

N년 후 당신에게 투자하세요: 배당투자로 연금 마련하기

발표자 박인서

초저금리 시대, 깜깜한 주식시장
당신과 함께 성장할 배당주를 찾는 방법

Keywords: 배당주, 생존분석, 다층모형

1. 왜 주식 투자를 해야 할까?

- 차원축소 기법을 활용한 금융지표 요약

2. 왜 배당주일까?

- 한국 기업의 배당현황



3. 배당주를 고르는 방법

[1단계] 생존 분석을 활용한 배당금 지급중지 설명

[2단계] 다층 모형을 활용한 연도별 배당변화 패턴 분석

[결론] 배당투자 유망종목

1. 왜 주식 투자를 해야 할까?

분석 목적 및 배경

세상은 두 부류의 사람이 있다.
주식을 하는 사람,
그리고 주식을 하지 않는 사람

금융 시장은 참여자와 비참여자의 구분이 명확한 영역
그러나 저금리 시대에 접어들면서 주식 투자의 필요성 증대
주식을 하지 않는 사람들을 위한, 쉽고도 확실한 제안

분석 목적

Who?

주식을 하지 **않는** 사람

Why?

저금리 · 저성장 시대, 주식 투자는 미래를 대비하는
전략이 될 수 있으므로

How?

함께 성장할 **배당주**를 찾아 투자한다!

1. 왜 주식 투자를 해야 할까?

분석 목적 및 배경

누가 주식 투자를 하는가?

비참여자의 관점에서 왜 시장에서 참여하지 않는가에 주목
⇒ **불확실성**에 대한 태도 차이로부터 발생

근본적으로 주식 투자에 대해 **위험 >> 기회**로 인식
손실을 입는 대상이 자신이 될 수 있음을 인지한 순간 더 멀어질 것
결국 시장 참여를 위해 선결되어야 할 문제는 무엇인가?

해결 방안

- 1) 시장이 점점 더 좋아져야 한다.
 - 2) 투자한 기업이 성장해야 한다.
- ⇒ 후자가 더 현실적이고 장기적인 방향성

- Q. 내가 얻은 수익이 누군가의 손실로부터 비롯된 것은 아닌가?



- 투자자와 기업이 **함께 성장**할 수 있는 방법 모색

1. 왜 주식 투자를 해야 할까?

분석 목적 및 배경

잠깐, 주식 투자를 꼭 해야 할까?

물론 아니다. 본인의 선택이니까.

다만 명심해야 할 것은 **금리가 낮고, 앞으로도 낮아질 것**이라는 점이다.

저성장 시대, 투자에서 고려해야 할 것은 무엇인가?

정책금리마저 1%대, 은행에 예금을 하면 그만이었던 과거와는 다르다.

주식 투자를 꼭 해야 하는 것은 아니지만, 분명 **미래에 대한 대비책**이 될 수 있을 것



Figure 1.1
연도별 국채 금리

1. 왜 주식 투자를 해야 할까?

차원 축소 기법을 활용한 금융지표 요약

저금리 기조에 따른 경제 흐름 읽기

금융시장에서 가장 두드러진 현상 중 하나는 **공동변화(co-movement)** 즉 주가, 금리, 환율 등 다양한 금융지표들이 서로 영향을 주고받으며 함께 움직임
⇒ 하나 이상의 변수들을 설명하는 **기저 차원을 가정**

주성분 분석 (Principal Component Analysis)

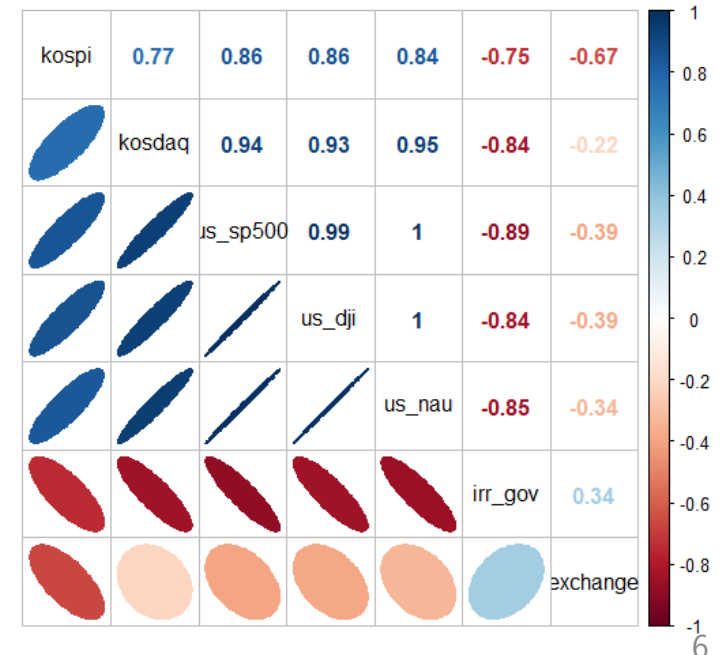
: 상관관계에 기반을 두어, 변수들의 변동을 설명할 수 있는 새로운 축을 상정

네이버 금융 최상단에 노출되는 지표들 대상: **코스피, 코스닥, S&P 500 지수, 다우지수, 나스닥 종합지수, 국채이자율, 미국 달러 환율**

강한 다중공선성이 발견되어, VIF(Variance Inflation Factor) 기준으로 S&P 500 지수, 나스닥 종합지수 제외; 총 6개 변수

주성분의 수는 카이저 규칙, 스크리플롯과 평행분석을 통해 **2개**로 선정
이로써 2개의 차원으로 기존 변동의 **90%** 이상을 설명

Figure 1.2
주요 금융지표들 간 상관계수



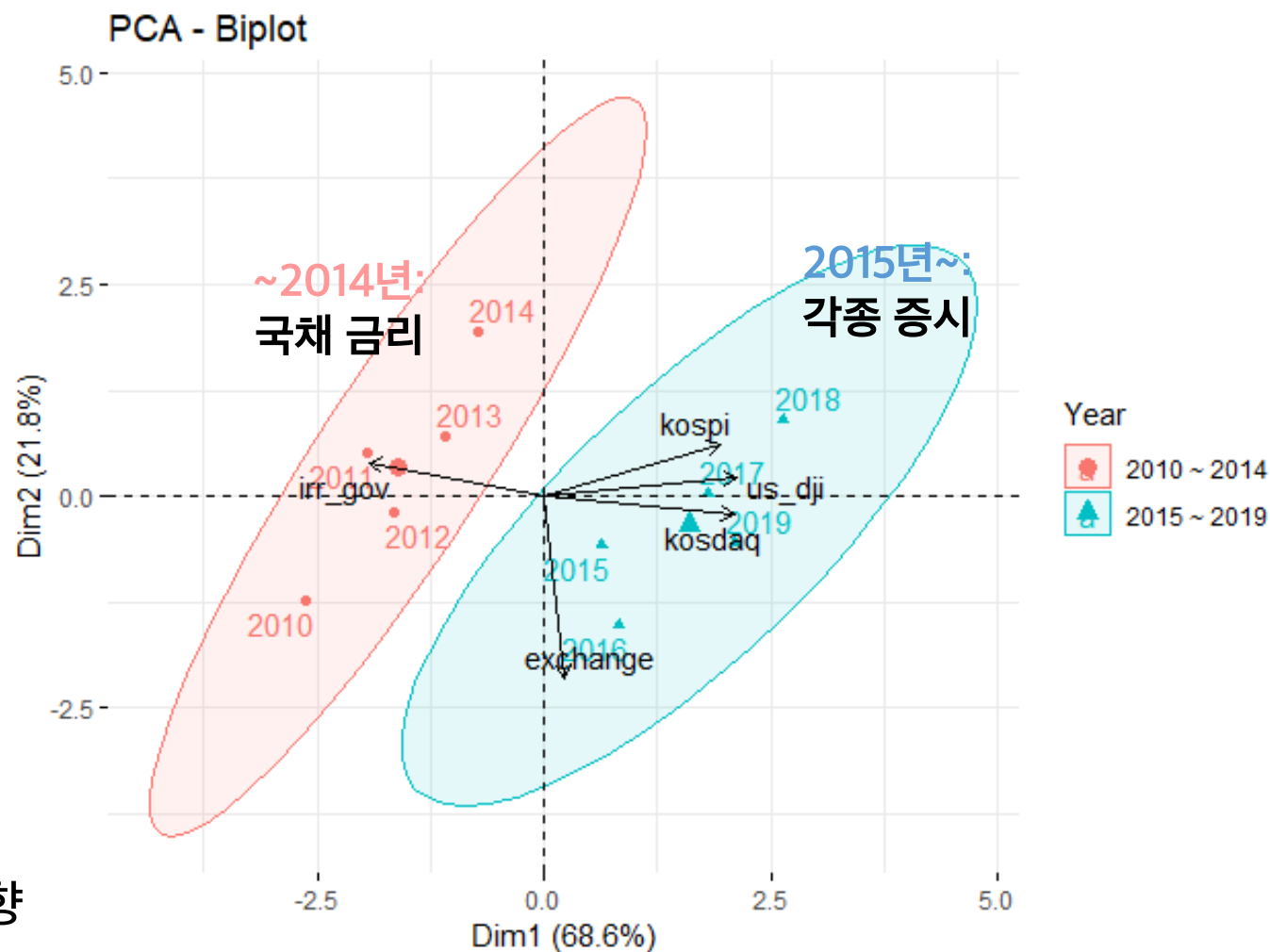
1. 왜 주식 투자를 해야 할까?

차원 축소 기법을 활용한 금융지표 요약

Figure 1.3
주성분 분석 결과 (Biplot)

PC 1	PC 2
+코스피	코스피
+코스닥	환율
+다우지수	
-국채 금리	

- (1) 국채 금리와 증시의 방향성은 **반대**
- (2) 2015년 이후 시장 지표들은 **증시** 성향



2. 왜 배당주인가?

신호 이론으로 본 배당주의 강점

배당주 (Dividend Stock)

쉽게 말해, 배당금(기업이 주주에게 나누어 주는 돈)을 주는 주식
국내 기업은 일반적으로 연 1회 현금 배당하며, 12월 말 배당금의 규모를 발표한 후 4월경 지급
주주가 직접 주식을 팔지 않더라도, 기준일까지 주식을 보유한 만큼 **현금 배당**

배당은 신호다

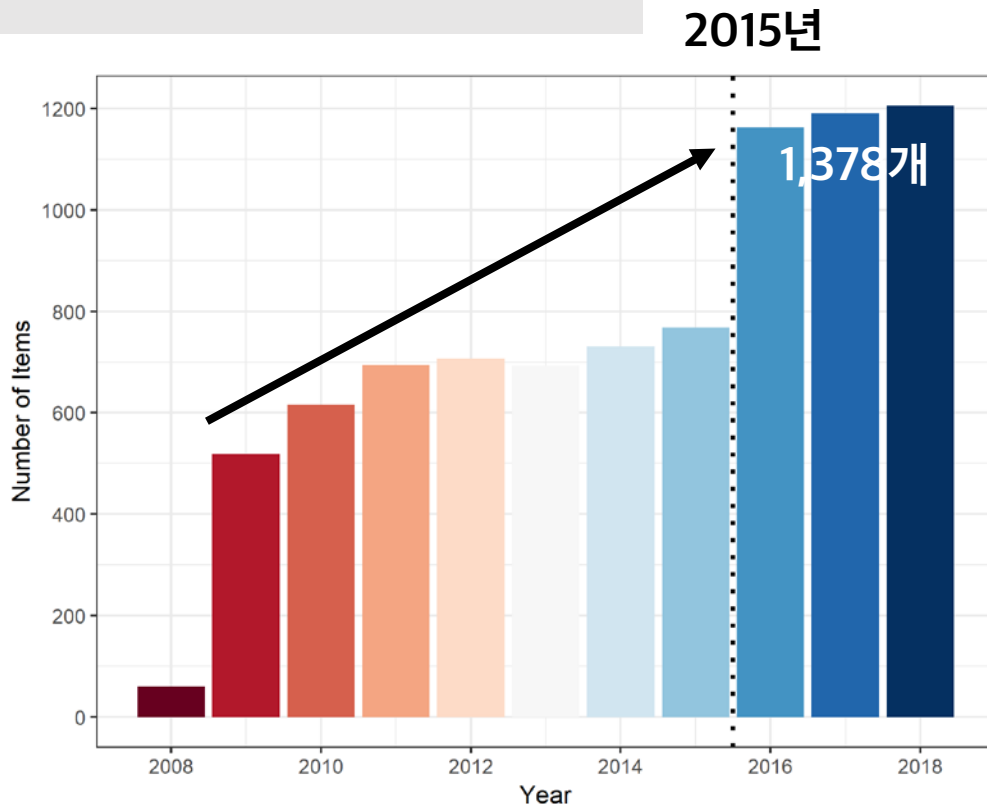
현금 배당은 기업이 실제로 시장에 푸는 돈이라는 점에 주목
불확실한 금융 시장에서 기업이 배당금을 주는지 여부와 배당금의 규모는 확실한 신호가 됨
즉 신호 이론의 측면에서 보면 기업이 건실하고 안정적이라는 증거로 삼을 만함

■ 위험 기피 성향이 강한 **시장 비참여자들에게 적합**

2. 왜 배당주인가?

한국 기업의 배당 현황

Figure 2.1
연도별 배당기업의 수 변화



- 배당 규모 증가

배당을 하는 기업 수도, 배당금 규모도 증가
KOSPI 14년' 15조 원(483개) → 18년' 30조 원(546개)

- 정부의 친배당 정책

한국 주식시장의 배당성향은 원래 전 세계 평균보다 낮은 편
2015년 배당소득 증대세제, 기업소득 환류세제 등 정책으로
투자환경이 점차 개선

- 국민연금의 배당압박

자금력이 큰 연기금 투자자들의 배당주 선호
국민연금에서도 배당주를 산다, 그것도 아주 많이!
국내주식 투자액 107조 원 중 91조 원 (*2018.12 기준)
⇒ 전체 84.2%에 달하며, 계속적으로 배당 확대 압박

3. 배당주를 고르는 방법

[1단계] 생존 분석을 활용한 배당금 지급중지 설명



안정적으로 배당금을 지급할 수 있는 기업

배당투자자에 있어서 판단의 기준이자, 제 1의 목적은 배당금을 받는 것
그러나 시장의 불확실성에 의해 배당금을 지급하지 못하는 사건이 발생한다면?



Figure 3.1
미래에셋대우
(#006800)

3. 배당주를 고르는 방법

[1단계] 생존 분석을 활용한 배당금 지급중지 설명

데이터 준비

2016~2018년 사이 배당을 지급한 이력이 있는 **1,378개** 코스피 및 코스닥 종목을 대상으로 2009~2018년까지의 **일별 시세, 매매동향** 데이터를 구축
tidyquant 패키지를 사용하여 Yahoo Finance로부터 배당 관련 정보 수집

코스피로 상장된 18개 종목을 포함하여 기존 4,077개에 5,314개 행 추가
최초 관측 시점은 2007년이며, 지급 횟수는 최소 1회에서 최대 10회

Cox 비례위험모형 (Cox Proportional Hazards model)

관심 변수는 첫 배당금 지급 이후 중지까지의 시간
관측 종료시까지 사건이 발생하지 않는 중도절단(censoring) 사례 포함
⇒ 대표적인 **생존분석** 기법 중 하나인 **Cox 비례위험모형** 이용

변수들의 측정 단위는 1년으로, 관측 기간 내내 동일한 것이 아니라 매년 다른 값
비례위험가정에 위배되므로 **시변 공변량**(time-varying covariates) 투입

Table 3.1

Cox 비례위험모형 투입 데이터 형식

*매 시점에서 지급중지가 발생하면 death=1, 아니면 0

itemcode	itemname	tstart	tstop	death	der_dps_lag1
000020	동화약품	2009	2010	0	0
000020	동화약품	2010	2011	0	100
000020	동화약품	2011	2012	0	100
000020	동화약품	2012	2013	0	100

*동화약품(2009년 배당 시작)의 경우 2010년 배당횟수는 1

*주가변동성은 변동계수(CV) = 표준편차/평균으로 계산

`coxph(Surv(tstart, tstop, death) ~
작년배당금+(이전)배당횟수+주가변동성
+작년배당금*주가변동성)`

3. 배당주를 고르는 방법

[1단계] 생존 분석을 활용한 배당금 지급중지 설명

Figure 3.2

Cox 비례위험모형 추정 결과

(1단위 증가에 따른 위험 증가율)

작년배당금

-18%

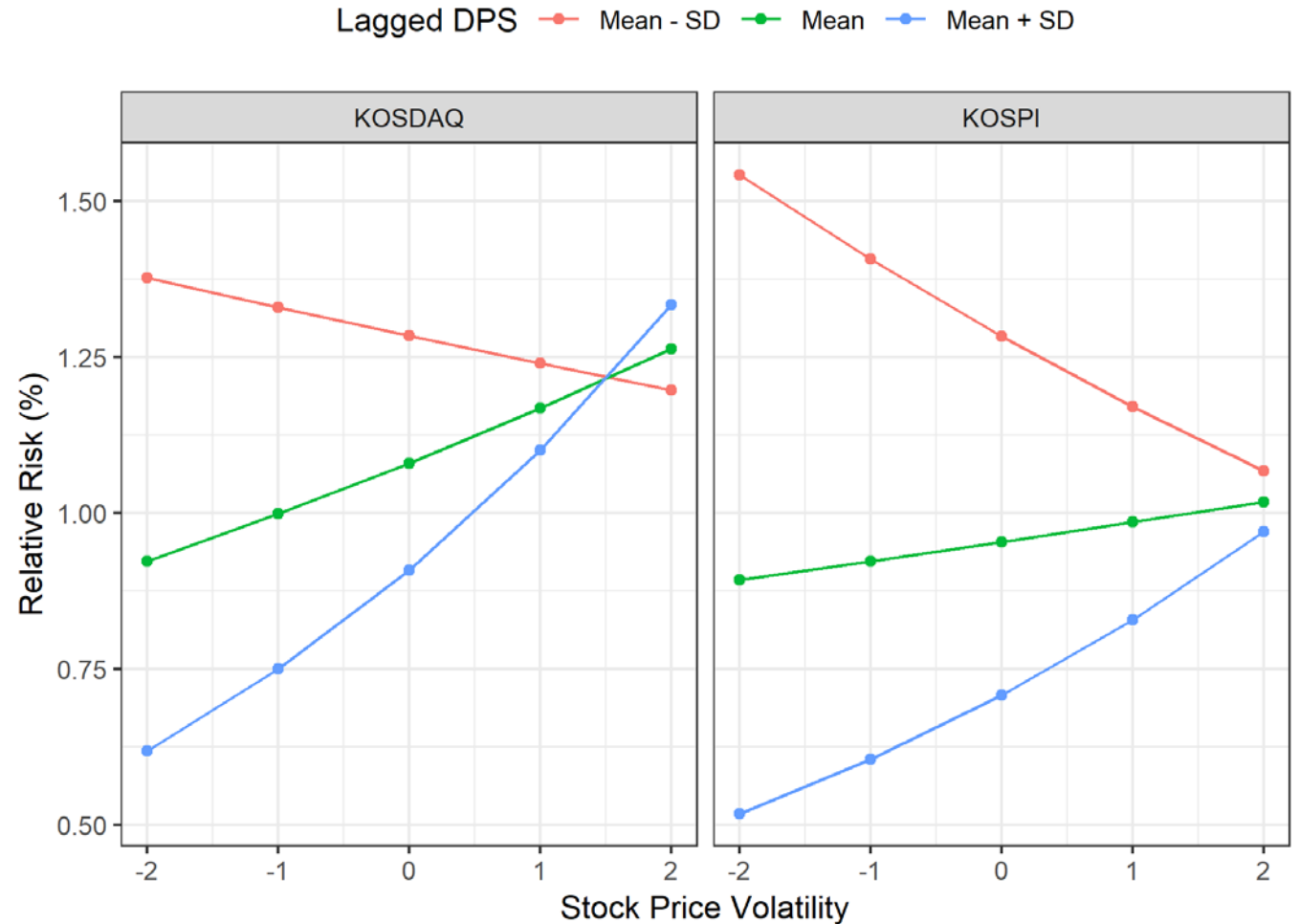
배당횟수

-15%

주가변동성

+13%

- (1) 배당금 규모가 **크면** 주가변동이 **악재**,
작으면 호재
- (2) 양쪽 기울기를 비교하니 코스피에서 **위험 증가율은 높고, 감소율은 낮음**
- (3) 코스닥 << 코스피 시장이 안정적



3. 배당주를 고르는 방법

[1단계] 생존 분석을 활용한 배당금 지급중지 설명

Figure 3.3
LTRCtrees 추정 결과

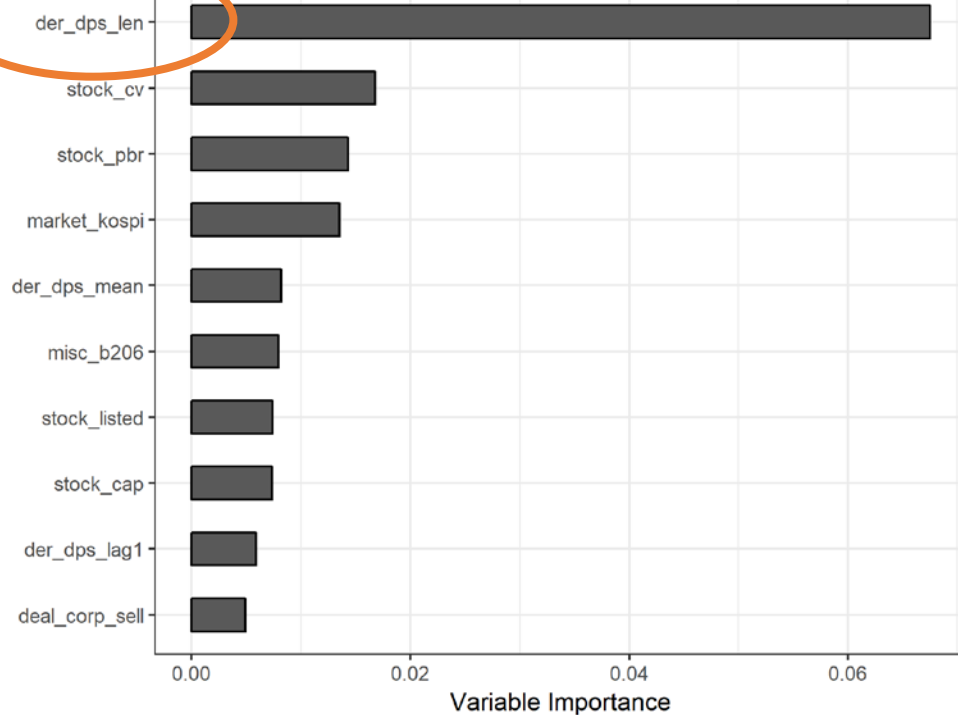
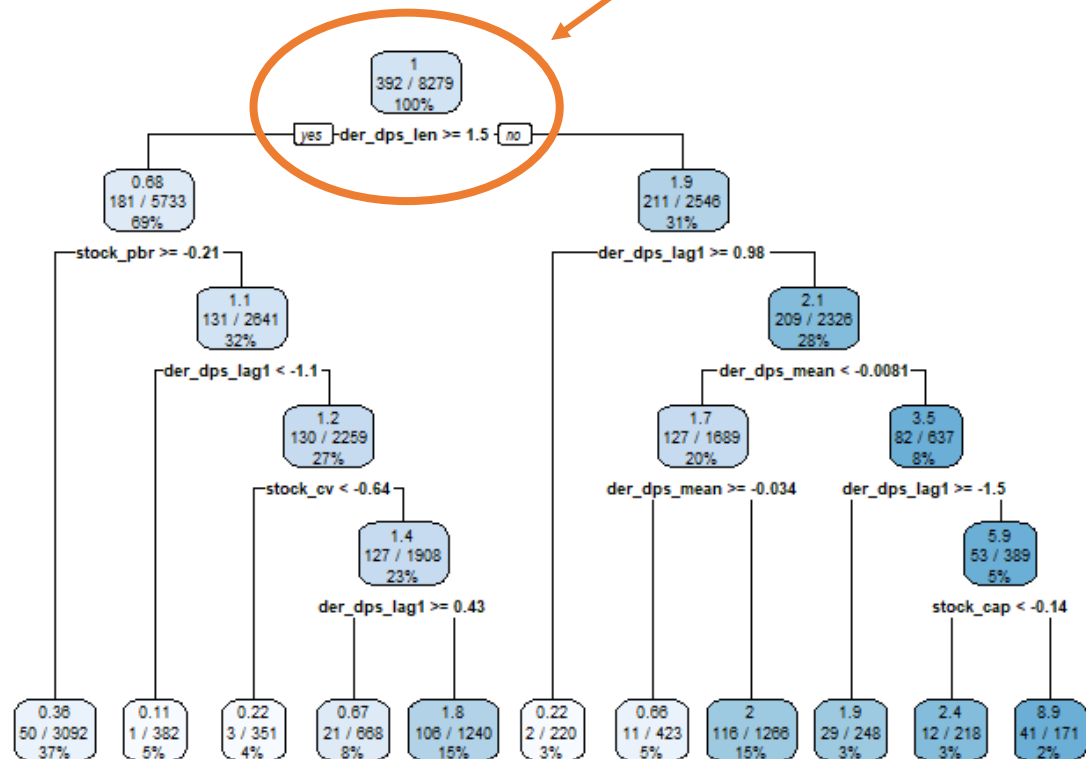


Figure 3.4
랜덤 포레스트-생존 모형 추정 결과

3. 배당주를 고르는 방법

[2단계] 다층 모형을 활용한 연도별 배당변화 패턴 분석



지속적으로 성장해왔고, 앞으로도 성장할 기업

안정성이라는 hurdles을 넘었다면, 이제 수익성을 줄 수 있는 배당주를 고를 차례
주당배당금과 시가배당률 모두를 고려하여 성장세인 종목을 선정

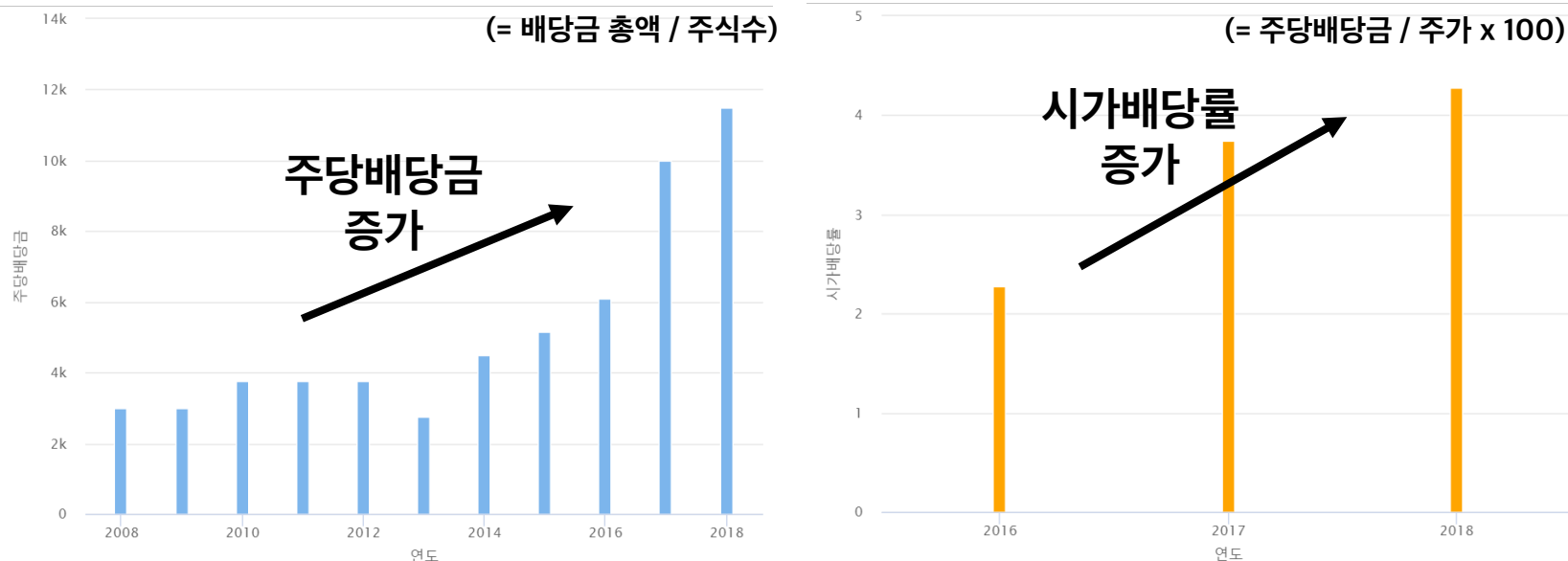


Figure 3.5
삼성화재
(#001140)

3. 배당주를 고르는 방법

[2단계] 다층 모형을 활용한 연도별 배당변화 패턴 분석

배당변화 패턴 분석

2016~2018년 **배당 이력**과 **기업실적**을 이용해 배당금, 시가배당률 추이 분석

1) 배당금: **선형** 회귀를 적합한 뒤 기울기가 양수

2-1) 시가배당률: **선형** 회귀를 적합한 뒤 기울기가 양수

2-2) 이차 회귀를 적합한 뒤 18년' 기울기(즉 미분값)이 양수

그 결과 배당금, 시가배당률 모두 증가한 **300**여개의 종목들이 선택됨

다층 모형 (Multilevel model)

2016~2018년 (**시간 수준**까지 종목별로 (**종목 수준**) 각 3회 반복 측정

⇒ 위계적(hierarchical) 데이터 분석에 적합한 **2수준 다층 모형**

300여개 종목들을 가지고 배당금, 시가배당률 각각에 대해 다층모형 적합
정규분포 가정을 충족시키고자 상용로그변환 실시

이때 그룹변수로 사용한 것은 앞서 주요한 변수로 지목한 바 있는 **배당횟수**

Table 3.2

다층 모형 투입 데이터 형식

itemcode	itemname	tstart	tstop	dps	der_dps_lag1
000020	동화약품	2016	2017	110	80
000020	동화약품	2017	2018	150	110
000020	동화약품	2018	2019	120	150
000050	경방	2016	2017	180	125

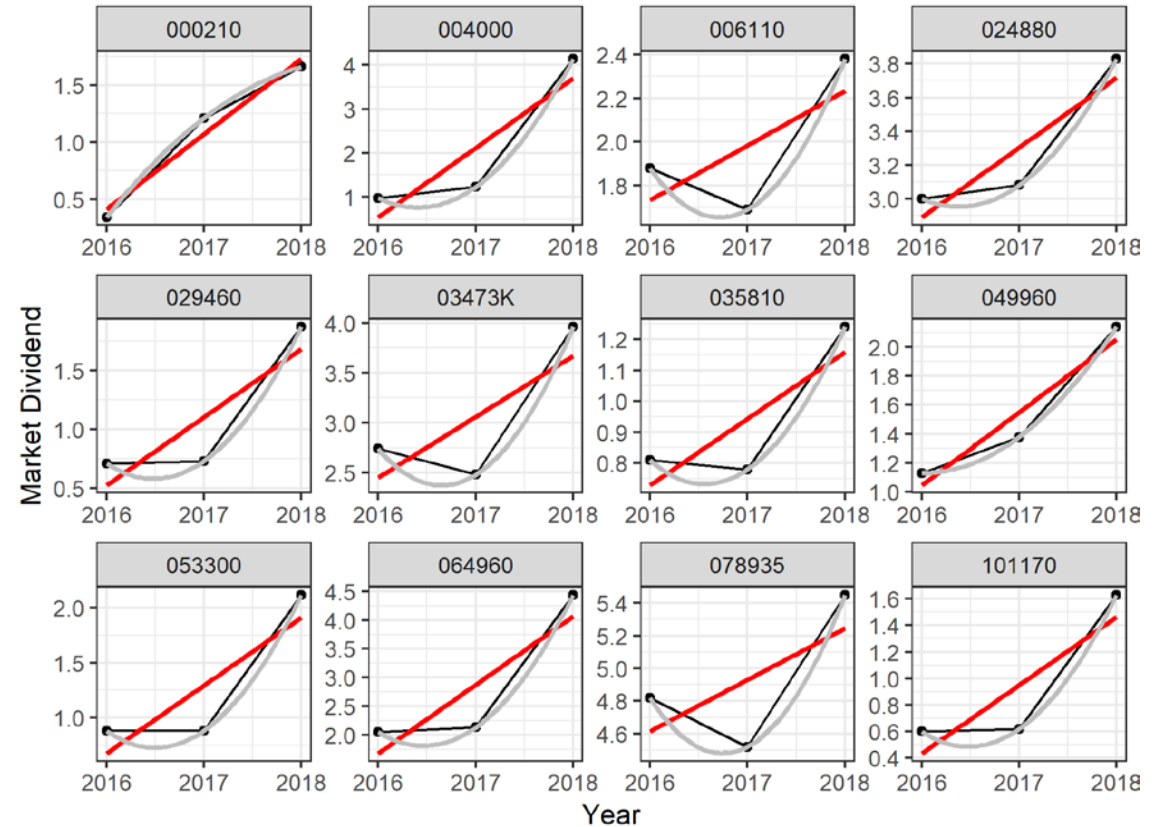
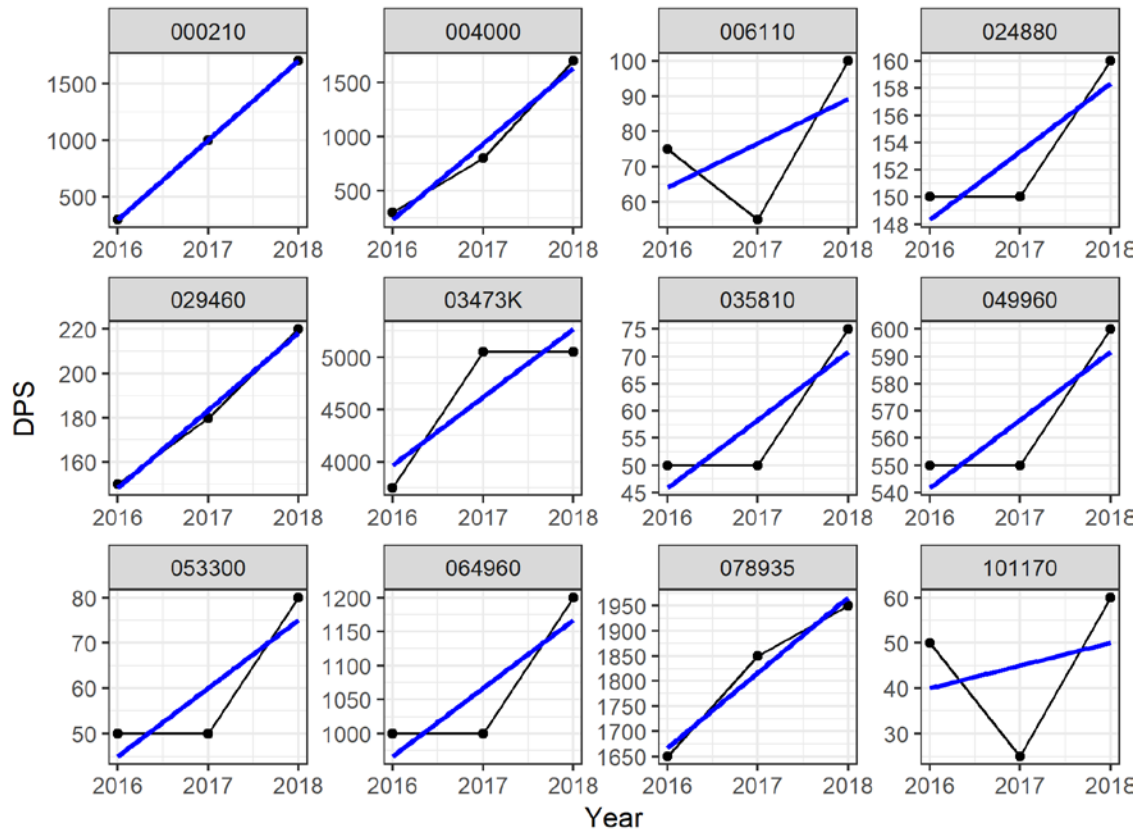
`lmer(배당금 or 시가배당률 ~ (배당횟수
+KOSPI 여부)*시간+(시간|종목))`

3. 배당주를 고르는 방법

[2단계] 다층 모형을 활용한 연도별 배당변화 패턴 분석

Figure 3.6
주당배당금 증가 패턴

1) **선형** 회귀를 적합한
뒤 기울기가 양수



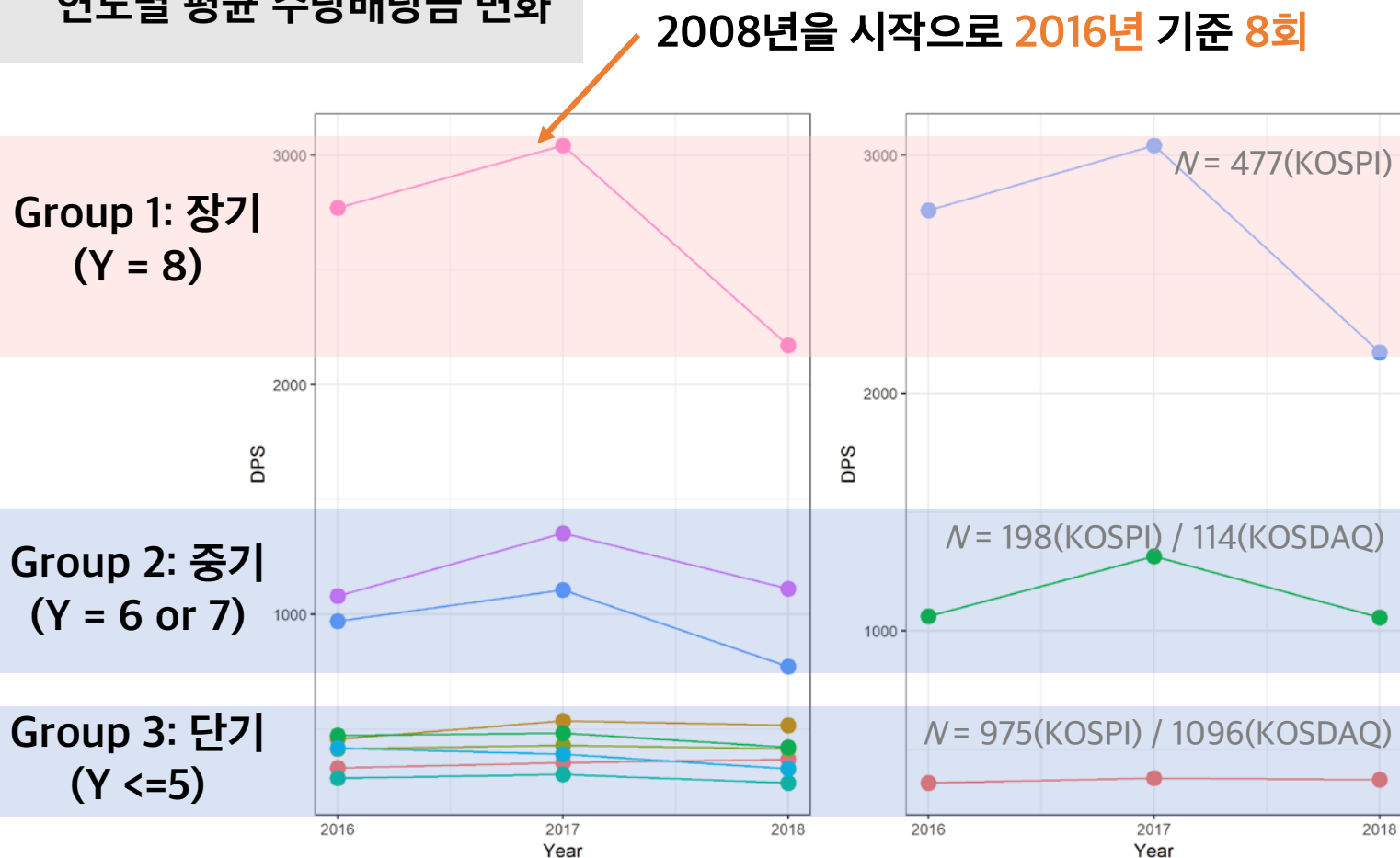
2-1) **선형** 회귀를 적합한 뒤 기울기가 양수
2-2) **이차** 회귀를 적합한 뒤 18년' 기울기
(즉 미분값)이 양수

Figure 3.7
시가배당률 증가 패턴

3. 배당주를 고르는 방법

[2단계] 다층 모델을 활용한 연도별 배당변화 패턴 분석

Figure 3.8
연도별 평균 주당배당금 변화



배당횟수 그룹의 설명력

기본 모형

$\text{Imer}(\text{배당금} \sim (1|\text{종목}))$

ICC(Intra-class correlation) = 0.93

⇒ 전체 분산의 93%가 종목별 차이 때문

배당횟수(및 KOSPI 여부) 투입 모형

$\text{Imer}(\text{배당금} \sim (\text{배당횟수} + \text{KOSPI 여부}) * \text{시간} + (\text{시간}|\text{종목}))$

PRE(Proportional Reduction in Error) = 0.60

⇒ 종목 수준에서의 분산이 60% 감소

우리 분류가 종목별 차이의 60%를 설명

3. 배당주를 고르는 방법

[2단계] 다층 모형을 활용한 연도별 배당변화 패턴 분석

주당배당금

시작점도, 기울기도 더 높음

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	2.133368	0.049266	289.025814	43.303	< 2e-16	***
g2Y = 6 or 7	0.210418	0.075666	289.025783	2.781	0.00578	**
g2Y = 8	0.310442	0.162993	289.025790	1.905	0.05782	.
market_kospi	0.444936	0.076852	289.025790	5.789	1.84e-08	**
time	0.106119	0.009082	289.011703	11.685	< 2e-16	**
g2Y = 6 or 7:time	-0.006005	0.013948	289.011704	-0.431	0.66712	
g2Y = 8:time	0.080290	0.030046	289.011707	2.672	0.00796	**
market_kospi:time	-0.007560	0.014167	289.011703	-0.534	0.59400	

시가배당률

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	0.161941	0.027023	288.998940	5.993	6.13e-09	**
g2Y = 6 or 7	0.004666	0.041504	288.998934	0.112	0.9106	
g2Y = 8	0.190524	0.089404	288.998935	2.131	0.0339	*
market_kospi	0.048423	0.042155	288.998935	1.149	0.2516	
time	0.150734	0.009767	288.998194	15.434	< 2e-16	**
g2Y = 6 or 7:time	0.006612	0.015000	288.998194	0.441	0.6597	
g2Y = 8:time	0.060354	0.032312	288.998193	1.868	0.0628	.
market_kospi:time	-0.032970	0.015235	288.998194	-2.164	0.0313	*

8회(최대 배당횟수) 배당한
기업들의 성장세가 뚜렷함

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목



안정적으로 배당금을 지급할 수 있는 기업

[1단계] 작년배당금, 배당횟수, 주가변동성 그리고 시장 종류를 고려하여 지급중지 회피



지속적으로 성장해왔고, 앞으로도 성장할 기업

[2단계] 최근 3년간 주당배당금과 시가배당률이 성장세 & 8회 배당한 기업

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목

$N = 2,353$

배당주인가?

KOSPI/KOSDAQ 상장기업 중 2016~2018년 배당이력이 있는 1,368개

$N = 1,368$

[1단계] 안정적으로 지급할 수 있는가?

작년배당금, 배당횟수, 주가변동성 그리고 시장 종류를 고려하여 지급중지 회피
지급중지 이력이 있는 592개 종목 제외

$N = 776$

$N = 293$

[2단계] 성장가능성이 있는가?

3년간 주당배당금과 시가배당률이 성장세인 293개 종목
최대 배당횟수(2016년 기준 8회)를 배당한 9개 종목을 최종 선정

$N = 9$

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목

최종 선정된 종목 리스트

대신증권	삼성전자	삼성전자우
삼성증권	삼성화재	삼성화재우
에스엘	한국금융지주	한독

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목

Figure 3.9 주당배당금 증가 패턴

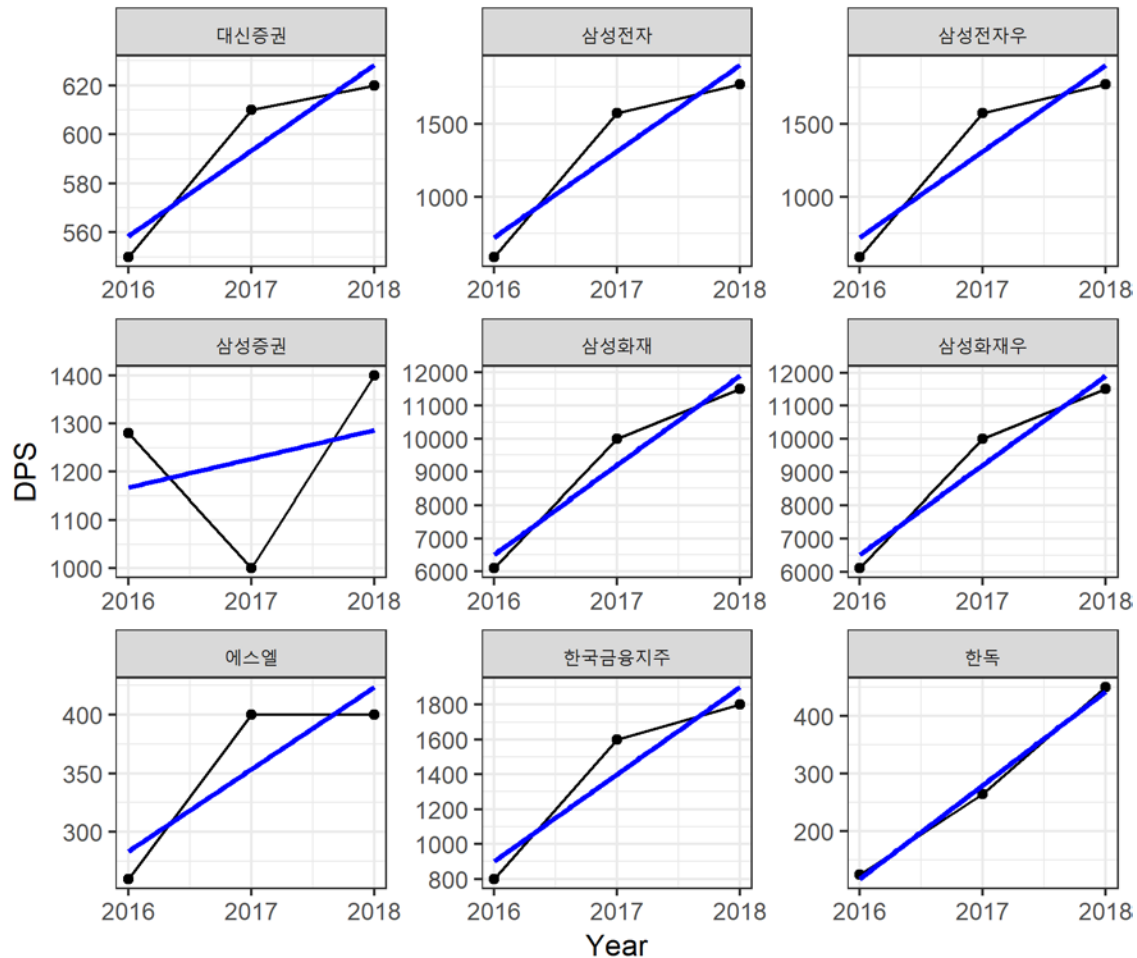
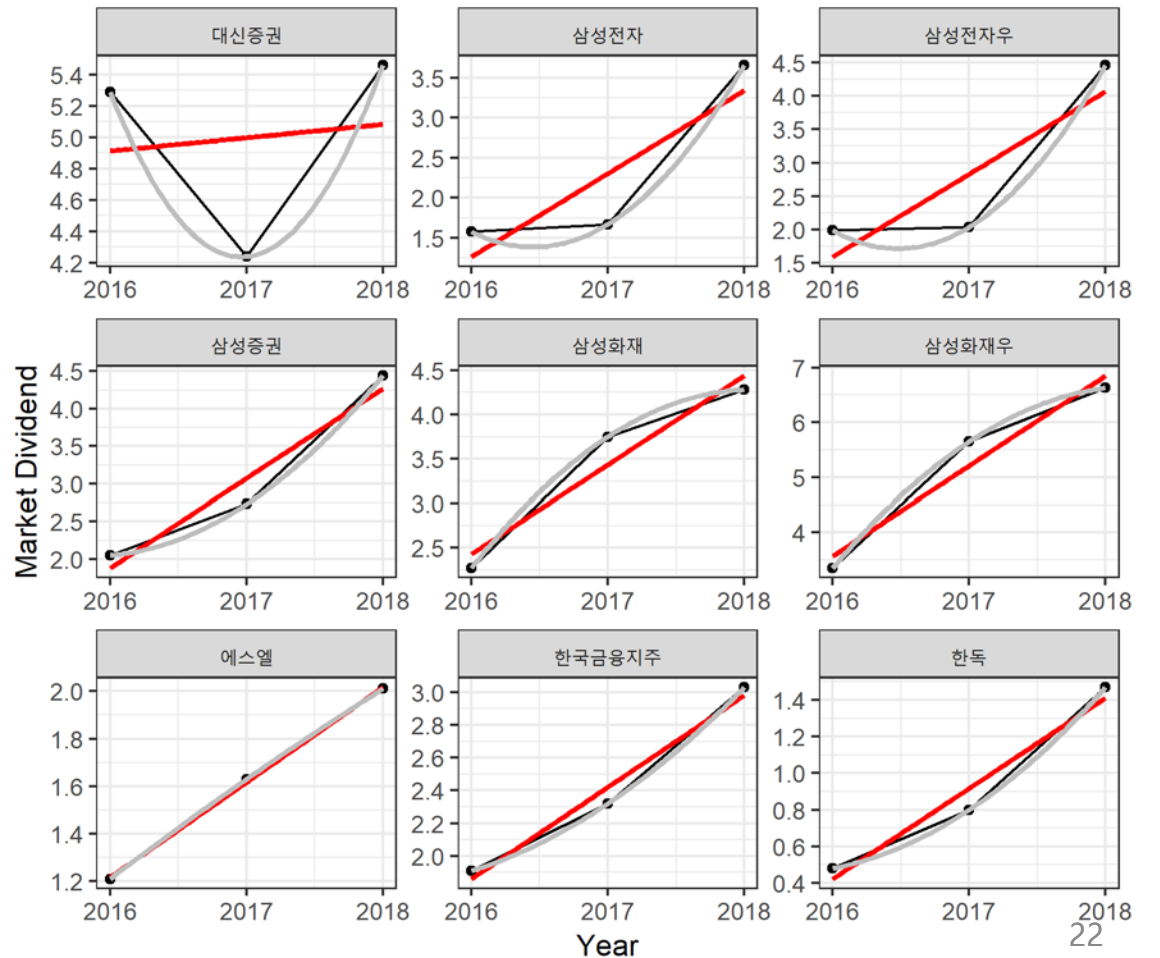


Figure 3.10 시가배당률 증가 패턴



3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목

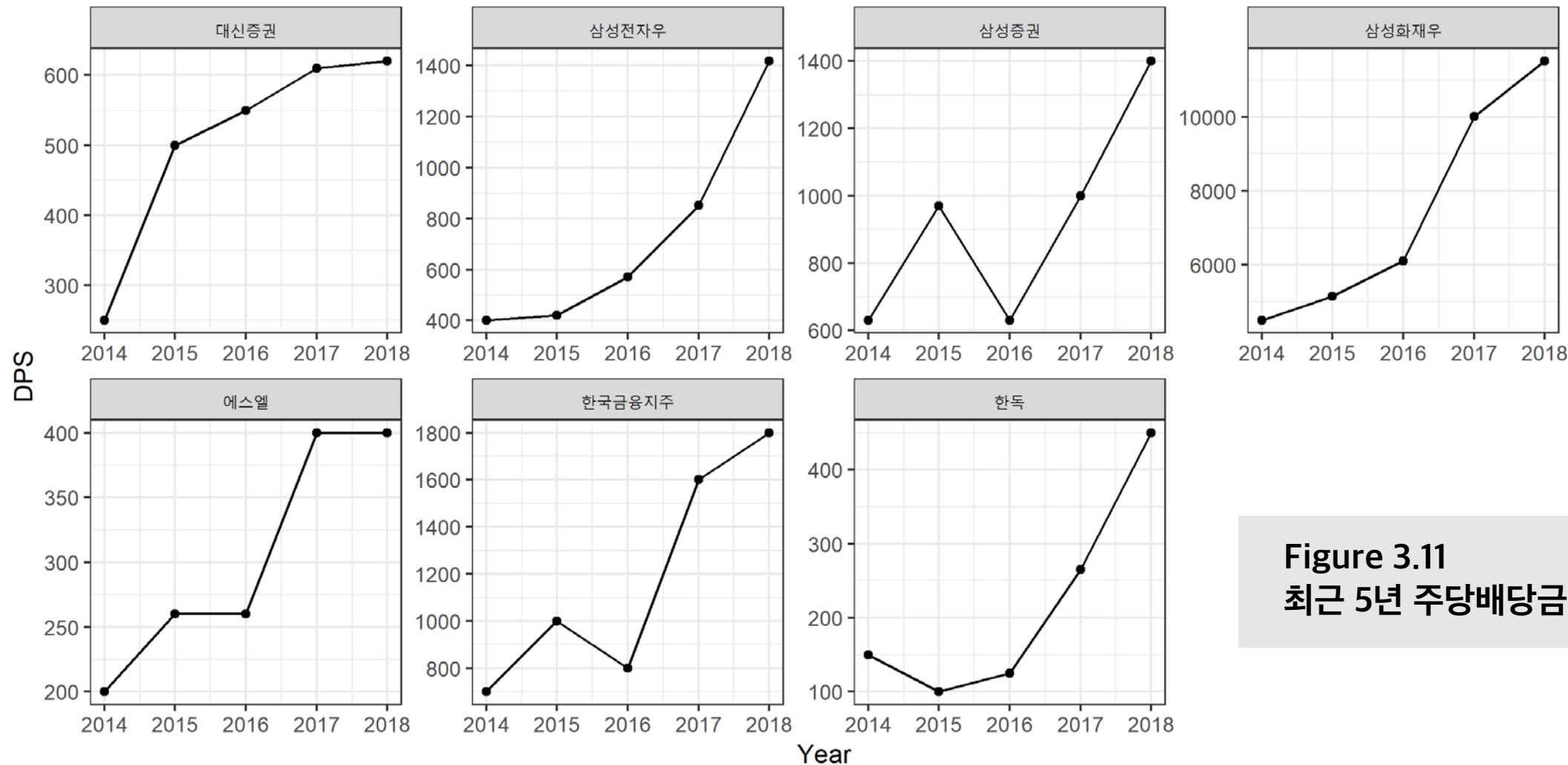
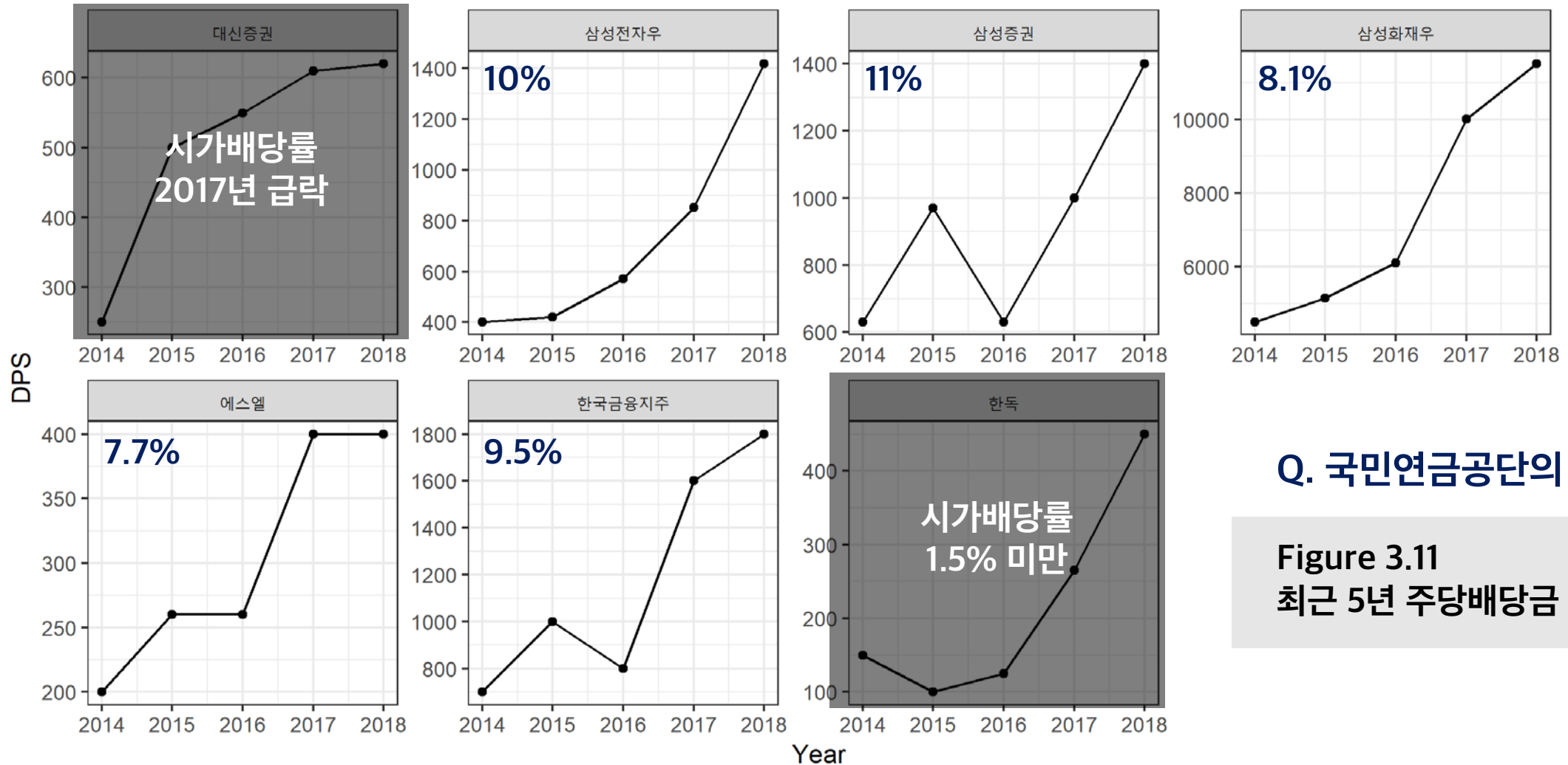


Figure 3.11
최근 5년 주당배당금 증가 패턴

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목



Q. 국민연금공단의 지분율?

Figure 3.11
최근 5년 주당배당금 증가 패턴

3. 배당주를 고르는 방법

[결론] 배당투자 유망 종목

자, 이제 실전이다!

주식 시장은 너무도 복잡하다

- 안전성과 성장가능성을 두 축으로 배당주 종목을 선정

무엇을, 언제, 얼마나 살 것인가에 대해 충분히 고민해야

- 수익성: 주당순이익(EPS), 당기순이익
- 투자금 회수: 주가, 주가수익비율(PER)
- 거시적인 금융시장 상황: 주가, 금리, 환율
- 배당소득세 15.4%(지방소득세 1.4% 포함)

⇒ 수많은 지표들이 존재하지만, 무엇보다 중요한 것은 자신의 판단!

NAVER Data Science Competition 2019

감사합니다.

발표자 박인서

Appendix

부가 설명 자료

Cox 비례위험모형 결과

```
Call:
coxph(formula = Surv(tstart, tstop, death) ~ der_dps_lag1 + der_dps_len +
      stock_cv + der_dps_lag1 * stock_cv, data = data)
```

n= 8279, number of events= 392

	coef	exp(coef)	se(coef)	z	Pr(> z)	
der_dps_lag1	-0.16621	0.84686	0.05129	-3.240	0.00119	**
der_dps_len	-0.19442	0.82332	0.02592	-7.500	6.36e-14	***
stock_cv	0.12334	1.13126	0.03096	3.983	6.79e-05	***
der_dps_lag1:stock_cv	0.11084	1.11721	0.03586	3.091	0.00200	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
der_dps_lag1	0.8469	1.1808	0.7659	0.9364
der_dps_len	0.8233	1.2146	0.7825	0.8662
stock_cv	1.1313	0.8840	1.0647	1.2020
der_dps_lag1:stock_cv	1.1172	0.8951	1.0414	1.1986

Concordance= 0.693 (se = 0.013)

Likelihood ratio test= 131.2 on 4 df, p=<2e-16

Wald test = 123 on 4 df, p=<2e-16

Score (logrank) test = 133.4 on 4 df, p=<2e-16

Likelihood ratio test 결과

Likelihood ratio test

Model 1: Surv(tstart, tstop, death) ~ der_dps_lag1 + der_dps_len + stock_cv + der_dps_lag1 * stock_cv

Model 2: Surv(tstart, tstop, death) ~ der_dps_lag1 + der_dps_len + stock_cv + stock_mean + stock_sd + stock_cap + stock_pbr + stock_listed +

market_kospi + market_kq2ks + deal_corp_buy + deal_foreign_buy +

deal_corp_sum + deal_foreign_sum + deal_corp_sell + deal_foreign_sell +

der_dps_lag1:stock_cv

#Df LogLik Df Chisq Pr(>Chisq)

1 4 -2577.5

2 17 -2565.0 13 25.068 0.02262 *

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- 모형 추정에는 R의 survival 라이브러리의 coxph() 함수를 이용
- [Time dependent covariates](#)

Appendix

부가 설명 자료

2수준 다층 모형 결과: 주당배당금

Formula: $dps \sim (g2 + market_kospi) * time + (time | itemcode)$
Data: dps_sub_long

REML criterion at convergence: -79.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.9954	-0.3354	0.0316	0.3355	5.7293

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
itemcode	(Intercept)	0.263687	0.51350	
	time	0.004140	0.06434	-0.33
Residual		0.009865	0.09932	

Number of obs: 879, groups: itemcode, 293

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	2.133368	0.049266	289.025814	43.303	< 2e-16	***
g2Y = 6 or 7	0.210418	0.075666	289.025783	2.781	0.00578	**
g2Y = 8	0.310442	0.162993	289.025790	1.905	0.05782	.
market_kospi	0.444936	0.076852	289.025790	5.789	1.84e-08	***
time	0.106119	0.009082	289.011703	11.685	< 2e-16	***
g2Y = 6 or 7:time	-0.006005	0.013948	289.011704	-0.431	0.66712	
g2Y = 8:time	0.080290	0.030046	289.011707	2.672	0.00796	**
market_kospi:time	-0.007560	0.014167	289.011703	-0.534	0.59400	

기본 모형: 주당배당금

Formula: $dps \sim (1 | itemcode)$
Data: dps_sub_long

REML criterion at convergence: 343.2

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.2827	-0.3700	0.0360	0.3707	6.0577

Random effects:

Groups	Name	Variance	Std.Dev.
itemcode	(Intercept)	0.3438	0.5864
	Residual	0.0246	0.1568

Number of obs: 879, groups: itemcode, 293

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	2.51134	0.03466	291.99999	72.45	<2e-16	***

- 모형 추정에는 R의 lme4 라이브러리의 lmer() 함수를 이용
- 모형 추정방법으로는 제한적 최대우도법(REML) 사용
- 고정효과의 자유도(df)는 새터스웨이트(Satterthwaite)의 제안에 따라 조정

Appendix

부가 설명 자료

2수준 다층 모형 결과: 시가배당률

Formula: $md \sim (g2 + market_kospi) * time + (time | itemcode)$
Data: md_sub_long

REML criterion at convergence: -289.1

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.7242	-0.3652	0.0303	0.4393	4.8260

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
itemcode	(Intercept)	0.07573	0.2752	
	time	0.00360	0.0600	-0.42
Residual		0.01378	0.1174	

Number of obs: 879, groups: itemcode, 293

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	0.161941	0.027023	288.998940	5.993	6.13e-09 ***
g2Y = 6 or 7	0.004666	0.041504	288.998934	0.112	0.9106
g2Y = 8	0.190524	0.089404	288.998935	2.131	0.0339 *
market_kospi	0.048423	0.042155	288.998935	1.149	0.2516
time	0.150734	0.009767	288.998194	15.434	< 2e-16 ***
g2Y = 6 or 7:time	0.006612	0.015000	288.998194	0.441	0.6597
g2Y = 8:time	0.060354	0.032312	288.998193	1.868	0.0628 .
market_kospi:time	-0.032970	0.015235	288.998194	-2.164	0.0313 *

기본 모형: 시가배당률

Formula: $md \sim (1 | itemcode)$
Data: md_sub_long

REML criterion at convergence: 139.8

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.8366	-0.4615	0.0003	0.5319	4.3866

Random effects:

Groups	Name	Variance	Std.Dev.
itemcode	(Intercept)	0.06982	0.2642
	Residual	0.03597	0.1897

Number of obs: 879, groups: itemcode, 293

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	0.20153	0.01671	292.00000	12.06	<2e-16 ***

- 모형 추정에는 R의 lme4 라이브러리의 lmer() 함수를 이용
- 모형 추정방법으로는 제한적 최대우도법(REML) 사용
- 고정효과의 자유도(df)는 새터스웨이트(Satterthwaite)의 제안에 따라 조정

Appendix

부가 설명 자료

Figure A.1 Lasso 정규화를 이용한 계수 추정: 주당배당금

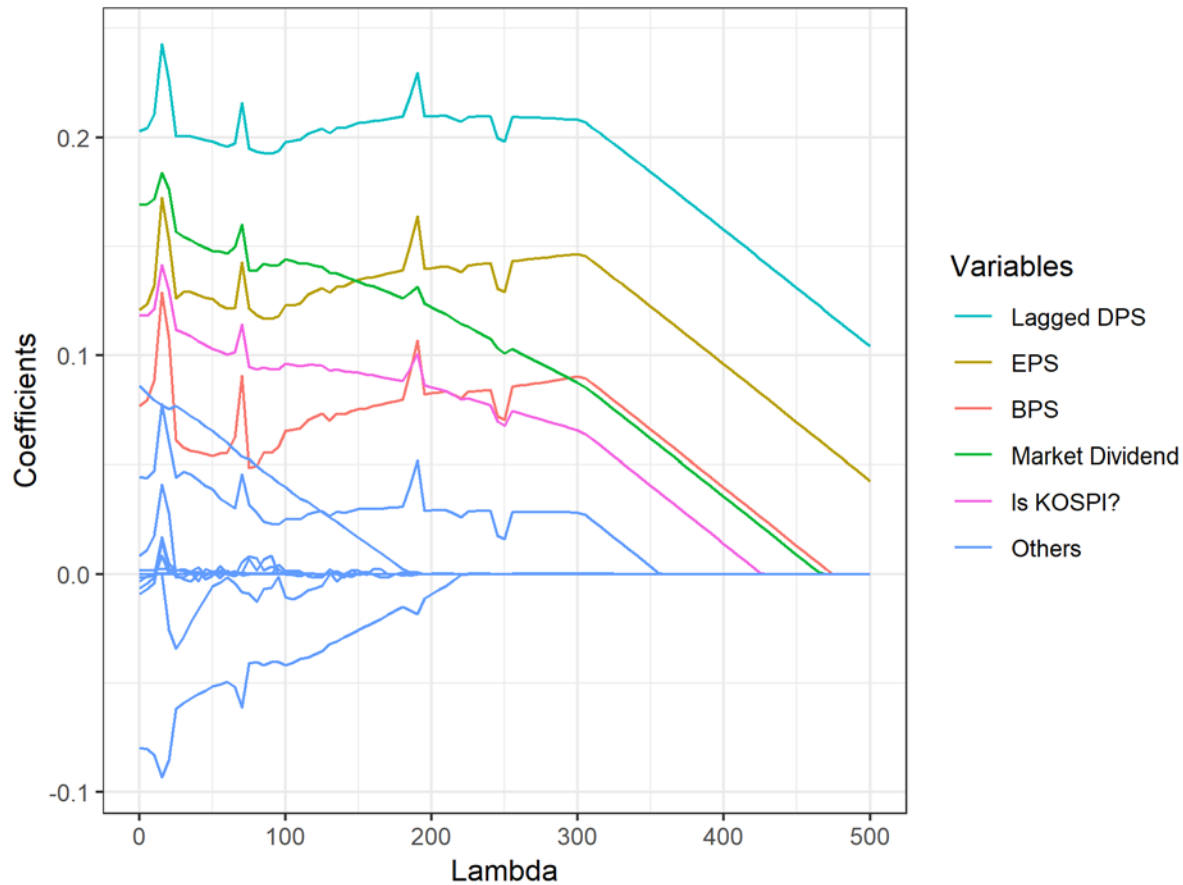


Figure A.2 Lasso 정규화를 이용한 계수 추정: 시가배당률

