

# 스타일 로테이션 모형 : AI(인공지능)와 퀀트 전략



### Contents

스타일 로테이션 모형: AI(인공지능)와 퀀트 전략	03
Intro: 통계학과 알파고, 그리고 스타일 로테이션	04
모형 개요	06
Step1. Style index 작성	07
Step2. 시장 데이터 선정	14
Step3, 시장 데이터의 주성분 분석(PCA)	15
Step4, Multinomial Logistic Regression	16
백테스트 결과 정리	19
결론 및 모델 개선 가능성	22

# 퀀트 Strategy

# 스타일 로테이션 모형: AI(인공지능)와 퀀트 전략

Quantitative Analyst 염동찬 • 02)6923-7314 • dongchan@ligstock.com

알파고와 이세돌 기사의 대결 이후 인공지능에 대한 관심이 높아졌습니다. 관심의 중심에는 인공지능 발전과 인간 활동 영역의 감소에 대한 우려가 있습니다. 하지만 현재 시점에서는 먼 미래의 시나리오를 상상하기보다, 인공지능이 Input과 Output 사이에서 어떤 모형에 근거해 판단하는지를 이해하는 노력이 중요할 것입니다.

코스피가 박스권 흐름을 지속하고 있지만, 스타일별 성과는 서로 다른 흐름을 지속하고 있습니다. 최근 3년간의 박스권에서 대형주 투자자와 중소형주 투자자가 체감하는 증시는 서로 달랐습니다. 본 자료에서는 인공 지능의 판단 모형 중 하나인 다항 로지스틱 회귀 모형(Multinomial Logistic Regression Model)에 기반하여 다음 달 가장 유망한 스타일 투자전략이 무엇인지를 판별하는 모형을 작성하였습니다.

결과를 먼저 언급하면 KOSPI가 CAGR 기준 9.1%의 수익률을 기록하는 동안 모형에 의한 투자 전략은 31.0%의 수익률을 기록했습니다. 승률 측면에서, 모형의 포트폴리오가 KOSPI를 Outperform할 확률은 60%였습니다.

다음 달 긍정적일 것으로 전망되는 스타일이 어떤 것인지, 해당 스타일에 포함되는 종목이 무엇인지를 체크하는 것은 액티브 투자자와 패시브 투자자 모두에게 의미있는 내용이 될 것으로 기대합니다. 앞으로 매 달 마지막 주에 모형과 관련된 전망과 성과 보고서를 발간할 예정입니다.

본 보고서는 독자가 기초통계학 이상의 지식을 갖추었다는 가정에 맞추어 작성했습니다. 수학 및 통계학의 이론적 배경보다는 아이디어와 결론이 궁금하신분은 Intro와 개요, 결론 부분만 읽어보시길 권합니다.



자료: LIG투자증권 리서치본부

#### Intro: 통계학과 알파고, 그리고 스타일 로테이션

로지스틱 회귀모형을 이용한 분석 배경

궁극적으로, 모든 지식은 역사학이라고 할 수 있다. 추상적으로, 모든 과학은 수학이라고 할 수 있다.. 근본적으로, 모든 판단은 통계학이라고 할 수 있다. - C. R. 라오

현재 생존해 있는 가장 위대한 통계학자 중 하나인 C. R. 라오는 모든 판단은 근본적으로 통계학이라고 말했다. 과거의 데이터를 정리하고 정보를 얻어, 이를 의사 결정의 근거로 사용한다는 점에서 판단과 통계학은 유사한 과정을 거친다. 실제로 인간의 무의식적인 판단과 의사 결정은 통계적 모델과 유사한 경우가 많다.

예를 들어 보자. 항아리 안에 붉은 색과 흰 색 공이 일정한 비율로 담겨 있다. 임의로 공을 10개 뽑았을 때 붉은 공이 6개, 흰 공이 4개 나왔다면 우리는 직관적으로 붉은 공과 흰 공의 비율이 6:4로 담겨 있다고 추론할 것이다. 이는 실현된 현상이 어떠한 모집단에서 가장 높은 확률로 발생하는가를 기반으로 추론했다는 점에서 최대우도추정(MLE: Maximum likelihood Estimator) 방법론과 유사하다고 설명할 수 있다.

최근 이세돌 기사와 구글의 인공지능 알파고의 대결이 이목을 끌었다. 이 대결은 인간과 기계의 경쟁이라는 측면에서 큰 관심을 받아왔다. 하지만 인공 지능의 영역이 발전하는 시대에 살아야 하는 우리가 현재 시점에 고민해야 할 것은, 인간과 기계 중 누가 더 위대한가 혹은 기계의 영역 확대 및 인간 영역의 축소에 대한 것이 아니다. 인공 지능의 판단과정을 조금이나마 이해하고 이를 적용하여 더 나은 의사결정을 위한 도구로 사용하려는 노력이 현재 시점에는 더욱 필요할 것이다.

이번 자료에서는 인공 지능의 판단 모형 중 하나인 로지스틱 회귀분석을 정리했다. 인공 지능은 통상 인풋과 아웃풋 사이에 여러 단계의 의사 결정 과정이 진행되는데, 로지스틱 회귀분석은 그 의사결정 과정 중 하나인 '판단에 관한 모형'이라고 설명할 수 있다. 당연히 복합적인 툴을 사용하는 인공지능의 의사 결정 모형과 비교한다면 매우 단순하지만, 결과나 모델링 과정 역시 쉽게 이해할 수 있다는 점에서 투자에 적용하기가 용이하다.

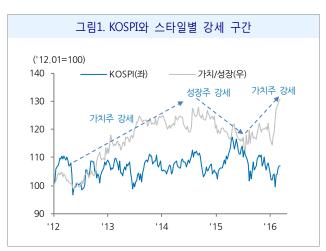
모델의 결론을 먼저 언급하면, 수익률과 승률 모든 측면에서 유의미한 결과를 얻었다. 이용자 입장에서도 패시브 펀드 운용역과 액티브 펀드 운용역 모두 참고할 수 있는 내용 이 될 것으로 판단한다. 지수 방향성 정체와 스타일별 수익률 차별화 2011년 이후 KOSPI는 박스권 흐름을 지속 중이다. 그렇지만 가치주 투자자와 성장주 투자자, 대형주 투자자와 중소형주 투자자의 심리 역시 박스권에 갖혀 있지는 않았을 것이다. 전체 인덱스는 좁은 범위에서 등락을 거듭했지만, 스타일별 인덱스는 서로 다른 흐름을 지속했기 때문이다.

장기적 관점에서 가치주의 성과는 국가를 막론하고 언제나 성장주 성과를 상회했다. 하지만 박스권 흐름을 보였던 2012년 한국 주식 시장만을 놓고 보면 가치주가 압도적으로 강한 흐름을 보여왔다고 말하기는 어렵다(그림1,2 참조).

대형주와 중소형주의 차이는 더욱 크다. 특히 2014년 이후 중소형주의 강세가 나타나면서 대형주 지수와 중소형주 지수 사이의 갭은 크게 벌어져 있는 상황이다. 2015년 중순 이후 대형주의 상대적 강세에도 불구하고, 그 동안의 갭은 메우지 못하고 있다(그림3,4 참조).

가치주/성장주에 비해 대형주/중소형주의 갭이 큰 만큼, 대형주 투자자는 중소형주 투자 자에 비해 체감하는 시장 상황이 다를 수밖에 없었다.

지수의 방향성이 정체되고 스타일별 성과가 차별화되는 시장에서, 투자 스타일의 선택은 포트폴리오의 성과에 유의미한 영향을 미치고 있다. 최근 투자 스타일 선택 모형에 대한 수요가 늘어났는데, 그 배경에는 스타일별로 판이하게 달랐던 성과가 있었다.



자료: Fnguide, LIG투자증권 리서치본부



자료: Fnguide, LIG투자증권 리서치본부



자료: Fnguide, LIG투자증권 리서치본부



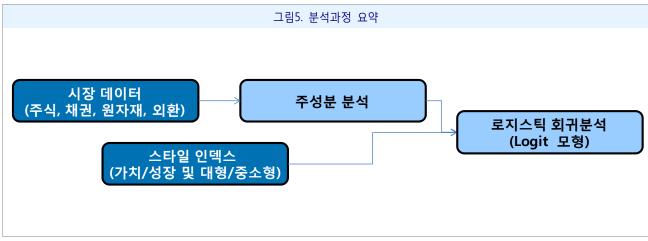
자료: Fnguide, LIG투자증권 리서치본부

#### 모형 개요

모델 도출 및 백테스팅 과정 요약 모델 도출 및 백테스팅 과정을 〈그림 5〉와 〈그림 6〉에 정리했다. 주식, 채권, 원자재, 외환 시장 데이터를 독립 변수로, 가치와 성장, 대형주와 중소형주를 혼합한 4개 인덱스의 수익률을 종속변수로 놓고 모델을 작성했다. 시장데이터의 경우 다중공선성 문제를 해결하기 위해 중간에 주성분 분석과정을 거쳤으며, 최종적으로 주성분 분석의 결과값을 로지스틱 회귀분석의 독립변수로 사용했다.

모델의 유효성을 검증하기 위하여 60개월간의 데이터를 이용하여 모델을 작성한 후, 61번째 상황에서 가장 유망할 것으로 보이는 스타일을 선택하는 시뮬레이션을 지속했다. 따라서, 독립변수는 2004년부터의 데이터를 이용했지만, 실질적인 백테스트 값은 2009년부터 확인할 수 있다.

즉, 1~60번 값을 이용하여 61번째 독립변수에 대한 결과를 예측한 후 실제 61번 값과 비교를 실시하는 것이다. 이후 다음 변수로 넘어가, 2~61번 값을 이용하여 62번째 결과를 예측하는 과정을 시계열의 마지막까지 반복하는 것이다.



자료: LIG투자증권 리서치본부



자료: LIG투자증권 리서치본부

#### Step1. Style index 작성

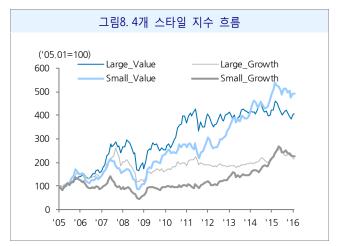
스타일 구분 방법

분석의 첫 번째 과정은 종속 변수라고 할 수 있는 스타일별 인덱스의 수익률 작성 과정이다. 대형주와 중소형주, 가치주와 성장주를 혼합한 4가지 스타일 지수를 만들었고, 해당지수의 월간 수익률을 모형의 종속 변수로 이용하였다.

가치주와 성장주는 월말 PBR 밸류에이션 순위를 기준으로 3개 그룹으로 나눈 후, 가장 밸류에이션이 높은 1분위를 성장주, 가장 밸류에이션이 낮은 1분위를 가치주로 정의하였다. 시가총액 역시 월말 시가총액을 기준으로 분류하였다. 이렇게 구분된 종목들로 월 마지막 거래일에 리밸런상해 동일 가중 방식으로 수익률을 계산하고, 최종적인 인덱스를 도출하였다.



자료: LIG투자증권 리서치본부



자료: WiseFn, LIG투자증권 리서치본부

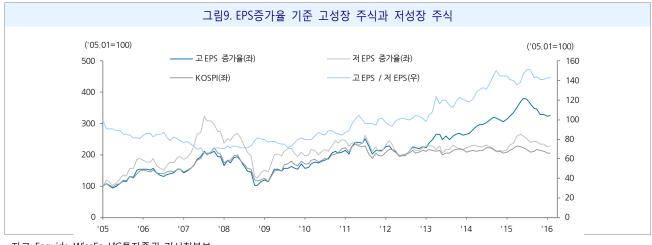
가치주와 성장주 구분 방법 이러한 스타일 인덱스 작성 방법에는 두 가지 정도의 이슈가 존재하는데, 첫 번째는 가치 주와 성장주의 구분 방법이다. 본 자료에서는 12개월 Forward PBR 값만을 이용하여 가 치주와 성장주를 구분하는 단순한 방법을 사용했다.

이는 실제 MSCI와 같이 가치주와 성장주 지수를 제공하는 업체들이 PER, PBR, 배당수익률, PSR, EPS 증가율, SPS 증가율 등 다양한 지표를 이용하여 점수를 매기고, 역탄젠트 함수 등을 이용해 가치주와 성장주를 분류하는 것과 비교하면 매우 간단한 방법이다.

이러한 간단한 방법을 사용한 이유는 국내 여건을 고려할 때 PBR이 다양한 종목의 데이터를 안정적으로 얻을 수 있는 지표이기 때문이다. 밸류에이션 지표 중 하나인 PER은 적자 기업에는 적용할 수 없고, OECD 국가 최저 수준의 배당수익률을 보여왔던 한국 기업들을 배당수익률로 구분하는데도 한계가 있다.

순이익 성장률로 기업들을 구분하는 것도 고성장/저성장 기업으로 구분하기 어려웠다. 순이익 증가율 컨센서스가 확인된 시점에서는 이미 주가가 기대감을 반영했던 경우가 많고, 실적이 적자에서 흑자로 돌아서는 등의 턴어라운드 주식은 순이익 증가율을 산정할 수 없어서 종목 분류에서 빠지는 경우가 많았기 때문이다.

EPS 증가율을 이용한 구분의 한계 실제로 고 EPS 증가율 / 저 EPS 증가율 종목간의 수익률 차이는 생각보다 크지 않았다. 12개월 Fwd EPS 증가율의 순위를 이용해 기업들을 다섯 개 그룹으로 분류하고 가장 상위 그룹과 하위그룹을 구분하여 인덱스를 작성해도 그 차이는 그렇게 크지 않다(그림9 참조). 이는 EPS 증가율 순위만을 이용해 고 성장주와 저 성장주를 구분할 경우 적자에서 흑자로 돌아서는 그룹은 고 성장주에 포함되지 않고, 반대로 적자에서 흑자로 전환되는 기업 역시 저 성장주에 포함되지 않기 때문이다.



자료: Fnguide, WiseFn, LIG투자증권 리서치본부

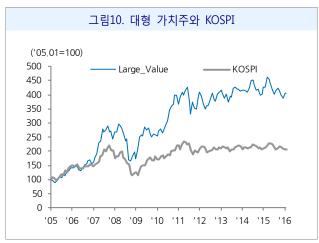
인덱스 산정 관련 일반 사항 인덱스 분류에 사용한 유니버스는 월말을 기준으로 시가총액과 12개월 Forward PBR이 모두 존재하는 기업들을 대상으로 했다. 즉, 시가총액 하위 33%는 전체 상장 기업 중 시가총액 하위 33%가 아니라, PBR 추정치와 시가총액 데이터가 모두 존재하는 데이터 중하위 33%라는 의미이다.

추정치가 존재하지 않는 초소형주는 분류에 포함되지 않았겠지만, 기관투자자 입장에서 시가총액이 너무 작은 종목을 편입하는 것은 현실적으로 쉽지 않다는 점에서 충분히 납득할 수 있는 분류일 것이다.

인덱스는 생존편향(Survivorship bias) 문제를 해결하기 위해 지금은 상장 폐지된 종목이라 할 지라도 당시에 PBR과 시가총액 데이터가 존재한다면 인덱스 산정에는 포함시켰다. 또한 동일 비중 방식을 사용했지만 리밸런싱은 월간 기준으로 했으므로, 월말에 해당 분류에 포함되는 종목들을 동일한 비중으로 매수한 후 다음달 말까지 유지한 수익률이라고 판단하면 될 것이다.

동일 비중 방식을 사용했다는 점에서 Fama-French의 중소형주 효과(SMB: Small Minus Big)가 발생한다는 점은 부정하기 어렵다. 다만 시가총액 가중방식을 사용할 경우 일부 구간에서는 그룹 내에서 시가총액 비중이 높은 기업이 전체 수익률을 좌우하는 경우가 있었다는 점에서, 동일 비중 방식의 인덱스 산정이 분석에는 더욱 효과적이라고 판단한다.

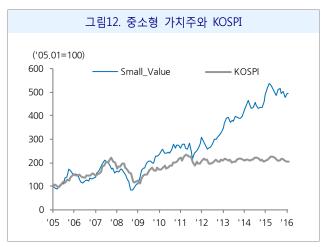
다음 페이지에는 4가지 스타일의 인덱스를 KOSPI와 비교한 차트와 각 스타일별로 2월말을 기준으로 포함 종목들을 정리했다. KOSPI와 가장 유사한 흐름을 보인 인덱스는 대형 성장주였다는 점 정도는 체크할 필요가 있어보인다(그림10~14, 표1~4 참조).



자료: Fnguide, WiseFn, LIG투자증권 리서치본부



자료: Fnguide, WiseFn, LIG투자증권 리서치본부



자료: Fnguide, WiseFn, LIG투자증권 리서치본부



자료: Fnguide, WiseFn, LIG투자증권 리서치본부

표1. 대형 가치주 포함 종목

코드	   종목	시가총액	12F	코드	   종목	시가총액	12F
<u> </u>	0 7	(억원) PBR (X) 고프 이득		0 =	(억원)	PBR (X)	
A005930	삼성전자	166,830	0.88	A010140	삼성중공업	2412.65	0.49
A015760	한국전력	33,767	0.46	A047040	대우건설	2302.55	0.76
A005380	현대차	29,297	0.52	A138930	BNK금융지주	2300.86	0.43
A012330	현대모비스	25,163	0.83	A037620	미래에셋증권	2154.29	0.61
A032830	삼성생명	22,000	0.78	A004170	신세계	2092.10	0.66
A000660	SK하이닉스	19,874	0.80	A008560	메리츠종금증권	1817.70	0.84
A055550	신한지주	18,304	0.58	A006360	GS건설	1796.30	0.5
A000270	기아차	18,302	0.67	A034020	두산중공업	1783.46	0.44
A005490	POSCO	15,563	0.34	A002550	KB손해보험	1782.00	0.78
A096770	SK이노베이션	12,113	0.70	A047050	대우인터내셔널	1753.69	0.67
A105560	KB금융	11,803	0.39	A000670	영풍	1753.62	0.58
A066570	LG전자	9,475	0.84	A012450	한화테크윈	1716.10	0.81
A034220	LG디스플레이	7,836	0.63	A000240	한국타이어월드와이드	1651.11	0.6
A023530	롯데쇼핑	7,463	0.40	A000150	두산	1629.35	0.54
A030200	KT	7,115	0.57	A028670	팬오션	1618.81	0.64
A009540	현대중공업	6,992	0.42	A003690	코리안리	1594.89	0.69
A004020	현대제철	6,519	0.40	A010060	DO	1581.21	0.53
A006400	삼성SDI	6,429	0.55	A011780	금호석유	1538.62	0.8
A024110	기업은행	6,399	0.42	A139130	DGB금융지주	1506.26	0.4
A086790	하나금융지주	6,334	0.27	A120110	코오롱인더	1487.06	0.82
A088350	한화생명	6,019	0.53	A010620	현대미포조선	1450.00	0.74
A000030	우리은행	5,949	0.30	A001740	SK네트웍스	1350.14	0.51
A078930	GS	4,683	0.68	A071840	롯데하이마트	1322.03	0.68
A002380	KCC	4,618	0.65	A003410	쌍용양회	1269.23	0.87
A139480	이마트	4,572	0.60	A006260	LS	1202.67	0.43
A032640	LG유플러스	4,226	0.87	A001120	LG상사	1197.68	0.77
A009830	한화케미칼	4,219	0.87	A011790	SKC	1135.19	0.7
A000720	현대건설	4,087	0.65	A082640	동양생명	1097.24	0.49
A036460	한국가스공사	3,448	0.31	A064350	현대로템	1032.75	0.66
A069960	현대백화점	2,996	0.77	A073240	금호타이어	987.46	0.75
A016360	삼성증권	2,874	0.67	A004000	롯데정밀화학	975.24	0.78
A011210	현대위아	2,855	0.76	A181710	NHN엔터테인먼트	935.21	0.67
A000880	한화	2,732	0.54	A037560	CJ헬로비전	929.36	0.88
A000210	대림산업	2,714	0.63	A001430	세아베스틸	907.31	0.56
A005940	NH투자증권	2,673	0.60	A020560	아시아나항공	890.64	0.87
A071050	한국금융지주	2,586	0.77	A042670	두산인프라코어	871.31	0.32
A006800	대우증권	2,558	0.57				

주: 1월말 시가총액과 PBR 기준

표2. 대형 성장주 포함 종목

<u> </u>		시가총액	12F			시가총액	12F
코드	종목	(억원)	PBR (X)	코드	종목	(억원)	PBR (X)
A090430	아모레퍼시픽	23,705	6.94	A006280	녹십자	2,501	2.34
A035420	NAVER	20,701	6.03	A105630	한세실업	2,396	4.44
A018260	삼성에스디에스	17,139	3.43	A192400	쿠쿠전자	2,338	3.28
A051900	LG생활건강	15,478	6.45	A111770	영원무역	2,260	1.73
A033780	KT&G	14,210	1.93	A005250	녹십자홀딩스	2,257	3.89
A068270	셀트리온	12,604	6.72	A161890	한국콜마	2,142	7.50
A002790	아모레G	11,849	4.14	A016170	로엔	2,011	6.98
A035250	강원랜드	8,814	2.55	A014820	동원시스템즈	1,922	5.72
A001040	Cl	8,565	2.04	A078340	컴투스	1,745	2.72
A086280	현대글로비스	7,650	2.12	A192820	코스맥스	1,683	11.21
A021240	코웨이	7,520	5.20	A214320	이노션	1,670	2.79
A128940	한미약품	7,315	9.73	A039030	이오테크닉스	1,533	4.15
A047810	한국항공우주	6,901	4.75	A049770	동원F&B	1,474	2.47
A035720	카카오	6,569	2.41	A114090	GKL	1,413	2.68
A009240	한샘	6,401	8.39	A214370	케어젠	1,383	15.50
A001800	오리온	6,155	3.29	A185750	종근당	1,350	3.85
A018880	한온시스템	5,605	2.83	A170900	동아에스티	1,295	2.36
A027410	BGF리테일	5,326	6.11	A048260	오스템임플란트	1,177	6.97
A036570	엔씨소프트	5,142	2.74	A145990	삼양사	1,168	1.72
A007070	GS리테일	5,044	2.53	A007570	일양약품	1,168	5.54
A007310	오뚜기	4,864	4.55	A036490	SK머티리얼즈	1,166	2.32
A051600	한전KPS	4,478	4.81	A068870	LG생명과학	1,150	4.15
A012750	에스원	3,876	3.05	A039130	하나투어	1,134	4.45
A000100	유한양행	3,714	2.27	A052690	한전기술	1,108	2.35
A130960	CJ E&M	3,567	2.00	A081660	휠라코리아	1,098	1.87
A026960	동서	3,126	2.62	A016450	한세예스24홀딩스	1,046	2.36
A086900	메디톡스	2,904	16.77	A096530	씨젠	1,005	7.85
A079160	CJ CGV	2,836	5.56	A000640	동아쏘시오홀딩스	999	1.79
A079550	LIG넥스원	2,794	4.23	A051500	CJ프레시웨이	939	4.28
A005610	삼립식품	2,766	8.48	A084110	휴온스	928	4.83
A030000	제일기획	2,600	2.37	A041510	에스엠	903	2.65
A008770	호텔신라	2,555	2.85				

주: 1월말 시가총액과 PBR 기준

표3. 소형 가치주 포함 종목

쿠드	조 ·	시가총액	12F	코드	종목	시가총액	12F
고드	종목	(억원)	PBR (X)	고드	古古	(억원)	PBR (X)
A090350	노루페인트	231	0.81	A084730	팅크웨어	126	0.81
A003070	코오롱글로벌	231	0.58	A049520	유아이엘	125	0.59
A023160	태광	223	0.48	A222800	심텍	123	0.56
A001940	KISCO홀딩스	220	0.24	A040300	YTN	119	0.65
A015360	예스코	212	0.30	A108380	대양전기공업	117	0.67
A077360	덕산하이메탈	207	0.88	A067310	하나마이크론	112	0.73
A044450	KSS해운	203	0.83	A058430	포스코강판	110	0.59
A004130	대덕GDS	197	0.46	A033500	동성화인텍	109	0.81
A084010	대한제강	193	0.40	A069510	에스텍	107	0.64
A010690	화신	193	0.50	A002230	피에스텍	107	0.80
A006220	제주은행	193	0.53	A003610	방림	104	0.54
A004980	성신양회	191	0.48	A008250	이건산업	97	0.53
A007660	이수페타시스	186	0.78	A106240	파인테크닉스	91	0.79
A092440	기신정기	182	0.74	A085810	알티캐스트	90	0.86
A052330	코텍	181	0.71	A023900	풍국주정	85	0.78
A084870	베이직하우스	179	0.66	A170920	엘티씨	84	0.87
A082740	두산엔진	178	0.26	A090460	비에이치	81	0.55
A004960	한신공영	177	0.68	A096240	청담러닝	81	0.88
A051370	인터플렉스	174	0.71	A049830	승일	79	0.55
A031980	피에스케이	173	0.88	A170030	현대공업	75	0.78
A049070	인탑스	169	0.40	A039340	한국경제TV	73	0.81
A004910	조광페인트	166	0.71	A062860	티엘아이	73	0.63
A100840	S&TC	163	0.65	A065350	신성델타테크	72	0.60
A065680	우주일렉트로	159	0.72	A049430	코메론	70	0.65
A009160	SIMPAC	157	0.55	A036010	아비코전자	70	0.86
A075180	새론오토모티브	153	0.70	A133820	화인베스틸	65	0.57
A088390	이녹스	152	0.80	A073490	이노와이어리스	55	0.65
A005950	이수화학	145	0.43	A140520	대창스틸	53	0.51
A058860	케이티스	144	0.85	A112240	에스에프씨	49	0.67
A001530	동일방직	136	0.19	A134580	디엠티	44	0.73
A039560	다산네트웍스	135	0.82	A005670	푸드웰	40	0.84
A079370	제우스	131	0.76				

주: 1월말 시가총액과 PBR 기준

표4. 소형 성장주 포함 종목

<u> </u>	001 = 0 0 7						
코드	종목	시가총액	12F	코드	종목	시가총액	12F
		(억원)	PBR (X)			(억원)	PBR (X)
A189980	흥국에프엔비	244	3.37	A104200	<u></u> 벅스	151	1.87
A114810	아이원스	242	2.75	A123330	제닉	147	2.17
A200780	비씨월드제약	242	3.00	A101490	에스앤에스텍	145	2.00
A067160	아프리카TV	241	4.34	A142280	녹십자엠에스	144	3.99
A018120	진로발효	234	2.76	A084650	랩지노믹스	140	4.62
A039200	오스코텍	229	15.03	A094850	참좋은레져	139	2.06
A138690	엘아이에스	222	2.85	A078140	대봉엘에스	137	2.23
A149980	하이로닉	218	3.34	A011930	신성솔라에너지	137	1.91
A214430	아이쓰리시스템	214	4.75	A160600	에스엔텍	123	4.06
A088130	동아엘텍	212	1.90	A001820	삼화콘덴서	122	2.19
A011090	에넥스	211	2.67	A052770	와이디온라인	121	3.38
A000040	KR모터스	210	2.50	A059100	아이컴포넌트	116	2.26
A225650	쿠첸	206	2.87	A092130	이크레더블	112	3.05
A041920	메디아나	203	4.44	A131390	피앤이솔루션	108	2.29
A086960	MDS테크	194	1.78	A053950	경남제약	103	4.70
A121600	나노신소재	193	2.34	A052220	iMBC	101	1.85
A086520	에코프로	188	2.11	A207760	미스터블루	101	3.72
A065510	휴비츠	185	2.30	A009520	포스코엠텍	99	1.99
A131030	디에이치피코리아	183	2.28	A036090	위지트	91	1.85
A214180	민앤지	181	2.95	A184230	SGA솔루션즈	78	2.71
A043610	KT뮤직	172	2.50	A130580	나이스디앤비	78	2.39
A086890	이수앱지스	172	3.77	A217820	엔에스	76	3.54
A040160	누리텔레콤	164	2.31	A182360	큐브엔터	57	2.29
A035900	JYP Ent.	154	2.34	A194610	우성아이비	31	2.58
A187220	디티앤씨	154	1.98				
		1	I		1		

주: 1월말 시가총액과 PBR 기준

#### Step2. 시장 데이터 선정

#### 독립변수 선정

독립변수로 사용할 시장데이터는 주식, 채권, 상품, 외환 시장 등 다양한 지표를 이용했다. 다만 모든 데이터는 월 마지막 금요일 전주 금요일의 데이터를 기준으로 했다(즉, 1개월 동안 금요일이 4번 있을 경우 3번째 금요일).

이는 모델의 현실성을 높여주기 위해 3주차까지의 데이터를 이용하여 4주차에 모델을 업데이트하고, 마지막 거래일 이전에 다음 달 액션 플랜을 만들기 위해서 설정했다. 마지막 거래일 종가를 확인하고 마지막 거래일 기준으로 다음 달의 비중확대 종목을 선정하는 것은, 현실성이 떨어지기 때문이다.

데이터는 수익률 데이터 정도를 제외하면 변동성을 완화시켜 주기 위해 월 평균 수치를 이용했다. 즉, 전월 4주차~이번달 3주차까지의 데이터의 평균이라고 할 수 있는데, 종속 변수인 수익률은 한 달 동안 발생한 수치인데 비하여, 독립변수는 기말만을 기준으로 한 다면 기간의 대응에 문제가 발생하기 때문이다.

분석에 사용한 31개 데이터는 하단에 정리되어 있다(표5 참조).

표5. 분석에 사용한 시장 데이터

No.	분류	데이터명	No.	대분류	데이터명
1	채권 시장	미국채 30년물 금리	17	해외 주식	MSCI AC World PER
2		미 30Y, 3M 스프레드	18		MSCI Emerging PER
3		미국 Break-Even 금리(5년)	19		MSCI AC World PBR
4		미국 AAA등급 회사채 스프레드	20		MSCI Emerging PBR
5		한국채 3년물 금리	21	국내 주식	MSCI Korea PER
6	상품 시장	WTI 선물 근월물 가격	22		MSCI Korea PBR
7		구리 선물 근월물 가격(COMEX 기준)	23		한국 PER / 전세계 PER
8		금 선물 근월물(COMEX 기준)	24		한국 PER / 이머징 PER
9		금/은 가격 비율	25		한국 PBR / 전세계 PBR
10		옥수수 근월물	26		한국 PBR / 이머징 PBR
11	외환시장	달러인덱스	27		한국 주식시장 Yield Gap
12		OITP 달러인덱스	28		KOSPI 배당수익률
13		원/달러	29		KOSPI 외국인 보유 비중
14		엔/달러	30		최근 4주 외국인 순매수 대금
15		달러/유로 ATM 3개월 내재변동성	31		최근 4주 수익률
16		엔/달러 ATM 3개월 내재변동성			

자료: LIG투자증권 리서치본부

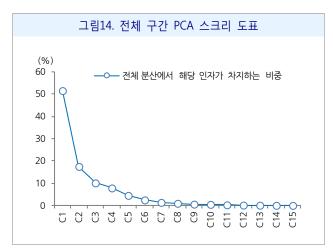
#### Step3. 시장 데이터의 주성분 분석(PCA)

주성분 분석의 개요와 모델에서의 사용 주성분 분석(PCA: Principal Component Analysis)는 두 가지 이유 때문에 사용한다. 첫 번째는 데이터의 차원(입력 변수의 갯수)을 축소시키기 위해서이고, 두 번째는 독립 변수의 직교 선형 변환을 이용하여 회귀분석에서의 다중공선성 문제를 완화시켜 주기 위해서이다. 다중공선성이란 독립변수간의 상관계수가 높을 경우 전체 모형의 분산이 커지고 모형 전체의 설명력이 낮아지는 문제이다.

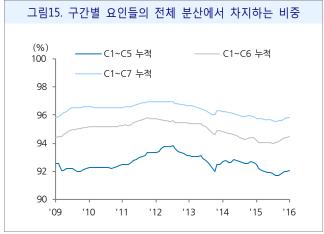
앞서 언급한 것처럼 본 자료에서는 1~60까지의 독립 변수와 2~61까지의 종속 변수를 이용해 모형을 작성하고, 작성된 모형에 61번째 독립 변수를 대입해 62번째 종속 변수를 예측하는 방식으로 진행되었다.

여기에 주 성분 분석 내용을 합쳐서 설명하면, 먼저 1~60까지의 독립변수를 주성분 분석한 후 변수 변환을 통해 얻은 변수로 모델을 만든다. 다음으로 61번째 독립변수를 변수변환하여 후 모델에 대입한 후 최종적으로 62번째 종속 변수를 예측하는 과정으로 풀어쓸 수 있다.

주 성분은 상위 7개 성분만을 사용했다. 주성분 분석 가이드 라인에서는 주성분의 분산이 급격하게 줄어드는 부분(스크리 도표의 팔꿈치, 그림14의 C3)까지의 성분만을 사용하는 것이 가장 효율적이라고 제시한다. 하지만 본 자료에서는 전체 분산에서 안정적으로 95% 이상을 확보할 수 있는 7번째 성분까지를 독립 변수로 사용했다(그림15 참조).



자료: LIG투자증권 리서치본부



자료: LIG투자증권 리서치본부

#### Step4. Multinomial Logistic Regression

로지스틱 회귀분석의 특징

단순 회귀분석과의 차이를 통해 로지스틱 회귀분석을 설명하면 크게 세 가지 특징을 꼽을 수 있다. 먼저 단순 회귀 모형이 독립 변수와 종속 변수간의 관계를 선형적인 관계로 설명한다면, 로지스틱 회귀 모형은 비선형 관계로 설명한다.

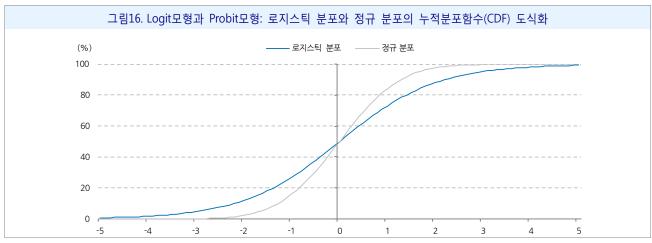
두 번째로, 모형을 추정할 때 회귀분석이 최소분산법(LSE: Least Square Estimator)를 사용한다면, 로지스틱 회귀분석은 최우추정법(MLE: Maximum Liklihood Estimator)을 사용한다. 회귀 분석과 달리 닫힌 형태의 계수 추정치가 존재하지 않기 때문에, 반복적인 처리를 통해 계수를 추정할 수밖에 없다.

세 번째가 가장 중요한데, 회귀분석의 경우 독립 변수와 종속 변수가 모두 연속형 변수이지만, 로지스틱 회귀분석의 경우에는 독립 변수는 연속형, 종속 변수는 이산형 변수를 사용한다.

예를 들어 중간 고사 성적(독립 변수)과 기말 고사 성적(종속 변수)과 같이 두 변수가 모두 연속형인 경우에는 회귀 분석을 이용할 수 있다. 하지만 몸무게(독립 변수)에 따른 당뇨 발생 여부(종속 변수)같은 상황을 가정해보자. 독립변수는 연속적이지만 종속변수는 1) 발생, 2) 발생하지 않음의 두 가지 상황만이 존재하며, 이러한 경우에는 단순 회귀분석이 아닌 로지스틱 회귀 분석이 적합하다.

로지스틱 회귀 분석의 예측은 '사건이 일어날 확률' 형태로 결과값이 나올 수밖에 없고, 0에서 1사이의 값만이 결과값으로 나오는 누적분포함수(CDF: Cumulattive Density Function)를 이용해 모형을 적합시킨다.

이 때 가장 많이 사용하는 분포는 로지스틱 분포와 정규 분포인데, 로지스틱 분포의 누적 분포함수에 적합시킨 모형을 Logit 모형, 정규 분포에 적합시킨 모형을 Probit 모형이라 한다. 정규분포의 경우 로지스틱 분포에 비해 꼬리가 얇은 형태를 보이기 때문에, 누적분 포 함수 역시 정규분포가 로지스틱 분포에 비해 빠르게 증가하는 형태를 보인다(그림16 참조). 본 자료에서는 Logit 모형을 사용하였다.



자료: LIG투자증권 리서치본부

이항 로지스틱 회귀모형(Binomial Logistic Regression) 개요 모델링에 사용한 다항 로지스틱 회귀모형(Multinomial Logistic Regression)을 작성하는 과정을 간략히 정리하였다. 먼저, 종속변수가 두 가지(1 or 0)만 존재하는 베르누이 공간을 가정하고, 이를 Logit 모형에 적합시키면 아래와 같이 표시할 수 있다.

$$P(Y_i = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_i)}}$$

$$P(Y_i = 0) = 1 - P(Y_i = 1) = \frac{e^{-(\beta_0 + \beta_1 X_i)}}{1 + e^{-(\beta_0 + \beta_1 X_i)}}$$

이를 다시 두 확률의 비율(Odds Ratio)로 표시하면 다음과 같다.

$$\frac{P(Y_i = 1)}{P(Y_i = 0)} = e^{(\beta_0 + \beta_1 X_i)}$$

마지막으로 양변에 로그를 취하면,

$$\ln\left(\frac{P(Y_i=1)}{P(Y_i=0)}\right) = \beta_0 + \beta_1 X_i$$

이러한 형태로 표시할 수 있다. 즉, 두 확률의 비율에 로그를 취한 값은 단순 선형 회 귀분석의 형태로 변환시킬 수 있다는 의미이다.

다만 여기서 좌변의 값은 한 개의 종속 변수가 아닌 두 종속 변수의 비율 형태로 표현되기 때문에 일반적인 회귀 분석에서 사용하는 최소분산법(LSE)를 사용하지 못한다. 로지스틱 회귀분석에서 최우추정법(MLE)를 사용하는 것은 이 때문이다.

다항 로지스틱 회귀모형(Multinomial Logistic Regression) 개요 다항 로지스틱 회귀분석의 경우에는 결과값이 세 개 이상인 경우를 의미한다. 다항 로지스틱 회귀분석 역시 확률 비율에 로그를 취한 후에 각각의 경우에 대하여 회귀 계수를 추정하는 방식으로 분석을 진행한다. 본 자료의 모형에서는, 대형 가치주의 한달 수익률이 가장 높았던 경우를  $P(Y_i=1)$ , 대형 성장주는  $P(Y_i=2)$ , 중소형 가치주는  $P(Y_i=3)$ , 중소형 성장주는  $P(Y_i=4)$ 로 표시한다. 중소형 성장주를 분모에 놓고 모형을 표시하면 다음과 같다.

$$\ln \frac{P(Y_i = 1)}{P(Y_i = 4)} = \beta_1 X_i$$

$$\ln \frac{P(Y_i = 2)}{P(Y_i = 4)} = \beta_2 X_i$$

$$\ln \frac{P(Y_i = 3)}{P(Y_i = 4)} = \beta_3 X_i$$

다항 로지스틱 회귀모형 유도 좌변에 각 경우의 확률만을 남겨놓고 정리하면 다음과 같다.

$$P(Y_i = 1) = P(Y_i = 4)e^{\beta_1 X_i}$$

$$P(Y_i = 2) = P(Y_i = 4)e^{\beta_2 X_i}$$

$$P(Y_i = 3) = P(Y_i = 4)e^{\beta_3 X_i}$$

좌변과 우변을 더해주고  $P(Y_i=1)+P(Y_i=2)+P(Y_i=3)=1-P(Y_i=4)$ 로 표시할 수 있으므로 이를 정리하면 아래와 같다.

$$1 - P(Y_i = 4) = P(Y_i = 4) \sum\nolimits_{k = 1}^3 {{e^{{\beta _k}{X_i}}}}$$

$$P(Y_i = 4) = \frac{1}{\sum_{k=1}^{3} e^{\beta_k X_i}}$$

마지막으로  $P(Y_i = 4)$ 를 각각의 경우에 대입해주면 최종적인 확률을 구할 수 있다.

$$P(Y_i = 1) = \frac{e^{\beta_1 X_i}}{\sum_{k=1}^3 e^{\beta_k X_i}}$$

$$P(Y_{i} = 2) = \frac{e^{\beta_{2}X_{i}}}{\sum_{k=1}^{3} e^{\beta_{k}X_{i}}}$$

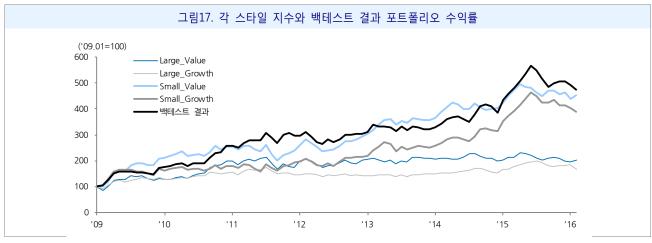
$$P(Y_i = 3) = \frac{e^{\beta_3 X_i}}{\sum_{k=1}^3 e^{\beta_k X_i}}$$

#### 백테스트 결과 정리

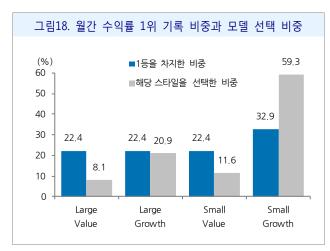
백테스트 결과와 스타일별 수익률 비교 로지스틱 회귀모형을 이용하여 도출한 최종적인 결과는 다음과 같다. 매월 마지막 금요일 전 금요일(즉, 한달에 금요일이 4번 있다면 3번째 금요일)까지의 데이터를 바탕으로 모형 을 만들고, 실제 비중 변경은 월 마지막 거래일 종가에 이루어졌다.

백테스트 포트폴리오는 매달 모델을 바탕으로 4개 중 가장 수익률이 높을 것으로 추정되는 스타일을 선정하고, 해당 스타일에 동일 비중으로 포트폴리오 투자를 하는 방식이다. 2009년 1월 마지막 거래일에 4가지 스타일 포트폴리오와 백테스트 포트폴리오에 100만 원씩을 투자했다면, 2016년 2월말에 대형 가치주는 203만원, 대형 성장주는 167만원, 중소형 가치주는 452만원, 중소형 성장주는 390만원이 되어 있고, 백테스트 포트폴리오는 475만원이 되어 있을 것이다(그림17 참조).

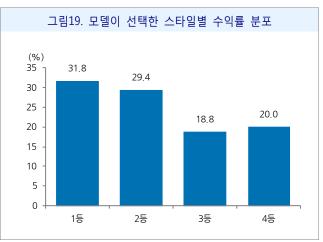
네 스타일 중 월간 수익률이 1등을 차지한 횟수가 가장 많았던 것은 소형 성장주였고, 모델 역시 다음달 유망한 스타일로 소형 성장주를 가장 많이 선택하였다(그림18 참조).



자료: WiseFn, LIG투자증권 리서치본부



자료: LIG투자증권 리서치본부



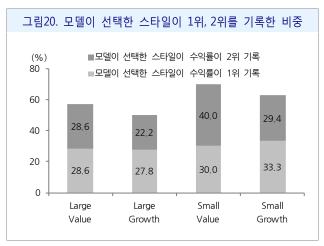
자료: LIG투자증권 리서치본부

스타일별 선택 비중과 R<sup>2</sup> 수익률 1등 스타일을 선택한 확률이 31.8%밖에 되지 않았지만, 3, 4위 스타일을 선택할 확률이 38.8%밖에 되지 않았기 때문에 전체 기간에서 긍정적인 수익률을 기록할 수 있었던 것으로 판단한다. 1등을 선택하는 것도 중요하지만, 3, 4등을 피해가는 확률 역시 중요하기 때문이다(앞페이지 그림19 참조).

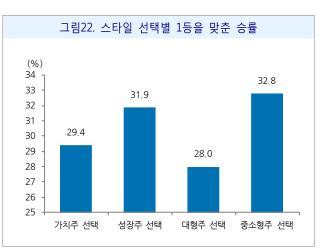
스타일 선택 측면에서 살펴보면 소형 가치주와 소형 성장주를 선택했을 때가 대형 가치주나 대형 성장주를 선택했을 때보다 승률이 좀 더 높았다. 가장 승률이 낮았던 것은 대형성장주를 선택했을 때인데, 승률은 50% 수준에 불과하였다(그림20,21 참조).

전체적으로 가치주보다는 성장주가, 대형주보다는 중소형주를 선택한 경우 높은 수준의 승률을 기록하였다(그림22 참조).

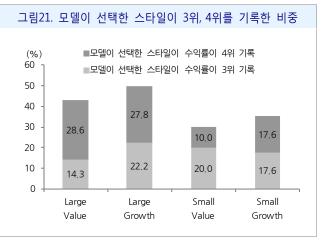
로지스틱 회귀모형의 적합도는 R²가 아닌 McFadden Pseudo R², Cox & Snell R², Nagelkerke R² 중 하나를 사용하는데, 본 자료에서는 Nagelkerke R²를 이용하여 적합도를 계산하였다. 30~50% 수준의 R²가 유지되는 모습을 보였으며, 이는 충분히 의미있는 수준의 모형이라고 판단한다.



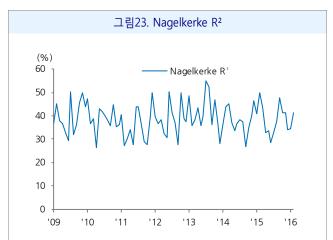
자료: LIG투자증권 리서치본부



자료: LIG투자증권 리서치본부



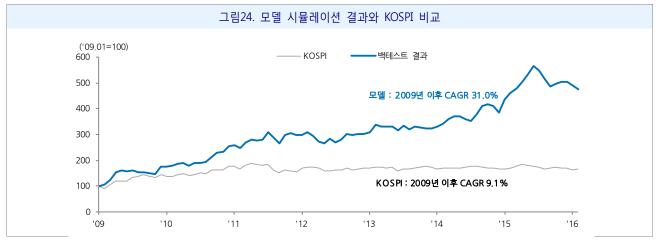
자료: LIG투자증권 리서치본부



자료: LIG투자증권 리서치본부

백테스트 결과 KOSPI와 비교 백테스트 결과를 KOSPI와 비교해보면 수익률 차이는 매우 크다. 2009년 이후 KOSPI 수 익률은 CAGR 9.1% 수준이었는데, 모형에 의한 포트폴리오는 31.0%를 기록했다. KOSPI 가 박스권 흐름을 보인 상황에서 스타일별 수익률 차이가 상이했던 만큼, 이에 따른 수익률 차이가 발생한 것으로 보인다.

KOSPI를 Outperform했느냐 Underperform했느냐로 나누어 보면, 월간 수익률 기준 60% 의 확률로 KOSPI 대비 높은 수익률을 기록했다. 지속적으로 KOSPI 수익률을 하회한 경우도 있었는데, 연속적으로 KOSPI 수익률을 하회한 기간은 최장 4개월이었다.



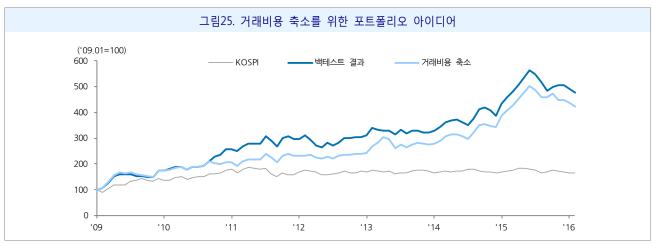
자료: Fnguide, LIG투자증권 리서치본부

거래비용 축소 아이디어 하지만 실제 투자에서 백테스트와 같은 결과를 100% 낸다는 것은 불가능하다. 이유는 1) 거래 비용이 발생하고, 2) 중소형주를 투자할 때와 대형주에 투자할 때 가격에 미치는 영향이 같다고 말하기 어려우며, 3) 시가총액 차이가 존재하는 60여 개의 종목을 동일 비중으로 모두 산다는 것 역시 현실적으로 쉽지 않다. 특히 투자하는 규모가 커질수록 포트폴리오 교체에 따른 유동성 위험은 피하기가 어렵다.

현실적인 거래 측면에서 거래비용을 축소시키기 위한 추가적인 포트폴리오를 구성해 보았다. 새로운 포트폴리오는 예측 모형에서 현재 보유하고 있는 스타일의 수익률을 가장 높은 확률로 제시하지 않더라도, 2번째로 높은 확률을 제시한다면 현재 포트폴리오를 유지했다.

즉, 예전 포트폴리오가 현재 보유 스타일을 다음달에도 가장 긍정적으로 판단할 때에만 스타일을 유지했다면, 새로운 포트폴리오는 1, 2위인 경우 모두 현재 스타일을 유지해 포트폴리오 변경을 감소시켰다. CAGR 수익률은 31.0%에서 28.4%로 감소하지만, 포트폴리오 변경 역시 1년에 평균 5.4번 변경에 3.2번 변경으로 줄어든다.

수익률은 2.5%p, 변경 횟수는 2.1번 감소한다는 점에서, 한 번 포트폴리오를 변경하는데 드는 비용이 1.2% 이상 든다면 두 번째 포트폴리오가 현실적으로 더 높은 성과를 제공한 다는 의미이다(다음페이지 그림25 참조).



자료: Fnguide, LIG투자증권 리서치본부

## 결론 및 모델 개선 가능성

모형의 의미 및 개선 가능성 로지스틱 회귀모형을 이용한 스타일 선택은 충분히 의미있는 전략이라고 판단한다. 하지만 실질적인 대형주와 중소형주간의 편드 변화를 단기간에 실시하는 것은 현실적으로 쉽지 않다는 점에서 상기의 수익률을 100% 달성하기는 어렵다.

다만 패시브 펀드의 경우 유망한 스타일을 비중 확대하고, 유망하지 않은 스타일을 비중 축소하는 등의 전략을 인덱스 플러스 알파 전략으로 활용할 수 있다. 액티브 펀드 입장에 서도 충분히 참고할 만한 내용이 될 수 있을 것이다. 우리는 지난 박스권 시장에서 스타 일 선택이 얼마나 성과를 다르게 만드는지를 확인했기 때문이다.

본 자료의 모형은 Journal of Asset Management<sup>1</sup>에 소개된 논문을 많이 참고하였는데, 논문의 모형은 주성분 분석을 실시하지 않았다는 점과 독립변수를 Granger Causality Test를 이용해 선정했다는 점에서 본 자료와 차이가 존재한다.

주성분 분석의 경우 다중공선성 문제는 해결하지만, 독립변수와 종속변수 간의 민감도에 대한 수치(예를 들면 "원자재 가격이 상승하는 구간에서 대형 가치주 성과가 긍정적이었다"와 같은 독립변수와 종속변수의 인과관계 해석)를 제공해주지 못한다는 점에서 직관적인 이해에 걸림돌이 된다.

앞으로 Granger Causality를 활용한 독립변수 선정이나 전월 데이터만이 아닌 2개월, 3개월 전의 데이터까지 사용한 모형으로의 확장도 고려할 수 있다.

향후 모델의 수정이 있을 경우 다시 보고서를 통해서 소개할 예정이며, 그 전까지는 매월 말에 스타일별 전망 및 확률을 점검하고 과거 성과를 표시한 자료를 발간할 예정이다.

<sup>&</sup>lt;sup>1</sup> Bala G. Arshanapalli, Lorne N. Switzer, and Karim Panju, "Equity-Style Timing: A Multi-Style Rotation Model for the Russell Large-Cap and Small-Cap Growth and Value Style Indexes", Journal of Asset Management vol. 8, no. 1 (May 2007):9–23



LIG Research

리서치본부장 지기호/기술적분석

02)6923-7330 khchi@ligstock.com

투자전략팀	Analyst	내선번호	E-Mail
이코노미스트(총괄)	김유겸	6923-7332	yukyum@ligstock.com
투자전략	윤영교	6923-7352	skyhum00@ligstock.com
계량분석	염동찬	6923-7314	dongchan@ligstock.com
채권/크레딧	유선응	6923-7353	carpelife@ligstock.com
시황	김예은	6923-7349	yekim@ligstock.com
경제 RA	문다솔	6923-7338	dasol.moon@ligstock.com
퀀트/채권 RA	한민우	6923-7331	mwh7733@ligstock.com
투자전략 RA	신현우	6923-7348	god5892@ligstock.com
기업분석팀	Analyst	내선번호	E-Mail
정유/화학(총괄)	박영훈	6923-7317	houn0715@ligstock.com
철강/금속	김윤상	6923-7319	yoonsang.kim@ligstock.com
지주회사/유통	김태현	6923-7312	thkim@ligstock.com
반도체/디스플레이	신현준	6923-7336	anthony0808@ligstock.com
스몰캡/통신서비스	김인필	6923-7351	ipkim01@ligstock.com
제약/바이오	이무진	6923-7343	c_in_c@ligstock.com
자동차	신재영	6923-7341	jyshin@ligstock.com
은행/증권	은경완	6923-7322	kw.eun@ligstock.com
엔터/레저/미디어	지인해	6923-7315	emalee89@ligstock.com
인터넷/게임	김성은	6923-7350	eunee714@ligstock.com
건설	김기룡	6923-7321	kirykim@ligstock.com
스몰캡	박상용	6923-7323	sean.park@ligstock.com
IT/스몰캡 RA	고의영	6923-7345	keuiy@ligstock.com
지주/유통/금융 RA	강수민	6923-7362	soomin913@ligstock.com
철강/제약 RA	원민석	6923-7313	ethanwon@ligstock.com
엔터/인터넷 RA	윤을정	6923-7318	eyun16@ligstock.com
정유/화학 RA	최선미	6923-7311	serena.choi@ligstock.com

#### ▶ Compliance

본 조사자료는 고객의 투자에 정보를 제공할 목적으로 작성되었으며, 어떠한 경우에도 무단 복제 및 배포 될 수 없습니다. 또한 본 자료에 수록된 내용은 당사가 신뢰할 만한 자료 및 정보로 얻어진 것이나, 그 정확성이나 완전성을 보장할 수 없으므로 투자자 자신의 판단과 책임하에 최종결정을 하시기 바랍니다. 따라서 어떠한 경우에도 본 자료는 고객의 주식투자의 결과에 대한 법적 책임소재의 증빙자료로 사용될 수 없습니다.

