으로 구성한 포트폴리오의 수익률을 설명하는 정도에 한계가 있기 때문에 이런 결과가 나온다고 해석할 수 있습니다.

그래프 1. 팩터스코어 기준 정렬 포트폴리오 초과수익률과 샤프지수



그래프 2. 팩터스코어 기준 정렬 포트폴리오의 FF5 모델에서의 Alpha값

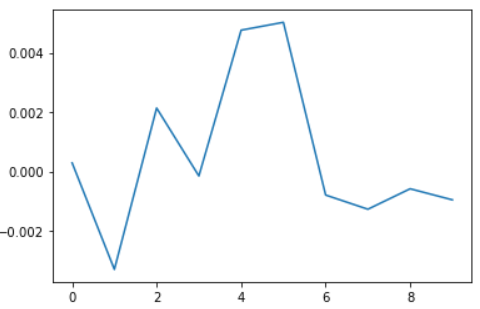
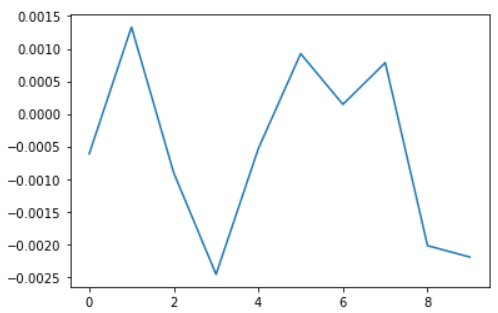


표 2. 팩터스코어 기준 정렬 포트폴리오

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | P1  (Low) | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 (High) |
| Panel A: Beta | | | | | | | | | | | |
| Excess Return | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | (-0.06) | (1.02) | (-0.17) | (-1.01) | (-0.04) | (0.25) | (0.03) | (0.13) | (-0.84) | (-0.61) |
| CAPM alpha | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | (0.09) | 1.62 | -0.03 | -1.25 | 0.16 | 0.58 | 0.31 | 0.47 | -1.06 | -0.66 |
| FF3 alpha | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | -1.27 | 0.48 | -1.51 | -2.43 | -0.99 | -0.33 | -0.98 | -0.38 | -1.92 | -1.44 |
| FF5 alpha | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | -0.74 | 1.17 | -0.75 | -1.90 | -0.38 | 0.71 | 0.13 | 0.62 | -1.29 | -0.88 |
| Sharpe | -0.01 | 0.09 | -0.02 | -0.09 | 0.00 | 0.02 | 0.00 | 0.01 | -0.08 | -0.06 |
| FF5 Adjusted R2 | 0.63 | 0.69 | 0.74 | 0.65 | 0.70 | 0.67 | 0.78 | 0.74 | 0.68 | 0.65 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | |
|  | P1  (Low) | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 (High) |
| Panel B: Quality | | | | | | | | | | | |
| Excess Return | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | 0.15 | -0.19 | 0.39 | 0.15 | 0.73 | 1.03 | 0.03 | -0.15 | -0.15 | -0.28 |
| CAPM alpha | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|  | -0.01 | -0.56 | 0.31 | 0.00 | 0.84 | 1.07 | -0.19 | -0.48 | -0.41 | -0.54 |
| FF3 alpha | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| FF5 alpha | -0.05 | -0.94 | 0.29 | -0.14 | 0.93 | 1.15 | -0.41 | -0.60 | -0.52 | -0.63 |
|  | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 |
| Sharpe | 0.07 | -0.79 | 0.49 | -0.04 | 1.34 | 1.25 | -0.20 | -0.37 | -0.23 | -0.29 |
| FF5 Adjusted R2 | 0.02 | -0.03 | 0.05 | 0.02 | 0.10 | 0.14 | 0.00 | -0.02 | -0.02 | -0.04 |

**Conclusion**

**Reference**

[1] Frazzini, A. & Pedersen, L.H., 2013, Betting against beta, Journal of Financial Economics, 111 (2014), pp. 1-25

[2] Asness, C.S., Frazzini, A. & Pedersen, L.H., 2018, Quality minus junk, Review of Accounting Studies, 24 (2019), pp. 34-112

**Appendix**

Appendix 1 : QMJ팩터 구성식의 항목들

Moggle

Sirius and Terrarium

**Gradient for ALL**

**Let’s play, create and farm**

1. 모멘텀 팩터는 다음에 출간할 Factor Series for All – III에서 다룰 예정입니다 [↑](#footnote-ref-1)
2. 한국 주식시장 데이터를 통해 검증하는 Verification에서는 NYSE기준을 KOSPI기준으로 대체하여 진행했습니다. [↑](#footnote-ref-2)
3. Sirius API는 저희가 만든 API로, 가공된 재무데이터를 손쉽게 불러올 수 있는 API입니다. 저희 홈페이지의 Sirius API Guideline을 참고해주시기 바랍니다. [↑](#footnote-ref-3)
4. 1% winsorization을 적용해서도 진행했으나, 한국 주식시장의 특성상 이상값이 많아 결과값이 특징을 보이지 않았고 5%가 더 전형적인 결과를 보여주었기에 5% winsorization의 결과를 채택했습니다. [↑](#footnote-ref-4)
5. Factor Series for All – I에 소개된 팩터들입니다. [↑](#footnote-ref-5)
6. Quality 팩터스코어 산출식에 들어있는 Growth Factor 산출식에서 5년을 Lag로 정하고 있기 때문에, 현재 Sirius에서 제공하는 2010년 4분기부터의 재무데이터를 활용하여 구한다면 2016년 4분기부터의 Quality 팩터스코어를 산출할 수 있습니다. [↑](#footnote-ref-6)