

개인 투자자 관심도 팩터를 이용한 Long-Hedge 모델 제안- KOSPI200 종목 중심으로

요약

본 연구에서는 2018년부터 2022년 1월까지 KOSPI200에 편입된 종목을 대상으로 개인의 관심도를 측정하여 팩터 모델에 활용하는 방안을 연구하였다. 팩터의 동태적인 특성을 포착하기 위해 마르코프 국면전환 모델과 자체적으로 개발한 개인 거래대금 팩터를 활용하여 국면을 판별하고, 포트폴리오 구성 과정에서 함께 고려하고자 하였다. 팩터의 강건성을 검증하기 위해 팩터의 매수 비중을 분위수를 기준으로 다양하게 시도하고, Long-Only, Long-Short, Long-Hedge 등 여러 투자 전략을 적용하여 관심도 팩터의 유효성을 검증하였다. 연구 결과 5%의 상위 팩터를 보유하는 Long-Hedge 투자 전략이 연 수익률 환산 시 벤치마크보다 약 9.36%~14.56% 높은 초과수익률을 더 낮은 변동성 아래에서 달성하고 있음을 확인하였다. 개인 투자자의 영향력이 증대되고 있는 한국 주식시장에서 이러한 연구 결과는 이러한 팩터를 바탕으로 하는 투자 상품을 개발하는데 기여할 수 있을 것이라고 기대된다.

목차

I. 서론 -----	3
i. 연구배경	
ii. 연구 목적	
iii. 연구 방법	
II. 연구 설계 -----	7
i. 데이터	
(1) 관심도 스코어	
(2) 개별 종목 주가 및 벤치마크 데이터	
ii. 국면 전환	
(1) 마르코프 전환 모델	
(2) 개인 거래 대금	
iii. 투자 전략 및 백테스팅	
(1) 투자 전략 -----	
(2) 백테스팅 -----	
III. 연구 결과 -----	17
i. 평가지표 -----	
ii. Long-Only 전략 -----	
iii. Long-Short 전략 -----	
IV. 결론 및 시사점 -----	24
V. 참고문헌 -----	29

I. 서론

i. 연구배경

한국 금융시장에서 개인투자자가 미치는 영향력은 코로나 19 이후 점점 증가하고 있다. 지난 2021년 자본시장연구원에서 발간한 보고서에 따르면, 코로나 19가 발생한 2020년도에는 개인투자자가 무려 63조 8천억 원 규모를 순매수하였고, 2020년에서 2021년 1년 동안 주식 계좌 또한 760만 개가 증가한 것으로 파악된다. 또한, 개인투자자 군의 매도 및 매수 비용은 각각 4,387조 원 및 4,323조 원으로, 한국 주식 시장 시가총액의 2.5배 규모인 것으로 드러났다. 이에 힘입어 KOSPI 지수 또한 코로나 19 이후 3000대까지 도달하여 오히려 주식 시장 과열 우려를 자아내기도 했다¹.

관련하여 동 연구소에서 2022년에 개인투자자의 투자 행태에 관하여 발간한 보고서에 따르면² 코로나 19 이후 유입된 개인투자자의 거래 행태를 분석했을 때 일정한 패턴이 보이는 것으로 드러났다. 보고서는 코로나 19 이후 거래에 참여한 개인투자자의 투자 행태를 몇 가지로 구분하였는데 이들은 자신이 보유한 주식에 과잉 확신하고 복권형 주식을 선호하며, 특히 일정한 기간을 거쳐 특정 종목을 같은 방향으로 매수, 매도하는 군집 거래행태를 보인다고 보고하였다. 즉, 개인투자자의 경우 기관투자자에 비해 한정된 정보만을 가지고 투자를 결정하기 때문에, 다수가 관심을 가지고 유행하는 주식에 투자할 가능성이 높다는 점을 지적한다. 이러한 보고서는 개인투자자의 매매 행위가 주가에 미칠 수 있는 영향력이 증가하고 있음을 시사한다. 본 연구는 이러한 개인 투자자의 영향력을 하나의 팩터로 활용할 수 있는지 입증하고자 한다.

ii. 연구 목적

¹ 김준석, 『코로나19 국면의 개인투자자』, 자본시장 포커스, 2021-4호, 2021

² 김민기, 김준석, 『국내 개인 투자자의 행태적 편의와 거래행태』, 자본시장연구원.71-90, 2022

한편 본 연구의 주요 관심 주제인 개인 투자자의 거래 행태는 행동 재무학 및 행동 경제학에서 꾸준히 연구되어왔다. Summers (1990)등은 개인투자자를 noise trader, 즉 일종의 비합리적인 투자자로 간주하여 특정 조건하에서는 주가에 유의미한 영향력을 미칠 수 있는 시장 참가자로 이해하였다. 이러한 선행 연구에서는 개인 투자자의 투자심리와 주가 수익률과 관련된 다양한 해석을 제시했는데, 류두진(2018), 이효정(2020) 등은 개인 투자자들이 선호하는 기업의 특성을 횡단면적으로 분석하여 이들이 선호하는 기업이 주로 소형주, 유형자산이 적고 시장가치 대 장부가치 높은 기업 등 비합리적인 선택을 했음을 분석했다. 투자심리를 측정하는 지표 또한 다양하게 발전했는데, 단일 지표로는 주로 경기 선행지수를 사용하거나 매도-매수 불균형 거래를 활용했고, 다요인 지표로는 Baker et al(2006)이 개발한 Investment Sentiment Index가 그 유용성을 인정받아 현재까지도 활용되고 있다. 한국에서는 김가람(2018)이 BW의 감정 지표를 한국의 주식 시장에 적용하여 그 효과를 입증한 바 있다.

하지만 이러한 다양한 선행연구에도 불구하고, 개인 투자자의 거래 행위를 하나의 팩터로 분류하여 동태적인 투자 전략에 활용한 연구는 아직 없는 것으로 알려져 있다. 본 연구는 개인투자자의 관심도가 주식 시장에 미칠 수 있는 영향력이 나날이 커짐에도 불구하고, 기존의 다요인 팩터 모델 중 개인의 주식에 대한 관심도를 계량화한 모형은 아직 없다는 문제의식에서 출발하였다. 본 연구에서는 개인의 투자 심리를 측정하기 위해 개인 투자자의 거래 행위를 가장 잘 표현할 수 있는 한국의 매체 중 하나인 네이버 주식 종목토론방을 활용하고자 한다. 개인의 관심도라는 팩터를 정의한 후, 주식 시장의 다양한 국면에 따라 초과 수익률을 올릴 수 있는 전략이 있는지 분석하여 주식 시장에서 달라지고 있는 개인 투자자의 위상을 점검해보고자 한다.

iii. 연구 방법

먼저 개인투자자의 특정 종목에 대한 관심도를 측정하기 위해 네이버의 주식

토론방에서 일정 기간 해당 종목이 언급된 횟수를 계량화하여 새로운 팩터를 추출하였다. 또한, 각 국면에 따라 개인 관심도가 주가에 미칠 수 있는 영향력이 다를 것이라는 가정 아래 마르코프 국면 전환 모델(Markov Regime Switching Model)을 활용하여 한국 주식 시장의 국면을 저변동 및 고변동 국면으로 예측하고자 하였다. 또한, 주식 시장 강세와 개인 거래 매매 강도 간의 세밀한 관계를 파악하기 위해 개인 거래대금 데이터를 활용하여 종합적으로 분석하였다.

이를 바탕으로 KOSPI200에 상장된 종목에 대해 관심도 팩터가 얼마만큼 유의미한 영향을 미칠 수 있을지 파악하기 위해 다양한 투자 전략을 수행하였다. 먼저 본 연구에서 가정한 국면 변수가 실제로 유의미한 차이를 도출하는지 분석하기 위해 그 유무에 따라 관심도 팩터의 수익률을 분석한다. 국면 변수가 유의함을 보인 뒤, 각 국면에서 관심도 팩터를 어떠한 방식으로 투자할지 결정하였다. 더 나아가 이를 Long 포지션만 취할 때와 Long, Short 포지션을 모두 취할 때의 포트폴리오 결과와 함께 제시하여 강건성을 검증하고자 하였다. 또한, Long-short모델이 다소 비용이 높은 점을 감안해 kodex inverse ETF로 short포지션을 대체한 현실적인 모델인 Long-Hedge model 역시 검증하고자 하였다.

관심도를 팩터로 한 투자 전략의 기대 효과는 다음과 같다. 우선 엄밀한 백테스팅을 통해 검증된 투자 전략은 높은 수익률을 안정적으로 제공할 수 있을 것으로 기대된다. 본 연구에서는 팩터 스코어 매수 비중의 변화 및 Long-Short 모델을 사용하여 주간 수익률을 기준으로 벤치마크보다 약 0.2% 높은 초과수익률을 달성하면서도, 동시에 더 낮은 변동성을 달성하는 투자 전략을 제안한다. 운용사의 현실적인 제약들을 고려하였기에, 실제 투자 현장에서도 새로운 상품으로 이용되는 등 그 범용성이 높을 것으로 기대된다.

또한, 본 연구는 이전의 개인 투자자의 투자 심리와 주식의 수익률 간 관계

를 보인 선행 연구에서 확장하여, 거시 경제적인 상황을 고려한 팩터 포트폴리오를 수립한다는 점에서 그 의의가 있다. 이러한 연구는 코로나 19 및 거시경제적 변화에 의해 개인이 주가에 미칠 수 있는 질적, 양적인 영향력이 향후 계속 증가할 것이라고 기대되는 상황 속에서 개인의 투자 행태를 밝히는 유의미한 연구가 될 것이라고 기대된다.

II. 연구 설계

i. 데이터

(1) 관심도 스코어

본 연구에서는 개인의 관심도를 측정하기 위해 네이버 주식 종목 토론방에서 해당 종목이 언급된 횟수를 활용하였는데, 종목 토론방이 주식 종목에 대하여 토론하는 가장 활발한 매체로 알려져 있기 때문이다. 수집 기간은 게시물이 처음 기록된 2017년부터 2022년 2월까지이며, 특정 종목에 대한 게시글이 과도하게 많을 경우 10 페이지 간격으로 샘플링하여 왜곡을 줄이고자 하였다. 이를 활용한 관심도 score는 다음과 같다.

$$\frac{\text{좋아요} + \text{싫어요}}{\text{해당 게시글의 조회수}} \quad (1)$$

박단비(2020) 등의 선행연구는 종목에 대한 감정 분석을 수행하기 위해 종목 토론방 내 게시물 중 해당 종목이 언급된 게시글의 좋아요 및 싫어요를 합한 횟수가 0 이상, 즉 종목에 대한 최소한의 관심이 있는 경우에만 데이터로 활용하였다. 본 연구에서는 종목에 대한 긍부정이 아닌 종목에 대한 관심도 그 자체를 구하고자 하였다. 이를 측정하기 위해 해당 종목을 언급한 게시글의 “좋아요”와 “싫어요” 수를 합쳐서 긍부정을 합친 개인들의 전체적인 관심도를 구하였는데, 이는 종목 토론방이라는 매체의 특성 상 광고성 글이 많았기 때문에 단순히 해당 종목을 언급한 게시글의 수만으로 스코어링하는 것은 노이즈를 제거하지 못할 수 있다고 판단했기 때문이다. 이렇게 구한 개인의 관심도 점수가 기업의 규모 및 인지도에 의해 점수가 왜곡되지 않도록 조회수로 정규화하였다. 게시물마다 해당 관심도 score가 산출되며 투자지표로 활용되기 위해서 투자시점에서 4주전까지의 게시물의 관심도 score의 평균을 산출해 최종 score로 지정한다. 특정 기간 동안 언급되지 않았던 종목도 수치화하기 위해, 만약 해당 기간의 게시글 데이터

가 없다면 팩터의 스코어가 0이 되도록 설정하였다.

날짜	제목	글쓴이	조회	공감	비공감
2022.01.10	달달하다 달달해π	vikt****	1277	0	0
2022.01.10	실손보험 지급기준 마무리 적용 초읽기	jy88****	950	1	0
2022.01.10	[강세업종] 보험 관련주 상승세...DB손해보험 4% ↑	sykw****	1072	0	0
2022.01.10	제2의 메리츠	topb****	1470	0	1
2022.01.10	db는 개미거의 없겠는데요? 유통물량 10%	dizn****	1339	2	0
2022.01.09	□□핵심 주간 증시일정□□	tnsg****	1100	1	0
2022.01.08	미국시장으로 본 시장 흐름.	5841****	1162	2	0
2022.01.08	이재명 "실손보험 청구 간편하게 제대로 바꾸겠다"	dbwn****	1046	1	1
2022.01.07	실손보험 지급기준	jy88****	1070	2	0
2022.01.07	이런 주식을	jsm8****	1005	3	0
2022.01.07	미국장의 특징-->보험주 버크셔 해서웨이 큰 상승	5841****	824	1	0
2022.01.06	주식 개고수 디비손보 입성	wjdd****	978	0	0
2022.01.06	상?가 만들어서	jsm8****	747	0	0

그림 II-1<종목 추출 예시-DB손해보험>

2018-02-02 00:00:00	2018-03-02 00:00:00	2018-03-30 00:00:00	2018-04-27 00:00:00
0.005280497	0.004650859	0.00286264	0.003070537
0.008078323	0.010096159	0.007915263	0.013050324
0.005587587	0.008550122	0.005862923	0.006806617
0.005893295	0.004074499	0.009026496	0.005308023
0.008652586	0.006520817	0.008688247	0.010667401
0.009010242	0.007233853	0.007073621	0.009616261
0.010945792	0.011587632	0.012438489	0.016569935
0.010617893	0.009616126	0.010134657	0.010951899
0.002537086	0.001894854	0.003564958	0.003346872
0.006035146	0.006230259	0.006491818	0.006451863
0.007784182	0.008229999	0.006921557	0.00826738
0.009640488	0.011075737	0.0096281	0.011543234
0.007511719	0.009243993	0.010103858	0.007268526
0.006972809	0.005270882	0.004934941	0.005797165
0.013744386	0.015508274	0.014342382	0.012403521

그림 II-2<관심도 팩터 스코어링 예시>

(2) 개별 종목 주가 및 벤치마크 데이터

벤치마크 데이터로는 KOSPI200을 선정하였으며, 개별 주식 및 벤치마크의 가격 데이터는 KRX Api를 통해 추출하여 데이터 프레임화 하였다. 또한, 본 연구는 백테스트 과정에서 과거의 실적이 부진한 기업이 퇴출되고 우수한 기업만 생존하여 백테스트의 수익률이 과대평가되는 생존자 편향이 생기는 것을 제거하기 위해, 실험 기간 동안 벤치마크를 구성하는 종목의 편입 편출까지 동태적으로 고려한 투자 유니버스를 사용하고자 했다. 아래 <그림 II-1>은 2018년도의 코스피 200

종목의 편입과 편출을 반영한 데이터프레임을 예시로 보이고 있다. 본 연구에서는 해당 작업을 2022년까지 수행한 데이터 사용하였다.

날짜	편입 종목 코드	편출 종목 코드
2018-03-09	068270	004710
2018-02-02	2018-03-09	
005930	005930	
004710	068270	
138930	138930	
...	...	

그림 II-1 <Kospi200 종목 편입 편출 반영한 데이터프레임 예시>

ii. 국면 판별

본 연구에서는 투자 국면의 복합적인 특성을 포착하기 위해 총 2가지 지표를 활용하여 국면을 판별하고자 한다. 먼저 주가 시장의 변동성 및 전체적인 강세, 약세장을 판별하기 위해 마르코프 국면 전환 모형을 활용하여 주식시장의 국면을 예측하였다. 또한, 개인의 관심도에 기반한 단일 팩터가 영향력이 생기는 시기를 고려하기 위해 개인 거래대금 데이터를 활용해 개인 투자자의 순매수금이 증가 또는 감소하는 구간으로 구분하였다. 이를 통해 향후 백테스팅 과정에서 두 지표 간의 관계를 총 3가지(강세-개인 거래 감소, 약세-개인 거래 증가, 그 외) 경우로 구분하여 상이한 투자 전략을 구축할 수 있도록 하였다.

(1) 마르코프 국면 전환 모델

여러 선행연구에서 거시 경제적 상황이나 시장 환경의 변화에 따라 팩터의 초과 수익률이 크게 변동한다는 사실을 보고하였다. 이준행(2017)은 마르코프 국면 전환 모형을 활용하여 가치주와 성장주의 비중을 시장의 불황 및 활황 국면에 따라 동태적으로 조정하는 포트폴리오를 제안했는데, 이는 두 팩터를 일정한 비율로 유지한 정태적인 포트폴리오보다 높은 수익률을 달성한다는 사실을 보였다.

본 연구에서는 이러한 선행 연구 결과를 참고하여 한국 주식시장의 국면을 변동성이 낮은 국면, 즉 호황 국면과 변동성이 높은 구간인 불황 국면으로 구별하고자 하였다.

마르코프 국면 전환 모델은 대표적인 State- space 모델로, 시점 t 에서의 관측 가능한 시계열 데이터 y_t 의 이면의 숨겨진 국면 S_t 가 있다고 가정하고, 그 둘 간의 관계를 설명하고자 한다. 즉, 국면 변수와 이전 기까지의 정보가 주어졌을 때 관측 변수의 조건부 확률 $\eta_{jt} = f(y_t | S_t = j, I_{t-1}; \theta)$ 및 전환확률 p_{ij} 를 활용하여 베이저안 추론을 반복적으로 수행하면, 이전 기까지의 정보 집합 I_t 및 모수 θ 에 대하여 국면의 조건부 확률인 $\xi_{j,t} = \Pr(S_t = j | I_t; \theta)$ 을 도출할 수 있다. 본 연구에서는 코스피 지수를 차분한 수익률 데이터의 변동성을 1차 자기회귀 마르코프 국면 전환 모델에 적용하여 국면 예측 변수를 생성하고자 하였는데, 수식은 아래와 같다.

$$y_t \sim N(0, \sigma_t^2) \quad (1)$$

$$y_t^2 = \mu_{St} + y_{t-1}\beta_{St} + \varepsilon_t \quad (2)$$

$$S_{kt} = 1 \text{ if } S_t = k \text{ and } S_{kt} = 0 \text{ otherwise, } k = 1, 2, 3, 4 \quad (3)$$

$$\Pr[S_t = j | S_{t-1} = i] = p_{ij} \quad i = 1, 2, 3, 4 \quad (4)$$

$$\sum_{j=1}^4 p_{ij} = 1 \quad (5)$$

y_t : KOSPI 지수의 수익률, S_t : 관측되지 않는 국면, p_{ij} : 전환확률

본 연구에서는 $S_t \in \{0, 1, 2, 3\}$, 즉 4개의 국면으로 분할하는데, 크게 변동성이 높은 국면과 낮은 국면으로 구분하고, 그 정도에 따라 하위 2개의 국면을 추가하였다. 이때 국면 간의 전환은 전환확률(transition probability)에 따라 발생하는데, 이는 1차 Markov 과정을 따른다고 가정하였다. Python에서 제공하는 statsmodels 패키지를 활용하여 해당 모델을 학습시켰는데, 임의의 초기값에 대

해 EM 알고리즘을 활용하여 모수를 안정적으로 추정할 수 있도록 한다고 알려져 있다. 반복 과정에서 전체 기간 동안의 상태 변수를 전체한 조건부 확률인 smoothed probability인 $\xi_{j,t} = \Pr(S_t = j | y_1, y_2, \dots, y_t, \dots, y_T; \theta)$ 을 사용하였다.

모델을 학습하여 도출한 4개의 국면에 대해 추가적인 조작을 가하여 변동성이 높은 국면과 낮은 국면이라는 직관적인 해석이 가능하도록 하였다. 이를 위해 총 4개의 국면을 변동성이 높아질수록 높은 숫자를 부여하는 방식으로 0에서부터 1까지의 숫자로 인코딩한 뒤, 시점 t에서 그 직전 주의 예측된 국면과 t에서의 국면을 비교하여 변화량이 양수이면 불황, 음수이면 호황으로 가정하였다. 또한, 불황이 유지되거나 호황이 유지되는 국면도 따지기 위해 변화량이 0이라 하더라도 0이 아닌 최근 이전 변화량이 양수면 불황으로 정의하고, 0이 아닌 최근 이전 변화량이 음수면 호황으로 정의하였다. 아래 <그림 II-1>은 마르코프 국면 전환 모델의 결과를 분석 기간인 4년 동안 고변동 구간인 0과 저변동 구간인 1으로 설정하여 시각화한 것이다.

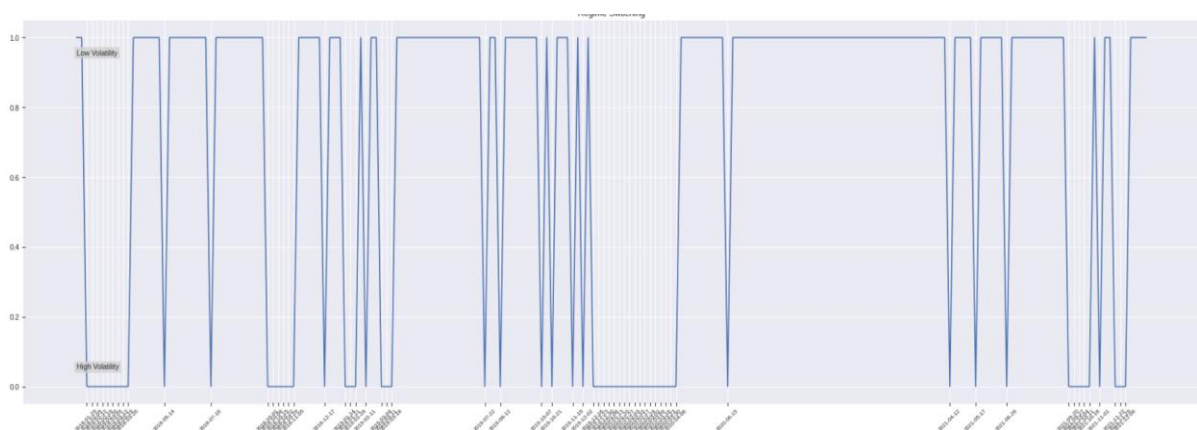


그림 II-3 <마르코프 국면 전환 모델 예측 결과>

(2) 개인거래대금

본 연구의 궁극적인 목표는 개인 투자자의 관심도가 주가 및 투자 전략에 얼마만큼 유의미한 영향을 미치는지 분석하는 것이므로, 단순히 주식 시장의 변동성

만을 예측하는 것은 효과적이지 못할 수 있다. 실제로 박단비(2020)에 따르면, 게시글에 매수의견이 많을수록 당일과 그 다음날까지의 수익률이 통계적으로 유의하게 상승하고 의견 불일치정도가 커질수록 당일과 다음날의 변동성이 유의적으로 증가한다는 것으로 나타났다. 이는 게시글에 대한 관심도가 시장 속 개인의 투자자금에 변화량에 따라 유의적으로 변화할 수 있음을 의미한다.

따라서 본 연구 또한 개인의 투자 참여도 정도를 수치화 하여 국면을 판별하는 지표로 활용하기로 하였다. 이를 위해 한국거래소에서 개인의 순매수 자금 데이터를 활용하여 누적 순매수 금액을 구하고, 시간에 따라 개인 거래대금이 변화하는 정도를 측정하여 개인 투자자의 거래 정도를 측정하고자 하였다. 즉, 지난 12주에서 4주간 개인의 순 거래자금 변화량 평균이 4주 전부터 현재까지 개인의 순 거래자금 변화량 평균보다 작으면 “개인 거래 자금 변화량이 감소하는 시기이자 투자에 소극적인 시기”, 클 경우 “개인 거래 자금 변화량이 증가하는 시기이자 투자에 적극적인 시기”로 정의하였다. 4주는 약 한달이므로 월마다 리밸런싱하는 백테스팅 과정에 적합한 변수설정이라고 판단하였다. 분석 기간은 위와 동일하게 2001년부터 2022년까지의 데이터를 활용하였다.

iii. 투자 전략 및 백테스팅

(1) 투자 전략

본 연구에서는 앞서 정의한 국면에 따라 투자 전략을 다르게 가지고 가는 이유는 주식 시장의 상태를 나타내는 개인투자자의 주식 참여도 변수와 미래 주식 시장의 예측을 암시하는 코스피 시장 국면 예측 변수의 조합에 따라 각 국면에서 개인의 행동이 다를 것으로 판단되기 때문이다. 본 절에서는 우선 시장 국면 변수와 개인 거래 대금 팩터를 조합하여 나올 수 있는 모든 전략의 수를 고려한 뒤, 이 중 선행연구에 의해 뒷받침될 수 있는 투자 전략 2가지를 선정하여 향후 포트폴리오 구축 과정에서 활용하고자 하였다.

	개인 거래 자금 변화량 증가(적극적)	개인 거래 자금 변화량 감소(소극적)
시장 강세 (저변동)	해당 경우, 시장이 상승할 때 상방 변동성에 투자를 하여 관심도가 큰 자산군에 베팅을 해 시장보다 더 큰 수익을 노리는 전략을 취할 수 있지만, 이는 다소 risk taking 포지션이기에 안정적인 변동성을 추구하는 본 연구의 포트폴리오 전략에는 적절치 않다고 판단하고 제외하였다.	이는 시장이 상승을 하지만 개인의 자금이 빠져나가면서 개인들이 관심이 많은 주식 중심으로 하방 변동성이 커지며 하락을 할 가능성이 높다. 이에 따라 하락위험이 적은, 관심도 팩터가 낮은 종목을 위주로 포트폴리오를 구성하기로 하였다
시장 약세 (고변동)	해당 경우, 개인들의 자금에 의해 상방 변동성이 커지며 관심도에 따라 시장대비 방어적인 하락 혹은 수익을 기대할 수 있다. 따라서 하락장에서 개인의 자금이 증가하는 경향이 보일 때 관심도 스코어가 높은 주식들을 중심으로 포트폴리오를 구성하였다	해당 경우, 시장이 하락할 때 개인자금 유입이 감소하면 이는 매수세가 약하기에 하방 변동성이 심화될 가능성이 크다. 하방 변동성에 관심이 많아진 종목을 short 하는 전략을 취할 수 있지만 이는 short 포지션의 비용과 리스크를 고려하여 전략에 포함시키지 않았다

표Ⅱ-1 <국면 변수와 거래 대금 팩터 조합에서 고려할 수 있는 모든 전략>

먼저 이윤정(2013)은 주가 변화와 게시물 수 변화 사이에 뚜렷한 상관관계는 나타나지 않지만 주가가 급락하거나 급등하는 시기에 게시물의 수가 증가하는 경향이 있으며, 주가와 게시물 수의 상관 계수가 강한 양이거나 강한 음을 나타내는 경우 주식수익률도 그에 따라 높거나 낮은 경향이 나타남을 밝혔다. 즉, 시장 국면에 따라 개인의 거래 대금으로 수치화한 개인의 관심도가 시장의 양의 변동성과 음의변동성을 표현할 수 있다. 예를 들어, 시장이 호황이지만 개인자금의 유입정도가 줄거나 혹은 감소하면 이는 개인의 관심도가 높은 종목 위주로 하락변동성이 커질 것이다. 반대로 시장이 불황이지만 개인자금의 유입정도가 늘면 이는 개인의 관심도가 높은 종목 위주로 상승 변동성이 커질 것이다. 본 연구는 시장의 국면과 개인자금의 변화에 따라 상방, 하방 변동성을 고려하여 위와 같이 투자전략을 제시한다. 한편, 시장 약세 국면에서 하방 변동성이 높아진 종목을 short 하는 전략은 리스크를 회피하기 위해 해당 전략은 고려하지 않았다.

아래 <표Ⅱ-2>은 이러한 근거에 따라 Long 포지션만을 취하는 Long-Only 포트폴리오에서 취한 전략을 정리하였다.

국면	투자 전략
시장 활황 및 개인 거래 감소	관심도 팩터 점수가 하위 10%인 종목 매수
시장 불황 및 개인 거래 증가	관심도 팩터 점수가 상위 10%인 종목 매수
그 외	Kodex200 매수

표Ⅱ-2 <Long- Only 투자 전략>

본 연구에서는 더 나아가 앞선 투자 전략을 확장하여 Long- Short을 수행할 때의 수익률도 분석하고자 하였다. Long- Short 포트폴리오는 관심도 팩터에 기반하여 Long 포지션과 Short 포지션을 함께 포트폴리오에 편입하는 전략이다. Long-Short 포트폴리오에서 Long 포지션을 취할 때와 Short 포지션을 취할 때의 수익률 간 차이가 존재하여야 이 둘의 포트폴리오 수익률이 통계적으로 유의하다고 입증할 수 있다³. 즉, 관심도 팩터가 주식의 수익률에 영향을 미친다는 근거를 강건하게 입증하여 해당 팩터의 효과를 검증할 수 있다는 점에 의의를 두고, 이러한 포트폴리오 전략을 구현하고자 하였다.

이에 더하여 실제 투자에서는 그 비용과 번거로움으로 인해 공매도를 현실적으로 수행하기 어렵다는 점을 고려해, KODEX 인버스 ETF로 숏 포지션을 대신한 Long-Hedge 포트폴리오를 사용하여 결과를 분석했다. 구체적인 투자 전략은 위에서 언급한 Long only 포트폴리오의 투자 전략과 동일하게 수행하며, Long-Hedge 포지션의 경우는 주식종목을 매도하는 대신 매수한 금액만큼 같은 비중으로 KODEX inverse ETF를 매수하는 전략을 취하였다.

국면	투자 전략
----	-------

³ 김은충, 『멀티 팩터에 기반한 원자재 포트폴리오 구성전략』, 충남대학교 경영경제연구소, 경영경제연구 제42권, 제 2호, 51-75, 2020

시장 활황 및 개인 거래 소극적	관심도 팩터의 하위 10% 주식 종목을 매수/ 상위 10% 주식 종목을 매도(inverse ETF 매 수)
시장 불황 및 개인 거래 적극적	관심도 팩터의 상위 10% 주식 종목을 매수/ 하위 10% 주식 종목을 매도(inverse ETF 매 수)
그 외	Kodex200 매수

표 II-3 <Long-Short / Long-hedge 투자 전략>

아래의 <그림 II-2>는 KOSPI200 지수와 본 연구에서 투자 전략에 활용하고자 하는 국면들 간 분포를 시각화 하였다. 파란색 선은 시장 국면이 불황이면서 동시에 개인 거래 자금 변화량은 증가하는 경우에 해당하고, 빨간색 선은 시장 국면은 호황이지만 개인 거래 자금 변화량은 감소하는 국면을 나타낸다. 그림에서 확인할 수 있듯이 KOSPI200지수의 상승 및 하락과 본 연구에서 활용하는 국면 변수의 시장 활황 및 불황이 완전하게 일치하지는 않는 것을 확인할 수 있다. 이는 시장 약세 및 강세와 개인 거래 자금 데이터를 결합하여 새로운 국면변수를 만들었기 때문이며, 이러한 세분화된 국면 변수는 이후 연구에서 상하위 관심도 팩터에 해당하는 종목을 선정하는데 유의미한 영향을 줄 것으로 기대된다.

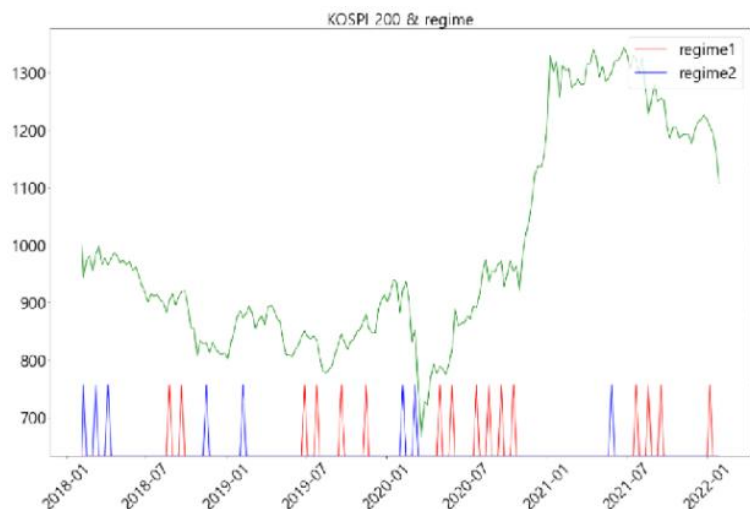


그림 II-4<Kospi200지수와 자체 국면 변수 간 관계>

(2) 백테스팅

종목 토론방이 활성화되기 시작한 시기부터 고려하기 위해 백테스팅 기간은 2018년 2월 2일부터 2022년 1월 28일까지 실시하였다. 백테스팅 과정에서 구성 종목들의 비중은 KOSPI200을 기초지수로 하는 KODEX200과 마찬가지로 시가총액 가중방식을 사용하였다. 세금 및 거래세는 모델을 구현하는 과정에서는 반영하지 않았지만, 연구 결과에서 추가적으로 고려한 분석을 제시했다. 리밸런싱 날짜는 매월 둘째 주 월요일에 실시하며, 4주 뒤 금요일에 종가로 매도하게끔 설계하였고, 만약 리밸런싱 날짜가 휴일인 경우 그 이후 휴일이 지난 바로 다음 날 매수하도록 하였다. 마지막으로 매도 날짜인 마지막 주 금요일이 휴일인 경우 휴일 바로 전 종가에 매도하도록 하였다.

이러한 기본적인 가정 이외에도 백테스팅 과정에서 발생하는 과적합을 최소화하기 위해 다음과 같은 방법으로 연구결과를 분석하였다. 먼저 본 연구에서 지정한 국면 변수가 있는 모델과 없는 모델 간 관심도 팩터가 얼마만큼 유의미한 차이를 보이는지 분석하였다. 두 번째로 국면 별 Long-Only 전략을 취할 때, 팩터 분위수를 기준으로 상위 5%, 10%, 상위 20%, 상위 30% 등 매수 비중을 달리해가며 포트폴리오가 얼마나 안정적인지 여러 지표를 통해 분석하였다. 팩터가 유의미하지 않다면 분위수를 달리할 경우 모델 값이 급격히 변하는 불안정한 모델이 관측될 것이다.

이후 Long-Only 투자 모형의 한계를 보완하기 위해 Long-Short 모델의 비중 변화를 통한 안정성 및 factor의 효과 실증 분석하였다. 마지막으로 실제 투자에서도 활용될 수 있도록 Long-Hedge 모델의 안정성 및 알파 창출의 정도를 분석하고자 하여 Long-Hedge 모델의 실용성을 입증하고자 하였다.

Ⅲ. 연구 결과

i. 평가 지표

백테스팅 결과를 분석하기 위해 수익률과 변동성 이외에도 Sharp ratio, Sharp ratio standard deviation, Sortino ratio, Maximum drawdown 등의 지표들을 종합적으로 활용했다. Sharp ratio에 사용될 무위험 이자율은 2018년부터 2021년까지의 3년 국채 수익률의 평균인 1.5%를 적용하였다. 또한, 52주 Sharp ratio를 이 동시킨 표준편차를 관찰하여 Sharp ratio의 일관성을 계산한 Sharp ratio standard deviation 변수도 산출하였다. 또한, 변동성의 상방과 하방을 모두 포함하고 있는 Sharp ratio의 단점을 보완하기 위해 Sortino 지수도 사용하였는데, 공식은 다음과 같다.

$$Sortino\ Ratio = \frac{R_i - R_f}{\sigma_i},$$

R_i : 포트폴리오 수익률, R_f : 무위험 이자율, σ_i : 마이너스 수익률의 표준편차

위 공식에서 볼 수 있듯이, Sortino 지수는 마이너스 수익률의 변동성만 고려하여 변동성을 계산하기 때문에 하락 위험도 대비 수익률을 비교할 수 있다. 마지막으로 하락의 정도를 비교하기 위해 Maximum drawdown(MDD) 와 Drawdown의 평균을 산출하였다. 아래에 이어지는 본 연구의 결과는 모두 주간 데이터에 기초하여 분석하였다.

ii. Long-Only 포트폴리오 결과

portfolio	Name	Return	Std D	Sharp	Sharp_std ⁴	Sortino	MDD(meanDD)
benchmark(kodex200)		0.088%	20.07%	0.189	1.04	0.2070	-21.82% (-3.8%)
국면 없음	High score 10% long	-0.08%	24.41%	-0.21	1.03	-0.13	-18.34% (-5.21%)

⁴ 52주 이동평균한 Sharp ratio의 표준편차

	Low score 10% long	0.02%	23.35%	-0.02	0.93	-0.08	-24.44%(- 4.64%)
국면 있음	10% long	0.104%	23.77%	0.194	0.83	0.2365	-18.34(-4.66%)
국면 있음 개인 자금 factor 포함	10%long Kodex매 수	0.16%	20.01%	0.39	0.911	0.498	-18.34 (-3.79%)

표Ⅲ-1 <국면 고려 여부 및 개인 자금 factor 유무에 따른 bactesting 결과>

국면 유무에 따른 Long- only 투자 전략의 백테스팅 결과를 <표Ⅲ-1>에 정리하였다. 마르코프 국면 전환 모델이 없는 경우 관심도 factor가 유의미한 결과를 보이지 않았지만, 마르코프 국면 전환 모델을 적용하면서 수익률, 전체적인 지표가 개선되었다. 또한, 개인의 영향력을 나타내기 위해 개인자금을 이용한 factor로 국면을 추가적으로 가려낼 경우, 주간 수익률은 벤치마크대비 0.08%가 개선이 되고, sharp ratio 역시 0.211만큼 개선이 되었다. Drawdown 지표 역시 안정적인 결과를 보이며 비교적 안정적인 모델이 되었다.

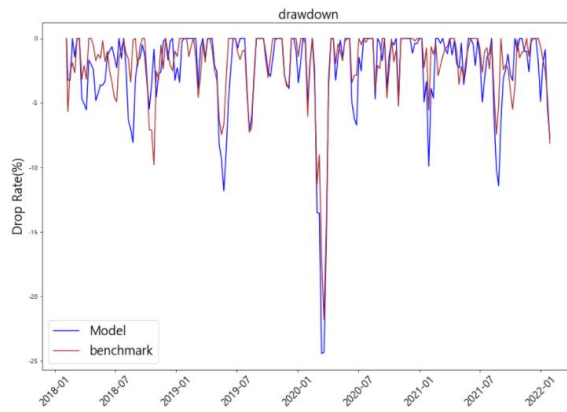
다음 <그림Ⅲ-1,2,3,4,5,6>은 국면이 없을 경우 성과가 그나마 가장 좋았던 Low-score long-only 전략과 국면 변수를 고려하여 long-only 전략을 취했을 때의 결과를 비교하여 시각화한 자료이다. 아래의 시각자료에서 알 수 있듯이 국면이 있는 모델이 전반적으로 안정적인 추세를 보이고 있으며, 샤프지수의 변화도 벤치마크를 전반적으로 약간 상회하는 것을 볼 수 있다.



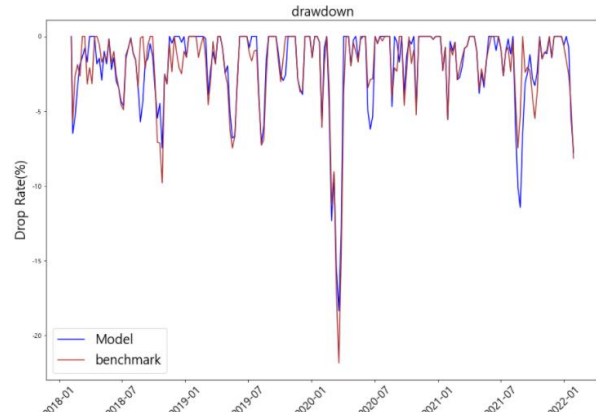
그림Ⅲ-1<low factor score 10% Long-Only 전략 백테스팅 결과>



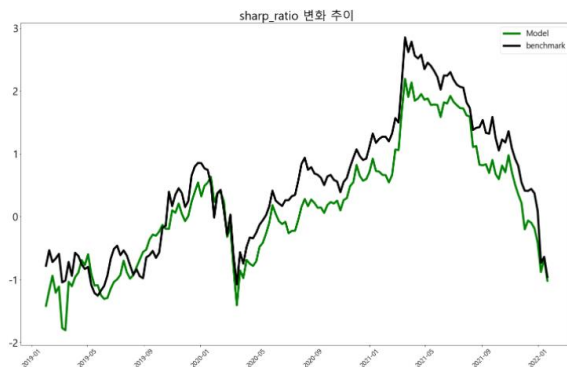
그림Ⅲ-2<국면 고려 10% Long-Only 전략 백테스팅 결과 >



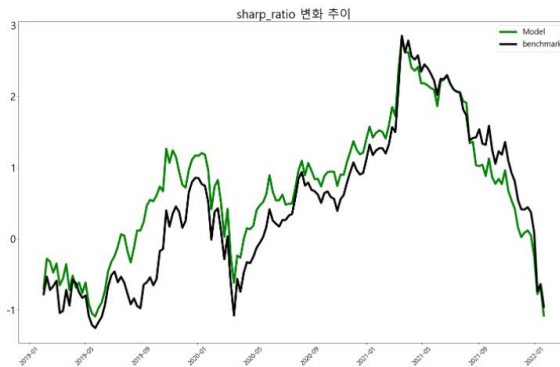
그림Ⅲ-3<low factor score 10% Long-Only 전략 Drawdown>



그림Ⅲ-4<low factor score 10% Long-Only 전략 Drawdown >



그림Ⅲ-5<low factor score 10% Long-Only 전략 52주 Sharp ratio 변화 >



그림Ⅲ-6<low factor score 10% Long-Only 전략 52주 Sharp ratio 변화 >

portfolio	Weight	Return	Std D	Sharp	Sharp_std	Sortino	MDD(meanDD)
Benchmark (KODEX200)	100% long	0.088%	20.07%	0.189	1.04	0.2070	-21.82%(-3.8%)
국면별 상위스코어 하위스코어 매수전략 (일반경우 kodex)	5% long	0.34%	21.56%	0.74	0.806	1.076	-14.60%(-3.77%)
	10% long	0.16%	20.01%	0.397	0.91	0.4982	-18.34%(-3.79%)
	20% long	0.080%	20.07%	0.165	1.12	0.1709	-21.53%(-3.84%)
	30% long	0.064%	20.01%	0.1287	1.09	0.119	-22.83%(-3.87%)

표Ⅲ-2 <국면 별 score 매수 비율에 따른 back-testing weekly 성과분석-Long Only 전략>

<표Ⅲ-2>는 국면별로 매수하는 비율을 변화해가며 투자했을 때의 성과 결과를 정리한 표이다. 표에서 알 수 있듯이, 국면별로 매수하는 전략의 성과는 매수하려는 스코어의 분위 수를 키울수록 일관되게 감소하는 것을 알 수 있다. 가장 매수 비중이 적은 5% 매수 비중 모델의 경우, 30% 매수 비중 모델보다 주간 수익률은 0.24% 높으며, Sharp ratio는 약 0.62%, Sortino ratio는 약 0.9051% 높은 것을 확인할 수 있다.

iii. Long-Short 포트폴리오 결과

위의 결과를 종합하면, 국면을 고려하여 팩터를 사용하는 것이 고려하지 않았을 때보다 더욱 효과적이며, 해당 국면에서 팩터에 맞게 소량의 비중을 들고 가는 것이 효과적임을 알 수 있다. 하지만 이러한 전략은 변동성 위험에 노출되어 있는데, <표Ⅲ-2>에서 적은 주식 종목을 운용하는 5% Long-Only 모델의 변동성이 다른 모델보다 상당히 큰 것을 확인할 수 있다. 따라서 위의 전략은 실제 운용사에서 운용을 하거나 투자를 하기에는 위험한 포트폴리오 전략이 될 수 있다. 이러한 문제점을 해결하고 실제 운용이 가능하도록 변동성을 줄이는 목적과 더불어, 연구설계에서 언급했듯 Long-Short 투자 전략이 Long 포지션과 Short 포지션에서의 수익률 차이를 검증하여 해당 팩터가 효과적임을 보여준다는 장점을 살

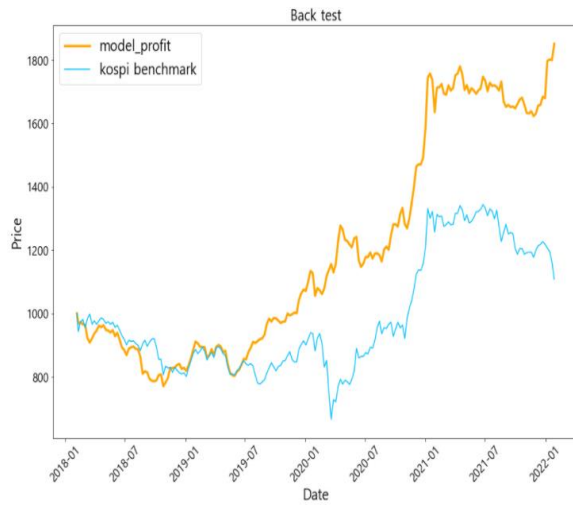
리기 위해 본 연구에서도 관심도 팩터에 대한 Long-Short 전략을 수행하였다.

portfolio	Name	Return	Std D	Sharp	MDD(meanDD)	P-value	alpha
Benchmark(kodex200)		0.088%	20.07%	0.189	-21.82%		
상하위 스코어 Long-short	5% long-short	0.292%	18.82%	0.765	-10.64%(-3.48%)	0.125	0.27% (유의x)
	10% long-short	0.281%	15.36%	0.894	-8.45%(-2.74%)	0.063	0.25% (10%유의)
	20% long-short	0.24%	13.76%	0.856	-6.78%(-2.38%)	0.06	0.21% (10%유의)
	30% long-short	0.25%	12.89%	0.9474	-6.78%(-2.11%)	0.04	0.22 (5%유의)

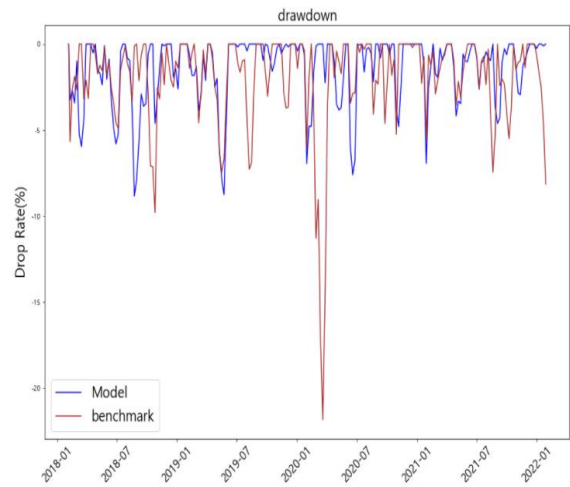
표Ⅲ- 3 <국면 별 score 매수 비율에 따른 back-testing weekly 성과분석-Long Short 전략>

<표Ⅲ-3>는 국면에 따른 투자 전략을 팩터 매수 비율을 달리하여 Long-short 전략을 취한 모델의 성과를 정리한 표이다. Long-Short 모델을 적용한 결과 전체적으로 벤치마크보다 성과가 좋음과 동시에 Long-Only 전략보다 변동성이 줄어들어 안정적인 결과를 확인할 수 있다. 이러한 주간 초과 수익률을 연으로 환산하면, 10% 유의수준 하에서 약 12.48% ~ 14.61% 초과 수익을 얻을 수 있음을 시사하고 있다. 더욱이 벤치마크의 주간 수익률과 회귀분석을 한 결과 5% 비중을 제외한 나머지 3개의 모델에서는 10% 유의수준 하에서 모두 알파가 존재함을 확인하였다. 또한, Long-Short 모델이 Long only 모델보다 비중에 큰 상관없이 더욱 안정적으로 초과수익을 이끄는데, 이는 개인의 관심도라는 팩터가 상승이 강한 종목보다는 하락이 강한 종목을 picking하는데 효과가 있다고 판단할 근거가 된다.

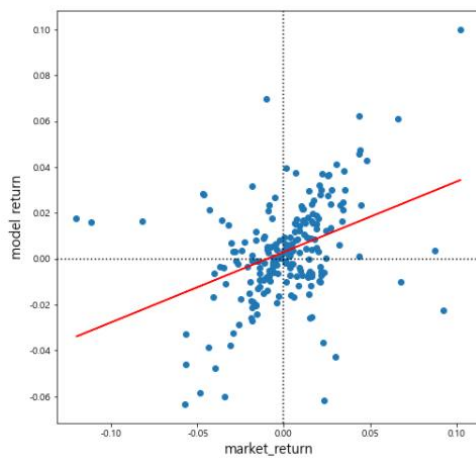
아래 <그림Ⅲ-7,8,9,10>는 결과가 가장 좋으면서 안정적인 비중 30% long-short 모델에 대한 백테스팅 결과 및 성과 지표를 시각화한 자료이다. 이를 통해 해당 전략이 벤치마크 지수보다 유의미한 초과수익률을 제공(그림Ⅲ-7, 9)하는 것 뿐만 아니라, 더 안정적(그림Ⅲ-8, 10)으로 제공하고 있음을 확인할 수 있다.



그림III-7 <팩터 비중 10% Long-Short 모델 백테스팅 결과>



그림III-8 <팩터 비중 10% Long-Short 모델 Drawdown>



그림III-9<팩터 비중 10% Long-Short 모델 OLS 결과>



그림III- 10<팩터 비중 10% Long-Short모델 52 주 Sharp ratio 변화 >

투자전략 \ 국면	시장국면 호황	시장국면 불황
	개인자금 변화율 감소	개인자금 변화율 증가
벤치마크 수익률 동일평균	-0.63%	-2.57%
팩터 상위 5개 종목 수익률 동일평균	-0.98%	-1.02%
팩터 하위 5개 종목 수익률 동일평균	-1.015%	-2.52%
연구 long- short 모델 수익률 동일평균	0.035%	1.5%

표Ⅲ-4 < 팩터값 상하위 5개 종목 수익률 동일평균의 평균>

위의 <표Ⅲ-4>는 Long-Short 모델이 제대로 구현되고 있는지 확인하기 위하여 상하위 팩터 값에 위치한 5개 종목을 대상으로 각각의 국면마다 수익률의 동일평균을 다시 평균한 표이다. 위의 표에서 알 수 있듯이 두 국면 모두 하락장을 잘 판별하였다. 뿐만 아니라 Long-Short 모델을 취할 시 팩터의 효과가 유지돼 알파의 수익률을 거둘 수 있다는 것을 확인할 수 있었다. 다만 연구 가설과는 다르게 시장국면이 호황이라 판단하고 개인자금의 변화율이 감소할 경우 팩터 상위 5개 종목의 수익률이 벤치마크보다 높지 않음을 확인할 수 있었다. 이는 개인 투자자가 하락을 방어한다고 해도 기관 및 외국인 투자자의 매도 같은 다른 요인이 작동하고 있었다고 판단한다. 반대로 시장국면이 불황이고 개인 자금이 증가하는 경우는 하락장을 상당히 방어할 뿐더러 팩터와의 수익률 차이가 극명하게 드러나는 것을 볼 수 있다.

위에 결과에서 알 수 있듯이 Long-Short 모델은 매수비중이 10~30%인 범위에서 안정적인 수익을 얻을 수 있다. 하지만 현실적으로 운용사가 공매도를 취하기 위해서는 주식마다 각기 다른 비용을 지불해야 하며, 손실 리스크가 매우 높아 Long-Short 모델을 실현하는데 제약이 있을 수 있다. 이러한 현실적인 제약을 완화하기 위해 하위 스코어 주식들을 short하는 전략을 취하는 대신 코스피 inverse ETF를 매수하는 전략을 취하여 그 성과를 비교하고자 하였다.

portfolio	Name	Return	Std D	Sharp	Mean_DD	P-value	alpha
Benchmark(kodex200)		0.088%	20.07%	0.189	-3.85%		
Score stock Long_ Inverse ETF Long	5% long-hedge	0.327%	13.82%	1.175	-2.27%	0.013	0.30% (5%유익)
	20% long-hedge	0.206%	12.65%	0.783	-2.5%	0.086	0.18% (10%유익)

	40% long- hedge	0.20%	12.45%	0.775	-1.97%	0.085	0.17% (10%유 의)
--	-----------------------	-------	--------	-------	--------	-------	----------------------

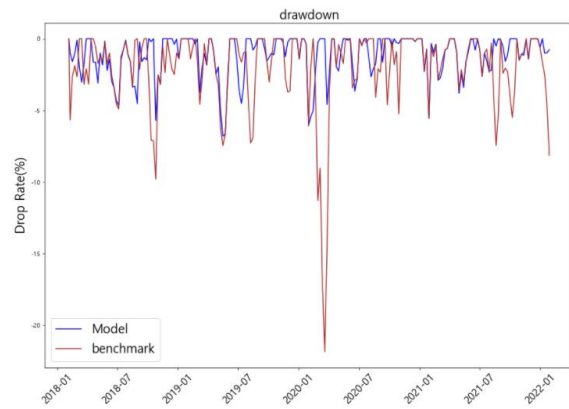
표Ⅲ-5<국면 별 score 매수 비율에 따른 back-testing weekly 성과분석-Long Hedge 전략>

<표Ⅲ-5> 에서 2가지 결과를 도출할 수 있다. 먼저, 5% Long-Hedge모델에서 비중을 늘리면 늘릴수록 sharp ratio와 초과수익률이 감소하는 것을 확인할 수 있다. 이는 Long-Only 모델 결과에서 언급했듯이 상위 스코어의 주식이 다른 주식보다 더 나은 수익률을 줄 수 있음을 암시한다. 두번째로는 국면전환 모델자체 play 방식에서 알파가 도출됨을 알 수 있다. 이는 비중을 다양하게 바꿔도 알파의 유의수준은 10%내에서 지속적으로 유의하든 사실에서 도출할 수 있다. 하락장을 잘 구별해낸 개인자금 팩터와 마르코프 국면 전환 모델을 결합한 모델이 하락장과 상승장을 잘 구분함에 따라 전체적으로 알파가 나옴을 알 수 있다. 수익률의 관점에서, 주간수익률이 0.18%이므로 연간 수익률로 환산하면 약 9.36%~14.56% 초과수익을 달성할 수 있음을 알 수 있다. 뿐만 아니라, short 포지션대신에 inverse ETF를 매수하는 전략을 취함으로써 수수료를 낮춰 초과수익률을 보전할 수 있는 결과를 가져올 수 있다.

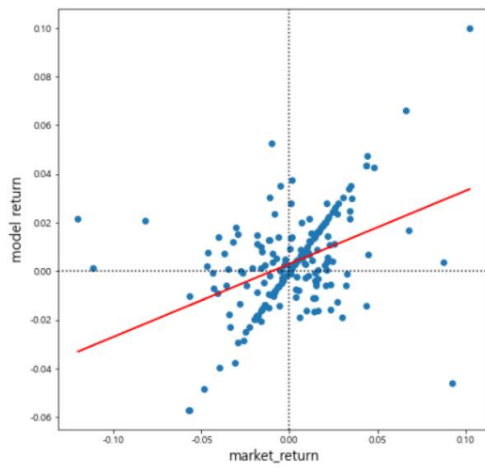
마지막으로 아래 <그림Ⅲ-11,12,13,14>에서 inverse ETF 매수 long-short 전략 중 가장 성과가 좋았던 5% long-hedge model의 결과를 시각화 하여 해당 전략의 효과를 보였다.



그림Ⅲ-11 <팩터 비중 5% Long-Hedge 모델 백테스팅 결과>



그림Ⅲ-12 <팩터 비중 5% Long-Hedge 모델 Drawdown>



그림Ⅲ-13 <팩터 비중 5% Long-Hedge 모델 OLS 결과>



그림Ⅲ- 14 <팩터 비중 5% Long-Hedge 52주 Sharp ratio 변화 >

IV. 결론

한국 금융시장에서 개인투자자가 미치는 영향력은 코로나 19 이후 점점 증가하고 있다. 개인 소매 투자의 비중이 증가하면서 개인의 투자자금은 더 이상 무시할 만한 수치가 아니게 되었으며, 이와 같은 자금들로 인해 개인의 영향력은 더욱 증가했다. 더욱이 인터넷과 기술이 발달하면서 전문적인 투자의 문턱이 낮아지면서, 이러한 개인의 투자 행태는 이전으로 줄어들 가능성이 희박해 보인다.

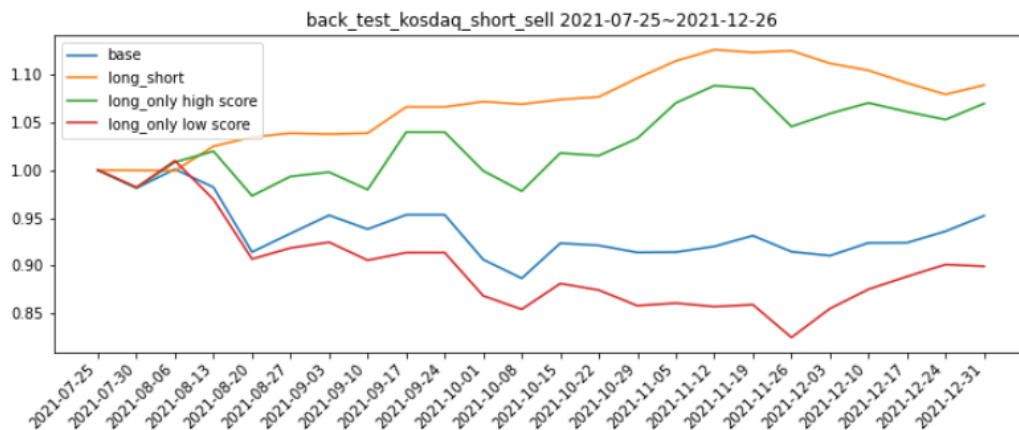
이에 따라 개인투자자의 심리를 판단하는 지표는 필수적으로 고려할 요소가 되었다. 하지만, 기존의 팩터 모델 중 개인 투자자의 관심도를 계량화한 모형은 단순히 텍스트데이터 분석, 수치적 분석, 횡단면적 분석에 그쳤다. 이에 본 연구는 네이버 주식 종목 토론방을 활용하여 개인의 관심도를 계량화한 뒤, 마르코프 국면 전환 모형을 통해 주식시장의 국면이 변화할 때 개인의 관심도라는 팩터가 거시경제적 상황에 따라 수익률에 얼마만큼 영향을 미칠 수 있는지 분석하고자 하였다. 이를 위해 국면 변수의 유무, Long-Only 전략, Long-Short 전략, Long-Hedge 전략 등 다양한 전략을 활용하여 강건성을 검증하고자 하였다.

본 연구의 주요 결과를 요약하면 다음과 같다. 첫째, 국면을 고려하지 않은 채 게시판의 관심도 팩터만 사용하면 유의미한 결과가 없었지만, 시간의 흐름에 따른 국면 변화를 고려하기 위해 국면전환모델과 개인자금 팩터를 적용한 포트폴리오 전략을 세우면 기존의 KOSPI200을 기초지수로 하는 ETF보다 비교적 좋은 성과를 보이는 모델을 구축할 수 있다. 둘째, 이를 근거로 국면에 따른 Long-only 모델을 취할 시 주식 비중을 적게 가져가는 포트폴리오가 더 높은 수익률을 달성하는 것을 확인하였다. 하지만 이는 비교적 변동성이 클 수 있는 단점이 존재하였다.

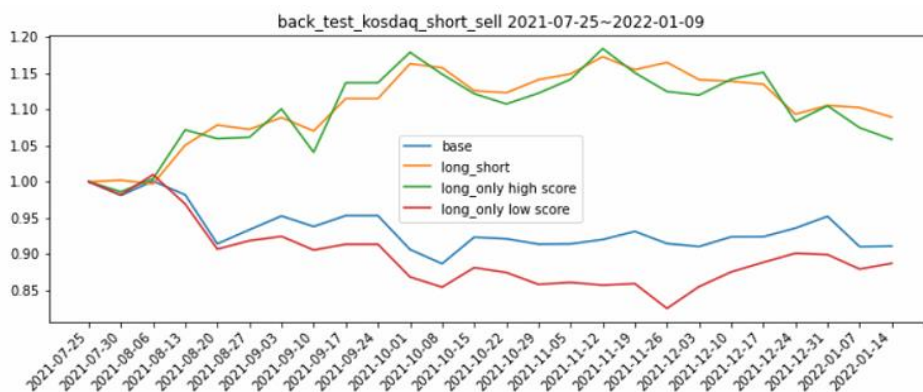
셋째, 국면에 따라 관심도 팩터를 Long-Short 전략을 취한다면 Long-Only 모델보다 변동성은 줄이면서 수익률은 증가하는 모델을 구축하였으며, 시장 수익률 대비 알파가 존재함을 보였다. 뿐만 아니라, Long-Short 모델의 통계적 유의성을 보여 관심도 팩터 자체의 효과가 있음을 검증하였다. 넷째, Short position 대신 KOSPI200 inverse ETF를 매수하는 포지션을 통해 국면에 따른 long-hedge 전략을 취했을 경우, 기존의 Long-Short 모델보다 수익률을 다소 감소하지만 안정적이며, Short 포지션을 취할 시 발생하는 거래비용을 고려한다면, 더욱 효과적인 모델이라는 것을 검증하였다. 다섯째, 비중을 달리하여 long-hedge 모델의 성과 분석을 함으로써 비중에 상관없이 전반적으로 long-hedge 모델이 시장수익률 대비 초과수익률을 거둔다는 것을 보이고 이를 통해 국면전환 모형 자체가 수익률에 있어서 유의미함을 검증하였다.

본 연구의 시사점 및 기대할 점은 크게 두가지가 있다. 첫번째로 운용 측면에 있어서 극적은 수익률은 아니지만 안정적인 수익을 추구하는 투자자에게 있어서 매력적인 투자전략이 될 수 있다. 더 나아가 ETF 운용에 있어서도 활용이 될 수 있음을 시사한다. 미국의 경우, SNS(소셜네트워크서비스)에서 화제가 된 주식만을 골라 투자하는 ‘VanEck Vectors Social Sentiment ETF(BUZZ)’이 인기를 뜨겁게 달구고 있다. 미국의 개인투자자가 이러한 ETF의 선호가 큰 것을 감안하였을 때, 본 논문의 전략을 더욱 발전시켜 밈 주식들을 위주로 포트폴리오로 편입시켜 ETF를 운용한다면 마케팅과 수익률을 모두 챙길 수 있는 ETF 역시 기대할 수 있다. 더 나아가 감정 분석 등을 통해 개인 투자자의 관심도뿐만 아니라 해당 종목에 대해 가지는 sentiment까지 분석하는 방식으로 확장할 수 있을 것이다. 실제로 Yang et al(2020)이 개발한 FinBert 등 도메인이 금융시장에 특화된 자연어 처리 모델 등 금융시장에서 자연어처리 분야의 위상이 점점 높아지고 있는 상황에서, 단어의 긍부정뿐만 아니라 국면에 따른 개인의 관심도까지 복합적으로 고려한 투자 상품까지 만들 수 있을 것이라 기대된다.

두번째로, 멀티팩터로 활용 가능성을 기대할 수 있다. 본 논문은 현재 국면에 따른 단일 팩터를 적용하여 포트폴리오 모델을 분석하였다. 하지만 실제 팩터 모델 운용은 더욱 다양한 팩터를 이용한다. 본 연구의 주요 주제는 아니지만, 다요인 팩터 모델을 구현하기 위해 KOSDAQ150에 있는 기업을 대상으로 관심도 팩터뿐만 아니라 추가적으로 공매도 수치를 팩터화 하여 백테스팅을 진행하였다. 관심도 팩터는 본 논문의 팩터를 ‘공매도’라는 키워드에 대한 게시물에 한해서 관심도를 측정했으며, 공매도 팩터의 경우 KRX 데이터에서 공매도 잔고값을 추출하여 28주 평균과 분산을 활용해 z-scoring한 값을 이용했다. 그림<IV-1>은 이러한 백테스팅 결과를 시각화하였다.



그림IV-1 <공매도 factor만 적용한 4가지 투자전략별 성과>



그림IV-2 <공매도 factor와 관심도 factor를 함께 적용한 4가지 투자전략별 성과>

초록색 선을 보면 알 수 있듯이, 단순히 공매도 factor를 적용한 결과보다 본 논문에서 제시한 관심도 factor를 함께 이용한다면 더욱 좋은 성과를 보임을 간략히 알 수 있다. 위의 예시처럼 다양한 팩터 모델에 본 논문에서 제시한 국면전환 모형과 관심도 factor를 이용한다면 더욱 유의미한 모델을 도출할 것이라고 기대된다.

본 연구의 한계점은 3가지가 있다. 첫 번째, 게시글의 데이터를 전부 가져오기에는 시간적 여유가 충분하지 못하여 게시물이 10,000개가 넘어가는 기업들 한에서는 샘플링을 통해 factor를 계산했다는 점이다. 본 연구에서 이용한 factor는 게시글 수와는 연관이 없지만, 더욱 정교한 factor 값을 도출하려면 게시글의 모든 데이터를 가져올 필요가 있다. 두 번째, 게시글 데이터의 추가적인 텍스트 분석을 하지 않은 점이다. 박단비 『온라인 주식게시판 정보를 활용한 주가분석』 연구에 의하면, 시장데이터나 설문 등을 이용하여 측정된 간접적인 투자심리지표보다 텍스트 데이터를 이용하여 측정된 투자심리지표가 주식수익률에 대해 더 높은 설명력을 갖는다고 실증 분석 결과가 제시되고 있다고 한다. 이는 코스피 200을 대상으로 모든 게시물의 내용을 가져와야 하므로 더욱 많은 시간을 필요로 한다. 추후 연구에 있어서 텍스트 감정분석, 네트워크와 같은 활동을 해준다면 더욱 개선될 여지가 있다. 세 번째, 국면에 따른 투자자의 행동을 세세히 분석하지 못한 점을 들 수 있다. 본 논문의 검증 방식은 백테스팅 방식으로 안정적인 수익을 가져오는 현실적인 모델에 초점이 맞춰져 있어, 팩터가 작용하는 근본적인 인과관계를 명확히 분석하지 못한 점을 한계로 들 수 있다. 향후 이러한 연구의 한계점을 보완한 연구를 통하여 더욱 개선된 모델을 구축할 수 있기를 기대한다.

V. 참고문헌

- 김가람, 류두진, 양희진.(2018).투자심리지수의 대용변수와 유용성: 개별기업 주식수익률에 미치는 영향을 바탕으로.경영학연구.제 47권 제 5호.1231~1260
- 김민기, 김준석.(2022).국내 개인 투자자의 행태적 편익과 거래행태. 자본시장연구원.71-90
- 김준석.(2021).코로나19 국면의 개인투자자.자본시장 포커스, 2021-4호
- 김민혁.(2013).한국 주식시장에서 개인투자자의 투자성과 및 행태 그리고 주식수익률과의 관계에 대한 연구.박사학위논문. 한국외국어대학교. 서울시
- 김은총.(2020). 멀티 팩터에 기반한 원자재 포트폴리오 구성전략. 충남대학교 경영경제연구소, 경영경제연구 제42권, 제 2호, 51-75
- 류두진.(2018).개별기업의 특성을 반영한 투자자 심리지수와 주식수익률.한국재무학회 재무연구 제31권 제 1호.1-38
- 박단비.(2020).온라인 주식게시판 정보를 활용한 주가분석. 석사학위논문. 조선대학교. 광주시
- 이윤정, 김건우, 우균.(2014).인터넷 주식 토론방 게시물과 주식시장의 상관관계 분석을 통한 투자 종목 선정 시스템.한국정보처리학회, 정보처리학회논문지, 소프트웨어 및 데이터 공학 제3권 제10호.441-450
- 이준행, 김류미(2017).국면전환모형을 활용한 동태적 스타일 배분전략의 유용성: 가치주 vs 성장.한국재무학회 재무연구 제 30권 제 3호.361-393
- 이효정.(2020). 투자심리가 횡단면 주식수익률에 미치는 영향. 한국재무학회 재무연구, 제33권 제 1호. 61-95
- Malcom Baker, Jeffrey Wurgler(2006).Investor Sentiment and the Cross-Section of Stock Returns. Journal of Finance. 61(4). 1645-1680

J. Bradford De Long, Andrei Shleifer, Lawrence H. Summers and Robert J. Waldmann.(1990). Noise trader risk in financial markets. Journal of Political Economy, 98(4), 703-738

본 연구의 모든 코드 및 데이터는 아래 깃허브에서 확인하실 수 있습니다.
[GitHub - wonsikd/private-investor_back_testing-DB-](#)