Università Commerciale
Luigi Bocconi

Lab 3: From ARMA to VAR

(Reminder: this is optional)

Prof. Massimo Guidolin

20192– Financial Econometrics

Winter/Spring 2019

The Goal

- In this lab we try to model and forecast real stock returns and real dividend growth contrasting what can be done in univariate vs. multi-variate (here, bivariate) models
- We extend the evidence to a famous fundamental stock market indicator, much debated by pundits and economists, the cyclically adjusted price-earnings ratio over 10 years (CAPE-10)
  - Important: CAPE has nothing to do with CAPEX!
- Because the data are kindly made available to the public by Prof. and Nobel laureate Robert Shiller and CAPE has been proposed by him and co-authors, this lab is inspired by his work
- Even though in this lab we use U.S. data only, there is ample international research on the forecasting power of CAPE for (inflation-adjusted) stock returns

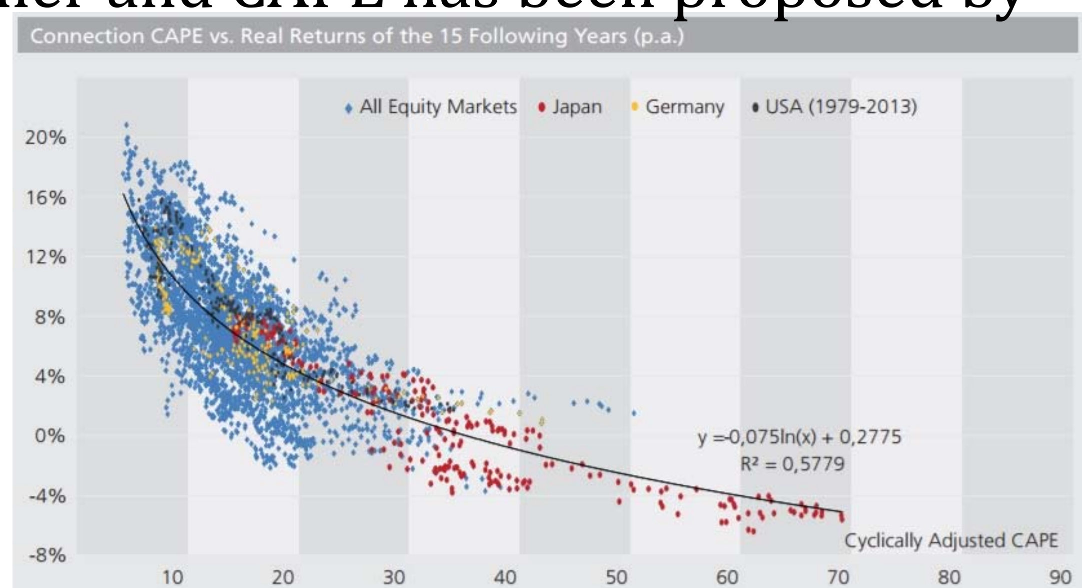
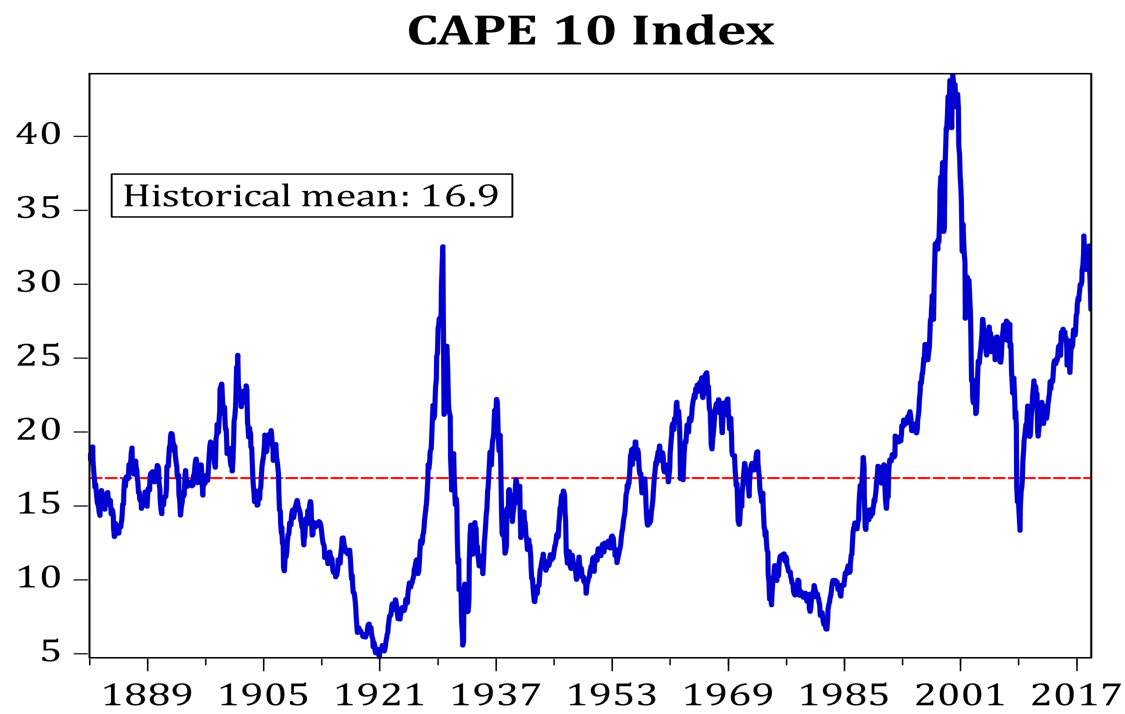


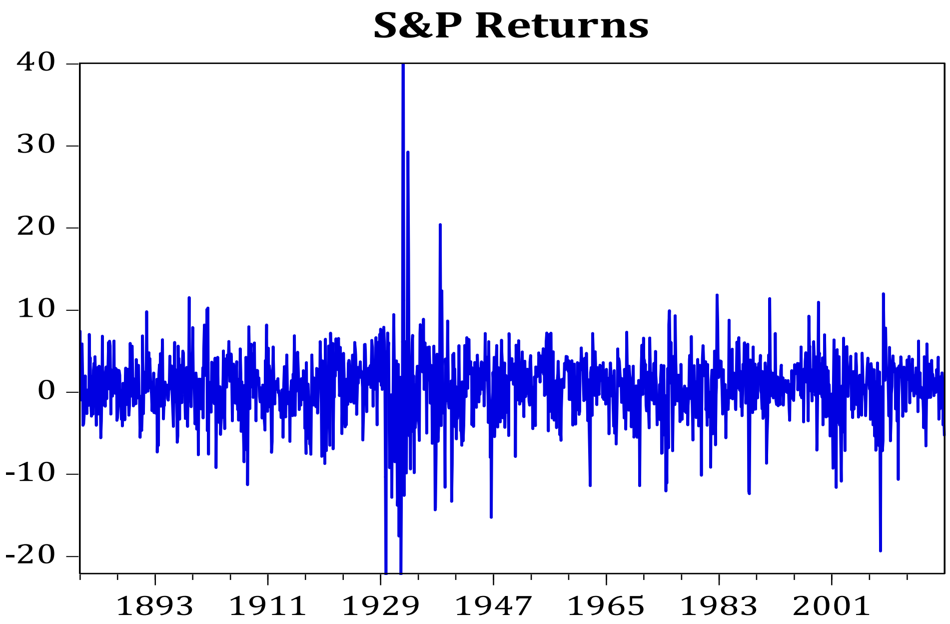
Figure 2: Connection between the CAPE and the returns of the 15 following years in the period 1881-2013 (US) and 1979-2013 (other markets). The USA, Japan and Germany are highlighted as examples in a single period 1979-2013. All returns inflation-adjusted, in local currency, incl. dividend income and annualized. Source: S&P 500: Robert J. Shiller, other countries: Worldscope, Thomson Reuters and own calculations.

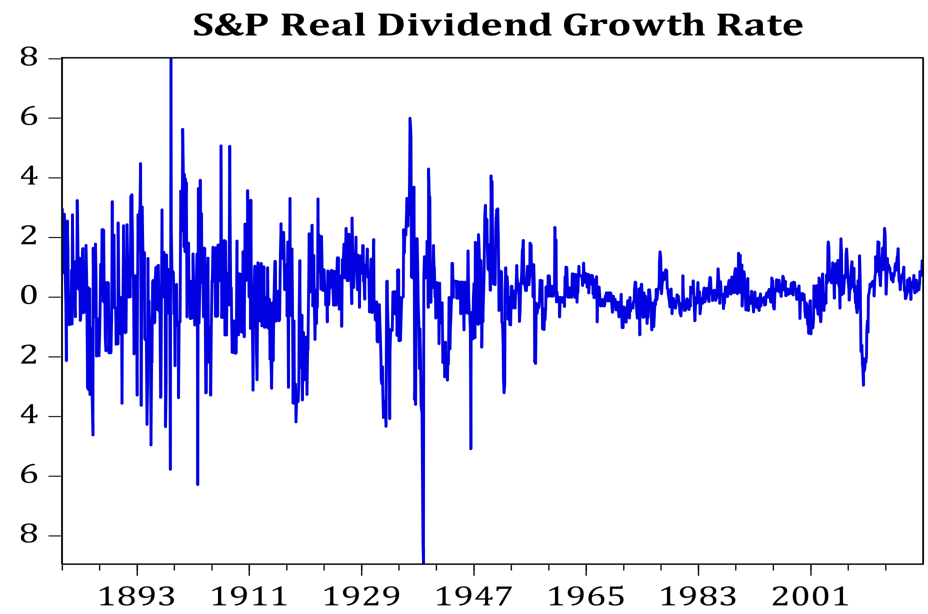
The Data

- We shall use a January 1881 – December 2018 monthly sample, for a total of 1,656 observations
  - Once more, not a big problem to obtain large samples in finances
  - We are on the verge of dealing with economic history here
- Three different series analyzed/commented
  - The discretely compounded (real) returns on the Standard & Poor's Composite index (that became S&P 500 in 1957) deflated using the CPI inflation index
  - The discretely compounded rate of growth of the (real) dividends paid by the companies in the S&P index, deflated using CPI inflation
  - The CAPE index, i.e., the ratio btw. the real S&P index and the average of the moving average of the last 10 years of real earnings reported by companies in the S&P index, a long-run, average real PE ratio
  - CAPE is use to forecast stock returns over timescales of 10 to 20 years, with higher than average CAPE values implying lower than average long-term annual average returns

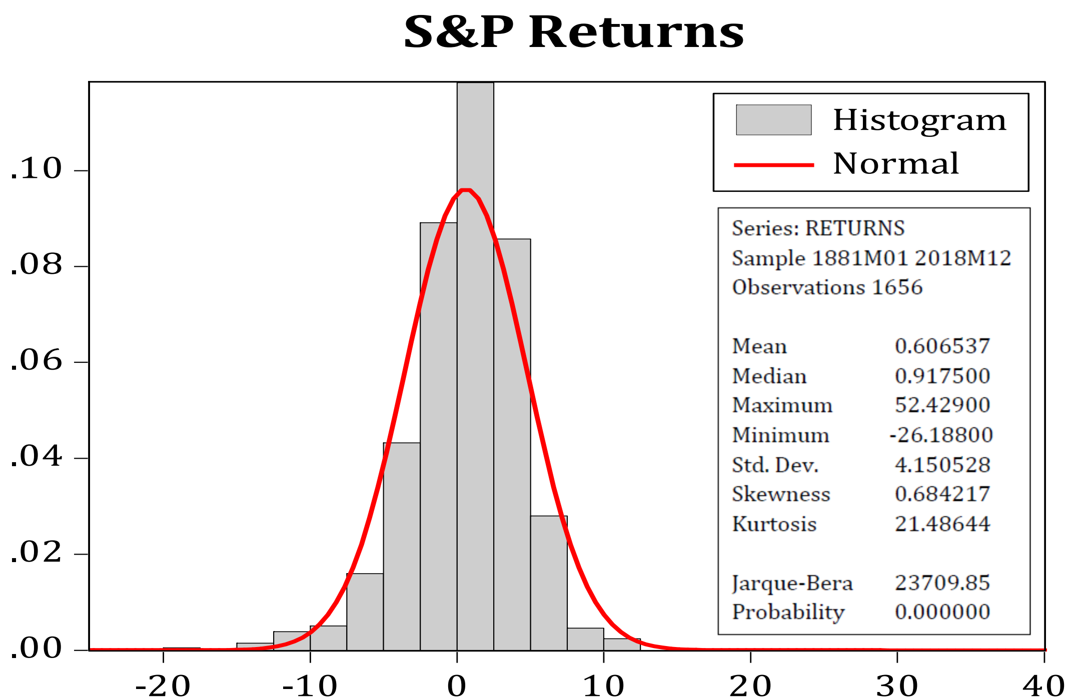
Step 0: The Data

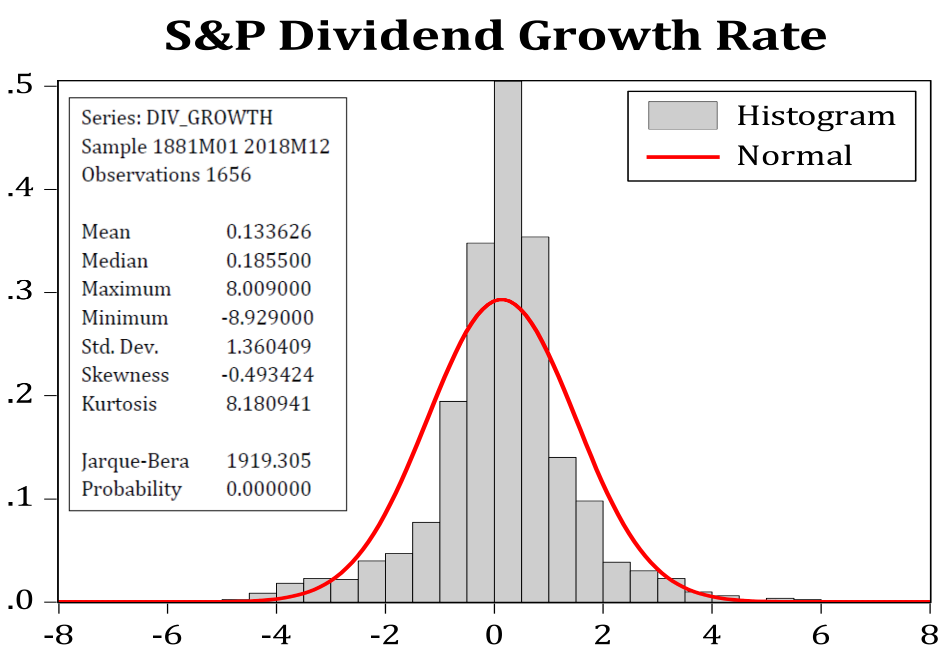


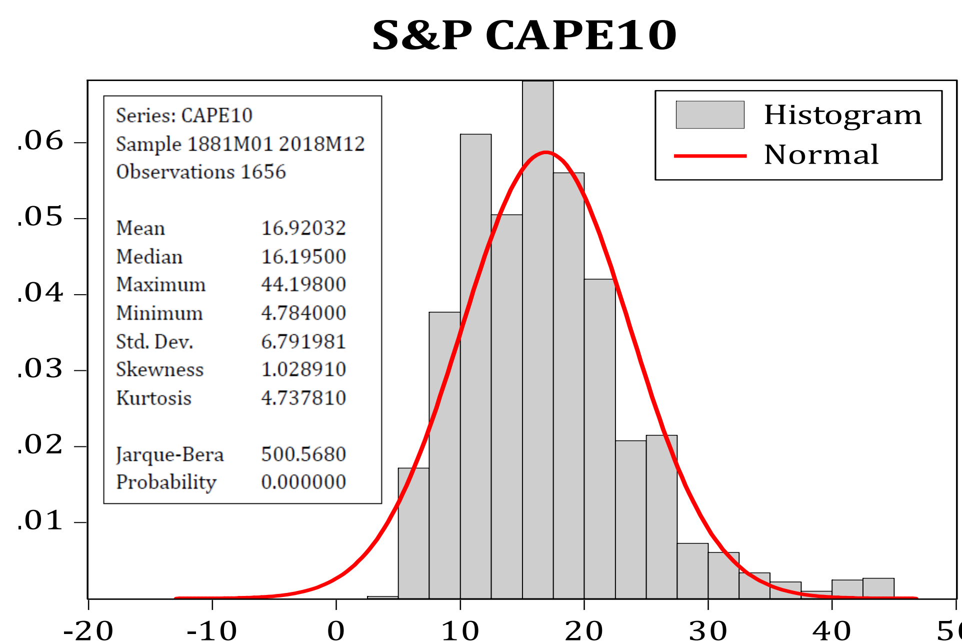




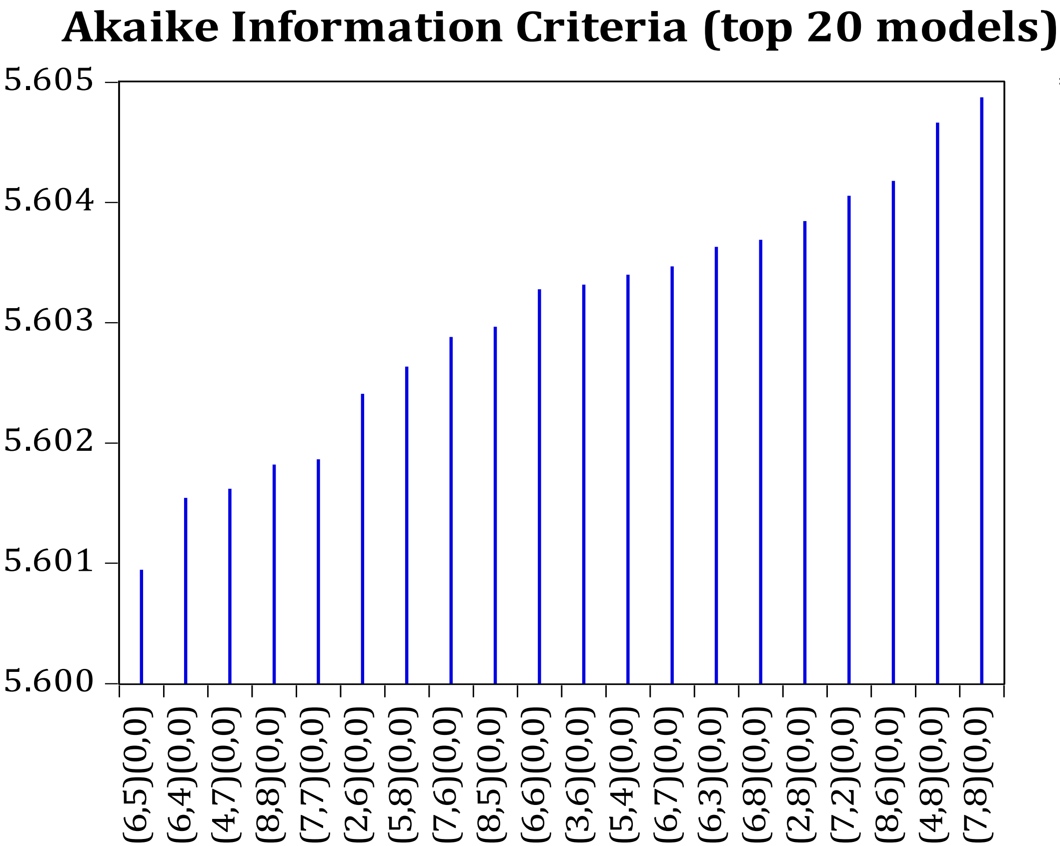
Step 0: The Data







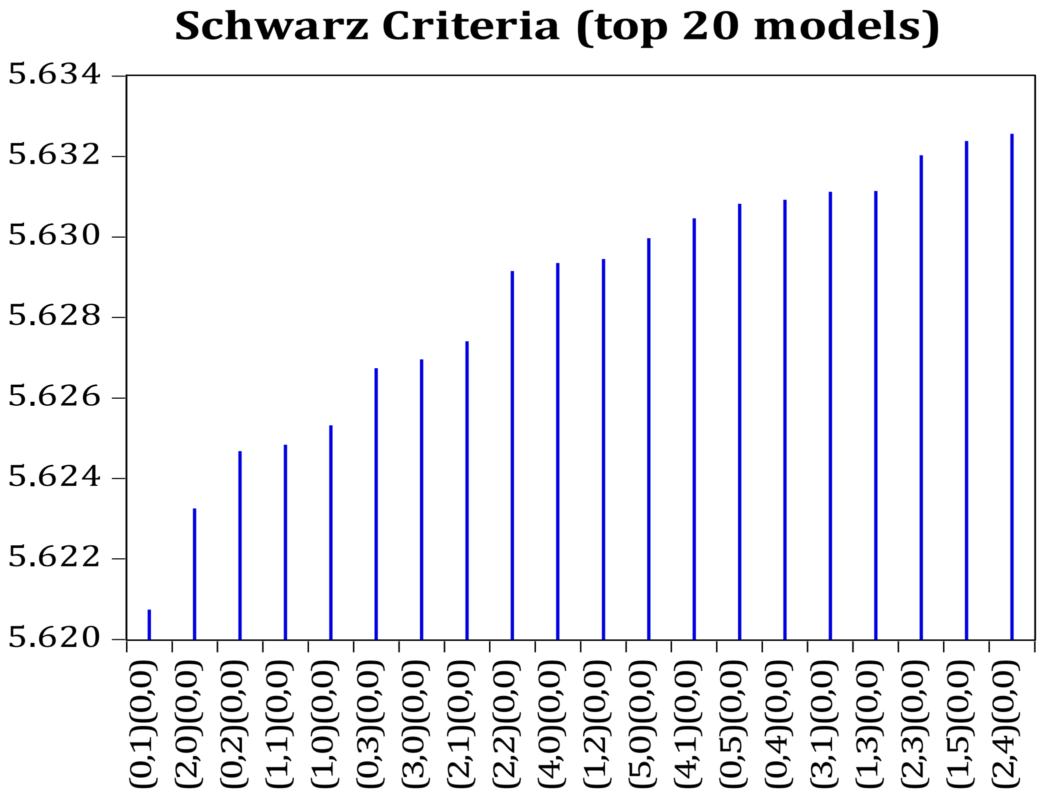
Step 1: Model Specification (S&P Returns)



Model	LogL	AIC*	BIC	HQ
(6,5)(0,0)	-4624.583957	5.600947	5.643434	5.616696
(6,4)(0,0)	-4626.078375	5.601544	5.640762	5.616082
(4,7)(0,0)	-4625.140904	5.601619	5.644106	5.617369
(8,8)(0,0)	-4620.308072	5.601821	5.660649	5.623628
(7,7)(0,0)	-4622.344438	5.601865	5.654157	5.621249
(2,6)(0,0)	-4628.795247	5.602410	5.635092	5.614525
(5,8)(0,0)	-4623.982693	5.602636	5.651659	5.620809
(7,6)(0,0)	-4624.186315	5.602882	5.651905	5.621055
(8,5)(0,0)	-4624.255452	5.602966	5.651989	5.621138
(6,6)(0,0)	-4625.515283	5.603279	5.649034	5.620240
(3,6)(0,0)	-4628.547532	5.603318	5.639269	5.616645
(5,4)(0,0)	-4628.615356	5.603400	5.639351	5.616727
(6,7)(0,0)	-4624.671032	5.603467	5.652491	5.621640
(6,3)(0,0)	-4628.808010	5.603633	5.639583	5.616959

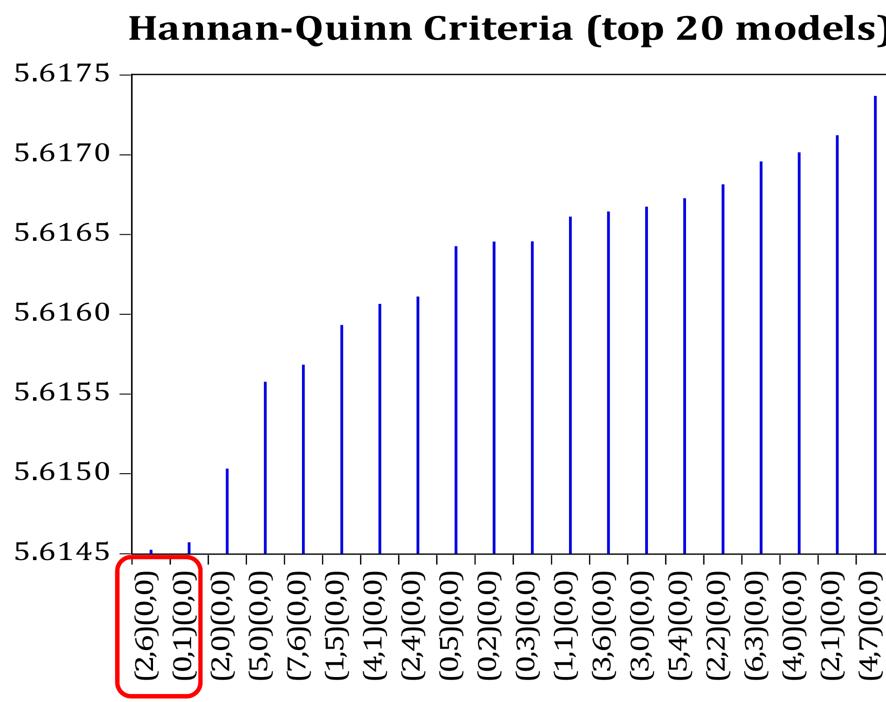
- A systematic search across ARIMA(p, d, q) with $p \leq 8$, $d \leq 1$, $q \leq 8$ using AIC as a selection criteria points to large ARMA(6, 5) model
  - However, the very table reveals that BIC and H-Q would lead to the selection of much smaller models
  - Such evidence in favor of large, rich ARMA models is unusual but we need to remember that we are using 138 years of data

Step 1: Model Specification (S&P Returns)



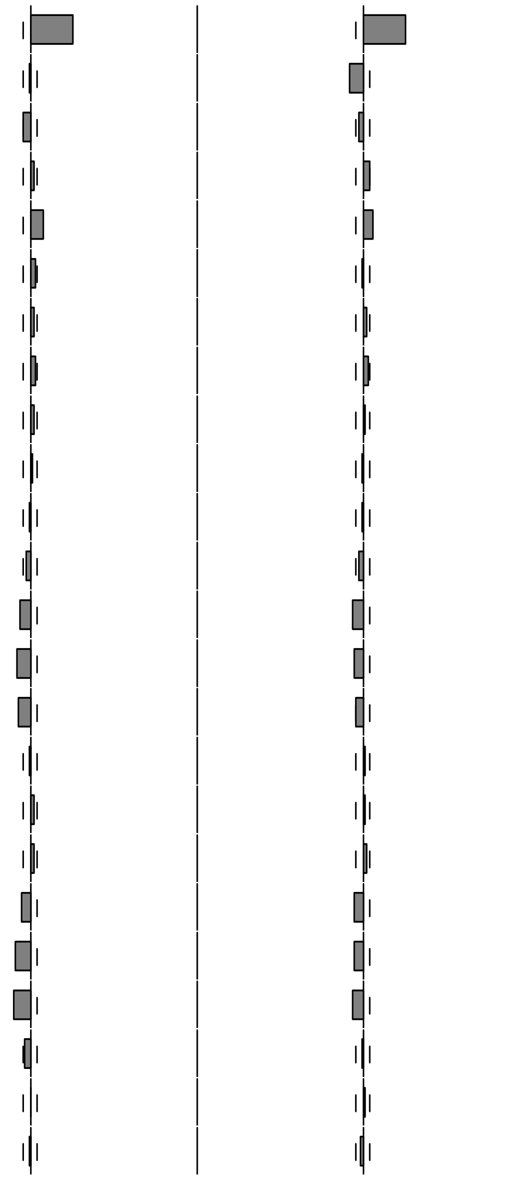
Model	LogL	AIC	BIC*	HQ
(0,1)(0,0)	-4642.855337	5.610936	5.620741	5.614571
(2,0)(0,0)	-4641.234039	5.610186	5.623259	5.615032
(0,2)(0,0)	-4642.412536	5.611609	5.624682	5.616455
(1,1)(0,0)	-4642.542116	5.611766	5.624839	5.616612
(1,0)(0,0)	-4646.647519	5.615516	5.625321	5.619151
(0,3)(0,0)	-4640.410543	5.610399	5.626740	5.616457
(3,0)(0,0)	-4640.591772	5.610618	5.626959	5.616676
(2,1)(0,0)	-4640.961736	5.611065	5.627406	5.617122
(2,2)(0,0)	-4638.704271	5.609546	5.629155	5.616815
(4,0)(0,0)	-4638.869992	5.609746	5.629356	5.617015
(1,2)(0,0)	-4642.659017	5.613115	5.629456	5.619172
(5,0)(0,0)	-4635.677053	5.607098	5.629975	5.615578

- In this case, the best ranked model is a simple MA(1)!
- If you repeat the sequential selection using H-Q you obtain another model, a ARMA(2, 6)
    - However, according to H-Q, also MA(1) ranks fairly high



Step 1: Model Specification (S&P Returns)

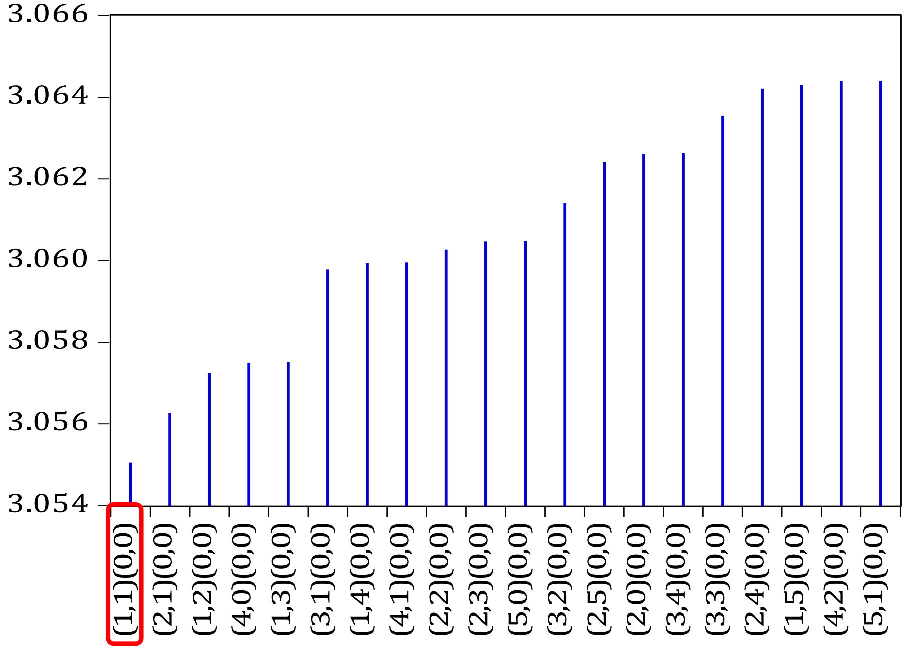
- Time to take a look at the SACF/SPACF
  - The obvious insight is that these data may have been generated by a MA(1)
  - However, SACF is also (borderline) significant at lags 13-15 and 19-21, accor-ding to a sinusoidal pattern
  - Such a pattern also charac-terizes the SPACF and this points towards a ARMA with coefficients of alterna-ting signs
  - This may be consistent with a ARMA(2, q) if the AR has complex roots



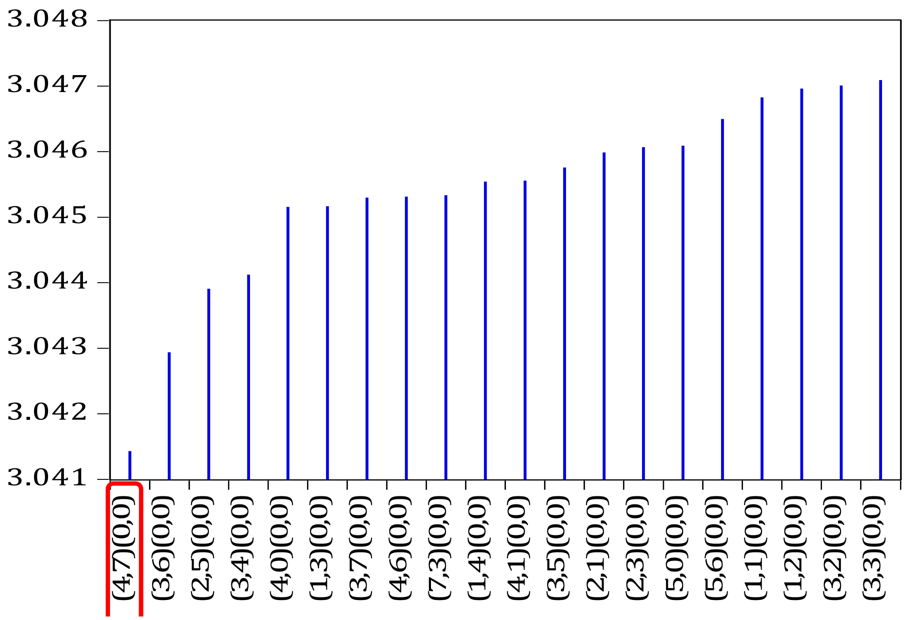
	AC	PAC	Q-Stat	Prob
1	0.263	0.263	114.76	0.000
2	-0.006	-0.080	114.81	0.000
3	-0.049	-0.028	118.79	0.000
4	0.022	0.046	119.56	0.000
5	0.079	0.062	129.84	0.000
6	0.038	-0.001	132.24	0.000
7	0.027	0.026	133.42	0.000
8	0.041	0.037	136.15	0.000
9	0.031	0.011	137.78	0.000
10	0.011	-0.002	137.97	0.000
11	-0.005	-0.005	138.00	0.000
12	-0.019	-0.021	138.64	0.000
13	-0.062	-0.064	145.14	0.000
14	-0.078	-0.055	155.22	0.000
15	-0.068	-0.045	163.06	0.000
16	-0.004	0.016	163.08	0.000
17	0.023	0.014	163.96	0.000
18	0.032	0.031	165.66	0.000
19	-0.050	-0.055	169.89	0.000
20	-0.096	-0.056	185.22	0.000
21	-0.101	-0.060	202.47	0.000
22	-0.038	-0.001	204.89	0.000
23	0.008	0.011	205.00	0.000
24	-0.004	-0.009	205.03	0.000

Step 1: Model Specification (S&P Dividend Growth)

Schwarz Criteria (top 20 models)



Hannan-Quinn Criteria (top 20 models)



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.558	0.558	516.61	0.000
		2	0.437	0.183	833.89	0.000
		3	0.333	0.048	1018.5	0.000
		4	0.318	0.108	1186.2	0.000
		5	0.273	0.037	1309.8	0.000
		6	0.236	0.018	1402.1	0.000
		7	0.202	0.014	1469.8	0.000
		8	0.187	0.027	1528.3	0.000
		9	0.145	-0.022	1563.3	0.000
		10	0.103	-0.032	1581.2	0.000
		11	0.097	0.017	1596.9	0.000
		12	0.065	-0.028	1604.0	0.000
		13	0.044	-0.019	1607.3	0.000
		14	0.076	0.067	1616.9	0.000
		15	0.038	-0.040	1619.3	0.000
		16	0.049	0.021	1623.3	0.000
		17	0.029	-0.006	1624.7	0.000
		18	0.008	-0.032	1624.8	0.000
		19	-0.010	-0.021	1625.0	0.000
		20	-0.054	-0.068	1629.9	0.000
		21	-0.026	0.038	1631.0	0.000
		22	-0.047	-0.036	1634.8	0.000
		23	-0.055	-0.020	1639.8	0.000
		24	-0.093	-0.049	1654.3	0.000

- Also in this case there is considerable heterogeneity in the models selected by different ICs

Step 2: Estimation

Dependent Variable: RETURNS
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1881M01 2018M12
Convergence achieved after 56 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.608992	0.119383	5.101150	0.0000
AR(1)	1.857814	0.047859	38.81860	0.0000
AR(2)	-0.888484	0.048093	-18.47430	0.0000
MA(1)	-1.583632	0.052399	-30.22236	0.0000
MA(2)	0.381928	0.050780	7.521271	0.0000
MA(3)	0.174042	0.037710	4.615225	0.0000
MA(4)	0.111555	0.033423	3.337674	0.0009
MA(5)	0.024096	0.035079	0.686903	0.4922
MA(6)	-0.073754	0.019955	-3.696004	0.0002
SIGMASQ	15.67879	0.239330	65.51118	0.0000

Dependent Variable: DIV_GROWTH
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1881M01 2018M12
Convergence achieved after 346 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.109536	0.027732	3.949780	0.0001
AR(1)	0.438053	0.025266	17.33755	0.0000
AR(2)	0.853659	0.018479	46.19650	0.0000
AR(3)	0.550911	0.019118	28.81631	0.0000
AR(4)	-0.850299	0.023986	-35.44987	0.0000
MA(1)	0.000917	0.775694	0.001182	0.9991
MA(2)	-0.720756	0.766401	-0.940442	0.3471
MA(3)	-0.889527	0.276931	-3.212088	0.0013
MA(4)	0.465789	0.476619	0.977278	0.3286
MA(5)	0.113664	0.112426	1.011009	0.3122
MA(6)	0.097457	0.036346	2.681398	0.0074
MA(7)	-0.067542	0.073374	-0.920518	0.3574
SIGMASQ	1.185044	0.919163	1.289264	0.1975

- Below models selected by BIC:

Dependent Variable: RETURNS
Method: ARMA Maximum Likelihood (OPG - BHHH)
Sample: 1881M01 2018M12
Convergence achieved after 20 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.606841	0.126755	4.787531	0.0000
MA(1)	0.276194	0.018176	15.19576	0.0000
SIGMASQ	15.94887	0.212700	74.98278	0.0000

Dependent Variable: DIV_GROWTH

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.024928	0.017566	1.419119	0.1561
DIV_GROWTH(-1)	0.805743	0.014454	55.74437	0.0000
MA(1)	-0.382138	0.023630	-16.17144	0.0000
SIGMASQ	1.218382	0.020472	59.51380	0.0000

Step 3: Diagnostic Checks (S&P Returns)

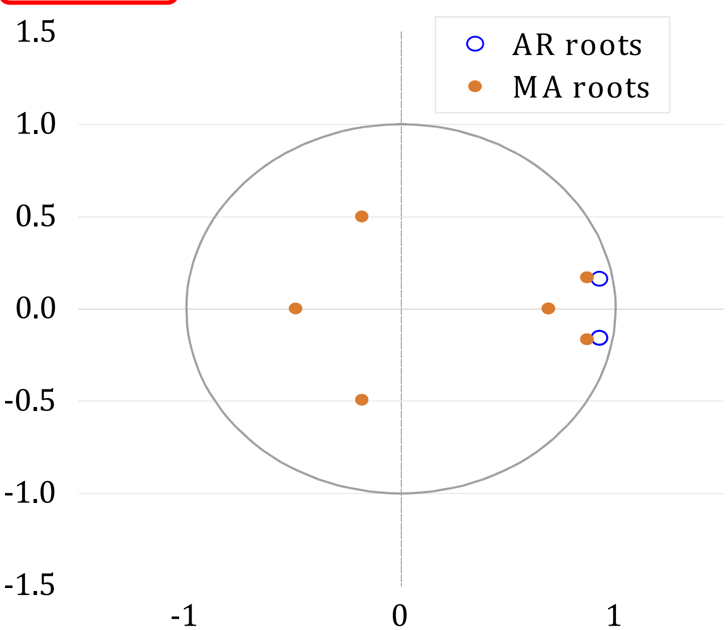
AR Root(s)	Modulus	Cycle
0.928907 ± 0.160050i	0.942594	36.82480

No root lies outside the unit circle.
ARMA model is stationary.

MA Root(s)	Modulus	Cycle
0.870085 ± 0.168875i	0.886322	32.77505
0.690517	0.690517	
-0.179409 ± 0.496281i	0.527714	3.276441
-0.488237	0.488237	

No root lies outside the unit circle.
ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.000	0.000	1.E-05	
		2	0.000	0.000	0.0004	
		3	0.003	0.003	0.0120	
		4	0.001	0.001	0.0135	
		5	-0.005	-0.005	0.0502	
		6	-0.014	-0.014	0.3604	
		7	-0.003	-0.003	0.3747	0.540
		8	0.020	0.020	1.0592	0.589
		9	0.019	0.019	1.6794	0.642
		10	-0.000	-0.000	1.6794	0.794
		11	-0.006	-0.006	1.7425	0.884
		12	-0.010	-0.010	1.9006	0.929
		13	-0.047	-0.047	5.6306	0.583
		14	-0.047	-0.046	9.3022	0.317
		15	-0.048	-0.048	13.211	0.153
		16	0.010	0.010	13.381	0.203
		17	0.010	0.010	13.551	0.259
		18	0.037	0.037	15.850	0.198
		19	-0.037	-0.038	18.133	0.153
		20	-0.055	-0.057	23.213	0.057
		21	-0.072	-0.074	31.996	0.006
		22	-0.024	-0.022	32.951	0.008
		23	0.006	0.011	33.011	0.011
		24	-0.007	-0.004	33.084	0.016

Bad luck?

- Further (JB) tests reveal that the residuals are not normally distributed, but never said they should be

Step 3: Diagnostic Checks (Dividend Growth Rate)

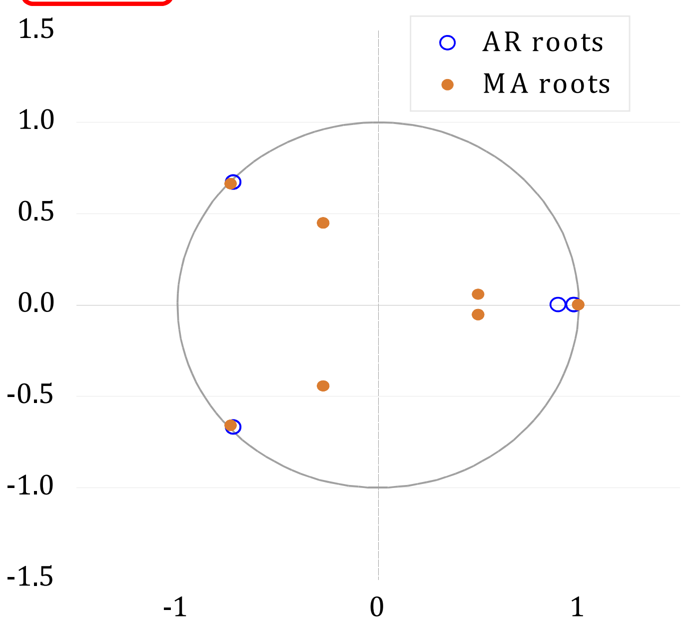
AR Root(s)	Modulus	Cycle
-0.719339 ± 0.670759i	0.983548	2.627709
0.977675	0.977675	
0.899055	0.899055	

No root lies outside the unit circle.
ARMA model is stationary.

MA Root(s)	Modulus	Cycle
1.000000	1.000000	
-0.732088 ± 0.662046i	0.987045	2.611039
-0.270294 ± 0.445755i	0.521302	2.969526
0.501923 ± 0.056391i	0.505081	56.15931

No root lies outside the unit circle.
ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)

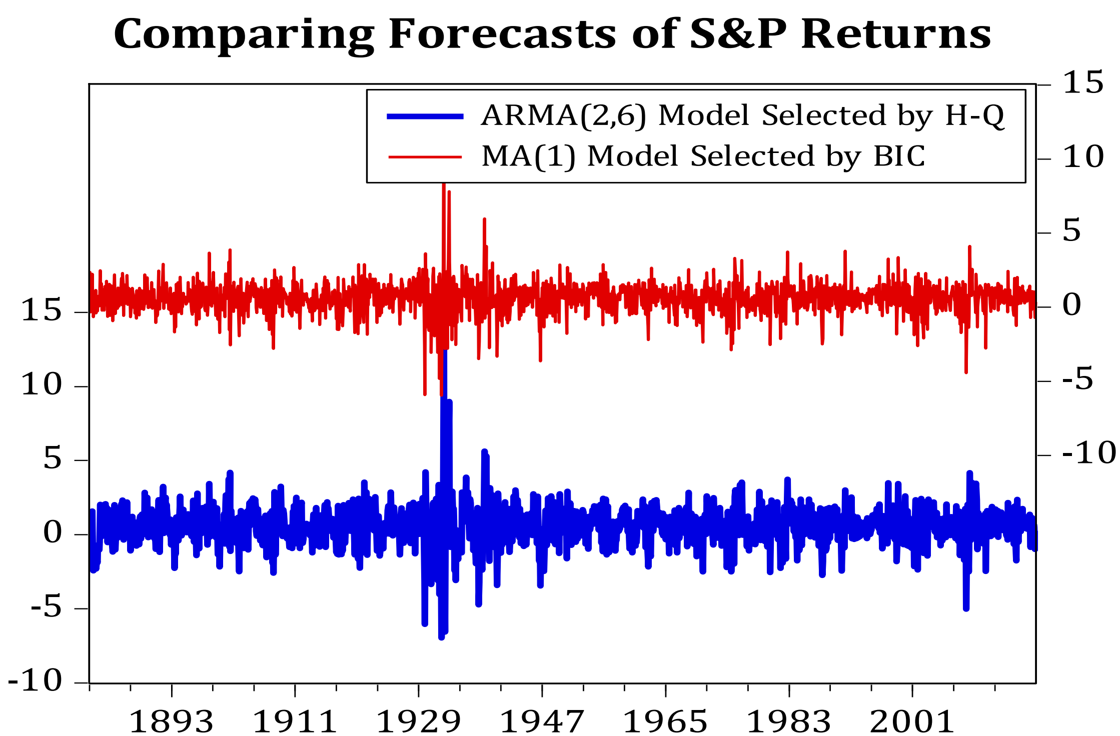


Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.000	0.000	4.E-05	
		2	-0.001	-0.001	0.0023	
		3	0.004	0.004	0.0229	
		4	-0.001	-0.001	0.0240	
		5	0.001	0.001	0.0245	
		6	-0.023	-0.023	0.9285	
		7	0.016	0.017	1.3807	
		8	0.027	0.027	2.5775	0.108
		9	-0.003	-0.003	2.5946	0.273
		10	-0.015	-0.015	2.9578	0.398
		11	-0.001	-0.001	2.9587	0.565
		12	-0.005	-0.006	3.0081	0.699
		13	-0.039	-0.038	5.5607	0.474
		14	0.044	0.045	8.7708	0.270
		15	0.003	0.002	8.7849	0.361
		16	0.018	0.017	9.3492	0.406
		17	0.015	0.015	9.7239	0.465
		18	0.020	0.021	10.392	0.496
		19	-0.009	-0.011	10.517	0.571
		20	-0.053	-0.050	15.280	0.290
		21	0.026	0.027	16.431	0.288
		22	-0.022	-0.025	17.268	0.303
		23	0.034	0.033	19.154	0.261
		24	-0.064	-0.065	26.061	0.073

- Further (JB) tests reveal that the residuals are not normally distributed, but never said they should be

Step 4: Forecasting (S&P Returns)

- Even though the two sets of forecasts look very similar, there is some evidence that ARMA(2,6) performs best
- This despite the fact that ARMA(2,6) implies estimating 10 parameters instead of 3
- Now we ask: can we gain forecast power for stock returns from dividends or the CAPE index?



Forecast: RETURNSF_ARMA	
Actual: RETURNS	
Forecast sample: 1881M01 2018M12	
Adjusted sample: 1881M03 2018M12	
Included observations: 1654	
Root Mean Squared Error	3.956392
Mean Absolute Error	2.766042
Mean Abs. Percent Error	227.7741
Theil Inequality Coef.	0.708173
Bias Proportion	0.000002
Variance Proportion	0.532786
Covariance Proportion	0.467213
Theil U2 Coefficient	0.850339
Symmetric MAPE	137.6465

MA(1)

Forecast: RETURNSF	
Actual: RETURNS	
Forecast sample: 1881M01 2018M12	
Included observations: 1656	
Root Mean Squared Error	3.993859
Mean Absolute Error	2.798597
Mean Abs. Percent Error	246.6588
Theil Inequality Coef.	0.732555
Bias Proportion	0.000000
Variance Proportion	0.582041
Covariance Proportion	0.417959
Theil U2 Coefficient	0.823977
Symmetric MAPE	140.4549

Unrestricted vs. Unrestricted VARs

- Let's build bivariate VAR(p) models to forecast stock returns
- Of course, in the process, we'll also forecast dividend growth rate

Lag	LogL	LR	AIC	SC	HQ
0	-7522.845	NA	9.132094	9.138656	9.134527
1	-7142.176	759.9527	8.674971	8.694657	8.682270
2	-7106.110	71.91234	8.636056	8.668867*	8.648222*
3	-7102.717	6.756488	8.636793	8.682729	8.653825
4	-7090.297	24.70484	8.626574	8.685635	8.648473
5	-7084.019	12.47320*	8.623809*	8.695994	8.650574
6	-7081.281	5.433102	8.625341	8.710650	8.656972
7	-7077.532	7.428617	8.625646	8.724080	8.662143
8	-7074.353	6.292698	8.626642	8.738201	8.668006

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

- While VAR(2) models are plausible in the light of earlier evidences, VAR(5) may represent a way to surrogate complex VARMA

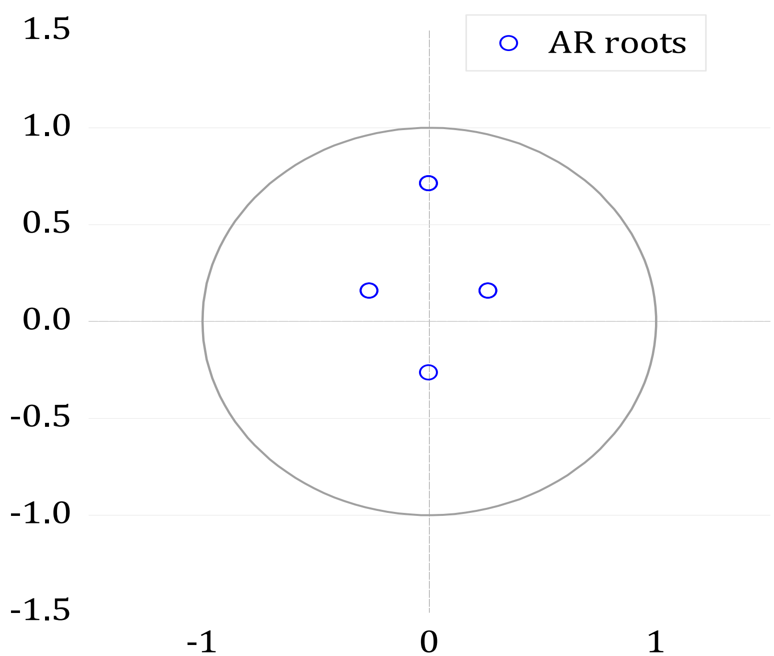
A VAR(2) Model

- Possible to present OLS, equation-by-equation estimates of a VAR(2)

Root	Modulus
0.710027	0.710027
0.155758 - 0.262144i	0.304926
0.155758 + 0.262144i	0.304926
-0.265956	0.265956

No root lies outside the unit circle.
VAR satisfies the stability condition.

Inverse Roots of VAR Polynomials



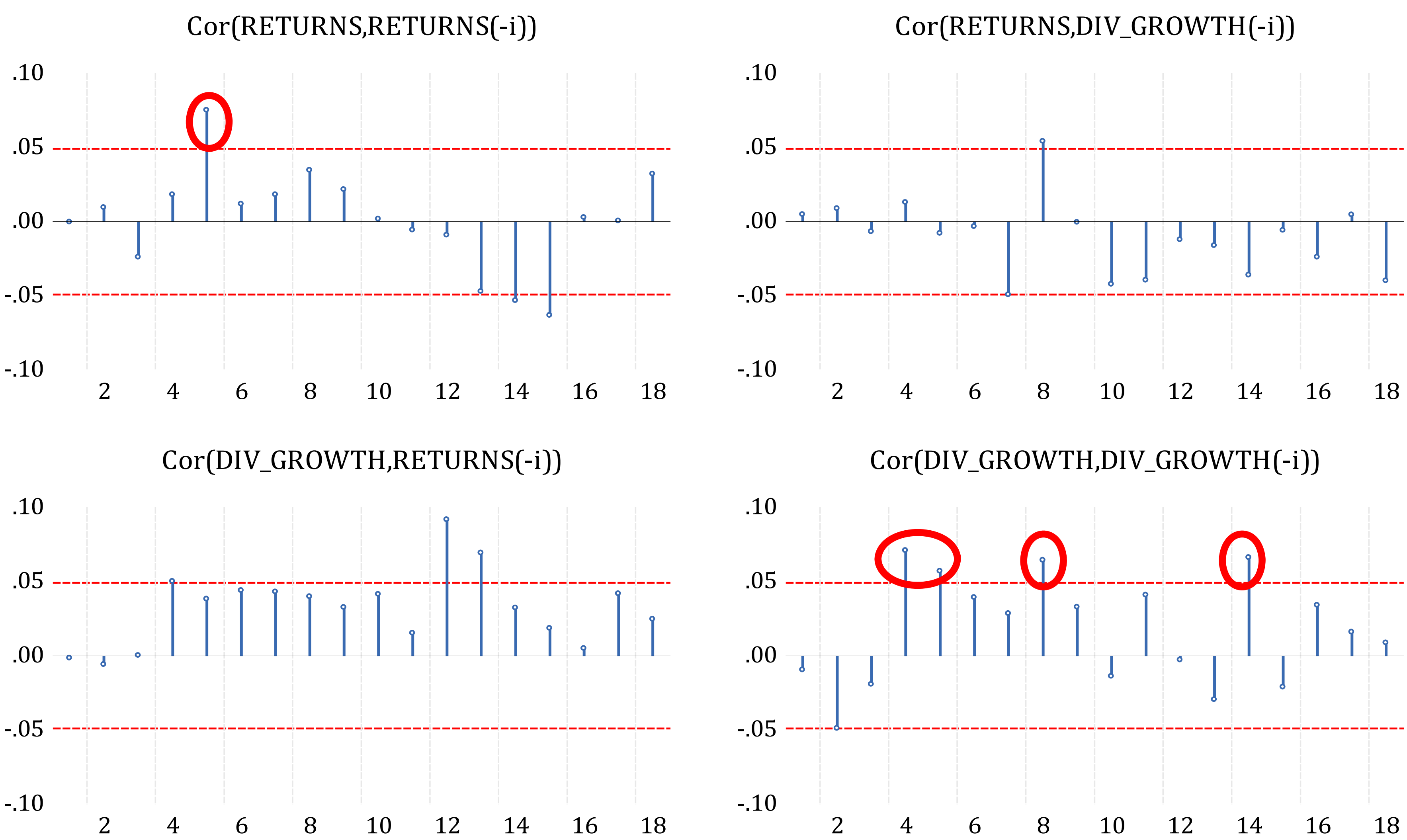
Vector Autoregression Estimates
Included observations: 1654 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	RETURNS	DIV_GROWTH
RETURNS(-1)	0.288218	-0.030126
	(0.02462)	(0.00681)
	[11.7049]	[-4.42054]
RETURNS(-2)	-0.089024	0.007453
	(0.02475)	(0.00685)
	[-3.59733]	[1.08810]
DIV_GROWTH(-1)	-0.120068	0.467370
	(0.08789)	(0.02432)
	[-1.36614]	[19.2141]
DIV_GROWTH(-2)	0.234069	0.177632
	(0.08761)	(0.02425)
	[2.67178]	[7.32602]
C	0.467934	0.060318
	(0.10015)	(0.02772)
	[4.67231]	[2.17612]
R-squared	0.079760	0.343184
Adj. R-squared	0.077528	0.341591
Log likelihood	-4631.285	-2506.573
Akaike AIC	5.606149	3.036969
Schwarz SC	5.622506	3.053326
Log likelihood		-7130.563
Akaike information criterion		8.634296
Schwarz criterion		8.667011
Number of coefficients		10

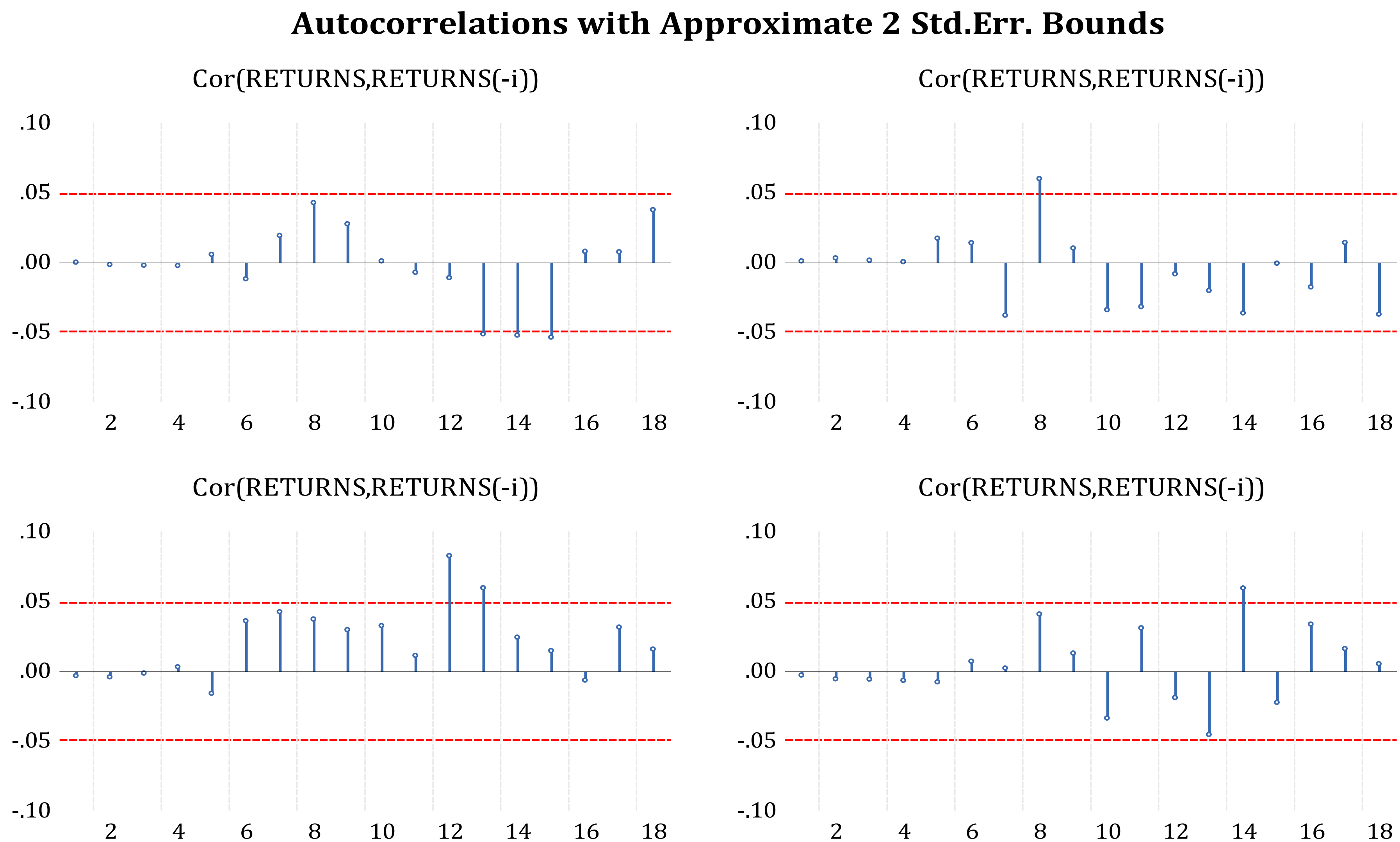
Model Diagnostics for the VAR(2)

- There are some residual concerns on the correct specification of the VAR(2) model, to which the VAR(5) puts remedy
  - Yet concerns more about the dividend growth series than for returns

Autocorrelations with Approximate 2 Std.Err. Bounds



Model Diagnostics for the VAR(5)



- Let's now compare the forecast performance of VAR(2) and VAR(5)
- A VAR(5) outperforms, although not massively, a VAR(2), i.e., not always smaller models outperform (not with 1,656 obs.)

Forecasting Performance of VAR(5) vs VAR(2)

Forecast Evaluation: VAR(2)

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DIV_GROWTH	1656	1.101335	0.696515	408.6220	0.507025
RETURNS	1656	3.979327	2.788962	1403.930	0.722252

Forecast Evaluation: VAR(5)

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DIV_GROWTH	1656	1.090412	0.689539	600.2749	0.499479
RETURNS	1656	3.965169	2.775006	1271.105	0.713692

Forecast: RETURNSF_ARMA
Actual: RETURNS
Forecast sample: 1881M01 2018M12
Adjusted sample: 1881M03 2018M12
Included observations: 1654

Root Mean Squared Error	3.956392
Mean Absolute Error	2.766042
Mean Abs. Percent Error	227.7741
Theil Inequality Coef.	0.708173
Bias Proportion	0.000002
Variance Proportion	0.532786
Covariance Proportion	0.467213
Theil U2 Coefficient	0.850339
Symmetric MAPE	137.6465

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

- However, by inspecting the VAR(5) estimates, we detect 13 estimated coefficients out of 22 that fail to be statistically significant in tests of size 5% or less
    - We record slight worsening vs. ARMA(2,6): multivariate may not pay!
- Can we estimate restricted VAR models imposing restrictions?
    - The answer is yes, of course, but by using MLE

A Restricted VAR(5)

- The model estimated is:

RETURNS = C(1)+C(2)*RETURNS(-1)+C(3)*RETURNS(-5)+C(4)*DIV_GROWTH(-2)

DIV_GROWTH = C(6)*RETURNS(-1)+C(7)*DIV_GROWTH(-1)+C(8)*DIV_GROWTH(-2)+C(9)*DIV_GROWTH(-4)

- The ML estimation output looks as follows

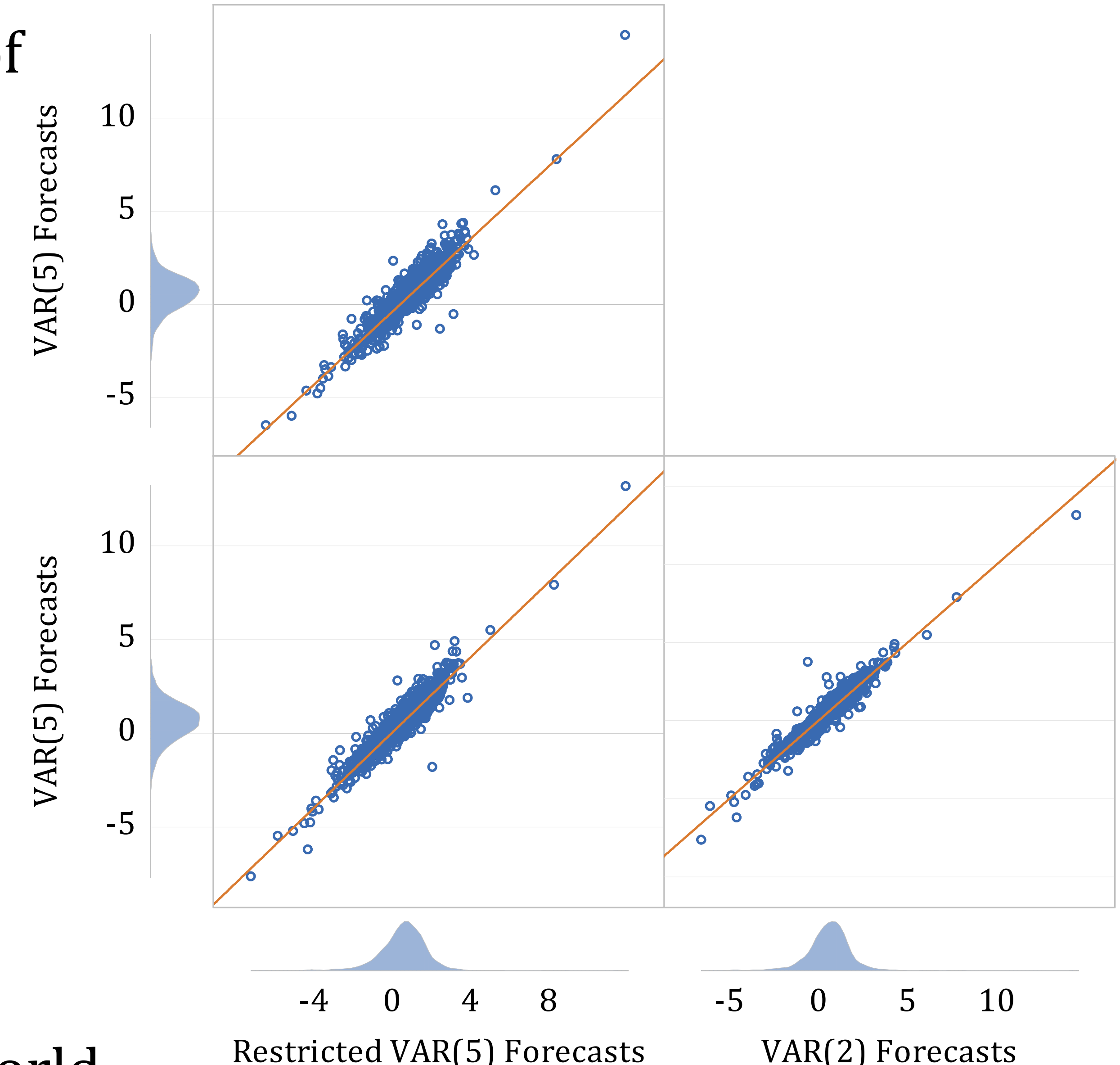
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Sample: 1881M06 2018M12
Included observations: 1651
Convergence achieved after 0 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.363588	0.102291	3.554451	0.0004
C(2)	0.262280	0.019393	13.52419	0.0000
C(3)	0.068010	0.020928	3.249664	0.0012
C(4)	0.145699	0.067321	2.164234	0.0304
C(6)	-0.026285	0.006438	-4.083030	0.0000
C(7)	0.454493	0.014314	31.75270	0.0000
C(8)	0.139912	0.016128	8.674920	0.0000
C(9)	0.110144	0.017475	6.302753	0.0000

Log likelihood	-7111.650	Schwarz criterion	8.650862
Avg. log likelihood	-2.153740	Hannan-Quinn criter.	8.634369
Akaike info criterion	8.624652		

A Restricted VAR(5)

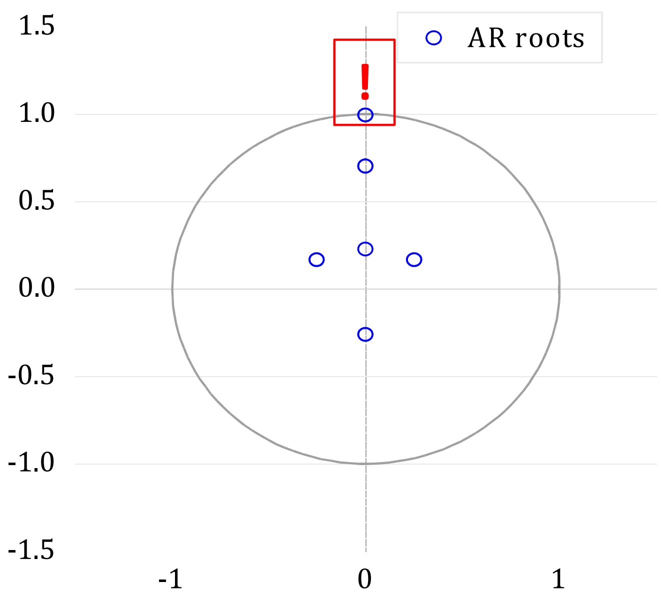
- The resulting forecasts of S&P returns are similar but not completely identical
- As a last step, given the glamour around its role in valuation theory, we try to augment the VAR model with CAPE10
- Surprisingly, we find that the one for CAPE almost represent an autonomous regression equation
- CAPE seems to be in a world apart – how is that possible?



Expanding the VAR models to include CAPE

- CAPE10 just depends on its own lags and (weakly) on returns at t-2
- Neither stock returns or dividend growth depends on CAPE
- Yet, the R-square of CAPE exceeds 99%
- Finally, the VAR is close to being non-stationary: what is going on?
- Stay tuned!

Inverse Roots of VAR Polynomials



	RETURNS	DIV_GROWTH	CAPE10
RETURNS(-1)	0.334065 (0.05457) [6.12169]	-0.029625 (0.01511) [-1.96044]	0.012443 (0.00925) [1.34460]
RETURNS(-2)	-0.085545 (0.02476) [-3.45438]	0.006456 (0.00686) [0.94141]	-0.009375 (0.00420) [-2.23243]
DIV_GROWTH(-1)	-0.101673 (0.08803) [-1.15491]	0.463108 (0.02438) [18.9966]	-0.014618 (0.01493) [-0.97914]
DIV_GROWTH(-2)	0.248219 (0.08769) [2.83079]	0.173840 (0.02428) [7.15928]	0.038406 (0.01487) [2.58278]
CAPE10(-1)	-0.329965 (0.32181) [-1.02533]	0.003639 (0.08912) [0.04083]	1.187009 (0.05457) [21.7501]
CAPE10(-2)	0.294168 (0.32208) [0.91333]	0.005574 (0.08919) [0.06249]	-0.193320 (0.05462) [-3.53935]
C	1.041243 (0.26684) [3.90212]	-0.094084 (0.07389) [-1.27325]	0.106540 (0.04525) [2.35436]
R-squared	0.083655	0.345235	0.990175
Adj. R-squared	0.080317	0.342849	0.990139
Log likelihood		-7457.888	
Akaike information criterion		9.043395	
Schwarz criterion		9.112095	
Number of coefficients		21	

Università Commerciale
Luigi Bocconi

Lab 3: From ARMA to VAR

(Reminder: this is optional)

Prof. Massimo Guidolin

20192– Financial Econometrics

Winter/Spring 2019

The Goal

- In this lab we try to model and forecast real stock returns and real dividend growth contrasting what can be done in univariate vs. multi-variate (here, bivariate) models
- We extend the evidence to a famous fundamental stock market indicator, much debated by pundits and economists, the cyclically adjusted price-earnings ratio over 10 years (CAPE-10)
    - Important: CAPE has nothing to do with CAPEX!
- Because the data are kindly made available to the public by Prof. and Nobel laureate Robert Shiller and CAPE has been proposed by him and co-authors, this lab is inspired by his work
- Even though in this lab we use U.S. data only, there is ample international research on the forecasting power of CAPE for (inflation-adjusted) stock returns

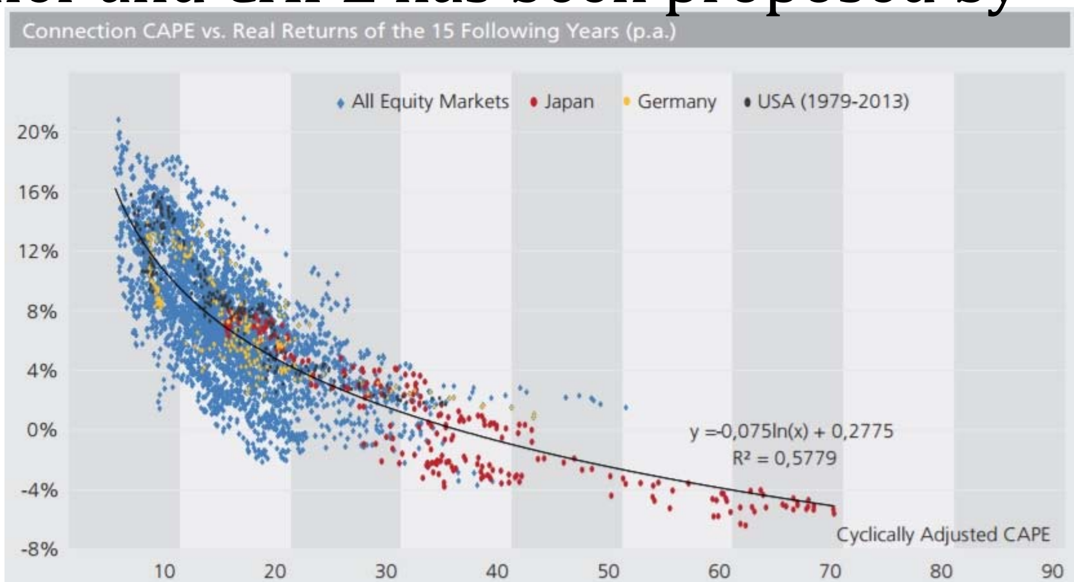
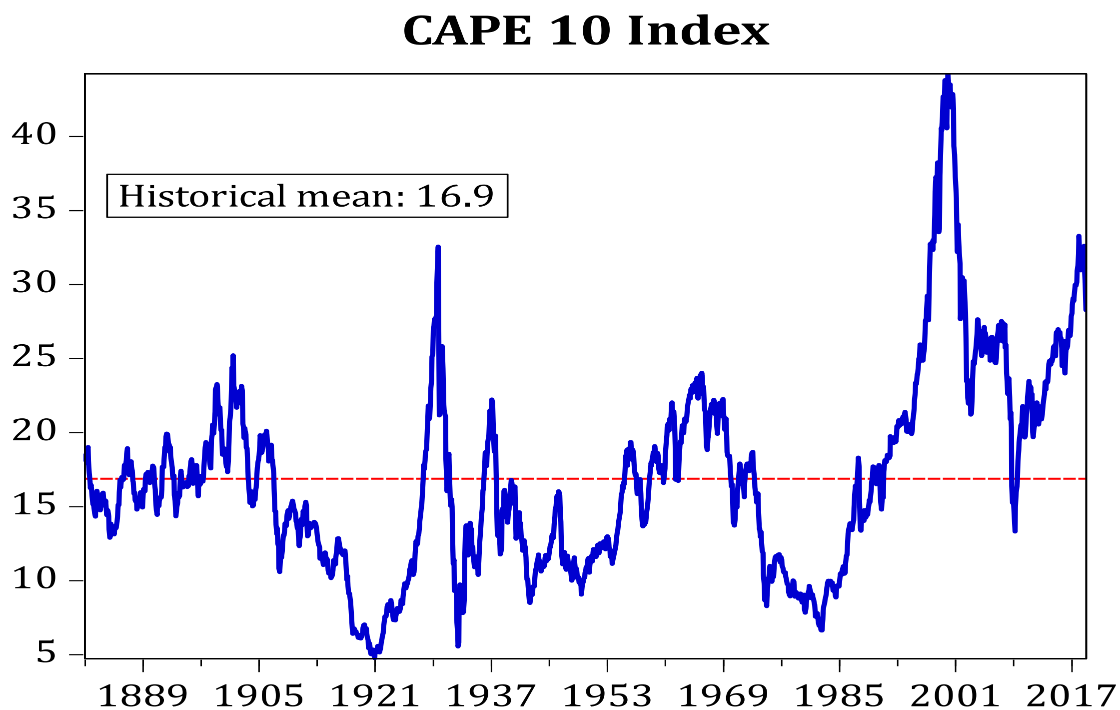


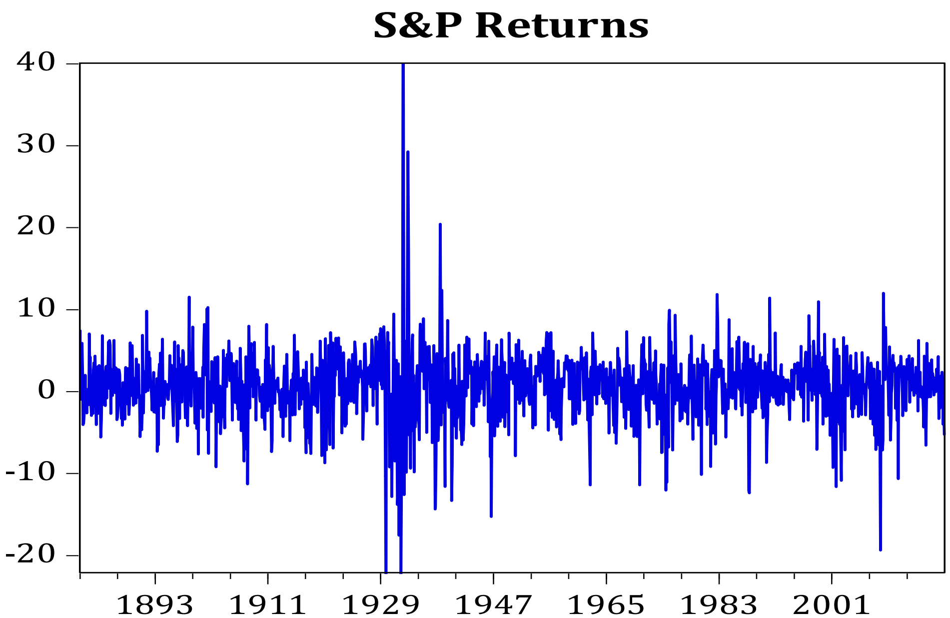
Figure 2: Connection between the CAPE and the returns of the 15 following years in the period 1881-2013 (US) and 1979-2013 (other markets). The USA, Japan and Germany are highlighted as examples in a single period 1979-2013. All returns inflation-adjusted, in local currency, incl. dividend income and annualized. Source: S&P 500: Robert J. Shiller, other countries: Worldscope, Thomson Reuters and own calculations.

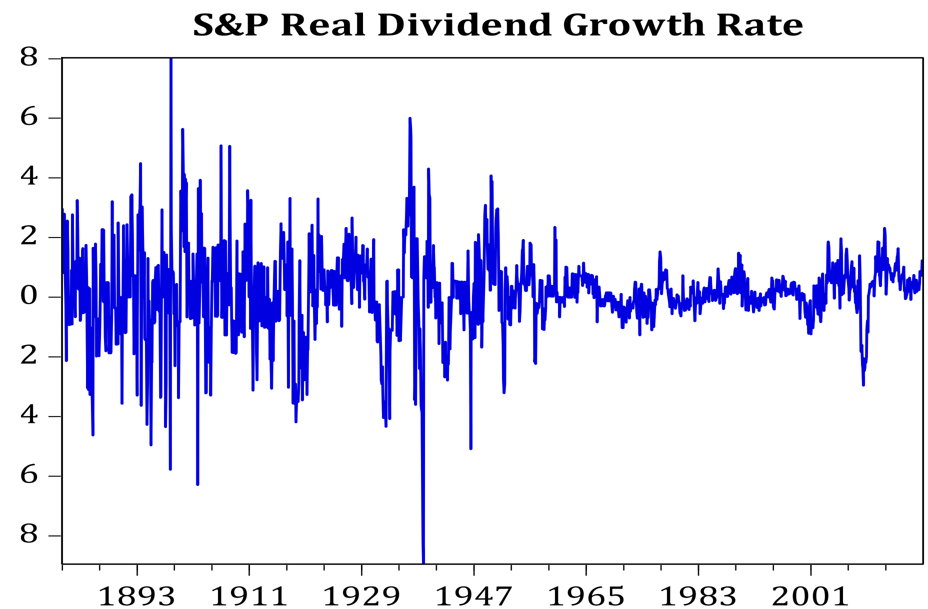
The Data

- We shall use a January 1881 – December 2018 monthly sample, for a total of 1,656 observations
  - Once more, not a big problem to obtain large samples in finances
  - We are on the verge of dealing with economic history here
- Three different series analyzed/commented
  - The discretely compounded (real) returns on the Standard & Poor's Composite index (that became S&P 500 in 1957) deflated using the CPI inflation index
  - The discretely compounded rate of growth of the (real) dividends paid by the companies in the S&P index, deflated using CPI inflation
  - The CAPE index, i.e., the ratio btw. the real S&P index and the average of the moving average of the last 10 years of real earnings reported by companies in the S&P index, a long-run, average real PE ratio
  - CAPE is use to forecast stock returns over timescales of 10 to 20 years, with higher than average CAPE values implying lower than average long-term annual average returns

Step 0: The Data

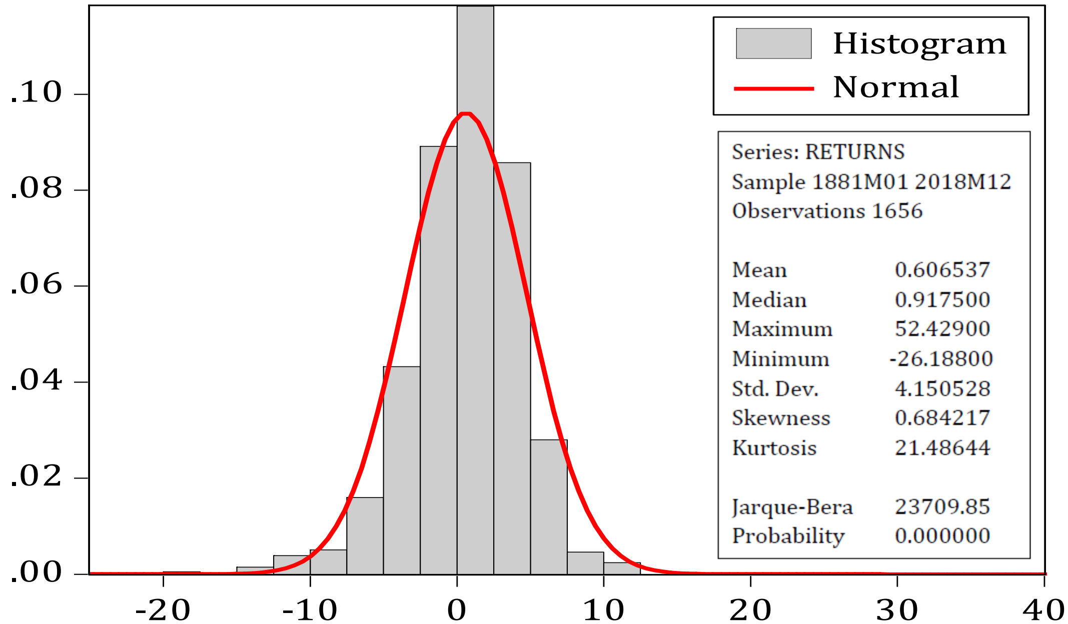




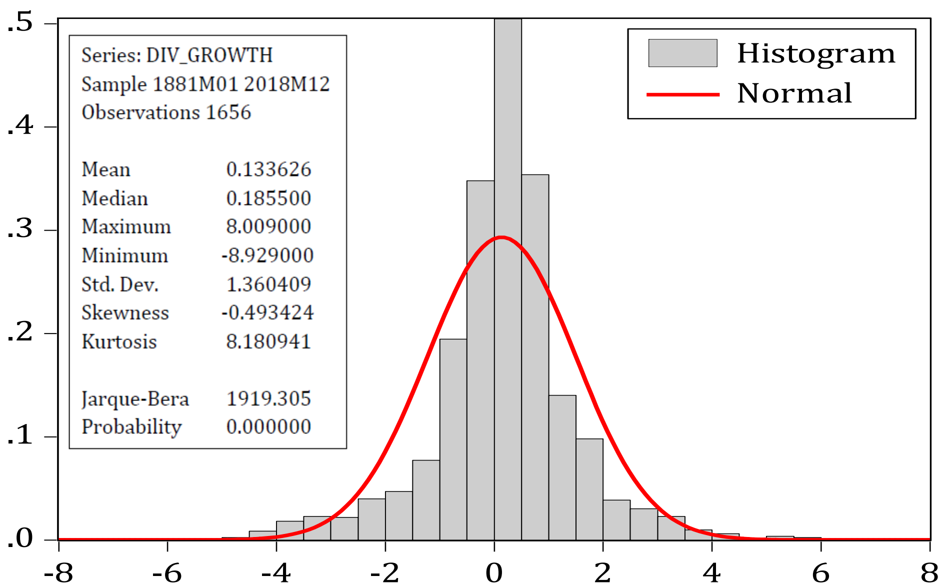


Step 0: The Data

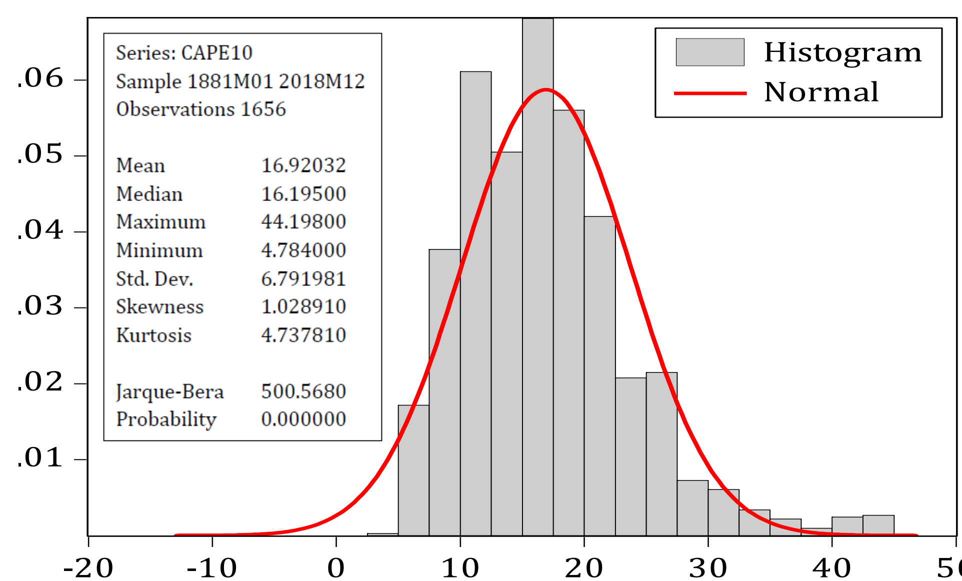
S&P Returns



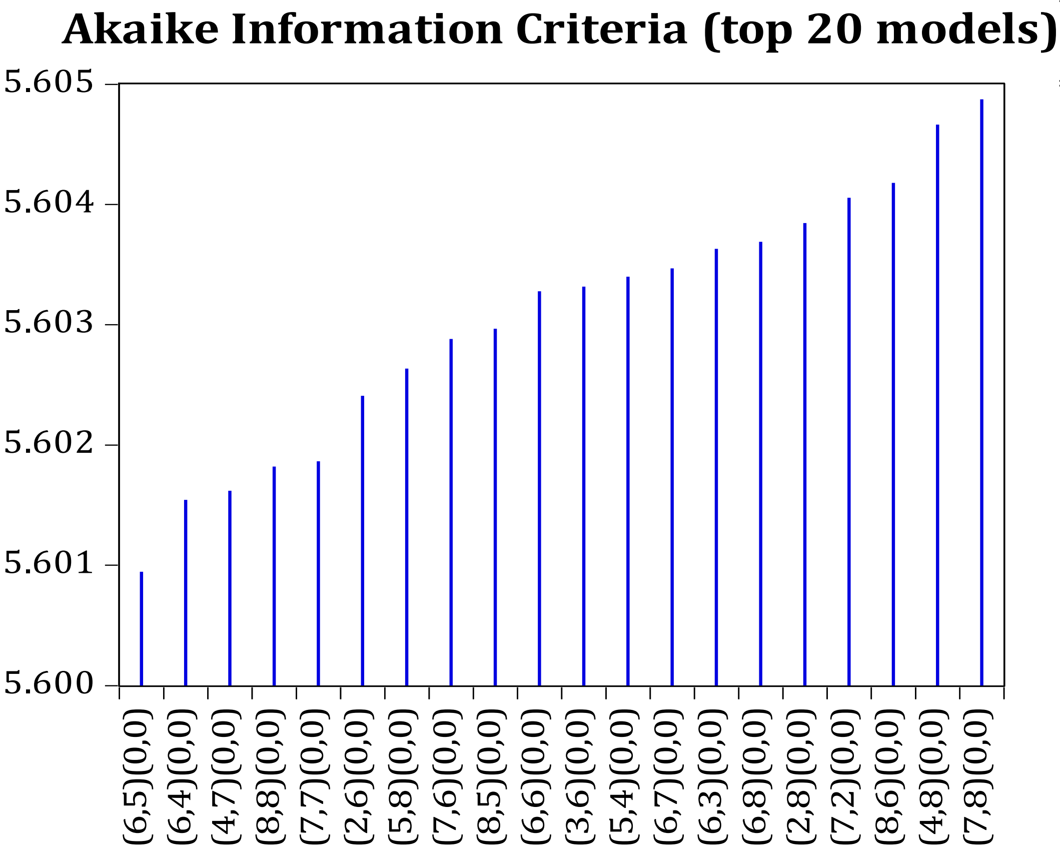
S&P Dividend Growth Rate



S&P CAPE10



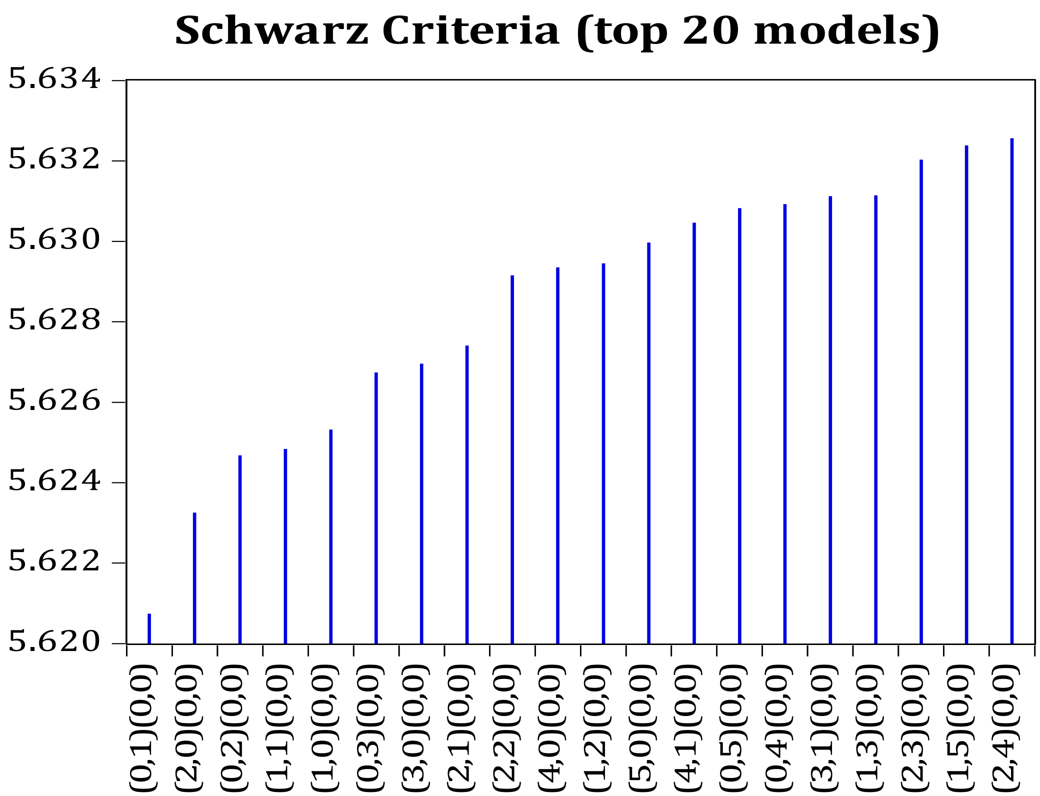
Step 1: Model Specification (S&P Returns)



Model	LogL	AIC*	BIC	HQ
(6,5)(0,0)	-4624.583957	5.600947	5.643434	5.616696
(6,4)(0,0)	-4626.078375	5.601544	5.640762	5.616082
(4,7)(0,0)	-4625.140904	5.601619	5.644106	5.617369
(8,8)(0,0)	-4620.308072	5.601821	5.660649	5.623628
(7,7)(0,0)	-4622.344438	5.601865	5.654157	5.621249
(2,6)(0,0)	-4628.795247	5.602410	$\boxed{5.635092}$	$\boxed{5.614525}$
(5,8)(0,0)	-4623.982693	5.602636	5.651659	5.620809
(7,6)(0,0)	-4624.186315	5.602882	5.651905	5.621055
(8,5)(0,0)	-4624.255452	5.602966	5.651989	5.621138
(6,6)(0,0)	-4625.515283	5.603279	5.649034	5.620240
(3,6)(0,0)	-4628.547532	5.603318	5.639269	5.616645
(5,4)(0,0)	-4628.615356	5.603400	5.639351	5.616727
(6,7)(0,0)	-4624.671032	5.603467	5.652491	5.621640
(6,3)(0,0)	-4628.808010	5.603633	5.639583	5.616959

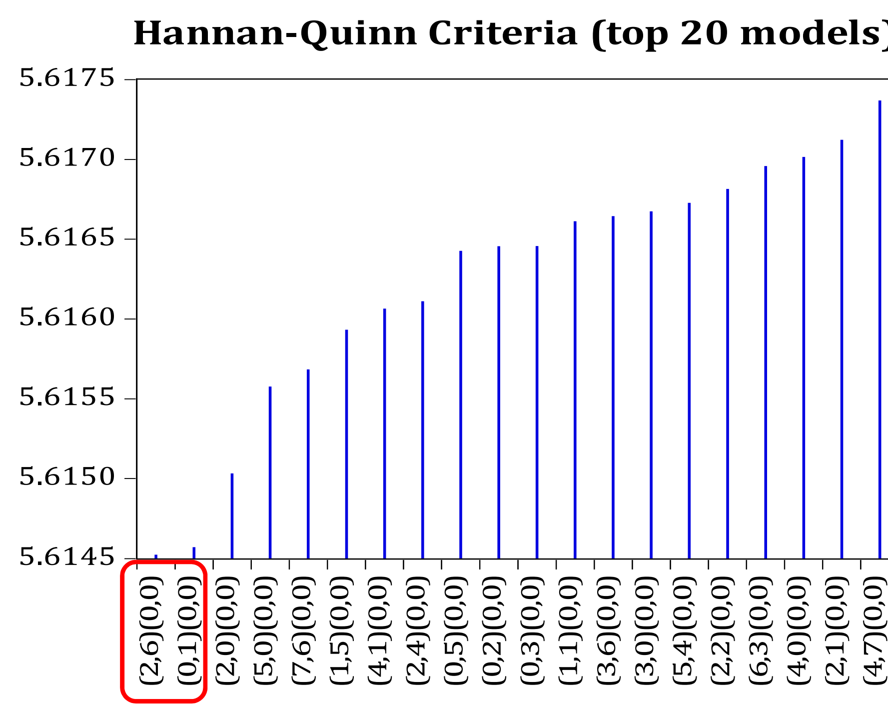
- A systematic search across ARIMA(p, d, q) with $p \leq 8$, $d \leq 1$, $q \leq 8$ using AIC as a selection criteria points to large ARMA(6, 5) model
  - However, the very table reveals that BIC and H-Q would lead to the selection of much smaller models
  - Such evidence in favor of large, rich ARMA models is unusual but we need to remember that we are using 138 years of data

Step 1: Model Specification (S&P Returns)



Model	LogL	AIC	BIC*	HQ
(0,1)(0,0)	-4642.855337	5.610936	5.620741	$\boxed{5.614571}$
(2,0)(0,0)	-4641.234039	5.610186	5.623259	5.615032
(0,2)(0,0)	-4642.412536	5.611609	5.624682	5.616455
(1,1)(0,0)	-4642.542116	5.611766	5.624839	5.616612
(1,0)(0,0)	-4646.647519	5.615516	5.625321	5.619151
(0,3)(0,0)	-4640.410543	5.610399	5.626740	5.616457
(3,0)(0,0)	-4640.591772	5.610618	5.626959	5.616676
(2,1)(0,0)	-4640.961736	5.611065	5.627406	5.617122
(2,2)(0,0)	-4638.704271	5.609546	5.629155	5.616815
(4,0)(0,0)	-4638.869992	5.609746	5.629356	5.617015
(1,2)(0,0)	-4642.659017	5.613115	5.629456	5.619172
(5,0)(0,0)	-4635.677053	$\boxed{5.607098}$	5.629975	5.615578

- In this case, the best ranked model is a simple MA(1)!
- If you repeat the sequential selection using H-Q you obtain another model, a ARMA(2, 6)
  - However, according to H-Q, also MA(1) ranks fairly high

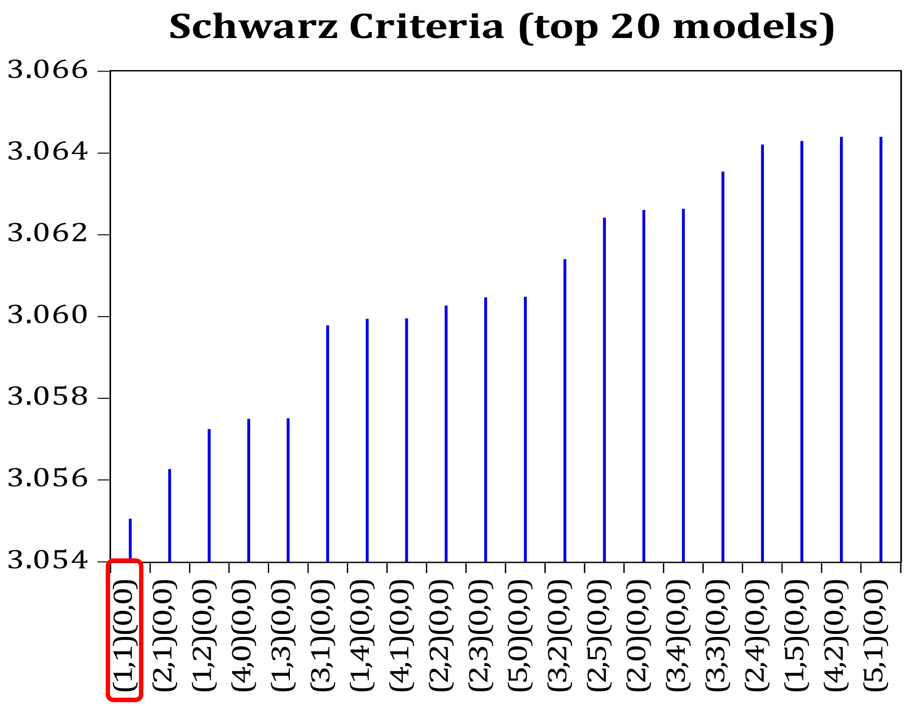


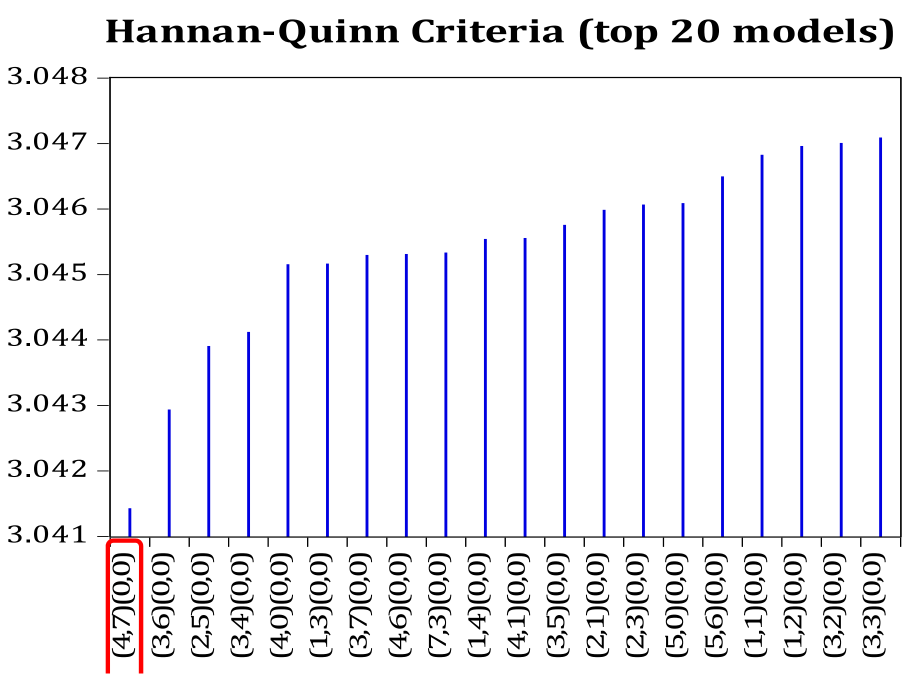
Step 1: Model Specification (S&P Returns)

- Time to take a look at the SACF/SPACF
  - The obvious insight is that these data may have been generated by a MA(1)
  - However, SACF is also (borderline) significant at lags 13-15 and 19-21, according to a sinusoidal pattern
  - Such a pattern also characterizes the SPACF and this points towards a ARMA with coefficients of alternating signs
  - This may be consistent with a ARMA(2, q) if the AR has complex roots

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.263	0.263	114.76	0.000
		2	-0.006	-0.080	114.81	0.000
		3	-0.049	-0.028	118.79	0.000
		4	0.022	0.046	119.56	0.000
		5	0.079	0.062	129.84	0.000
		6	0.038	-0.001	132.24	0.000
		7	0.027	0.026	133.42	0.000
		8	0.041	0.037	136.15	0.000
		9	0.031	0.011	137.78	0.000
		10	0.011	-0.002	137.97	0.000
		11	-0.005	-0.005	138.00	0.000
		12	-0.019	-0.021	138.64	0.000
		13	-0.062	-0.064	145.14	0.000
		14	-0.078	-0.055	155.22	0.000
		15	-0.068	-0.045	163.06	0.000
		16	-0.004	0.016	163.08	0.000
		17	0.023	0.014	163.96	0.000
		18	0.032	0.031	165.66	0.000
		19	-0.050	-0.055	169.89	0.000
		20	-0.096	-0.056	185.22	0.000
		21	-0.101	-0.060	202.47	0.000
		22	-0.038	-0.001	204.89	0.000
		23	0.008	0.011	205.00	0.000
		24	-0.004	-0.009	205.03	0.000

Step 1: Model Specification (S&P Dividend Growth)





Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.558	0.558	516.61	0.000
		2	0.437	0.183	833.89	0.000
		3	0.333	0.048	1018.5	0.000
		4	0.318	0.108	1186.2	0.000
		5	0.273	0.037	1309.8	0.000
		6	0.236	0.018	1402.1	0.000
		7	0.202	0.014	1469.8	0.000
		8	0.187	0.027	1528.3	0.000
		9	0.145	-0.022	1563.3	0.000
		10	0.103	-0.032	1581.2	0.000
		11	0.097	0.017	1596.9	0.000
		12	0.065	-0.028	1604.0	0.000
		13	0.044	-0.019	1607.3	0.000
		14	0.076	0.067	1616.9	0.000
		15	0.038	-0.040	1619.3	0.000
		16	0.049	0.021	1623.3	0.000
		17	0.029	-0.006	1624.7	0.000
		18	0.008	-0.032	1624.8	0.000
		19	-0.010	-0.021	1625.0	0.000
		20	-0.054	-0.068	1629.9	0.000
		21	-0.026	0.038	1631.0	0.000
		22	-0.047	-0.036	1634.8	0.000
		23	-0.055	-0.020	1639.8	0.000
		24	-0.093	-0.049	1654.3	0.000

- Also in this case there is considerable heterogeneity in the models selected by different ICs

Step 2: Estimation

Dependent Variable: RETURNS
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1881M01 2018M12
Convergence achieved after 56 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.608992	0.119383	5.101150	0.0000
AR(1)	1.857814	0.047859	38.81860	0.0000
AR(2)	-0.888484	0.048093	-18.47430	0.0000
MA(1)	-1.583632	0.052399	-30.22236	0.0000
MA(2)	0.381928	0.050780	7.521271	0.0000
MA(3)	0.174042	0.037710	4.615225	0.0000
MA(4)	0.111555	0.033423	3.337674	0.0009
MA(5)	0.024096	0.035079	0.686903	0.4922
MA(6)	-0.073754	0.019955	-3.696004	0.0002
SIGMASQ	15.67879	0.239330	65.51118	0.0000

Dependent Variable: DIV_GROWTH
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1881M01 2018M12
Convergence achieved after 346 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.109536	0.027732	3.949780	0.0001
AR(1)	0.438053	0.025266	17.33755	0.0000
AR(2)	0.853659	0.018479	46.19650	0.0000
AR(3)	0.550911	0.019118	28.81631	0.0000
AR(4)	-0.850299	0.023986	-35.44987	0.0000
MA(1)	0.000917	0.775694	0.001182	0.9991
MA(2)	-0.720756	0.766401	-0.940442	0.3471
MA(3)	-0.889527	0.276931	-3.212088	0.0013
MA(4)	0.465789	0.476619	0.977278	0.3286
MA(5)	0.113664	0.112426	1.011009	0.3122
MA(6)	0.097457	0.036346	2.681398	0.0074
MA(7)	-0.067542	0.073374	-0.920518	0.3574
SIGMASQ	1.185044	0.919163	1.289264	0.1975

- Below models selected by BIC:

Dependent Variable: RETURNS
Method: ARMA Maximum Likelihood (OPG - BHHH)
Sample: 1881M01 2018M12
Convergence achieved after 20 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.606841	0.126755	4.787531	0.0000
MA(1)	0.276194	0.018176	15.19576	0.0000
SIGMASQ	15.94887	0.212700	74.98278	0.0000

Dependent Variable: DIV_GROWTH

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.024928	0.017566	1.419119	0.1561
DIV_GROWTH(-1)	0.805743	0.014454	55.74437	0.0000
MA(1)	-0.382138	0.023630	-16.17144	0.0000
SIGMASQ	1.218382	0.020472	59.51380	0.0000

Step 3: Diagnostic Checks (S&P Returns)

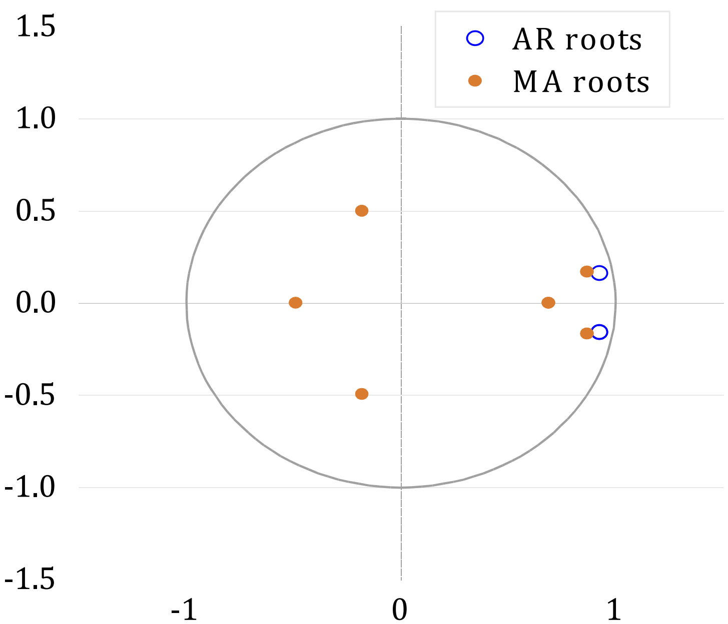
AR Root(s)	Modulus	Cycle
0.928907 ± 0.160050i	0.942594	36.82480

No root lies outside the unit circle.
ARMA model is stationary.

MA Root(s)	Modulus	Cycle
0.870085 ± 0.168875i	0.886322	32.77505
0.690517	0.690517	
-0.179409 ± 0.496281i	0.527714	3.276441
-0.488237	0.488237	

No root lies outside the unit circle.
ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.000	0.000	1.E-05	
		2	0.000	0.000	0.0004	
		3	0.003	0.003	0.0120	
		4	0.001	0.001	0.0135	
		5	-0.005	-0.005	0.0502	
		6	-0.014	-0.014	0.3604	
		7	-0.003	-0.003	0.3747	0.540
		8	0.020	0.020	1.0592	0.589
		9	0.019	0.019	1.6794	0.642
		10	-0.000	-0.000	1.6794	0.794
		11	-0.006	-0.006	1.7425	0.884
		12	-0.010	-0.010	1.9006	0.929
		13	-0.047	-0.047	5.6306	0.583
		14	-0.047	-0.046	9.3022	0.317
		15	-0.048	-0.048	13.211	0.153
		16	0.010	0.010	13.381	0.203
		17	0.010	0.010	13.551	0.259
		18	0.037	0.037	15.850	0.198
		19	-0.037	-0.038	18.133	0.153
		20	-0.055	-0.057	23.213	0.057
		21	-0.072	-0.074	31.996	0.006
		22	-0.024	-0.022	32.951	0.008
		23	0.006	0.011	33.011	0.011
		24	-0.007	-0.004	33.084	0.016

Bad luck?

- Further (JB) tests reveal that the residuals are not normally distributed, but never said they should be

Step 3: Diagnostic Checks (Dividend Growth Rate)

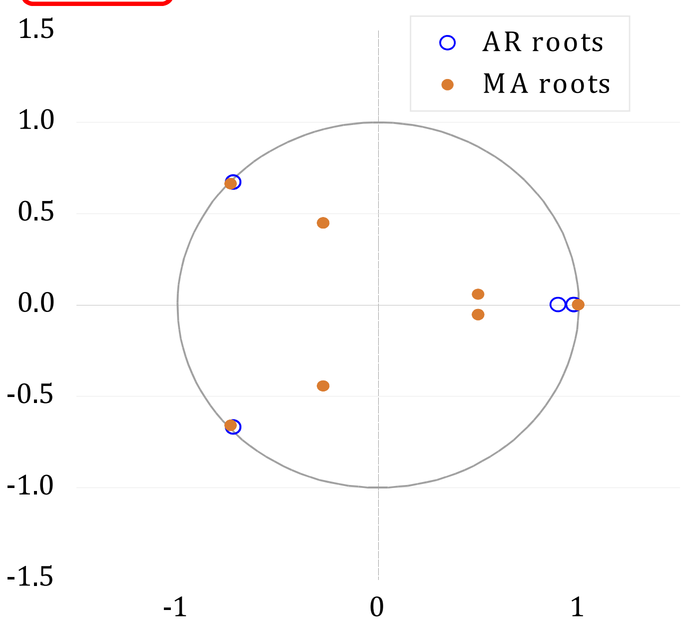
AR Root(s)	Modulus	Cycle
-0.719339 ± 0.670759i	0.983548	2.627709
0.977675	0.977675	
0.899055	0.899055	

No root lies outside the unit circle.
ARMA model is stationary.

MA Root(s)	Modulus	Cycle
1.000000	1.000000	
-0.732088 ± 0.662046i	0.987045	2.611039
-0.270294 ± 0.445755i	0.521302	2.969526
0.501923 ± 0.056391i	0.505081	56.15931

No root lies outside the unit circle.
ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)



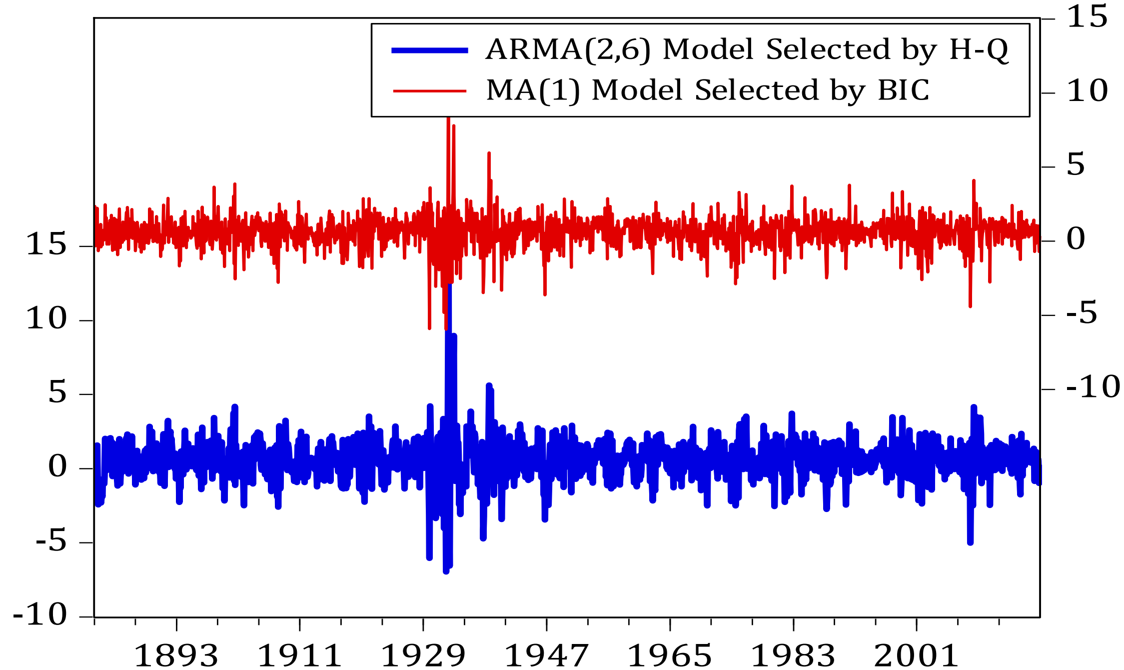
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.000	0.000	4.E-05	
		2	-0.001	-0.001	0.0023	
		3	0.004	0.004	0.0229	
		4	-0.001	-0.001	0.0240	
		5	0.001	0.001	0.0245	
		6	-0.023	-0.023	0.9285	
		7	0.016	0.017	1.3807	
		8	0.027	0.027	2.5775	0.108
		9	-0.003	-0.003	2.5946	0.273
		10	-0.015	-0.015	2.9578	0.398
		11	-0.001	-0.001	2.9587	0.565
		12	-0.005	-0.006	3.0081	0.699
		13	-0.039	-0.038	5.5607	0.474
		14	0.044	0.045	8.7708	0.270
		15	0.003	0.002	8.7849	0.361
		16	0.018	0.017	9.3492	0.406
		17	0.015	0.015	9.7239	0.465
		18	0.020	0.021	10.392	0.496
		19	-0.009	-0.011	10.517	0.571
		20	-0.053	-0.050	15.280	0.290
		21	0.026	0.027	16.431	0.288
		22	-0.022	-0.025	17.268	0.303
		23	0.034	0.033	19.154	0.261
		24	-0.064	-0.065	26.061	0.073

- Further (JB) tests reveal that the residuals are not normally distributed, but never said they should be

Step 4: Forecasting (S&P Returns)

- Even though the two sets of forecasts look very similar, there is some evidence that ARMA(2,6) performs best
- This despite the fact that ARMA(2,6) implies estimating 10 parameters instead of 3
- Now we ask: can we gain forecast power for stock returns from dividends or the CAPE index?

Comparing Forecasts of S&P Returns



Forecast: RETURNSF_ARMA
Actual: RETURNS
Forecast sample: 1881M01 2018M12
Adjusted sample: 1881M03 2018M12
Included observations: 1654

Root Mean Squared Error		3.956392
Mean Absolute Error		2.766042
Mean Abs. Percent Error		227.7741
Theil Inequality Coef.	0.708173	
Bias Proportion	0.000002	
Variance Proportion		0.532786
Covariance Proportion		0.467213
Theil U2 Coefficient		0.850339
Symmetric MAPE		137.6465

Forecast: RETURNSF MA(1)
Actual: RETURNS
Forecast sample: 1881M01 2018M12
Included observations: 1656

Root Mean Squared Error		3.993859
Mean Absolute Error		2.798597
Mean Abs. Percent Error		246.6588
Theil Inequality Coef.	0.732555	
Bias Proportion	0.000000	
Variance Proportion		0.582041
Covariance Proportion		0.417959
Theil U2 Coefficient		0.823977
Symmetric MAPE		140.4549

Unrestricted vs. Unrestricted VARs

- Let's build bivariate VAR(p) models to forecast stock returns
- Of course, in the process, we'll also forecast dividend growth rate

Lag	LogL	LR	AIC	SC	HQ
0	-7522.845	NA	9.132094	9.138656	9.134527
1	-7142.176	759.9527	8.674971	8.694657	8.682270
2	-7106.110	71.91234	8.636056	8.668867*	8.648222*
3	-7102.717	6.756488	8.636793	8.682729	8.653825
4	-7090.297	24.70484	8.626574	8.685635	8.648473
5	-7084.019	12.47320*	8.623809*	8.695994	8.650574
6	-7081.281	5.433102	8.625341	8.710650	8.656972
7	-7077.532	7.428617	8.625646	8.724080	8.662143
8	-7074.353	6.292698	8.626642	8.738201	8.668006

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

- While VAR(2) models are plausible in the light of earlier evidences, VAR(5) may represent a way to surrogate complex VARMA

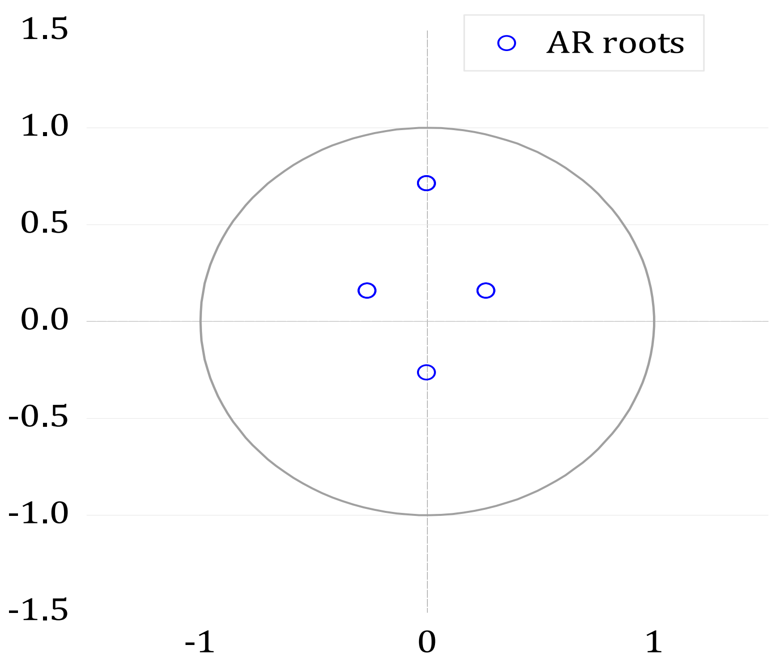
A VAR(2) Model

- Possible to present OLS, equation-by-equation estimates of a VAR(2)

Root	Modulus
0.710027	0.710027
0.155758 - 0.262144i	0.304926
0.155758 + 0.262144i	0.304926
-0.265956	0.265956

No root lies outside the unit circle.
VAR satisfies the stability condition.

Inverse Roots of VAR Polynomials



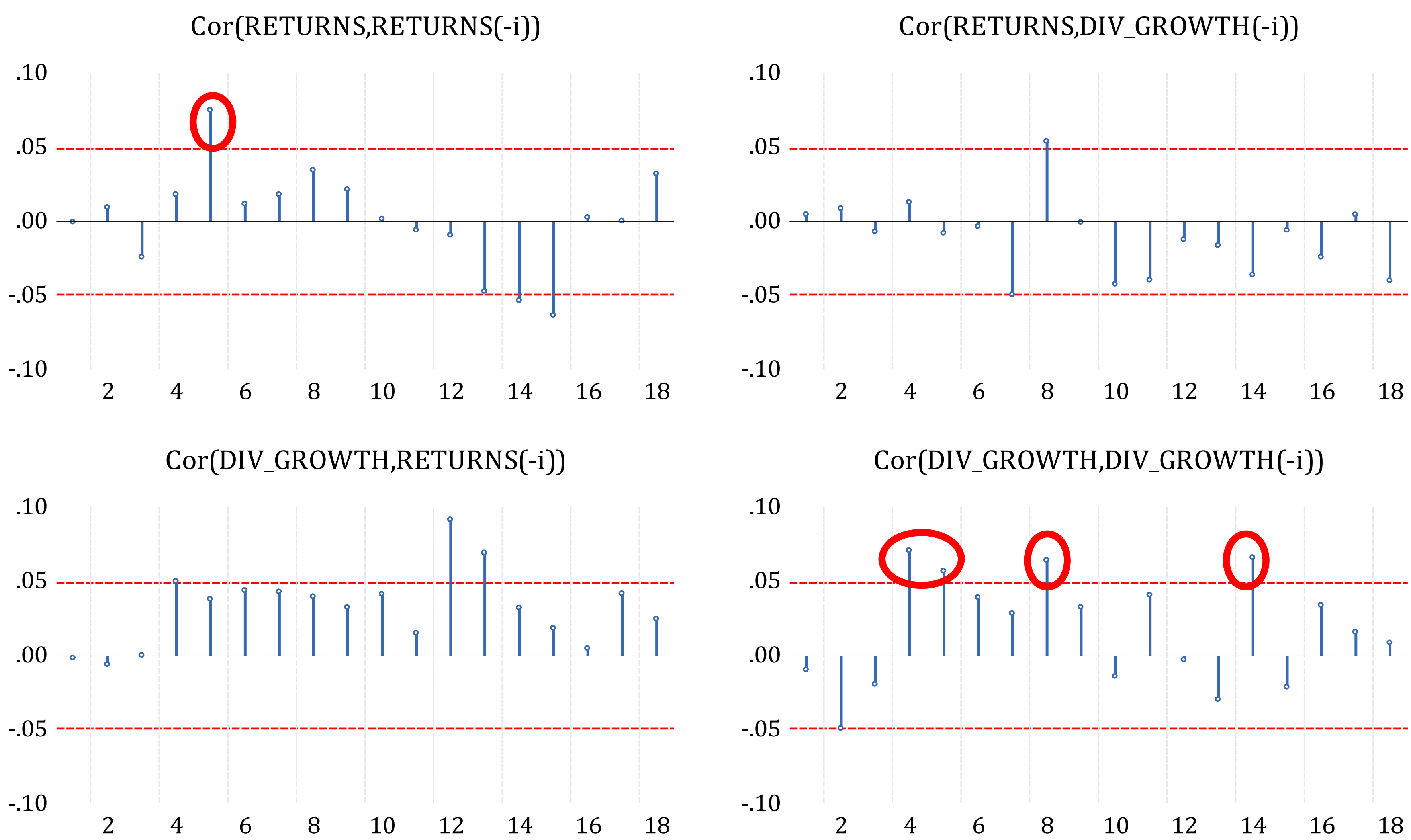
Vector Autoregression Estimates
Included observations: 1654 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	RETURNS	DIV_GROWTH
RETURNS(-1)	0.288218	-0.030126
	(0.02462)	(0.00681)
	[11.7049]	[-4.42054]
RETURNS(-2)	-0.089024	0.007453
	(0.02475)	(0.00685)
	[-3.59733]	[1.08810]
DIV_GROWTH(-1)	-0.120068	0.467370
	(0.08789)	(0.02432)
	[-1.36614]	[19.2141]
DIV_GROWTH(-2)	0.234069	0.177632
	(0.08761)	(0.02425)
	[2.67178]	[7.32602]
C	0.467934	0.060318
	(0.10015)	(0.02772)
	[4.67231]	[2.17612]
R-squared	0.079760	0.343184
Adj. R-squared	0.077528	0.341591
Log likelihood	-4631.285	-2506.573
Akaike AIC	5.606149	3.036969
Schwarz SC	5.622506	3.053326
Log likelihood		-7130.563
Akaike information criterion		8.634296
Schwarz criterion		8.667011
Number of coefficients		10

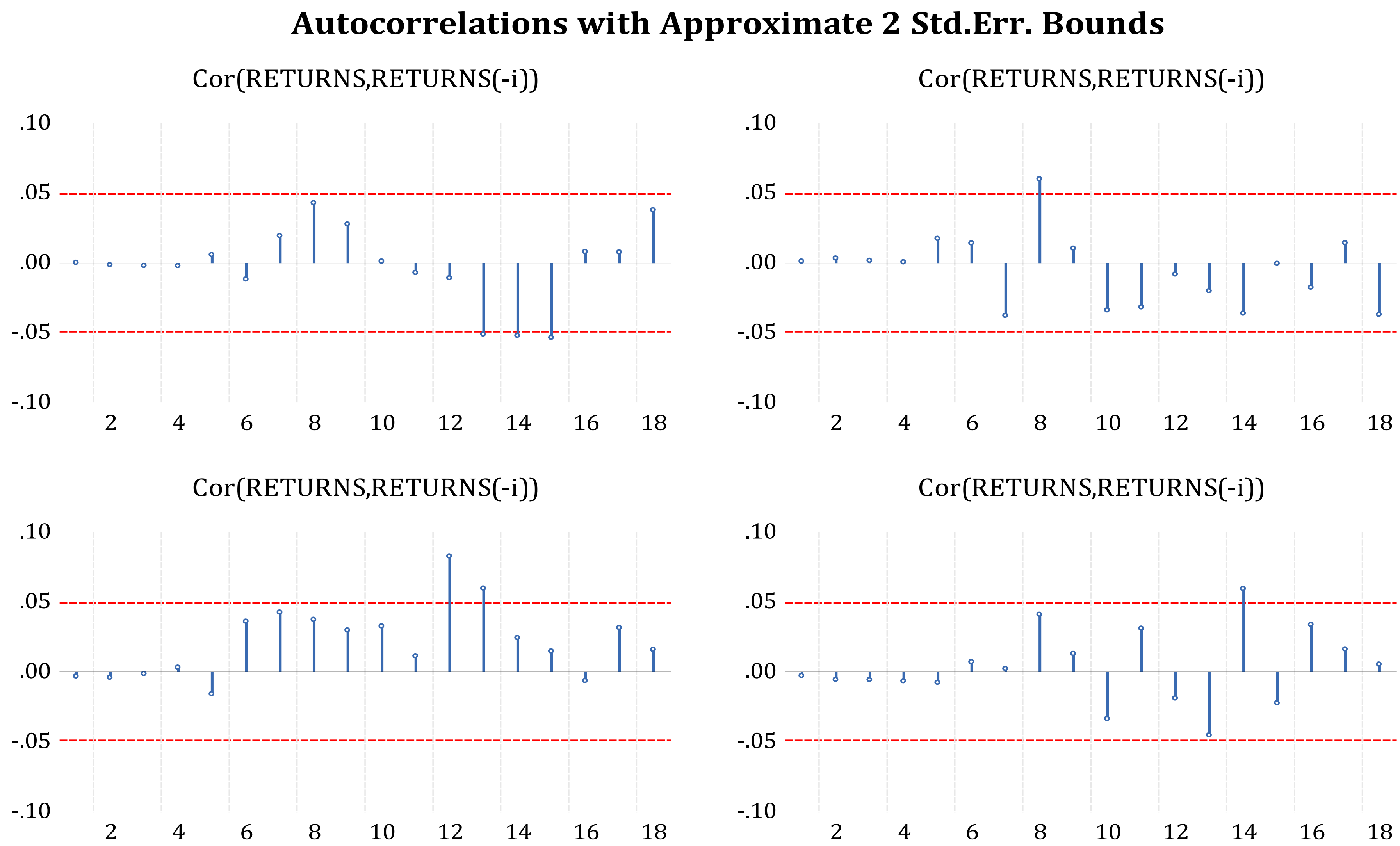
Model Diagnostics for the VAR(2)

- There are some residual concerns on the correct specification of the VAR(2) model, to which the VAR(5) puts remedy
  - Yet concerns more about the dividend growth series than for returns

Autocorrelations with Approximate 2 Std.Err. Bounds



Model Diagnostics for the VAR(5)



- Let’s now compare the forecast performance of VAR(2) and VAR(5)
- A VAR(5) outperforms, although not massively, a VAR(2), i.e., not always smaller models outperform (not with 1,656 obs.)

Forecasting Performance of VAR(5) vs VAR(2)

Forecast Evaluation: VAR(2)

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DIV_GROWTH	1656	1.101335	0.696515	408.6220	0.507025
RETURNS	1656	3.979327	2.788962	1403.930	0.722252

Forecast Evaluation: VAR(5)

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DIV_GROWTH	1656	1.090412	0.689539	600.2749	0.499479
RETURNS	1656	3.965169	2.775006	1271.105	0.713692

Forecast: RETURNSF_ARMA
Actual: RETURNS
Forecast sample: 1881M01 2018M12
Adjusted sample: 1881M03 2018M12
Included observations: 1654

Root Mean Squared Error	3.956392
Mean Absolute Error	2.766042
Mean Abs. Percent Error	227.7741
Theil Inequality Coef.	0.708173
Bias Proportion	0.000002
Variance Proportion	0.532786
Covariance Proportion	0.467213
Theil U2 Coefficient	0.850339
Symmetric MAPE	137.6465

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

- However, by inspecting the VAR(5) estimates, we detect 13 estimated coefficients out of 22 that fail to be statistically significant in tests of size 5% or less
  - We record slight worsening vs. ARMA(2,6): multivariate may not pay!
- Can we estimate restricted VAR models imposing restrictions?
  - The answer is yes, of course, but by using MLE

A Restricted VAR(5)

- The model estimated is:

RETURNS = C(1)+C(2)*RETURNS(-1)+C(3)*RETURNS(-5)+C(4)*DIV_GROWTH(-2)

DIV_GROWTH = C(6)*RETURNS(-1)+C(7)*DIV_GROWTH(-1)+C(8)*DIV_GROWTH(-2)+C(9)*DIV_GROWTH(-4)

- The ML estimation output looks as follows

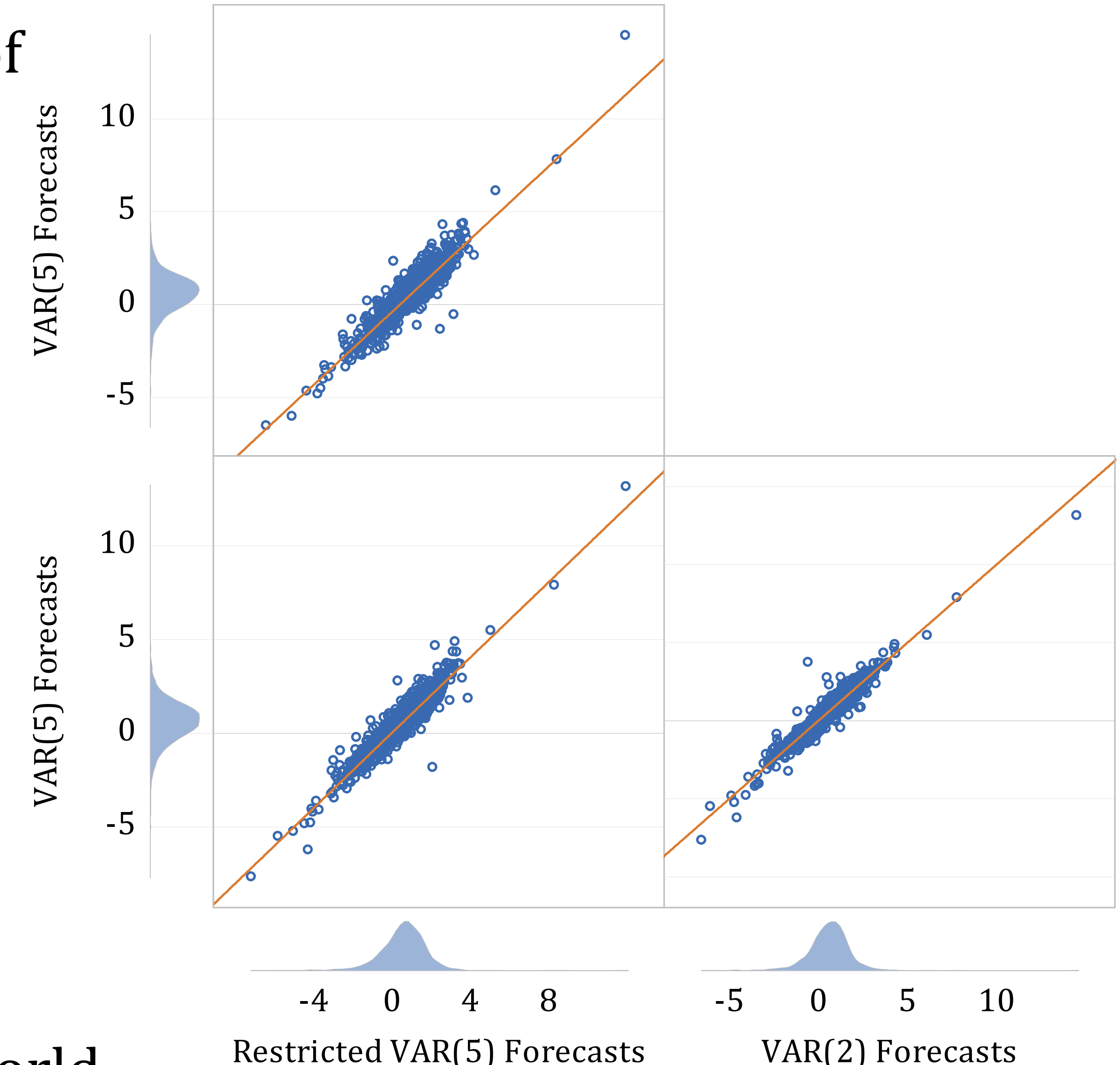
Estimation Method: Full Information Maximum Likelihood (BFGS / Marquardt steps)
Sample: 1881M06 2018M12
Included observations: 1651
Convergence achieved after 0 iterations
Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.363588	0.102291	3.554451	0.0004
C(2)	0.262280	0.019393	13.52419	0.0000
C(3)	0.068010	0.020928	3.249664	0.0012
C(4)	0.145699	0.067321	2.164234	0.0304
C(6)	-0.026285	0.006438	-4.083030	0.0000
C(7)	0.454493	0.014314	31.75270	0.0000
C(8)	0.139912	0.016128	8.674920	0.0000
C(9)	0.110144	0.017475	6.302753	0.0000

Log likelihood	-7111.650	Schwarz criterion	8.650862
Avg. log likelihood	-2.153740	Hannan-Quinn criter.	8.634369
Akaike info criterion	8.624652		

A Restricted VAR(5)

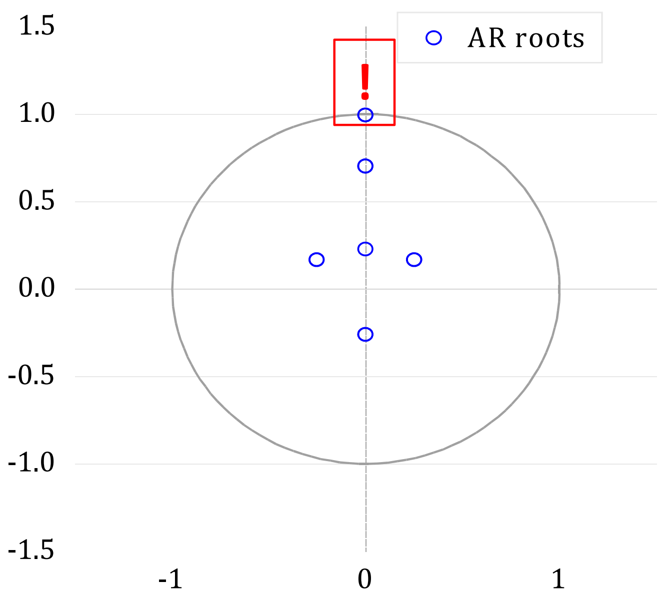
- The resulting forecasts of S&P returns are similar but not completely iden-tical
- As a last step, given the glamour around its role in valuation theory, we try to augment the VAR model with CAPE10
- Surprisingly, we find that the one for CAPE almost represent an autonomous regression equation
- CAPE seems to be in a world apart – how is that possible?



Expanding the VAR models to include CAPE

- CAPE10 just depends on its own lags and (weakly) on returns at t-2
- Neither stock returns or dividend growth depends on CAPE
- Yet, the R-square of CAPE exceeds 99%
- Finally, the VAR is close to being non-stationary: what is going on?
- Stay tuned!

Inverse Roots of VAR Polynomials



	RETURNS	DIV_GROWTH	CAPE10
RETURNS(-1)	0.334065	-0.029625	0.012443
	(0.05457)	(0.01511)	(0.00925)
	[6.12169]	[-1.96044]	[1.34460]
RETURNS(-2)	-0.085545	0.006456	-0.009375
	(0.02476)	(0.00686)	(0.00420)
	[-3.45438]	[0.94141]	[-2.23243]
DIV_GROWTH(-1)	-0.101673	0.463108	-0.014618
	(0.08803)	(0.02438)	(0.01493)
	[-1.15491]	[18.9966]	[-0.97914]
DIV_GROWTH(-2)	0.248219	0.173840	0.038406
	(0.08769)	(0.02428)	(0.01487)
	[2.83079]	[7.15928]	[2.58278]
CAPE10(-1)	-0.329965	0.003639	1.187009
	(0.32181)	(0.08912)	(0.05457)
	[-1.02533]	[0.04083]	[21.7501]
CAPE10(-2)	0.294168	0.005574	-0.193320
	(0.32208)	(0.08919)	(0.05462)
	[0.91333]	[0.06249]	[-3.53935]
C	1.041243	-0.094084	0.106540
	(0.26684)	(0.07389)	(0.04525)
	[3.90212]	[-1.27325]	[2.35436]
R-squared	0.083655	0.345235	0.990175
Adj. R-squared	0.080317	0.342849	0.990139
Log likelihood		-7457.888	
Akaike information criterion		9.043395	
Schwarz criterion		9.112095	
Number of coefficients		21	

Università Commerciale
Luigi Bocconi

Lab 3: Debriefing

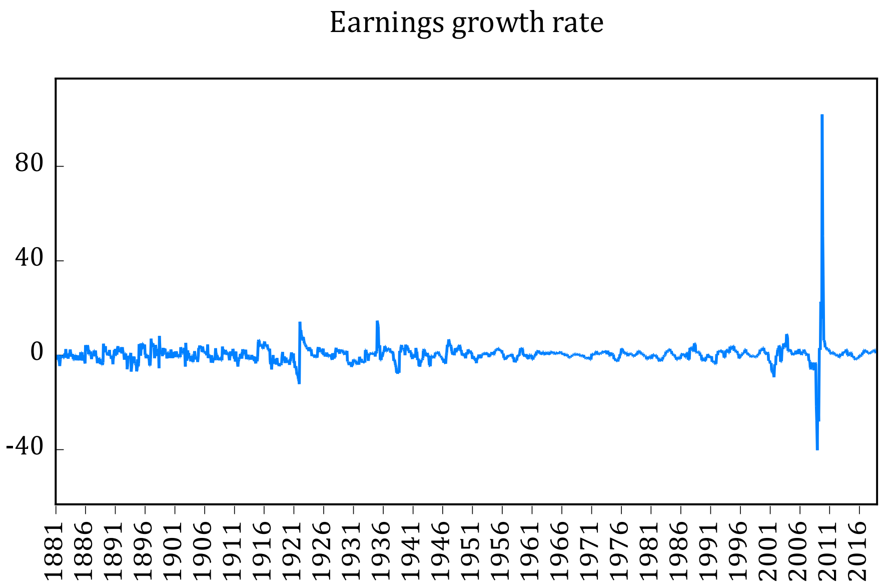
(Reminder: this is optional)

Prof. Massimo Guidolin

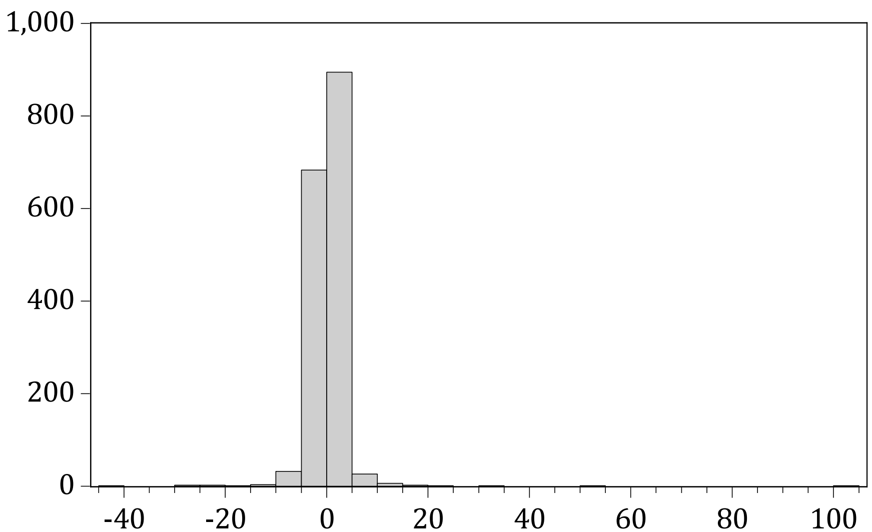
20192– Financial Econometrics

Winter/Spring 2019

A new variable: the earnings growth rate



- Very large spike in the aftermath of the financial crisis
- The data are very far from being normal (both because of positive skewness and massive kurtosis)



Series: EARN_GROWTH
Sample 1881M01 2018M12
Observations 1656

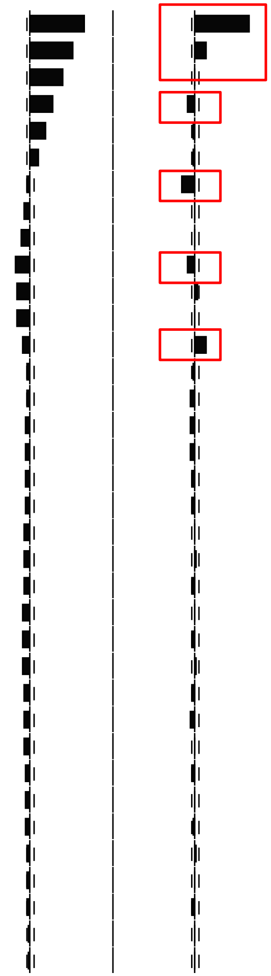
Mean	0.217824
Median	0.245000
Maximum	101.9590
Minimum	-40.42200
Std. Dev.	4.197110
Skewness	8.979636
Kurtosis	234.1309
Jarque-Bera	3708338.
Probability	0.000000

A first inspection to the data generating process

- A simple SACF/SPACF analysis applied to the earnings growth rate shows that
  - the series is stationary
  - but shows persistence
  - SACF decays slowly towards zero (displaying a sinusoidal pattern)
  - Interpretation of the SPACF is less clear
  - From the inspection of the correlogram it is difficult to distinguish whether data come from an AR or an ARMA

	ACF	PACF
AR(p)	Decays towards zero.	Cuts off after lag p.
MA(q)	Cuts off after lag q.	Decays towards zero.
ARMA(p, q)	Decays towards zero starting at lag q.	Decays towards zero starting at lag p.

Date: 03/11/19 Time: 18:13
Sample: 1881M01 2018M12
Included observations: 1656

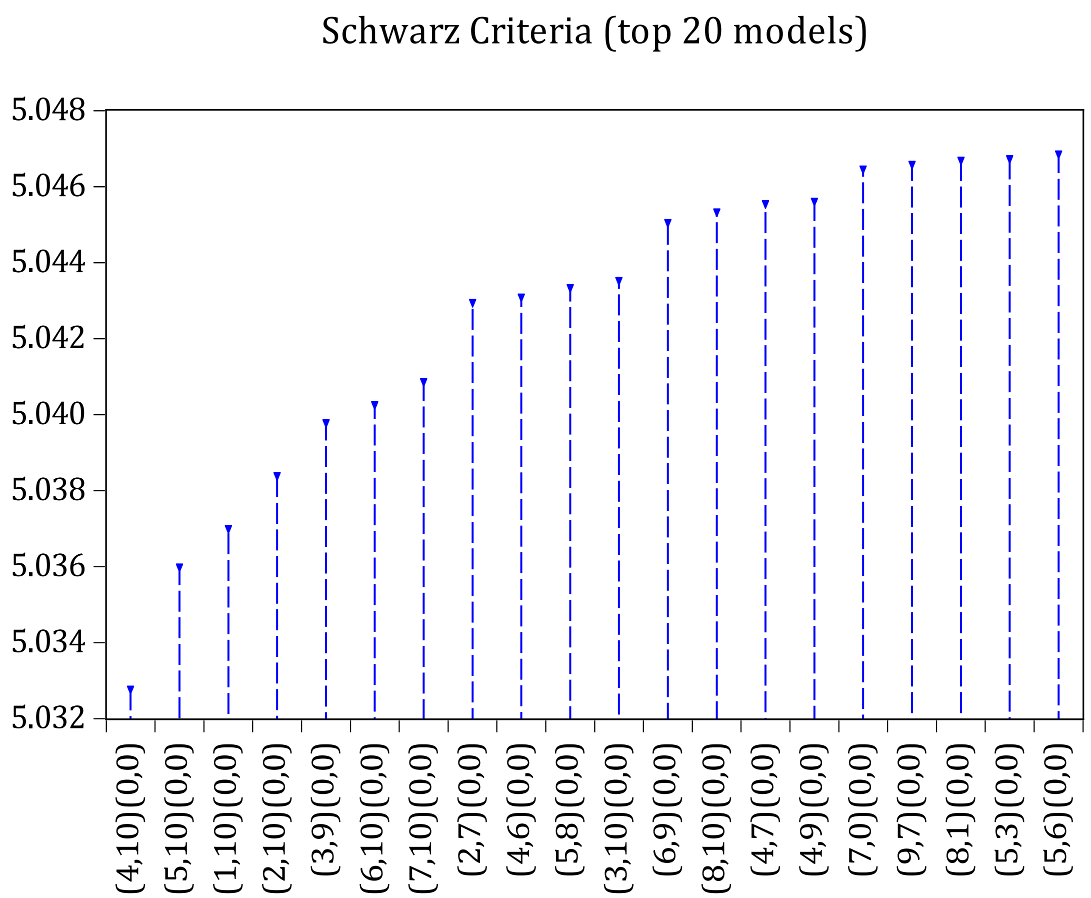


	AC	PAC	Q-Stat	Prob
1	0.689	0.689	787.79	0.000
2	0.546	0.136	1283.2	0.000
3	0.421	0.004	1577.9	0.000
4	0.282	-0.086	1710.3	0.000
5	0.190	-0.021	1770.0	0.000
6	0.112	-0.023	1791.0	0.000
7	-0.025	-0.170	1792.1	0.000
8	-0.078	-0.008	1802.2	0.000
9	-0.112	0.003	1822.9	0.000
10	-0.178	-0.091	1875.6	0.000
11	-0.169	0.026	1923.2	0.000
12	-0.162	0.006	1966.9	0.000
13	-0.084	0.143	1978.7	0.000
14	-0.047	-0.017	1982.4	0.000
15	-0.039	-0.057	1985.0	0.000
16	-0.049	-0.058	1989.0	0.000
17	-0.055	-0.052	1994.1	0.000
18	-0.064	-0.032	2001.0	0.000
19	-0.066	-0.032	2008.2	0.000
20	-0.069	0.003	2016.2	0.000
21	-0.071	0.009	2024.6	0.000
22	-0.078	-0.032	2034.9	0.000
23	-0.085	-0.002	2047.0	0.000
24	-0.095	-0.026	2062.3	0.000
25	-0.085	0.019	2074.5	0.000
26	-0.080	-0.034	2085.2	0.000
27	-0.079	-0.049	2095.8	0.000
28	-0.066	-0.008	2103.2	0.000
29	-0.064	-0.025	2110.1	0.000
30	-0.055	0.001	2115.3	0.000
31	-0.051	-0.017	2119.6	0.000
32	-0.041	0.009	2122.4	0.000
33	-0.035	-0.003	2124.5	0.000
34	-0.028	-0.027	2125.8	0.000
35	-0.021	-0.000	2126.6	0.000
36	-0.011	-0.002	2126.8	0.000

Select a model for the earnings growth rate

Automatic ARIMA Forecasting
Selected dependent variable: EARN_GROWTH
Date: 03/11/19 Time: 19:37
Sample: 1881M01 2018M12
Included observations: 1656
Forecast length: 0

Number of estimated ARMA models: 121
Number of non-converged estimations: 0
Selected ARMA model: (4,10)(0,0)
SIC value: 5.03276797658



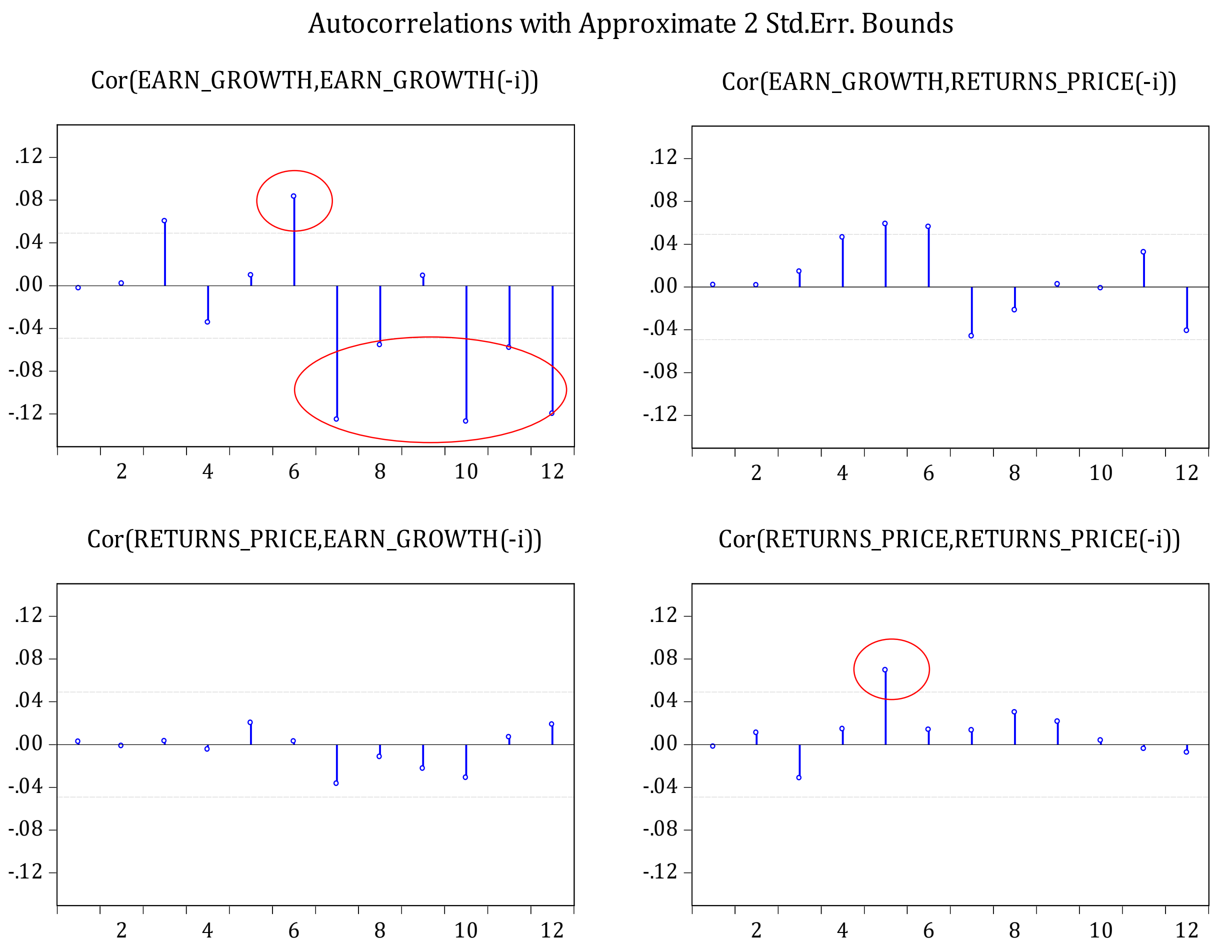
- According to the BIC (typically the most parsimonious of the criteria) we have to estimate a rather rich ARMA(4, 10) model (which is the same selected by HQ, as well)
- Note that the AIC would have selected a ARMA(7, 10) model

A VAR model for returns and earnings growth

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-9316.241	NA	303.6084	11.39149	11.39809	11.39394
1	-8730.867	1168.601	149.1589	10.68077	10.70057	10.68811
2	-8703.543	54.48117	144.9660	10.65225	10.68526*	10.66450
3	-8702.222	2.630680	145.4415	10.65553	10.70174	10.67267
4	-8692.095	20.14210	144.3563	10.64804	10.70745	10.67008
5	-8687.128	9.868019	144.1857	10.64686	10.71947	10.67379
6	-8685.399	3.429990	144.5866	10.64963	10.73545	10.68146
7	-8657.795	54.70121	140.4742	10.62078	10.71980	10.65751*
8	-8656.490	2.584110	140.9377	10.62407	10.73630	10.66570
9	-8656.014	0.941395	141.5462	10.62838	10.75381	10.67490
10	-8648.834	14.17531	140.9972	10.62449	10.76312	10.67591
11	-8646.891	3.831758	141.3523	10.62701	10.77884	10.68332
12	-8645.930	1.891653	141.8786	10.63072	10.79576	10.69194
13	-8625.233	40.71004	139.0122	10.61031	10.78855	10.67642
14	-8620.210	9.868864*	138.8386	10.60906	10.80050	10.68007
15	-8615.527	9.189035	138.7229*	10.60822*	10.81287	10.68413
16	-8611.887	7.132991	138.7842	10.60866	10.82651	10.68947
17	-8609.548	4.577627	139.0666	10.61069	10.84175	10.69640
18	-8607.672	3.666756	139.4286	10.61329	10.85755	10.70389
19	-8604.135	6.905630	139.5078	10.61386	10.87131	10.70935
20	-8600.940	6.230762	139.6455	10.61484	10.88550	10.71523

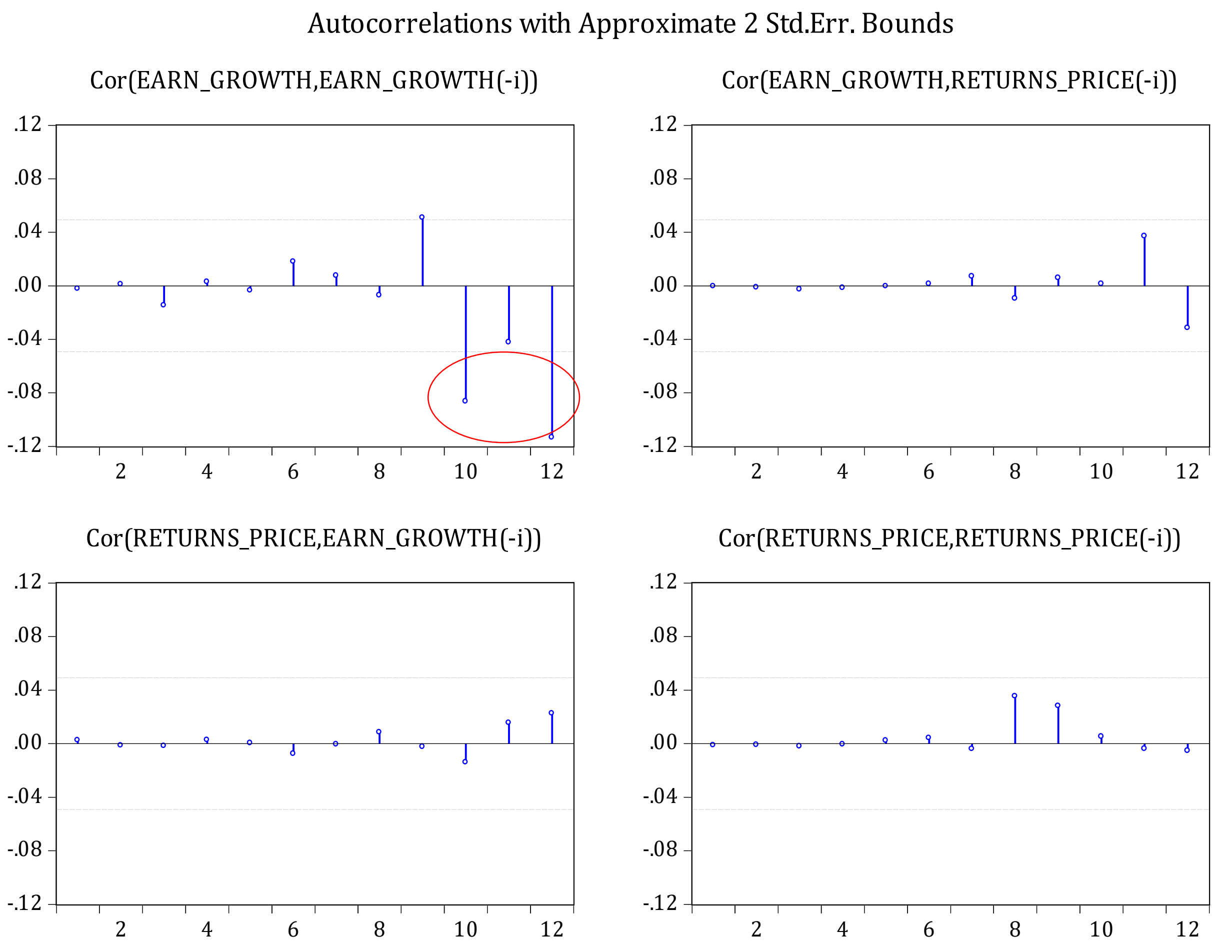
- As often happens, different criteria lead to different conclusions: the parsimonious SC chooses a VAR(2), while the AIC and HQ point towards more richly parametrized models

Model diagnostic for the VAR(2)



- Similar to the case that we have analyzed before, there are some concerns about the model specification
- HINT: try with the more richly parametrized one and compare

Model diagnostic for the VAR(7)



- Indeed VAR(7) seem to do much better, even if it does not solve the problem
- But which one has the best predictive accuracy?

Forecasting performance

VAR(2) Model – Forecasting performance

Variable	Inc. obs.	RMSE	MAE	MAPE
EARN_GROWTH	1656	3.004112	1.101981	417.0383
RETURNS_PRICE	1656	3.977785	2.790286	3288.828

VAR(7) Model – Forecasting performance

Variable	Inc. obs.	RMSE	MAE	MAPE
EARN_GROWTH	1656	2.936189	1.171655	2008.540
RETURNS_PRICE	1656	3.960802	2.774881	1478.742

- Similar forecasting performance of the two models

VAR(2) Model – Forecasts