

금융·거시경제변수들의 한국 주식수익률 예측가능성 검정

- 표본 내 검정과 표본 외 검정 결과 비교를 중심으로 -

Predicting Korean Stock Market Return with Financial and Macro variables

- Focusing on In-sample and Out-of-sample Tests -

전 성 주*

Sungju Chun

본 연구는 한국 주식시장에서 12개의 주요 거시경제변수 및 금융변수들을 이용하여 미래 장·단기 주식수익률을 예측할 수 있는지 실증적으로 검정하였다. 특히, 선행연구에서 많이 사용되었던 표본 내 예측가능성 검정뿐만 아니라 표본 외 예측가능성 검정을 함께 시행함으로써 각 예측변수들의 예측력을 보다 강건하게 검정하였다. 이를 위해 내포모형(Nested model)의 예측치를 검정할 수 있도록 McCracken (2007)이 제안한 MSE-F 검정기법과 Clark and McCracken (2001)이 제안한 ENC-NEW 검정기법을 사용하였다. 이와 함께, 부트스트랩을 통한 재표본추출(Resampling through bootstrapping)을 통해 임계치와 p-value를 산출함으로써 주식수익률 예측에서 일어나는 소표본 편차(Finite-sample bias)와 장기수익률 잔차항의 자기상관성(Autocorrelation) 문제를 해결하였다. 검정 결과, 주가순자산비율(Book-to-market ratio) 변수가 표본 내 검정과 표본 외 검정에서 모두 주식수익률 예측력을 갖고 있는 것으로 나타나 가장 일관성 있는 예측변수로 나타났다.

국문 색인어: 주식수익률 예측가능성, 거시경제변수, 금융변수, 표본 내 검정, 표본 외 검정

한국연구재단 분류 연구분야 코드: B050704, B030603

* 가천대학교 경영대학 경영학과 교수(sjchun@gachon.ac.kr), 제1저자

논문 투고일: 2019. 11. 28, 논문 최종 수정일: 2020. 1. 24, 논문 게재 확정일: 2020. 2. 21

I. 서론

본 연구는 한국 주식시장에서 주요 거시경제변수 및 금융변수들을 이용하여 미래 주식수익률을 예측할 수 있는지 실증적으로 검증하는 것을 목표로 한다. 재무학계에서 주식시장의 수익률 예측가능성을 검증하는 연구는 Cowles (1933)에 의해 시작된 이후 수많은 이론적 연구와 이를 검증하는 실증 분석 연구가 이루어졌다. 실제 금융업계에서는 주식수익률 예측을 통해 직접적인 이윤 추구를 시도하기도 한다. 하지만, 주식수익률은 대부분 예측할 수 없는 요소로 구성되어 있기 때문에 가장 뛰어난 예측모형이라도 주식수익률의 매우 작은 부분만을 설명할 수 있는 한계를 넘기 어렵다. 그럼에도 불구하고, 학문적으로 금융·거시경제 변수들의 주식수익률 예측가능성 검증이 의미를 갖는 이유는 금융시장의 효율성 가설 검증과 함께 더 현실적인 이론인 조건부 자산가격결정모형(Conditional asset pricing model)의 구축에 필수적인 상태 변수(State variable) 선정에 이용되기 때문이다.

Fama (1970)에 의해 제시된 금융시장의 효율성 가설은 약형(Weak-form), 준강형(Semi-strong form), 강형(Strong-form) 효율성으로 나뉘는데, 기본적으로 금융시장이 효율적이라면 주식가격은 무작위보행(Random walk) 혹은 마팅게일(Martingale) 프로세스를 따르게 되어 주식수익률은 예측 불가능하다. 따라서, 과거 주식가격에 담겨 있는 정보나 공개된 정보(거시경제변수 혹은 금융변수)를 활용하여 주식수익률이 예측 가능한 경우 금융시장의 약형 혹은 준강형 효율성 가설이 위배되는 것이다. 이 경우, 금융시장의 비효율성으로 인한 주식수익률 예측가능성은 수익 창출의 기회가 될 수 있다(Friedman (1953); Samuelson (1965)).

그러나, 최근에 제안된 합리적 기대(Rational expectation)에 기반한 조건부 자산가격 결정모형 이론은 주식수익률의 구성요소 가운데 시간변이 총체적 위험 프리미엄(Time-varying aggregate risk premium)이 존재하며, 이 프리미엄과 동반하여 움직이는 변수가 있다면 효율적인 금융시장하에서도 주식수익률이 예측 가능할 수 있다고 제시한다. 대표적으로 Campbell and Shiller (1988a)는 배당수익률의 장기 평균으로부터의 편차(Deviation)가 기대 미래 배당성장률(Expected future dividend growth rates)과 기대 미래 주가수익률(Expected future stock returns)의 변화를 예측할 수 있다는 점을

이론적으로 증명하였다. 특히 후자는 시간변이 할인율(Time-varying discount rate)을 결정하기 때문에 이와 관련된 금융변수들이 주식수익률을 예측할 수 있다는 점을 내포한다. 한편, Fama and French (1989), Campbell and Cochrane (1999), Cochrane (2007, 2011)은 금융자산 수익률이 실물경제의 상태 변수들(State variables)에 의해 영향을 받고 실물경제는 경기순환에 따라 변동하기 때문에 실물경제의 상태(State of the real economy)를 측정하거나 예측하는 변수들도 주식수익률의 예측에 도움을 줄 수 있다는 점을 보였다.

따라서, 본 연구에서는 한국 실물경제의 상태를 측정하는 12개의 대표적인 거시경제 변수들과 시간변이 기대 수익률을 측정하는 금융변수들이 한국 주식시장의 장·단기 미래 수익률을 예측할 수 있는지 실증적으로 검정하고자 한다. 특히, 본 연구에서는 기존의 선행 연구에서 많이 살펴보았던 표본 내 예측가능성(In-sample predictability) 검정과 함께 표본 외 예측가능성(Out-of-sample predictability) 검정을 시행함으로써 표본 내 검정에서 나타날 수 있는 극단치(Outlier)나 데이터 마이닝(Data-mining) 문제를 완화하고 실제 상황에서 수익률을 예측할 때 이뤄지는 패턴을 그대로 구현함으로써 예측가능성을 좀 더 강건하게 검정하고자 한다. 특히, 표본 외 예측가능성 검정에 있어서 예측오차를 이용하여 통계적 추론이 가능하도록 McCracken (2007)과 Clark and McCracken (2001)이 제안한 MSE-F 검정치와 ENC-NEW 검정치를 적용하였다. 또한, 예측회귀식 추정에 있어서 일어날 수 있는 소표본 편차(Small-sample bias) 문제를 해결하고 장기수익률 예측에서 일어나는 자기상관성(Autocorrelation) 문제를 해결하기 위해 부트스트랩을 통한 재표본 추출(Resampling through bootstrapping)을 통해 임계치와 p-value를 산출하여 추론하였다. 표본 내 예측가능성 검정 결과 채무불이행 수익률 스프레드(Credit spread)와 주가순자산비율(Book-to-market ratio), 배당가격비율(Dividend-price ratio), 배당수익률(Dividend yield), 주가이익비율(Earnings-to-price ratio) 등이 유의한 주가수익률 예측력을 갖는 것으로 나타났으며, 표본 외 예측가능성 검정 결과 주가순자산비율과 인플레이션이 주식수익률 예측력을 갖고 있는 것으로 나타났다. 특히, 주가순자산비율 변수는 표본 내 검정과 표본 외 검정에서 모두 주식수익률 예측력을 갖고 있는 것으로 나타나 가장 일관성 있는 예측변수로 나타났다.

본 연구의 구성은 다음과 같다. 제 2장에서는 선행연구를 개관하고 제 3장에서는 예측계량모형을 소개하고 표본 내 예측가능성과 표본 외 예측가능성 검정방법에 대해 자세히 살펴본다. 제 4장에서는 자료 분석과 함께 실증분석결과를 제시하였다. 제 5장에서는 연구 결과와 함의를 요약하고 향후 연구 과제를 제시하였다.

II. 선행연구

주식수익률 예측가능성에 대한 연구는 대부분 미국의 자료를 바탕으로 많은 연구가 진 행되었으며, 다양한 변수들이 주식수익률을 예측할 수 있는 것으로 밝혀졌다. 배당수익률 (Rozeff (1984); Campbell and Shiller (1988a, 1998); Fama and French (1988, 1989)), 이익주가비율(Campbell and Shiller (1988b, 1998)), 장부-시장가치 비율 (Kothari and Shanken (1997); Pontiff and Schall (1998)), 명목 이자율(Fama and Schwert (1977); Breen et al. (1989)), 이자율 스프레드(Campbell (1987); Fama and French (1989)), 인플레이션(Nelson (1976); Campbell and Vuolteenaho (2004)) 등 이 주식수익률을 예측하는 것으로 알려진 대표적인 거시경제 혹은 금융변수들이다. 이러한 결과들은 대부분 표본 내 회귀(In-sample regression) 추정을 통해 발견된 것들이 대부분인데, 일각에서는 표본 내 추정을 통해 예측가능성을 검정할 경우 추정치에 편차가 발생하거나 데이터 마이닝 문제가 발생할 수 있음을 지적하고 있다. Nelson and Kim (1993)과 Stambaugh (1999)는 대부분의 예측변수들이 단위근에 가까운 자기회귀계수를 갖고 있기 때문에 기울기 매개변수를 추정함에 있어 통계적인 편차(Bias)가 발생할 수 있음을 지적하였다. 또한, Foster et al. (1997)과 Ferson et al. (2003)은 기존의 문헌들이 많은 수의 예측변수들의 검증 결과를 직·간접적으로 인지하고 있음에도 불구하고 오직 통계적으로 유의성을 갖는 변수만을 반복적으로 보고함으로써 데이터 마이닝(Data mining) 문제가 발생함을 지적하였다. 특히, 이러한 데이터 마이닝 문제는 비슷하거나 중복된 표본 기간을 사용하는 표본 내 검정에서 심화되는 것으로 알려져 있다. 따라서, 이러한 문제들을 극복하기 위해 본 연구에서는 표본 외 회귀(Out-of-sample regression) 추정을 함께

살펴 봄으로써 주식수익률 예측가능성을 보다 강건하게 검증하고 부트스트랩(Bootstrap)을 이용한 재표본추출(Resampling)을 통해 임계치(Critical value)와 p-value를 계산함으로써 통계적인 편차를 감안할 수 있도록 하였다.

글로벌 주식시장에 대한 수익률 예측가능성 연구는 대표적으로 Wohar et al. (2005), Hjarmarsson (2010), Schrimpf (2010), Jordan et al. (2014), Charles et al. (2017) 등이 있다. Wohar et al. (2005)은 12개 선진국 주식시장을 대상으로 9개의 주요 거시경제변수들의 예측력을 살펴 보았으며, Hjarmarsson (2010)은 24개 선진국과 16개 개발도상국의 주식시장을 대상으로 배당가격비율, 주가순이익비율, 단기이자율, 금리 스프레드의 변수들이 예측력을 갖는지 살펴보았다. Schrimpf (2010)는 5개 주요국의 주식시장 수익률을 대상으로 9개의 금융 및 거시경제변수들의 예측력을 살펴보았으며 Jordan et al. (2014)은 14개 유럽 국가들의 주식시장 수익률 예측가능성을 살펴보았다. Charles et al. (2017)은 21개 유럽 국가들과 16개 아시아 국가들의 수익률에 대해 금융변수와 기술적인 지표(Technical indicator), 단기이자율 등이 예측력을 가지는지 살펴 보았다. 이들 연구들 가운데 한국 주식시장 수익률 예측가능성을 살펴 본 연구는 Charles et al. (2017)로 이들은 예측회귀식에서 일어나는 편차를 수정할 수 있는 추정법을 사용하여 Wild bootstrap 방법을 사용하여 예측력을 검증하였다. 물론 이들의 연구에서도 Theil's U-비율(Theil's U-ratio)을 통해 표본 외 예측력을 살펴보고는 있으나, 후술하는 바와 같이 U-비율은 예측오차의 평균제곱근오차(RMSE)의 비율만을 살펴보는 것이기 때문에 표본 외 예측력을 강건한 통계적 검정치인 MSE-F와 ENC-NEW를 통해 검증하는 본 연구와 차이가 있다고 할 수 있다. 이와 더불어, 본 연구에서는 이들이 살펴본 변수들보다 더 많은 12개의 금융변수와 거시경제변수들의 예측력을 살펴 본다는 점에서 보다 포괄적인 예측변수들의 예측력을 검증한다는 의미가 있다.

거시경제변수 혹은 금융변수와 같은 일반적으로 공개된 정보를 사용하여 한국 주식시장 수익률 예측가능성을 검증한 연구로는 김규영·김영빈 (2004), 김인무·박성근 (2009), 정정현·김수경 (2010), 정동준 (2012), 김현석·여효석 (2017) 등이 있다. 김규영·김영빈 (2004)은 배당수익률 및 이익주가비율(E/P ratio)과 같은 금융변수가 주가수익률 예측력을 갖는지 검증하였는데 금융변수의 자기상관계수에 담겨있는 정보를 활용하는 Lewellen

(2004)의 계량방법을 이용한 결과 이익주가비율이 예측력을 갖는 것을 발견하였다. 김인무·박성근 (2009)은 기존의 시계열 분석과 달리 한국 주식시장의 경제적 시간을 실현된 분산(Realized volatility)으로 추정하여 계산하는 변동성 시계로 측정하여 관측치를 추출한 변동성 시계표본을 사용하여 배당수익률, 이익주가비율, 이자율 등의 금융변수가 주식 수익률에 예측력을 가지는지 실증분석하였다. 이를 통해 이자율은 주가수익률에 예측력을 가지는 반면, 배당수익률과 이익주가비율의 금융변수는 표본 추출방법에 따라 예측력에 차이가 나타나는 것을 발견하였다. 정정현·김수경 (2010)은 배당수익률과 이익주가비율(E/P ratio)의 금융변수가 KOSPI 및 규모별·산업별 주가지수 수익률 예측가능성을 검정하였다. 이들은 특히 KOSPI 지수 수익률에 대해 금융변수가 예측력을 갖지 못하는 것을 발견하였으나 외환위기 이후 기간에서는 이익주가비율이 예측력을 가지는 것을 발견하였다. 정동준 (2012)은 금융변수들의 한국 주식수익률에 대한 장기 예측가능성을 분석하였는데 주식수익률, 배당수익률, 주가수익비율 등의 금융변수들이 장단기에 있어 예측력을 갖고 있으나 미국 주식시장에서 발견되었던 장기 수익률에 대한 예측력 증가 현상은 한국 주식시장에서 발견되지 않았으며, 과거의 실질이자율은 주식수익률 예측에 영향을 주지 못하였다. 김현석·여효석 (2017)은 배당수익률, 이익주가비율, 이자율, Yield Gap 등 금융변수들의 주가수익률 예측력을 검정하였다. 특히, 이들은 주가수익률 예측력 검정에서 일어날 수 있는 통계적 문제들을 처리할 수 있는 변동성 시계표본을 이용하였는데 그 결과 배당수익률이 주가수익률 예측력을 갖는 것을 발견한 반면, 새로운 변수인 Yield Gap의 예측력을 찾지 못하였다. 이와 같이 한국 주식시장 수익률 예측가능성에 관한 선행연구들은 주로 금융변수들의 예측력을 검정하는데 초점을 맞추고 있다.

본 연구에서는 다양한 거시경제변수들과 금융변수들을 활용하여 한국 주식시장의 수익률 예측가능성을 검정하고 표본 내 검정과 표본 외 검정을 동시에 시행함으로써 개별변수의 예측력을 보다 강건하게 검정하는데 연구 초점을 맞추고 있다.

III. 예측계량모형의 추정 및 검정 방법

일반적으로 주식수익률의 예측력 검정은 다음의 예측계량모형(Predictive regression model)을 이용하여 이루어져 왔다.

$$r_{t+1}^k = \alpha + \beta x_{i,t} + \gamma r_t + \epsilon_{t+1}^k \quad (1)$$

예측회귀식 (1)에서 r_t 는 $t-1$ 월부터 t 월까지의 초과 주식수익률(Excess stock return)을 나타내고 $r^{k_{t+1}} = r_{t+1} + \dots + r_{t+k}$ 은 t 월부터 $t+k$ 월까지 초과 주식수익률을 나타낸다. Fama and French (1988)와 같은 선행연구에서 주식수익률을 $r^{k_{t+1}}$ 과 같이 중복시킴으로써 장기 수익률(Long-horizon return)을 예측할 때 예측변수들의 예측력이 더 증가함을 보였다. 따라서, 본 연구에서는 1개월, 3개월, 1년과 2년의 장기수익률을 예측하는 경우를 모두 살펴보았다. $x_{i,t}$ 는 거시경제변수와 금융변수들로 구성된 K개의 잠재적인 예측변수들(Potential predictors) 가운데 i -번째 변수를 나타내고, $\epsilon^{k_{t+1}}$ 은 예측회귀식의 잔차항을 나타낸다. 보통 주식수익률의 예측가능성은 예측변수의 기울기를 나타내는 β 의 추정치인 $\hat{\beta}$ 이 통계적 유의성을 갖는지를 살펴보는 것으로 검정될 수 있다. ($H_0 : \beta = 0$ v.s. $H_A : \beta \neq 0$) Lettau and Ludvigson (2001)과 같은 선행연구들에서 전월 주식수익률을 예측회귀식의 통제변수로 사용하였던 점을 감안하여 예측변수의 예측력을 검증하는 통제변수로 전월 주식수익률인 r_t 도 포함한다.

1. 표본 내 예측가능성 검정(In-sample predictability)

예측회귀식 (1)에서 잠재적인 예측변수 x_t 의 예측력은 통상적인 선형회귀분석의 최소자승추정법(OLS)으로 구한 $\hat{\beta}$ 의 t-검정치(t-statistic)와 결정계수(R-square)를 통해 검정할 수 있다. 그러나, 주기와 관련이 깊은 금융변수들의 경우 예측회귀식 (1)의 과거 잔차항과 상관성(Correlation)이 높은 경우가 많으며, 이 때 $\hat{\beta}$ 의 추정치에 소표본 편차가 발생하기 쉽다(Nelson and Kim (1993); Stambaugh (1999)). 또한, 장기수익률($k > 1$) 예측 시 중복된 수익률 관측치를 사용하기 때문에 잔차항에 자기상관성(Serial correlation)이 높

아지며 t-검정치에 쓰이는 일반적인 표준편차 추정치는 실제 표준편차를 과소추정하는 오류에 빠지기 쉽다. 따라서 본 연구에서는 Newey and West (1987)에 제안된 이분산성과 자기상관성을 감안한 HAC(Heteroskedasticity and Autocorrelation Consistent) 공분산 추정방법을 사용한다.¹⁾ 하지만, 여전히 소표본 편차와 자기상관성에 의해 유의수준 왜곡(Size distortion) 현상이 일어날 수 있기 때문에 후술할 부트스트랩 절차(Bootstrap procedure)에 따른 재표본추출로 구한 임계치와 p-value를 통계적 추론(Statistical inference)에 사용한다.

2. 표본 외 예측가능성 검정(Out-of-sample predictability)

가. 주식수익률 예측치 및 예측오차 추정

본 연구에서는 표본 내 검정뿐만 아니라 표본 외 검정도 함께 실행한다. 표본 외 검정을 위해서 먼저 표본 전체 기간을 둘로 나눈다. 표본 전체 기간을 T , 첫 번째 기간을 R 이라고 할 때, 나머지 기간을 $P = T - R$ 이라고 하자. 이 때, 첫 번째 표본 외 예측치는 다음과 같이 구할 수 있다. 먼저, R 까지의 첫 번째 구간표본을 사용하여 예측회귀식 (1)의 매개변수 α , β 와 γ 를 OLS로 추정한다. 이들 OLS 추정치를 각각 $\hat{\alpha}_{1,R}$, $\hat{\beta}_{1,R}$, $\hat{\gamma}_{1,R}$ 이라고 하자. 이들 추정치와 함께 x_R 과 r_R 을 회귀식 (1)에 대입하여 다음기 주식수익률 예측치인 $\hat{r}_{1,R+1}^k = \hat{\alpha}_{1,R} + \hat{\beta}_{1,R}x_{i,R} + \hat{\gamma}_{1,R}r_R$ 과 예측오차 $\hat{\epsilon}_{1,R+1}^k = r_{R+1}^k - \hat{r}_{1,R+1}^k$ 을 추정할 수 있다. 이 예측오차는 대립가설(Alternative hypothesis)하에서 추정한 예측오차이다. 귀무가설(Null hypothesis)하에서도 예측변수의 예측력을 측정하는 매개변수인 $\beta = 0$ 으로 설정해 놓고 비슷한 절차를 통해 귀무가설하에서의 다음기 주식수익률 예측치 $\hat{r}_{0,R+1}^k = \hat{\alpha}_{0,R} + \hat{\gamma}_{0,R}r_R$ 와 예측오차 $\hat{\epsilon}_{0,R+1}^k = r_{R+1}^k - \hat{r}_{0,R+1}^k$ 을 추정할 수 있다.

두 번째 표본 외 예측치는 첫 번째 표본기간을 $R+1$ 로 늘린 후에 위의 과정을 반복함으로써 구할 수 있다. 귀무가설과 대립가설 하에서 먼저 매개변수들을 표본 내 OLS 추정법으로 추정한 후 x_{R+1} 과 r_{R+1} 을 이용하여 주식수익률 예측치 $\hat{r}_{1,R+2}^k$ 과 $\hat{r}_{0,R+2}^k$, 예측

1) 이 HAC 공분산 추정방법에 쓰인 커널(Kernel)은 Bartlette kernel을 사용하였으며, 커널의 시차를 결정하기 위해서 장기수익률 기간인 k 에 1.5배를 곱한 숫자에서 가장 가까운 정수를 사용한다.

오차 $\hat{\epsilon}_{1,R+2}^k$ 과 $\hat{\epsilon}_{0,R+2}^k$ 을 구할 수 있다. 이와 같은 과정을 전체 표본 마지막 관측치까지 반복함으로써 $T - R - k + 1$ 개의 예측오차 집합인 $\{\hat{\epsilon}_{1,t+1}^k\}_{t=R}^{T-k}$ 와 $\{\hat{\epsilon}_{0,t+1}^k\}_{t=R}^{T-k}$ 을 구할 수 있다.

나. 표본 외 검정

표본 외 예측가능성 검정은 각각 귀무가설과 대립가설하에서 추정된 예측오차를 비교함으로써 시행될 수 있다. 귀무가설하에서 미래수익률 추정은 $\beta = 0$ 의 설정을 통해 주가수익률 1차 시차만을 이용한 AR(1) 모형을 사용하여 추정된 것이며, 만약 예측변수를 사용한 미래수익률 추정치가 더 우월하다면 대립가설하에서 추정된 예측오차가 귀무가설하에서 추정된 예측오차보다 유의하게 더 작을 것이다. 이를 간단하게 비교하기 위해서 Theil's U-비율을 살펴볼 수 있는데, 이 비율은 대립가설 예측오차의 RMSE(Root-mean-squared error: 평균제곱근 오차)를 귀무가설 예측오차의 RMSE로 나눈 값이다. 만약 대립가설 예측오차가 더 작다면 $U < 1$ 일 것이다.

본 연구에서는 이들 예측오차를 통계적으로 정확히 비교하기 위해 McCracken (2007)이 제안한 MSE-F 검정치와 Clark and McCracken (2001)이 제안한 ENC-NEW 검정치를 산출한다. McCracken (2007)은 Diebold and Mariano (1995)와 West (1996)가 예측력 검정을 위해 손실 차이(Loss differential)에 기반하여 제안한 검정치를 개량한 검정법이다. 미래 주식수익률 예측오차의 손실차이를 $\hat{d}_{t+1}^k = (\hat{\epsilon}_{0,t+1}^k)^2 - (\hat{\epsilon}_{1,t+1}^k)^2$ 라고 하면, 귀무가설 및 대립가설하의 예측오차의 평균제곱오차는 각각 다음과 같다.

$$MSE_i = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} (\hat{\epsilon}_{i,t+1}^k)^2, \quad i = 0, 1 \tag{2}$$

이 때, $\bar{d} = MSE_0 - MSE_1$ 로 설정하면 MSE-F 검정치는 다음과 같다.

$$MSE - F = (T - R - k + 1) \times \bar{d} / MSE_1 \tag{3}$$

MSE-F 검정치는 월별 주가수익률($k = 1$) 예측력 검정과 같은 내포모형(Nested

model)의 예측치를 비교할 때 중심축량(Pivotal)을 갖게 되지만 브라운 운동(Brownian motion)을 이용한 확률적분(Stochastic integral)의 함수로 이루어진 비표준 극한분포(Non-standard limiting distribution)를 갖는다.²⁾ 특히 본 연구와 같이 장기수익률($k > 1$) 예측력을 검정하는 경우 Clark and McCracken (2002)는 MSE-F 검정치가 중심축량을 갖지 않는(Non-pivotal) 비표준 극한분포를 갖게 된다는 점을 보였다. 따라서, 이들은 부트스트랩에 기반한 임계치를 사용하도록 제안하고 있으며 본 연구에서는 후술할 부트스트랩 절차를 통해 MSE-F 검정치에 대한 통계적 추론을 이끌어 낸다.

ENC-NEW 검정치는 포괄적 예측(Forecast encompassing) 개념을 기반으로 예측력을 평가하기 위해 제안되었다. 귀무가설과 대립가설 하의 예측치를 최적으로 조합하여 포괄적 예측치를 구성한다고 할 때, 예측변수를 포함한 대립가설 예측치가 유용한 정보를 제공하지 않는다면 대립가설 예측치에 대한 포괄적 예측치의 가중치가 0에 가까울 것이다. 따라서, Clark and MacCracken (2001)은 다음과 같은 검정치를 제안하였다.

$$ENC-NEW = (T - R - k + 1) \times \bar{c} / MSE_1 \quad (4)$$

여기서 $\bar{c} = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \hat{c}_{t+1}^k$ 이고 $\hat{c}_{t+1}^k = \hat{\epsilon}_{0,t+1}^k (\hat{\epsilon}_{0,t+1}^k - \hat{\epsilon}_{1,t+1}^k)$ 이다.

이 ENC-NEW 검정치의 극한분포도 MSE-F 검정치의 극한분포와 유사한 성격을 갖는다. 다시 말해, 비표준 극한분포를 갖고 있으며 장기수익률에 대해서는 중심축량을 갖지 않는다. 따라서, ENC-NEW 검정치의 임계치 또한 부트스트랩 절차를 통해 얻어진 값을 사용한다. Clark and MacCracken (2001, 2002)에 따르면 MSE-F 검정치와 ENC-NEW 검정치는 주식수익률 예측치를 비교하는 경우와 같이 내포모형 예측치를 검정할 때 이전에 제안된 검정치들에 비해 유의수준 왜곡현상을 최소화하면서 더 강력한 검정력을 갖는 것으로 나타났다.

2) 주식수익률 예측치를 비교할 때 귀무가설하의 모형과 대립가설하의 모형이 독립된 다른 모형이 아니라 $\beta = 0$ 이라는 제약식을 부과함으로써 귀무가설이 성립하기 때문에 내포모형이 된다.

3. 부트스트랩 절차(Bootstrap procedure)

본 연구에서는 표본 내 예측가능성과 표본 외 예측가능성 검정을 위해 부트스트랩을 통한 재표본추출을 통해 임계치와 p-value를 산출하고 통계적 추론을 실행한다. 부트스트랩 절차는 Nelson and Kim (1995)과 Killian (1997)과 유사하게 예측변수가 예측력을 갖지 못한다는 귀무가설하에서 예측변수에 AR(p) 구조를 부여하는 다음의 자료형성과정(Data-generating process)을 통해 자료가 만들어졌다고 가정한다.³⁾

$$r_t = \alpha + \gamma r_{t-1} + \epsilon_{1,t} \tag{5}$$

$$x_t = a + \rho_1 r_{t-1} + \dots + \rho_p r_{t-p} + \epsilon_{2,t} \tag{6}$$

여기서 잔차항 벡터 $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})'$ 는 공분산 매트릭스 Σ 를 갖고 독립적이며 같은 분포를 따른다(i.i.d.). 부트스트랩은 먼저 회귀식 (5)와 (6)을 OLS로 추정하고 표본 잔차항 벡터 $\{\hat{\epsilon}_t\}_{t=p+1}^T$ 를 추정하는 것으로서 시작된다.⁴⁾ 부트스트랩을 위한 유사잔차(Pseudo-disturbance)를 생성하기 위해 OLS 표본 잔차항 $\{\hat{\epsilon}_t\}_{t=p+1}^T$ 으로부터 $T+100$ 개의 잔차를 무작위 복원추출(Random drawing with replacement)하여 유사잔차 벡터 $\{\hat{\epsilon}_t^*\}_{t=1}^{T+100}$ 를 생성해낸다. 특히, 원본자료의 동시적 상관성을 보존하기 위해 잔차항 벡터 $\hat{\epsilon}_t = (\hat{\epsilon}_{1,t}, \hat{\epsilon}_{2,t})'$ 를 한꺼번에 추출한다. 이렇게 추출한 잔차항과 회귀식 (5)와 (6)의 매개변수 추정치들을 이용하여 주식수익률과 예측변수의 유사표본 $\{\hat{r}_t^*, \hat{x}_t^*\}_{t=1}^{T+100}$ 를 생성해낼 수 있다. 이 때, 초기표본인 r_1 과 (x_1, x_2, \dots, x_p) 를 0으로 설정하는 대신 생성된 유사표본 가운데 첫 번째 100개의 추정치를 제거함으로써 초기값 설정의 영향을 최소화한다. 이렇게 생성된 유사표본을 이용하여 표본 내 회귀식 (1)의 β 에 대한 t-검정치를 추정하고 표본 외 회귀식의 검정치 (3)과 (4)를 추정한다. 이 과정을 5,000번 반복함으로써 표본 내 t-검정치와 MSE-F, ENC-NEW 검정치의 실증적 분포

3) 본 연구에서 채택한 부트스트랩 절차 이외에도 Block bootstrap과 같은 다양한 다른 부트스트랩 절차를 채택할 수 있으며, 다양한 부트스트랩 절차 채택에 따른 실증분석 결과 비교는 향후 연구과제로 남긴다.

4) 예측변수의 자기상관회귀차 p는 AIC를 이용하여 정해진다.

(Empirical distribution)를 구하고 p-값을 구할 수 있다.⁵⁾

IV. 주식수익률 예측가능성에 관한 실증분석 결과

1. 자료

본 연구에서는 월별 주가수익률과 이를 누적시킨 장기수익률의 예측가능성을 검정하기 때문에 예측변수로 사용하는 금융변수들과 거시경제변수들도 월별 자료를 이용한다. 코스피(KOSPI) 수익률과 배당수익률(Dividend yield), 배당가격비율(Dividend-price ratio), 주가이익비율(Earnings-price ratio), 주가순자산비율(Book-to-market ratio), 주가변동성(Stock variance)에 관한 자료는 데이터가이드에서 제공된 자료를 이용하였으며 단기이자율, 장기이자율, 채무불이행 수익률 스프레드(Default yield spread), 금리 스프레드(Term spread), 신용 스프레드(Credit spread), 소비자물가지수, 산업생산지수 등의 자료들은 한국은행 경제통계시스템을 사용하였다.

본 연구의 실증분석 기간은 모든 자료의 월별 자료가 수집 가능한 2000년 10월 31일부터 2017년 12월 31일까지의 기간이며 예측변수로 사용된 금융 및 거시경제 변수들은 다음과 같다.

1. 단기이자율(Short-term interest rate)
2. 장기이자율(Long-term interest rate)
3. 채무불이행 수익률 스프레드(Default yield spread)
4. 금리 스프레드(Term spread)
5. 신용 스프레드(Credit spread)
6. 인플레이션(Inflation)
7. 산업생산지수 성장률(Industrial production growth)

5) 각 검정치에 대한 p-값은 원본자료로부터 추정된 검정치보다 큰 부트스트랩 추정치들의 비율로 구해질 수 있다.

8. 배당가격비율(Dividend-price ratio)
9. 배당수익률(Dividend yield)
10. 주가이익비율(Earnings-price ratio)
11. 추가순자산비율(Book-to-market ratio)
12. 추가변동성(Stock variance)

〈Table 1〉은 초과 주식수익률과 금융변수 및 거시경제변수들의 평균과 표준편차를 보여주고 있다. 주가이익비율, 추가순자산비율, 추가변동성 변수를 제외한 모든 변수들의 값은 퍼센트로 표현되었다.

〈Table 1〉 Summary Statistics for Monthly Data

Variables	Mean	Standard Deviation
Excess stock return	-2.537	6.451
Short-term interest rate	3.508	1.355
Long-term interest rate	4.086	1.487
Default yield spread	4.793	1.270
Term spread	0.579	0.680
Credit spread	0.504	0.548
Inflation	0.205	0.357
Industrial production growth	0.325	2.211
Dividend-price ratio	1.588	0.411
Dividend yield	1.604	0.413
Earnings-price ratio	0.066	0.038
Book-to-market ratio	0.921	0.152
Stock variance	0.004	0.006

초과 주식수익률은 코스피 수익률과 무위험이자율인 단기이자율의 차이로 측정되며, 배당수익률과 배당가격비율은 전년도 배당금 총액을 전월말 및 해당월말 시가총액으로 나눈 비율을 사용하였다. 주가이익비율은 전년도 당기순이익을 해당월말 시가총액으로 나눈 비율을 사용하였으며 추가순자산비율은 자본금을 제외한 총자본을 해당월말 시가총액으로 나눈 비율을 사용하였다. 추가변동성은 코스피 지수의 매월 제곱일별수익률(Squared

daily returns)의 합으로 측정된다. 단기이자율은 무위험 이자율의 대응치로 사용되는 91일물 CD금리를 사용하였으며 장기이자율은 5년물 국고채금리를 사용하였다. 채무불이행 수익률 스프레드는 3년 만기 AA- 회사채 수익률과 BBB- 회사채 수익률의 차이로, 금리 스프레드는 장·단기 이자율의 차이로, 신용 스프레드는 AA- 회사채 수익률과 장기이자율의 차이로 측정하였다. 인플레이션과 산업생산지수 성장률은 소비자물가 총지수와 산업생산지수 성장률로 측정된다.

2. 표본 내 예측가능성 검정결과(In-sample predictability)

먼저 본 연구에서는 예측회귀식 (1)을 이용하여 표본 내 회귀분석을 통해 한국 주식시장 수익률 예측가능성을 검정하였다. <Table 2>와 <Table 3>은 표본 내 예측가능성 검정 결과를 보여주고 있다. 먼저 선행연구 결과들과 비슷하게 예측회귀식의 적합도(Goodness of fit)를 나타내는 R^2 값과 예측변수의 예측력을 나타내는 $\hat{\beta}$ 의 t-검정치는 수익률이 장기로 갈수록 더 증가하는 경향을 나타내고 있다. 이는 수익률의 전체 변동성 가운데 예측가능성의 근간이 되는 기대 수익률의 변동성이 장기로 갈수록 더 증가하기 때문이라고 할 수 있다(Fama and French (1988)). 만약 부트스트랩이 아닌 전통적인 t-검정치의 임계치를 적용하면 단기이자율과 장기이자율, 채무불이행 수익률 스프레드 변수는 모든 기간에 걸쳐 유의한 결과를 나타내고 있다고 판단할 수 있다. 그러나, 부트스트랩을 이용한 표본재추출을 통해 산출한 p-value를 살펴보면 단기이자율과 장기이자율은 모두 유의한 예측력을 갖지 못한 것으로 나타난다. 그리고 선행연구 결과들과 달리 오히려 비교적 단기인 1개월 수익률과 3개월 수익률을 예측할 수 있는 예측변수들이 더 많은 것으로 나타났다. 채무불이행 수익률 스프레드와 주가순자산비율의 경우 1개월 수익률과 3개월 수익률에 대해 귀무가설을 유의수준 5%하에서 기각하는 것으로 나타났다. 특히, 주가순자산비율은 12개월 수익률에 대해서도 귀무가설을 유의수준 10%하에서 기각하는 것으로 나타났다. 이외에도 배당가격비율과 배당수익률은 3개월 수익률에 대해 유의수준 5%하에서 귀무가설을 기각하며, 주가이익비율은 장기수익률인 24개월 수익률에 대해 유의수준 5%하에서 귀무가설을 기각하는 것으로 나타났다.

따라서, 표본 내 예측가능성 검정 결과 비교적 단기수익률인 월별 수익률이나 3개월 수익률에 대해서는 채무불이행 수익률 스프레드, 주가순자산비율, 배당가격비율과 배당수익률이 예측력을 갖는 것으로 나타났으며, 장기수익률인 24개월 수익률에 대해서는 주가이익비율이 예측력을 갖는 것으로 나타났다.

(Table 2) In-sample Predictability Test Results (1)

This table presents the in-sample regression results for Eq. (1). $\hat{\beta}$ and the t-statistic is the slope coefficient estimate and the t-statistic for the given predictive variable. R^2 is the goodness-of-fit measure for Eq. (1). The p-value is obtained through the bootstrap procedure. Bold entries indicate significance at the 5% level.

Horizon	1	3	12	24
A. Short-term interest rate				
$\hat{\beta}$	-1.561	-4.433	-13.127	-21.637
t-statistic	-3.479	-3.573	-4.276	-4.416
[p-value]	[1.000]	[0.998]	[0.993]	[0.984]
R^2	0.064	0.154	0.237	0.356
B. Long-term interest rates				
$\hat{\beta}$	-1.112	-3.634	-12.900	-21.752
t-statistic	-2.521	-2.909	-3.718	-4.895
[p-value]	[0.991]	[0.987]	[0.980]	[0.981]
R^2	0.038	0.118	0.248	0.366
C. Default yield spread				
$\hat{\beta}$	1.102	2.955	9.801	11.903
t-statistic	2.480	2.301	2.586	2.361
[p-value]	[0.009]	[0.031]	[0.061]	[0.119]
R^2	0.037	0.086	0.148	0.137
D. Term spread				
$\hat{\beta}$	0.497	0.315	-3.630	-3.843
t-statistic	1.112	0.215	-0.609	-0.584
[p-value]	[0.116]	[0.379]	[0.618]	[0.587]
R^2	0.014	0.026	0.022	0.017
E. Credit spread				
$\hat{\beta}$	-0.085	1.038	4.890	7.014

t-statistic	-0.191	0.773	1.607	1.211
[p-value]	[0.602]	[0.280]	[0.138]	[0.246]
R^2	0.008	0.033	0.041	0.057
F. Inflation				
$\hat{\beta}$	-0.638	-1.443	-3.406	-3.590
t-statistic	-1.457	-1.494	-1.810	-1.555
[p-value]	[0.925]	[0.923]	[0.958]	[0.883]
R^2	0.018	0.040	0.019	0.015

〈Table 3〉 In-sample Predictability Test Results (2)

This table presents the in-sample regression results for Eq. (1). $\hat{\beta}$ and the t-statistic is the slope coefficient estimate and the t-statistic for the given predictive variable. R^2 is the goodness-of-fit measure for Eq. (1). The p-value is obtained through the bootstrap procedure. Bold entries indicate significance at the 5% level.

Horizon	1	3	12	24
G. Industrial Production Growth				
$\hat{\beta}$	0.514	0.273	-2.593	-2.539
t-statistic	1.164	0.344	-1.597	-1.739
[p-value]	[0.134]	[0.368]	[0.913]	[0.914]
R^2	0.015	0.025	0.011	0.008
H. Dividend-price ratio				
$\hat{\beta}$	0.652	2.410	6.493	2.470
t-statistic	1.455	2.466	1.961	0.359
[p-value]	[0.127]	[0.044]	[0.175]	[0.588]
R^2	0.018		0.069	0.008
I. Dividend yield				
$\hat{\beta}$	0.562	1.970	5.353	1.292
t-statistic	1.282	2.068	1.714	0.191
[p-value]	[0.112]	[0.048]	[0.113]	[0.461]
R^2	0.016	0.053	0.049	0.003
J. Earnings-price ratio				
$\hat{\beta}$	0.323	0.924	7.264	12.848
t-statistic	0.733	0.617	2.383	3.489
[p-value]	[0.253]	[0.331]	[0.064]	[0.040]
R^2	0.011	0.031	0.091	0.192
K. Book-to-market ratio				

$\hat{\beta}$	1.405	4.501	11.874	9.606
t-statistic	3.216	3.391	2.934	1.243
[p-value]	[0.001]	[0.009]	[0.075]	[0.392]
R^2	0.056	0.169	0.232	0.098
L. Stock variance				
$\hat{\beta}$	0.241	1.161	-0.718	0.194
t-statistic	0.506	0.905	-0.127	0.022
[p-value]	[0.327]	[0.250]	[0.581]	[0.564]
R^2	0.009	0.033	0.001	0.001

3. 표본 외 예측가능성 검정결과(Out-of-sample predictability)

전체 표본기간 수는 207개이며 본 연구에서는 표본 외 예측가능성 검정을 위한 첫 번째 기간 $R = 120$ 으로 설정한다.⁶⁾ <Table 4>와 <Table 5>는 한국 주식시장 수익률 예측가능성에 대한 표본 외 검정결과를 나타내고 있다. 장·단기 수익률에 대해 가장 예측력을 많이 갖는 변수들은 주가순자산비율과 인플레이션이다. 이 두 변수들은 각각 1개월, 3개월, 12개월 혹은 24개월 수익률에 대해 MSE-F 검정치와 ENC-NEW 검정치 모두 유의수준 5% 이내에서 귀무가설을 기각하고 있다. 특히 주가순자산비율의 경우 대립가설과 귀무가설하의 Theil's U-비율이 모든 기간의 수익률에 대해 1보다 작은 값을 갖고 있으며, 24개월 수익률에 대해서도 MSE-F 검정치의 경우 유의수준 10% 이내에서는 귀무가설을 기각하고 있다.

산업생산지수 성장률은 장기수익률인 24개월 수익률에 대해 Theil's U-비율이 1보다 작고 MSE-F 검정치가 유의수준 5% 이내에서 유의한 결과를 나타내고 있다. ENC-NEW 검정치는 유의수준 5% 이내에서 유의하지는 않지만 유의수준을 10%로 높일 경우 유의해지기 때문에 장기수익률에 대해 예측력을 갖고 있다고 판단할 수 있다.

금융변수인 단기이자율과 장기이자율은 ENC-NEW 검정치에 대해 모든 기간 수익률에 대한 귀무가설을 유의수준 5% 이내에서 기각하고 있다. 그러나, 대부분의 경우 Theil's U-비율이 1을 초과하고 있고 MSE-F 검정치가 모두 유의하지 않다. 채무불이행 수익률 스프레드 또한 단기인 1개월 수익률과 3개월 수익률에 대해 ENC-NEW 검정치가 유의수준

6) 첫 번째 표본기간인 R 을 110 혹은 130으로 설정해도 실증분석 결과의 유의성은 달라지지 않는다.

5% 이내에서 유의하지만 Theil's U-비율이 1을 초과하고 MSE-F 검정치가 유의하지 않기 때문에 추가수익률 예측력이 높다고 판단하기 어렵다.

따라서, 표본 외 예측가능성 검정결과 추가순자산비율과 인플레이션이 장·단기에 걸쳐 추가수익률에 대한 예측력을 갖고 있다고 판단되며, 특히 추가순자산비율은 모든 검정치에 대해 가장 유의한 표본 외 검정결과를 나타내고 있다.

〈Table 4〉 Out-of-sample Predictability Test Results (1)

This table presents the out-of-sample regression results for Eq. (1). U ratio is the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the restricted model. MSE-F and ENC-NEW are the out-of-sample statistics given in Eqs. (3) and (4). Bold entries indicate significance at the 5% level.

Horizon	1	3	12	24
A. Short-term interest rate				
U ratio	1.042	1.087	0.950	0.899
MSE-F	-6.911	-13.025	8.257	15.153
[p-value]	[0.995]	[0.979]	[0.110]	[0.113]
ENC-NEW	12.622	31.854	60.230	66.460
[p-value]	[0.000]	[0.000]	[0.004]	[0.015]
B. Long-term interest rates				
U ratio	1.043	1.285	1.538	1.337
MSE-F	-7.050	-33.530	-43.876	-28.207
[p-value]	[0.995]	[1.000]	[0.992]	[0.881]
ENC-NEW	8.266	17.044	27.246	43.703
[p-value]	[0.000]	[0.002]	[0.029]	[0.031]
C. Default yield spread				
U ratio	1.011	1.021	1.210	1.061
MSE-F	-1.967	-3.392	-24.124	-7.095
[p-value]	[0.825]	[0.644]	[0.873]	[0.506]
ENC-NEW	4.981	11.846	9.906	13.490
[p-value]	[0.004]	[0.008]	[0.122]	[0.164]
D. Term spread				
U ratio	1.029	1.016	0.948	0.983
MSE-F	-4.790	-2.584	80520	2.186
[p-value]	[0.986]	[0.686]	[0.088]	[0.235]
ENC-NEW	-1.173	-0.719	4.738	1.538
[p-value]	[0.967]	[0.665]	[0.157]	[0.314]
E. Credit spread				
U ratio	1.001	1.047	1.154	1.272
MSE-F	-0.114	-7.484	-18.971	-24.417
[p-value]	[0.268]	[0.942]	[0.944]	[0.941]

ENC-NEW	-0.056	-3.068	-7.185	-10.193
[p-value]	[0.405]	[0.972]	[0.970]	[0.987]
F. Inflation				
U ratio	0.986	1.038	0.960	0.976
MSE-F	2.491	-6.119	6.382	3.139
[p-value]	[0.022]	[0.991]	[0.001]	[0.015]
ENC-NEW	2.472	-0.718	4.116	1.670
[p-value]	[0.016]	[0.878]	[0.003]	[0.041]

〈Table 5〉 Out-of-sample Predictability Test Results (2)

This table presents the out-of-sample regression results for Eq. (1). U ratio is the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the restricted model. MSE-F and ENC-NEW are the out-of-sample statistics given in Eqs. (3) and (4). Bold entries indicate significance at the 5% level.

Horizon	1	3	12	24
G. Industrial Production Growth				
U ratio	1.016	1.015	0.996	0.988
MSE-F	-2.655	-2.482	0.649	1.581
[p-value]	[0.948]	[0.928]	[0.130]	[0.042]
ENC-NEW	-0.485	-1.015	1.021	0.937
[p-value]	[0.844]	[0.958]	[0.076]	[0.075]
H. Dividend-price ratio				
U ratio	1.052	1.216	1.502	1.583
MSE-F	-8.451	-27.475	-42.298	-38.450
[p-value]	[0.998]	[1.000]	[0.997]	[0.983]
ENC-NEW	-1.248	-2.725	-3.620	-9.328
[p-value]	[0.967]	[0.942]	[0.818]	[0.963]
I. Dividend yield				
U ratio	1.050	1.185	1.462	1.536
MSE-F	-8.069	-24.502	-40.466	-36.877
[p-value]	[0.999]	[1.000]	[0.993]	[0.977]
ENC-NEW	-1.315	-2.802	-3.705	-8.818
[p-value]	[0.978]	[0.962]	[0.833]	[0.961]
J. Earnings-price ratio				
U ratio	1.000	0.998	1.039	1.050
MSE-F	-0.002	0.331	-5.572	-5.967
[p-value]	[0.232]	[0.213]	[0.635]	[0.627]
ENC-NEW	0.044	0.296	-0.952	-0.528
[p-value]	[0.340]	[0.313]	[0.556]	[0.528]

K. Book-to-market ratio				
U ratio	0.956	0.879	0.663	0.884
MSE-F	8.193	24.950	96.923	17.864
[p-value]	[0.000]	[0.000]	[0.000]	[0.057]
ENC-NEW	6.179	21.574	65.283	10.011
[p-value]	[0.000]	[0.000]	[0.000]	[0.110]
L. Stock variance				
U ratio	1.012	1.063	1.021	1.079
MSE-F	-1.972	-9.757	-3.111	-9.061
[p-value]	[0.865]	[0.979]	[0.672]	[0.866]
ENC-NEW	-0.755	-3.547	-1.372	-4.054
[p-value]	[0.907]	[0.993]	[0.777]	[0.938]

V. 결론

본 연구에서는 12개의 주요 금융변수와 거시경제변수들을 이용하여 한국 주식시장 (KOSPI)의 수익률 예측가능성을 검정하였다. 월별 자료를 바탕으로 1개월, 3개월, 12개월과 24개월의 기간 수익률을 예측하는 상황을 상정하여 각 변수들의 장·단기 주식수익률 예측력을 검정하였다. 특히, 선행연구에서 많이 살펴보았던 표본 내 예측가능성 검정에 있어서 일어날 수 있는 문제인 소표본 편차와 자기상관성 문제를 최소화하기 위해 부트스트랩을 통한 재표본추출을 통해 임계치와 p-value를 구하여 통계적 추론을 시행하였으며, 최근에 제안된 통계 검정기법인 MSE-F 검정치와 ENC-NEW 검정치를 사용하여 표본 외 예측가능성을 보다 강건하게 검정하였다.

표본 내 예측가능성 검정 결과 단기수익률에 대해서 채무불이행 수익률 스프레드와 주가순자산비율, 배당가격비율, 배당수익률이 유의한 주가수익률 예측력을 갖는 것으로 나타났다. 장기수익률에 대해서 주가이익비율이 예측력을 갖는 것으로 나타났다. 표본 외 예측가능성 검정 결과 주가순자산비율과 인플레이션이 장·단기 주식수익률 모두에 대해 예측력을 갖고 있는 것으로 나타났다. 특히, 주가순자산비율 변수는 표본 내 예측력과 표본 외 예측력을 모두 갖고 있는 것으로 나타났으며 표본 외 예측력을 검정하는 모든 검정치에 대해 가장 유의한 결과를 보였다는 점에서 예측변수로서 가장 일관성 있는 특성을 보이고 있다고 판단된다.

한국 주식시장 수익률 예측가능성에 관한 문헌에서 본 연구의 기여점은 다양한 금융변수와 거시경제변수들의 예측력을 함께 살펴보았다는 점이다. 그러나, 이렇게 여러 가지 변수들을 동시에 고려할 때 Foster et al. (1997)과 Ferson et al. (2003)이 지적하였듯이 통계적 추론 결과에 데이터 마이닝 문제가 발생할 여지가 있다. 데이터 마이닝 문제는 많은 수의 예측변수들의 검증 결과를 살펴볼 때 오직 통계적으로 유의성을 갖는 변수만을 반복적으로 보고함으로써 발생하는데, 특히 비슷하거나 중복된 표본기간을 사용하는 표본 내 검정에서 심화되는 것으로 알려져 있다. 본 연구에서는 내포모형에 사용될 수 있는 표본 외 검정기법을 함께 적용함으로써 이 문제를 최소화하기 위해 노력하였으나, 데이터 마이닝 문제를 직접적으로 해결할 수 있는 통계적 기법을 적용하지 못했다는 점에서 연구결과 해석에 주의할 필요가 있다. 데이터 마이닝 문제를 직접 해결할 수 있는 통계기법으로는 Foster et al. (1997)에서 살펴보았듯이 Maximum statistics를 이용할 수 있으며, 이에 대한 연구는 향후의 연구과제로 남긴다.

참고문헌

- 김규영·김영빈 (2004), “한국 주식시장에서 수익률의 예측에 관한 연구”, **산업경제연구**, 제17권 4호, pp. 1255-1271.
- (Translated in English) Kim, K., and Y., Kim (2004). “Testing the Predictability of Stock Return in the Korean Stock Market”, *Journal of Industrial Economics and Business*, 17(4): 1255-1271.
- 김인무·박성근 (2009), “주식수익률의 예측 가능성과 변동성 시계 표본”, **경제학연구**, 제57집 제3호, pp. 195-221.
- (Translated in English) Kim, I., and S., Park (2009). “The Predictability of Korean Stock Returns and Volatility Clock Samples”, *Korean Economic Review*, 57(3):195-221.
- 김현석·여효성 (2017), “금융변수의 주가수익률 예측력 검증”, **금융연구**, 제31권 제1호, pp. 39-73.
- (Translated in English) Kim, H., and H., Yeo (2017). “An Empirical Test of Stock Return Predictability in Korean Stock Market”, *Journal of Money and Finance*, 31(1):39-73.
- 정동준 (2012), “한국 주식수익률의 장기 예측가능성과 분산분해분석에 관한 실증 연구”, **금융연구**, 제26권 제4호, pp. 61-91.
- (Translated in English) Jeong, D. (2012). “Empirical Study on the Long-run Predictability and the Variance Decomposition of Korean Stock Returns”, *Journal of Money and Finance*, 26(4):61-91.
- 정정현·김수경 (2010), “배당수익률과 E/P 비율의 주식수익률 예측가능성에 관한 연구”, **금융공학연구**, 제9권 제3호, pp. 61-87.
- (Translated in English) Chung, C., and S., Kim (2010). “An Investigation on the Stock Return Predictability of Dividend Yield and Earnings-Price Ratio”, *Study on Financial Engineering*, 9(3):61-87.

- Breen, W., L., Glosten and R., Jagannathan (1989). "Economic significance of predictable variation in stock index returns", *Journal of Finance*, 44:1177-1189.
- Campbell, J. (1987). "Stock returns and the term structure", *Journal of Financial Economics*, 18:373-399.
- Campbell, J., and J., Cochrane (1999). "By force of habit: a consumption-based explanation of aggregate stock market behavior", *Journal of Political Economy*, 107:205-251.
- Campbell, J., and R., Shiller (1988a). "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies*, 1:195-228.
- _____ (1988b). "Stock prices, earnings, and expected dividends", *Journal of Finance*, 43:661-676.
- _____ (1998). "Valuation ratios and the long-run stock market outlook", *Journal of Portfolio Management*, 24:11-26.
- Campbell, J., and T., Vuolteenaho (2004). "Inflation illusion and stock prices", *American Economics Review*, 94:19-23.
- Clark, T., and M., McCracken (2001). "Tests of equal forecast accuracy and encompassing for nested models", *Journal of Econometrics*, 105:85-110.
- _____ (2002). "Evaluating long-horizon forecasts", *FRB of Kansas City Research Working Paper*, No.01-14:1-35.
- Charles, A., O., Darné and J., Kim (2017). "International stock return predictability: Evidence from new statistical tests", *International Review of Financial Analysis*, 54:97-113.
- Cochrane, J. (2007). *Financial markets and the real economy*, In: Mehra, R. (Ed.) *Handbook of the Equity Premium*, Elsevier, Amsterdam.

- _____ (2011). "Presidential address: discount rates", *Journal of Finance*, 66:1047-1108.
- Cowles, A. (1933). "Can stock market forecasters forecast?", *Econometrica*, 1:309-324.
- Diebold, F., and R., Mariano (1995). "Comparing predictive accuracy", *Journal of Economic and Business Statistics*, 13:253-263.
- Fama, E. (1970). "Efficient capital markets: a review of theory and empirical work", *Journal of Finance*, 25:383-417.
- Fama, E., and K., French (1988). "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 22:3-25.
- _____ (1989). "Business conditions and expected returns on stocks and bonds", *Journal of Financial Economics*, 25:23-49.
- Fama, E., and G., Schwert (1977). "Asset returns and inflation", *Journal of Financial Economics*, 5:115-146.
- Ferson, W., S., Sarkissian and T., Simin (2003). "Spurious regressions in financial economics?", *Journal of Finance*, 58:1393-1413.
- Foster, F., T., Smith, and R., Whaley (1997). "Assessing goodness-of-fit of asset pricing models: the distribution of the maximum R²", *Journal of Finance*, 52:591-607.
- Friedman, S. (1953). "The case for flexible exchange rate", *In: Essays in Positive Economics*, Chicago, IL: University of Chicago Press:157-203.
- Hjarlmarsson, E. (2010). "Predicting global stock returns", *Journal of Financial and Quantitative Analysis*, 45:49-80.
- Jordan, S., A., Vivian, and M.. Wohar (2014). "Forecasting returns: new european evidence", *Journal of Empirical Finance*, 26:76-95.
- Kilian, L. (1999). "Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions?", *Journal of Applied Econometrics*,

14:491-510.

- Kothari, S., and J., Shanken (1997). "Book-to-market, dividend yield, and expected market returns: what do the data tell us?", *Review of Financial Studies*, 9:1033-1059.
- Lettau, M., and S., Ludvigson (2001). "Consumption, aggregate wealth, and expected stock returns", *Journal of Finance*, 56:815-849.
- McCracken, M. (2007). "Asymptotics for out-of-sample tests of Granger causality", *Journal of Econometrics*, 140:719-752.
- Nelson, C. (1976). "Inflation and the rates of return on common stock", *Journal of Finance*, 31:471-483.
- Nelson, C., and M., Kim (1993). "Predictable stock returns: the role of small sample bias", *Journal of Finance*, 48:641-661.
- Newey, W., and K., West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, 55:703-708.
- Pontiff, J., and L., Schall (1998). "Book-to-market ratios as predictors of market returns", *Journal of Financial Economics*, 49:141-160.
- Rozeff, M. (1984). "Dividend yields are equity risk premiums", *Journal of Portfolio Management*, 11:68-75.
- Samuelson, P. (1965). "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review*, 6(2):41-49.
- Schrimpf, A. (2010). "International stock return predictability under model uncertainty", *Journal of International Money and Finance*, 29: 1256-1282.
- Stambaugh, R. (1999). "Predictive regressions", *Journal of Financial Economics*, 54:375-421.
- West, K. (1996). "Asymptotic inference about predictive ability", *Econometrica*,

64:1067-1084.

Wohar, M., D., Rapach and J., Ranvid (2005). "Macro variables and international stock return predictability", *International Journal of Forecasting*, 21:137-166.

Abstract

This study evaluates the predictive power of 12 financial and macroeconomic variables for Korean stock market returns of different horizons. Both the return predictability of in-sample and out-of-sample tests are considered to examine each variable's predictive ability more robustly. For this purpose, this article employs the MSE-F statistic developed by McCracken (2007) and the ENC-NEW statistic developed by Clark and McCracken (2001) to compare nested forecast models. In addition, the bootstrapping procedure is applied for both in-sample and out-of-sample inferences to address the finite-sample bias and the autocorrelated disturbances from overlapping observations. As a result, the book-to-market ratio variable is found to be the most consistent and significant predictor as it rejects the null of no predictability for both in-sample and out-of-sample tests.

※ Key words: Stock Market Return predictability, Macroeconomic variables, Financial variables, In-sample predictability, Out-of-sample predictability