Demand forecasting using complex-valued autoregressive models

Sergey Svetunkov, Ivan Svetunkov

OR60

13/09/2018

POLYTECH
Peter the Great St.Petersburg Polytechnic University

Marketing Analytics & Forecasting

Lancaster University Management School

Introduction

- There are many pairs of interchangeable and complementary groups of products in retail:
  - Cheese and wine,
  - Cleanser and sponge,
  - Paint and brush,
  - etc.
- Sales of one group might influence sales of the other.

Motivation

- A possible decision in this context – vector models.
- For example, VAR model.
- It is flexible enough to model the dependencies,
- But it’s hard to estimate.
- In case of a pair of products VAR(p) has at least 4p parameters.
- When the sample is small, it might be difficult to fit VAR(p),
- A more compact model is needed.

Literature review

- Literature on VAR in retail forecasting is very sparse…
- Caines et al. (1980) estimate and analyse VAR model for a supermarket.
- Leeflang & Selva (2012) discuss effects of promotions in one categories on the revenues in the other ones.

Literature review

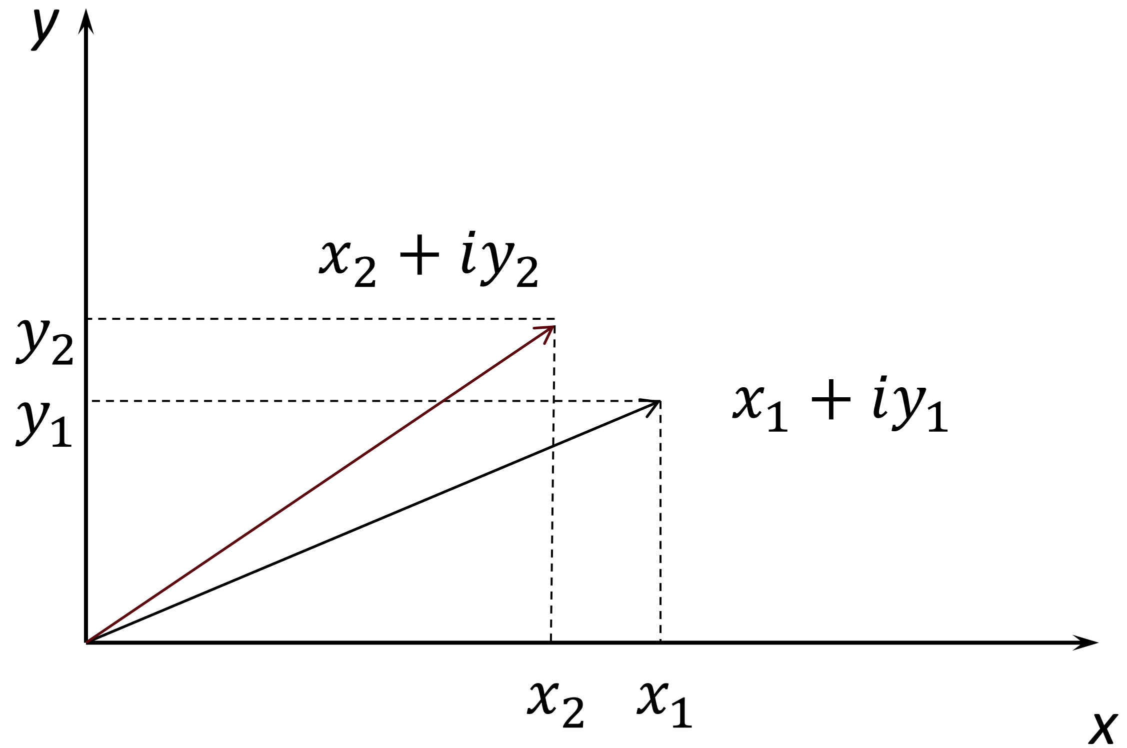
- Wilms et al. (2016) claim that "Retailers use VAR model as a standard tool to estimate the effects of prices, promotions and sales…"
- They jointly estimate cross-category effects for several VAR.
- But they focus on clustering and networks…
- Gelper et al. (2016) discuss a methodology to estimate parsimonious product category network, focusing on cross-category effects.

- So in general the idea is to estimate big VAR using some tricks…
- …And not much in forecasting direction.

Complex-valued approach

- We can use complex variables in this context.
- Variables x_t and y_t can be represented as a complex variable $x_t + iy_t$ when:
  - They have the same units,
  - They represent two parts of one process,
  - Their sum is another characteristic.
- Example: Sales of paint + Sales of brushes in pounds

Complex-valued approach



$$x_2 + iy_2 = (a_1 + ib_1)(x_1 + iy_1)$$

Complex Autoregressive Model

- The general CAR(p) can be written as:
  - $x_t + iy_t = \sum_{j=1}^{p}(a_j + ib_j)(x_{t-j} + iy_{t-j}) + e_{1,t} + ie_{2,t}$
    - where $x_t + iy_t$ is a complex variable of sales of two product groups,
    - $a_j + ib_j$ is a complex parameter of the model,
    - $e_{1,t} + ie_{2,t}$ is a complex error term.
    - It can be assumed that $e_{1,t} + ie_{2,t} \sim N(0, \Sigma)$, where Σ is a covariance matrix.

CAR(1)

The properties of these models differ from those of vector models.
For example CAR(1):

$$\hat{x}_{t+1} + i\hat{y}_{t+1} = (a_1 + ib_1)(x_t + iy_t)$$

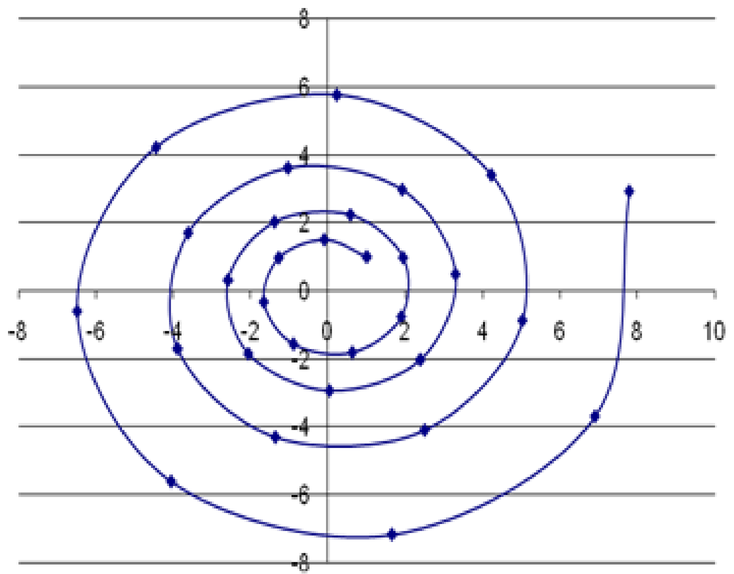
can be represented as a power function of horizon:

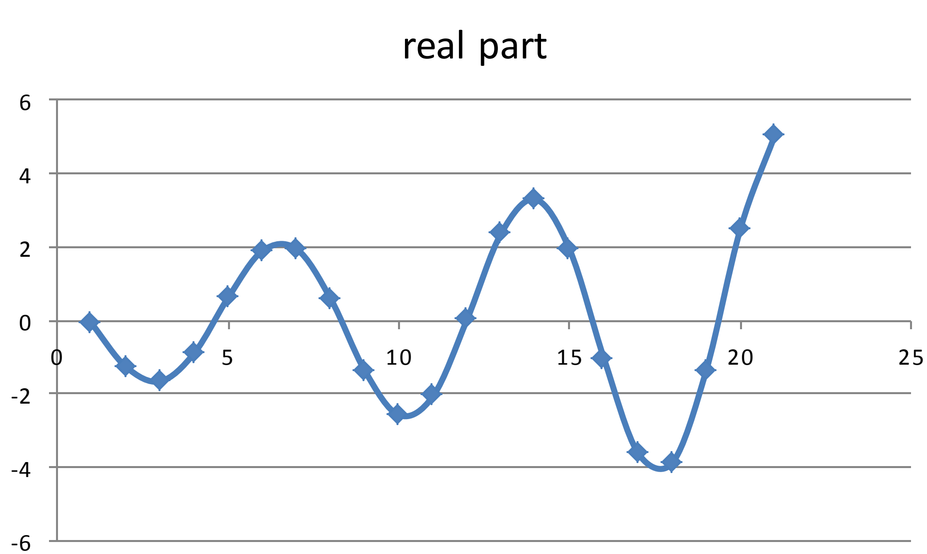
$$\hat{x}_{t+h} + i\hat{y}_{t+h} = (a_1 + ib_1)^h(x_t + iy_t)$$

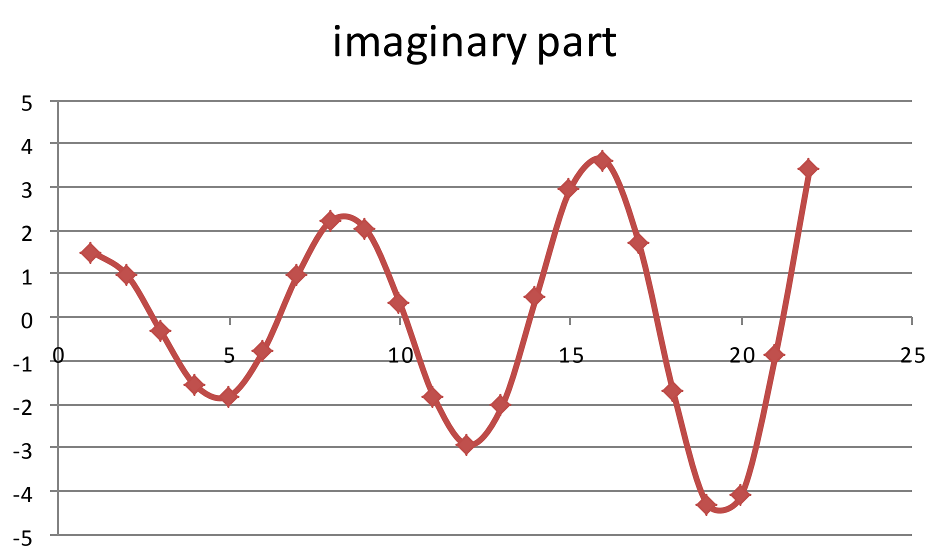
Power complex-value function:

- diverges in the form of a spiral if the magnitude of the complex variable is higher than one,
- converges along the spiral to zero if the magnitude of the complex variable is lower than one.

CAR(1) – graphical representation







Complex Autoregressive Model

- CAR can also be represented in the matrix form:
  - $\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \sum_{j=1}^{p} \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix} \begin{pmatrix} x_{t-j} \\ y_{t-j} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$
- So CAR(*p*) can be considered as a special case of VAR(*p*), which has 2*p* parameters to estimate instead of 4*p*.
- We can estimate CAR via maximising likelihood.
- We can select appropriate order *p* via information criteria.
- We can produce forecasts and prediction intervals.

Complex Autoregressive Model

- Maximising likelihood of multivariate normal distribution is equivalent to minimising $|\Sigma|$.
- In our case this means that the cost function is:
  - $\mathrm{CF} = \sigma_1^2 \sigma_2^2 - \sigma_{1,2}^2,$
    - where σ_1^2 is the variance of the $e_{1,t}$, σ_2^2 is the variance of the $e_{2,t}$ and $\sigma_{1,2}$ is the covariance between them.
- Or it can also be estimated using likelihood, assuming Complex Normal distribution:

$$f(z) = \frac{1}{\pi^k \sqrt{\det(\Gamma)\det(P)}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} (\overline{z} - \overline{\mu})^\intercal & (z-\mu)^\intercal \end{pmatrix} \begin{pmatrix} \Gamma & C \\ \overline{C} & \overline{\Gamma} \end{pmatrix}^{-1} \begin{pmatrix} z-\mu \\ \overline{z} - \overline{\mu} \end{pmatrix} \right\}$$

Complex Autoregressive Model

- Looks scary, but eventually has less parameters than the similar VAR(p).
- Allows using order selection.
- Allows constructing prediction intervals.
- Can be directly compared with VAR via information criteria.

Complex Autoregressive Model

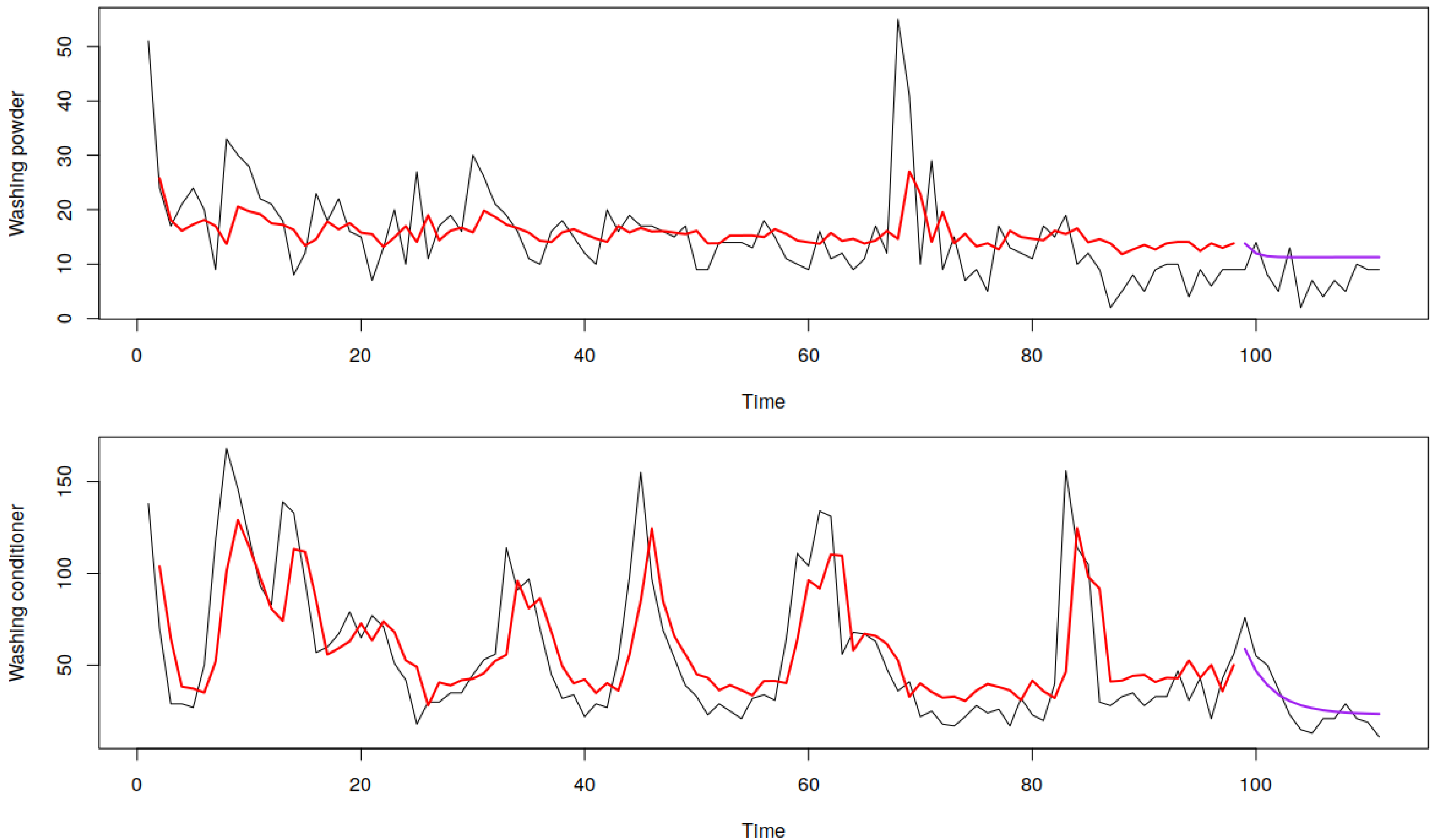
- So we have several options:
  - Unrestricted VAR with 4p parameters to estimate;
  - CAR(p) with 2p parameters, based on VAR and Multivariate Normal Distribution;
  - CAR(p) with 2p parameters with Complex Normal Distribution.

EXAMPLES OF APPLICATION

Example of application

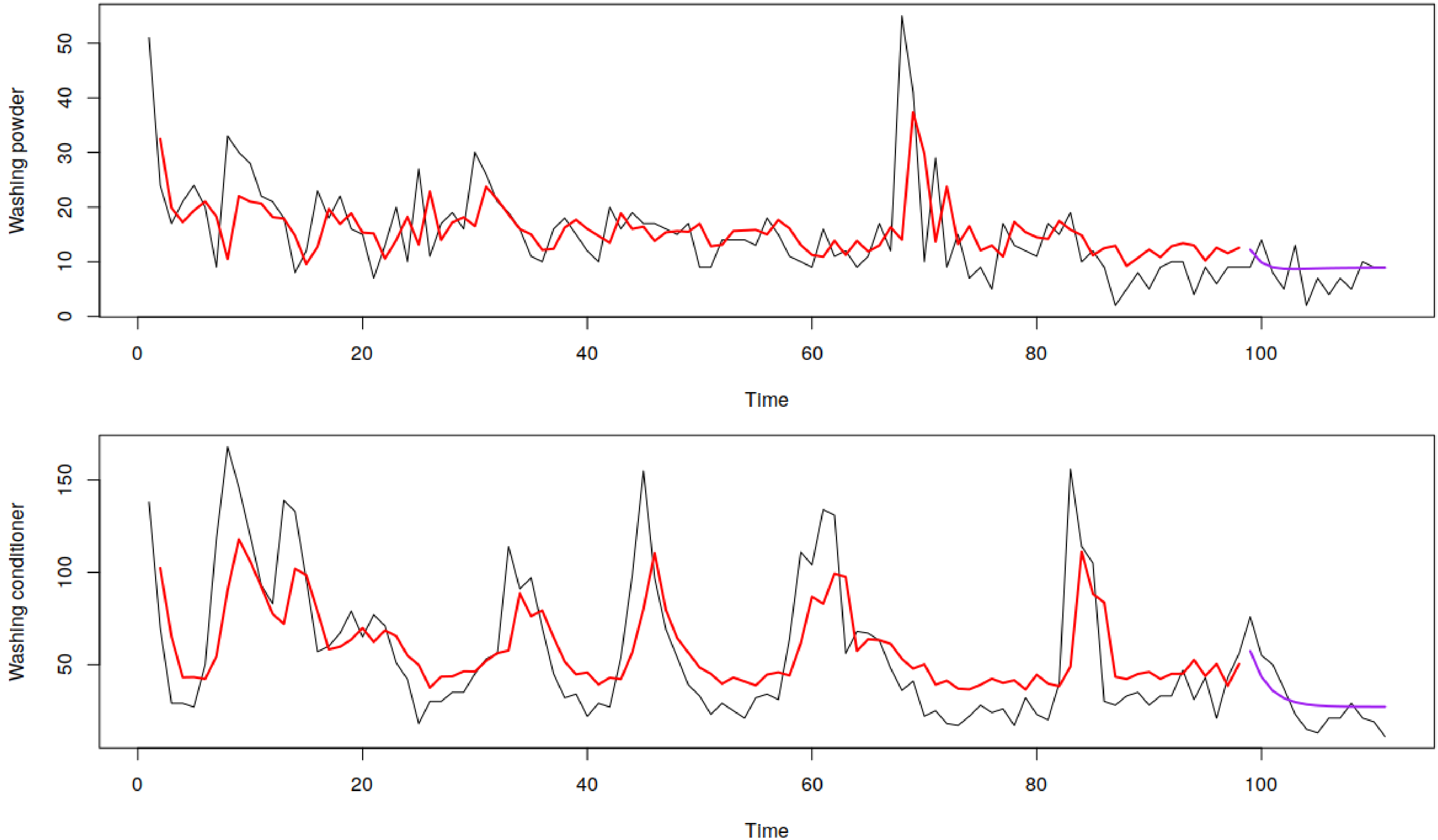
- Weekly data of sales of a retailer,
- Aggregated values for one shop,
- Two product categories:
  - Sales of washing powder,
  - Sales of conditioner.

VAR(1) model

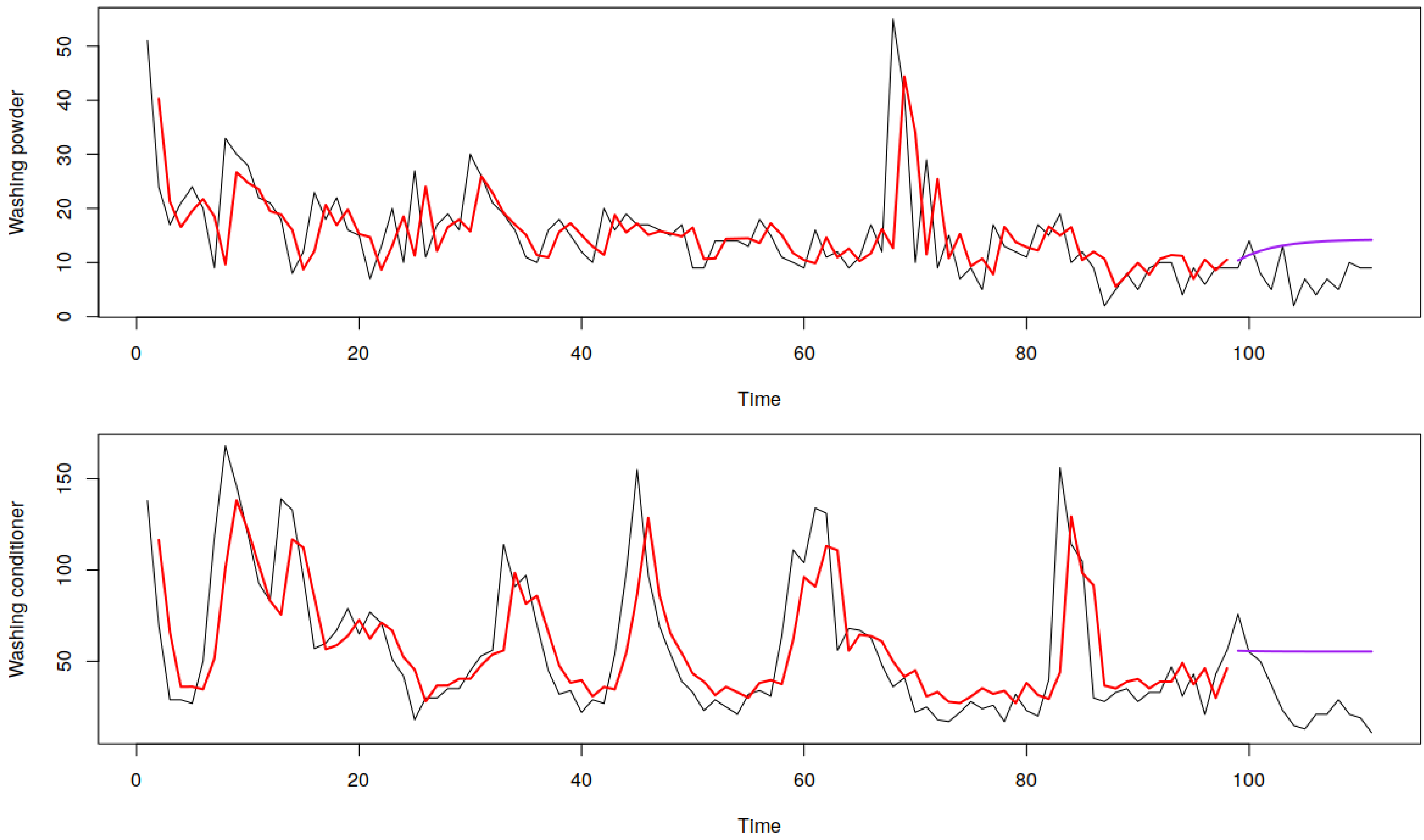


Washing powder
0
10
20
30
40
50
0
20
40
60
80
100
Time
Washing conditioner
50
100
150
0
20
40
60
80
100
Time

VAR(1) Restricted



CAR(1) with CND



Coefficients

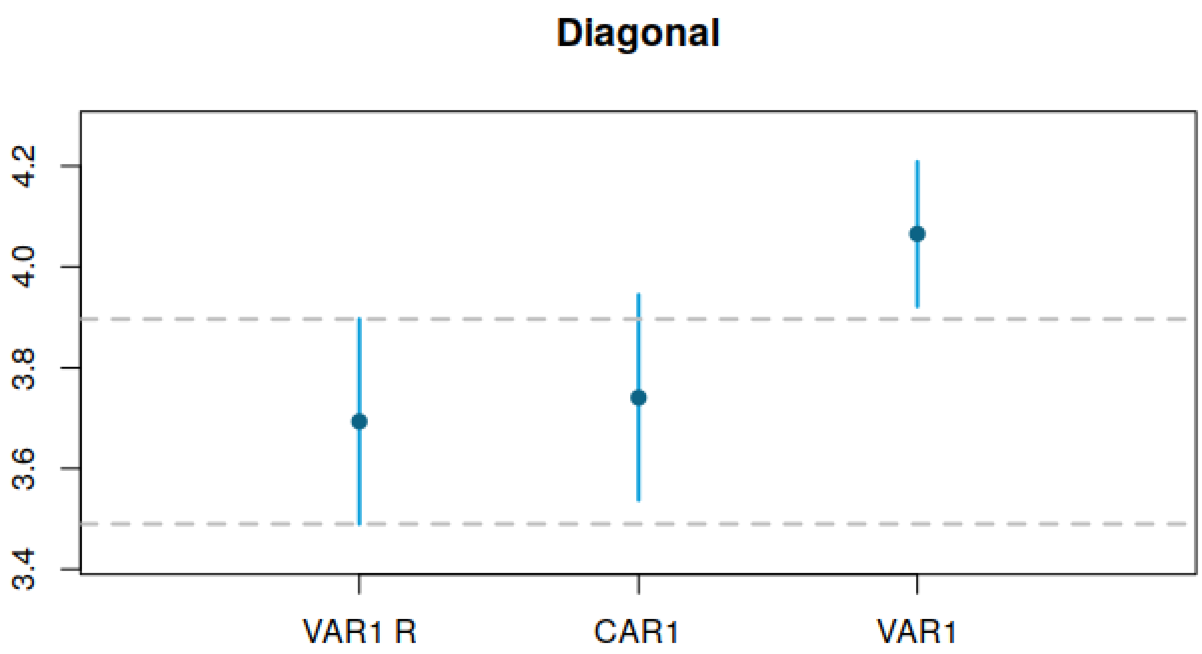
	Powder (t-1)	Conditioner (t-1)
	VAR(1)	
Powder (t)	0.613	-0.013
Conditioner (t)	-0.139	0.676
	VAR(1) Restricted	
Powder (t)	0.638	-0.004
Conditioner (t)	0.004	0.638
	CAR(1)	
Powder (t)	0.734	-0.012
Conditioner (t)	0.012	0.734

Simulation

- Investigate, in which circumstances CAR works.
- DGPs:
    1. VAR(1) with matrix of parameters $A = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.6 \end{pmatrix}$ (diagonal);
    2. VAR(1) with unrestricted $A = \begin{pmatrix} 0.8 & -0.3 \\ -0.2 & 0.6 \end{pmatrix}$;
    3. VAR(1) with CAR(1) parameters $A = \begin{pmatrix} 0.8 & -0.5 \\ 0.5 & 0.6 \end{pmatrix}$
- 1000 time series;
- 104 and 12 observations with respective holdouts of 13 and 4.
- MAE as error measure.

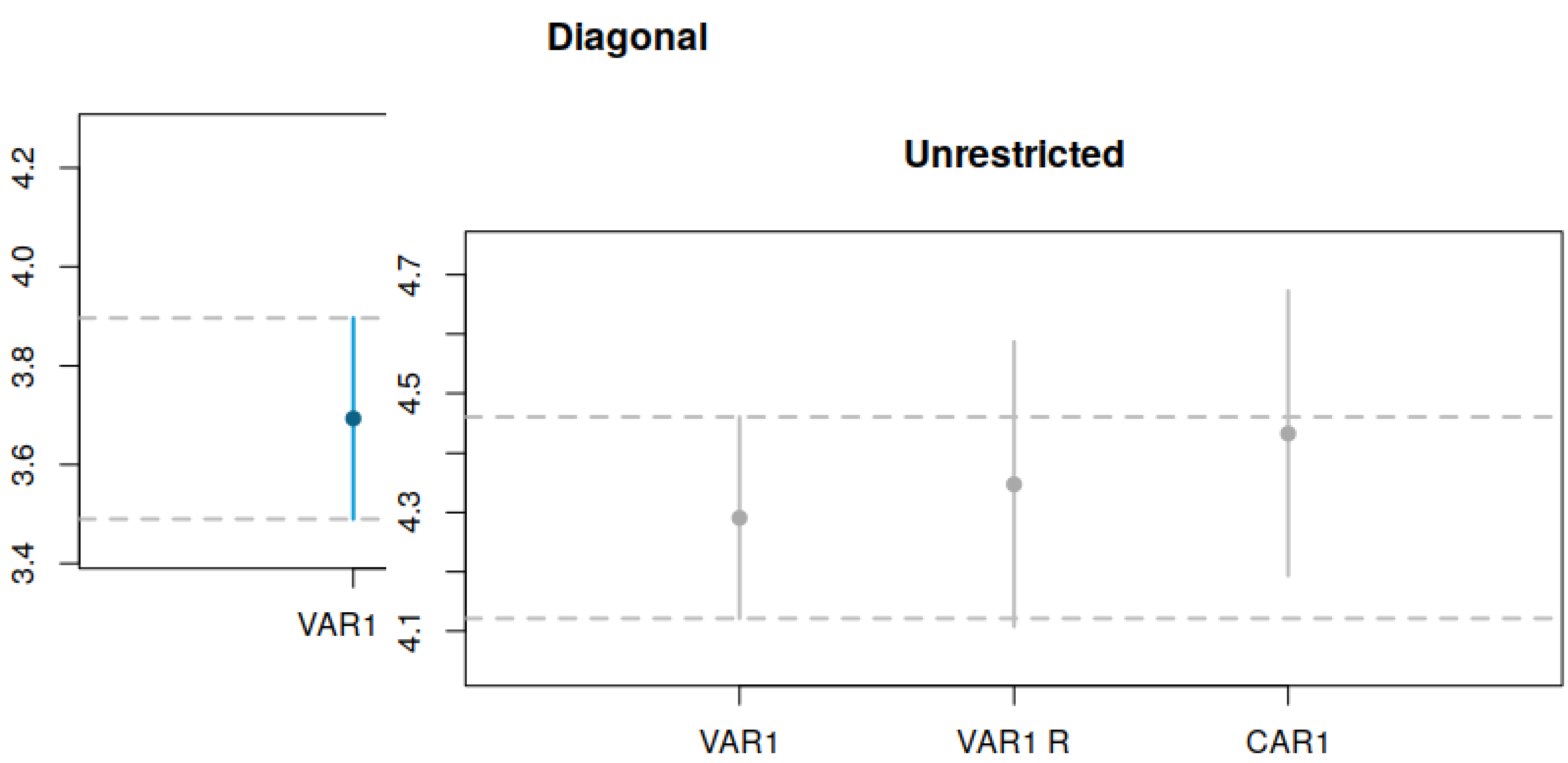
Simulation results, 12 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	4.08	**3.76**	3.80
Unrestricted	**4.29**	4.35	4.43
CAR	5.51	5.10	**4.66**



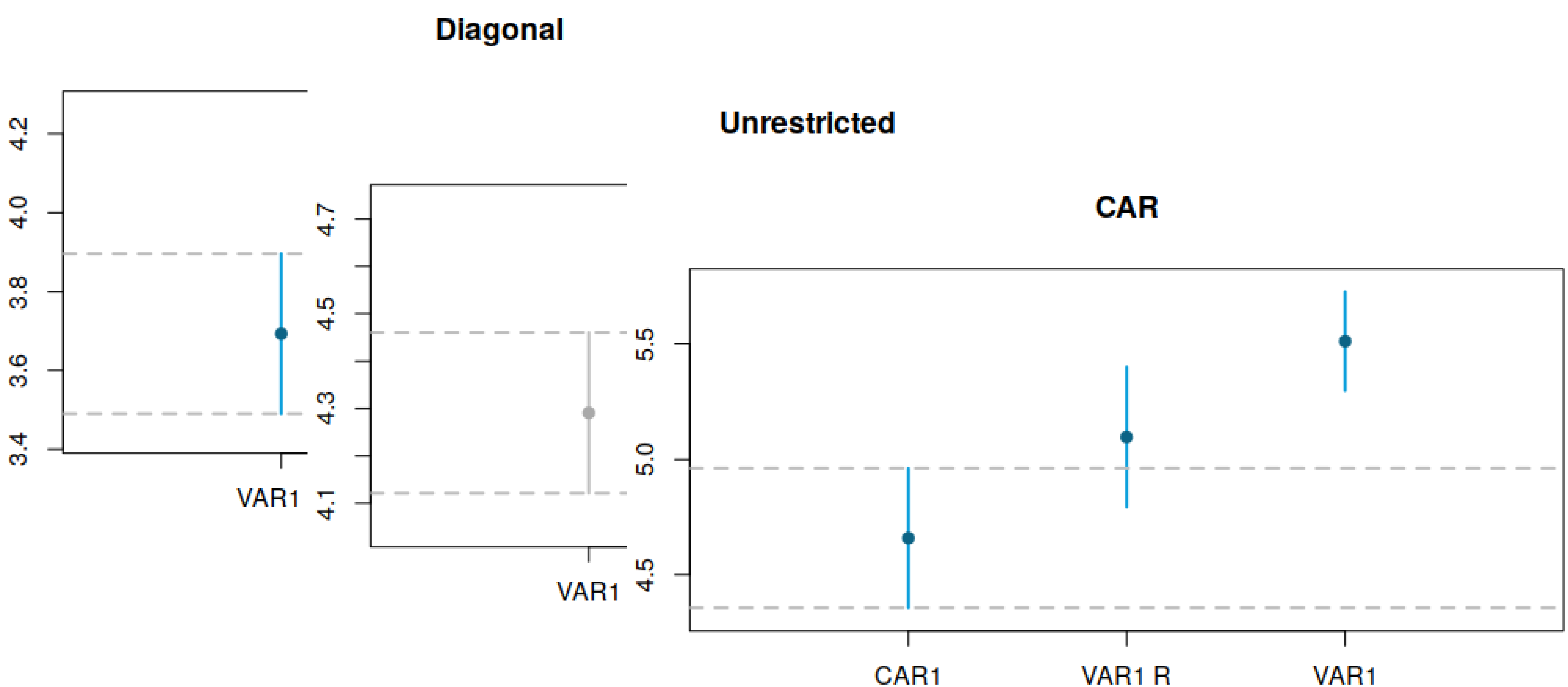
Simulation results, 12 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	4.08	**3.76**	3.80
Unrestricted	**4.29**	4.35	4.43
CAR	5.51	5.10	**4.66**



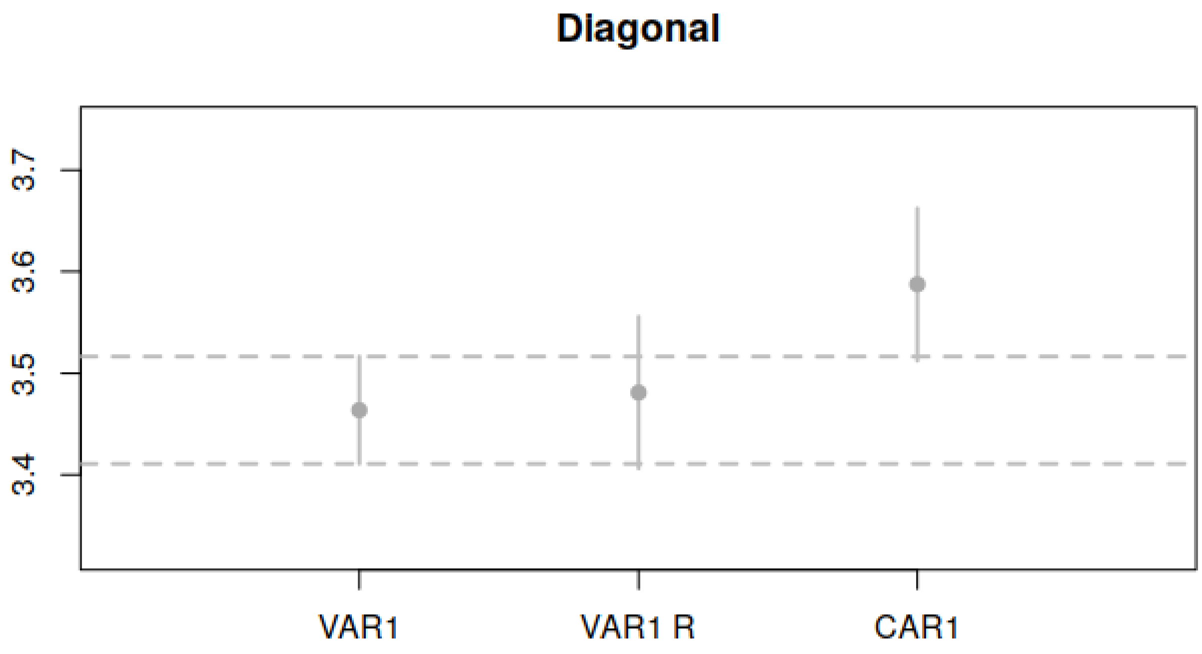
Simulation results, 12 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	4.08	**3.76**	3.80
Unrestricted	**4.29**	4.35	4.43
CAR	5.51	5.10	**4.66**



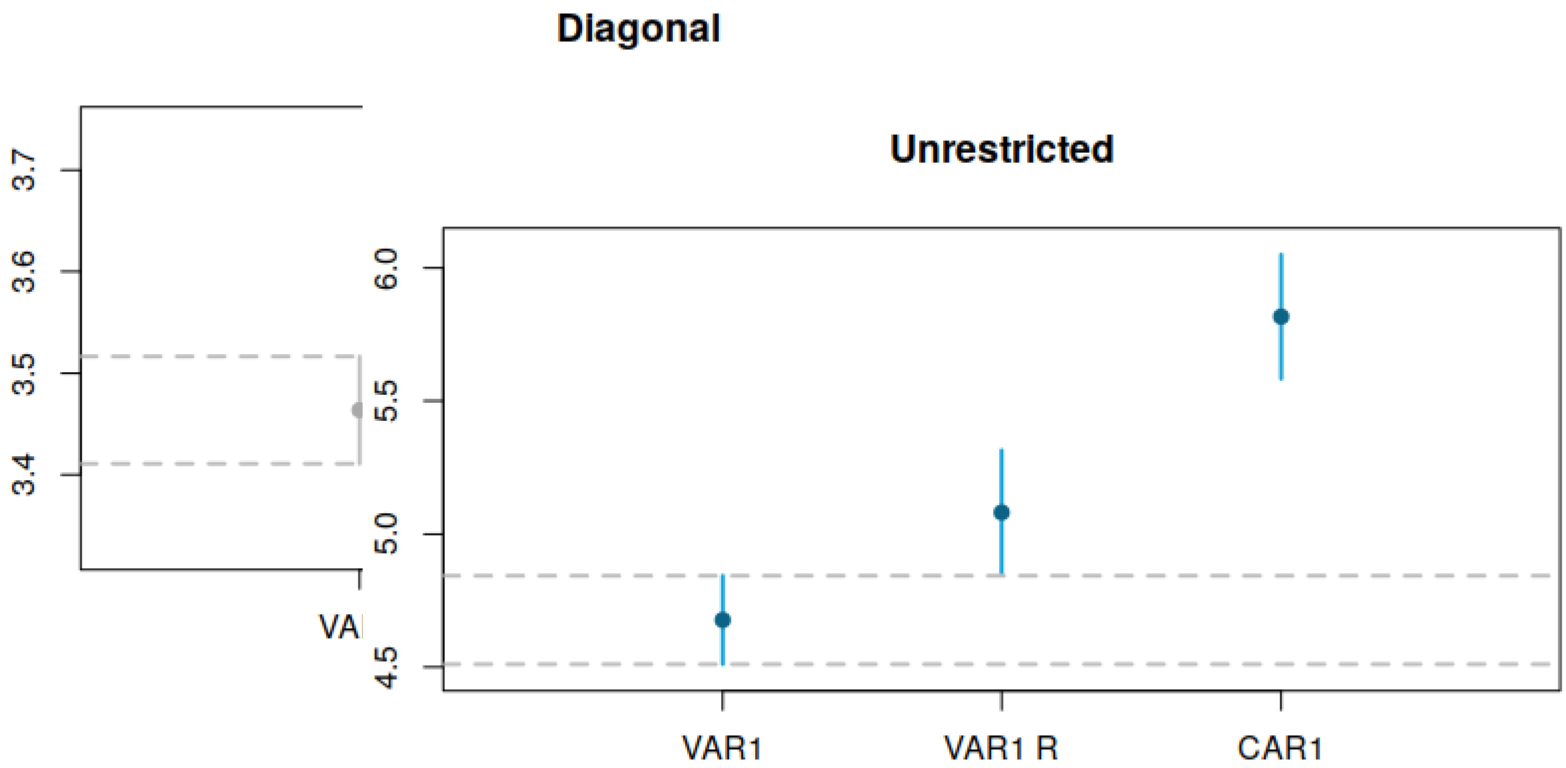
Simulation results, 104 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	**3.46**	3.48	3.59
Unrestricted	**4.69**	5.11	5.82
CAR	**4.42**	4.44	4.57



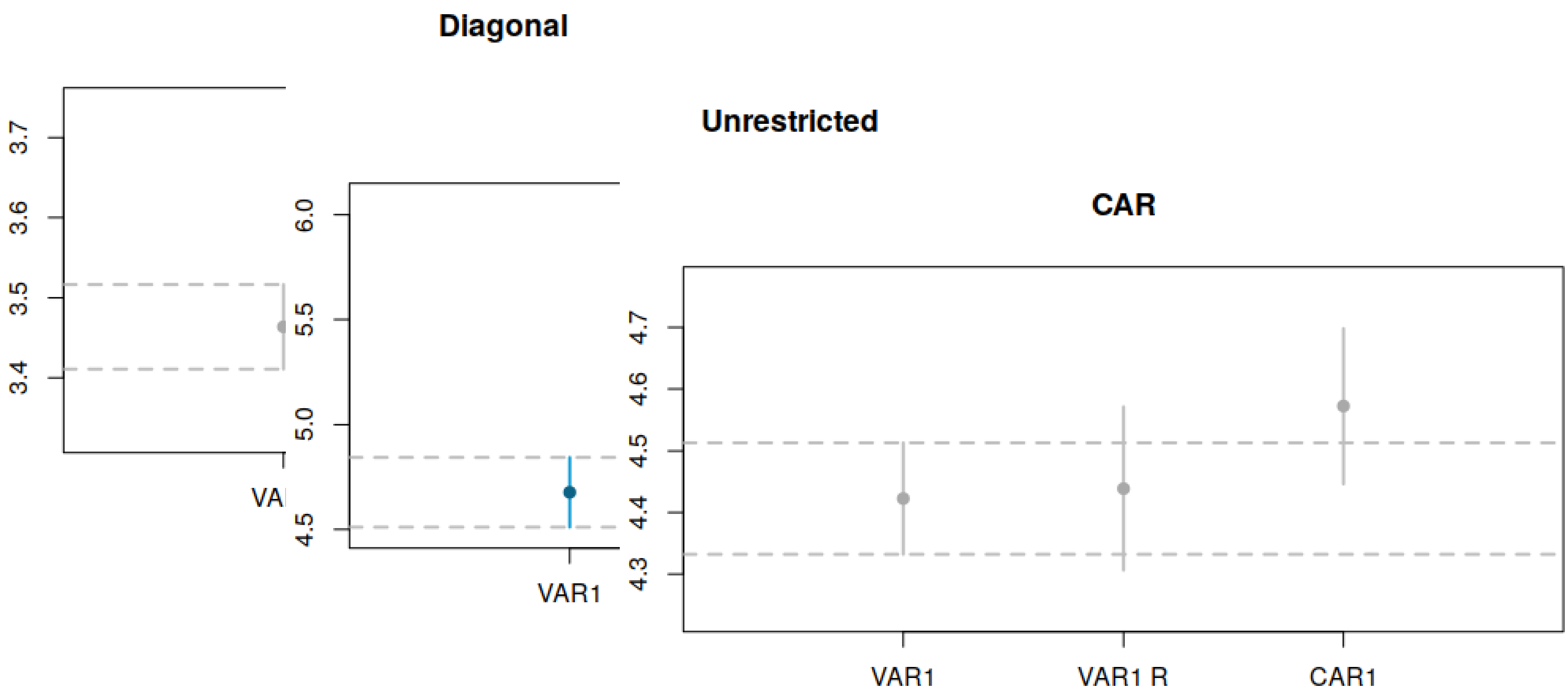
Simulation results, 104 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	**3.46**	3.48	3.59
Unrestricted	**4.69**	5.11	5.82
CAR	**4.42**	4.44	4.57



Simulation results, 104 obs

	VAR(1)	VAR(1) Restricted	CAR(1)
Diagonal	**3.46**	3.48	3.59
Unrestricted	**4.69**	5.11	5.82
CAR	**4.42**	4.44	4.57



CONCLUSIONS

Conclusions

- When there are pairs of product groups, CAR can be used
  - Forecast some elements of hierarchy?
- CAR needs less parameters than VAR;
- CAR can be used on small samples;
- The existing Complex Normal distribution works poorly;

Future research

- Study properties of CAR(p);
- Implement CAR with higher orders;
- Confidence intervals for parameters of CAR;
- Prediction intervals;
- Impulse response function;
- Compare CAR(p) with VAR(p) on real data:
  - Point values;
  - Distributions.
- Develop a new distribution for CAR;
  - Complex probability?

Thank you for your attention!

Prof. Sergey Svetunkov	Ivan Svetunkov
St. Petersburg Polytechnic University	Lancaster University
sergey@svetunkov.ru	ivan@svetunkov.ru

POLYTECH
Peter the Great
St.Petersburg Polytechnic
University

Marketing Analytics
& Forecasting

Lancaster University
Management School

Forecast results

Republic of Karelia	$y_{rt}+iy_{it}=(0,944-i0,079)+(0,601+i0,096)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Komi Republic	$y_{rt}+iy_{it}=(0,495+i0,058)+(0,869-i0,007)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Archangelsk region	$y_{rt}+iy_{it}=(1,119-i0,561)+(0,549+i0,329)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Nenets autonomous district	$y_{rt}+iy_{it}=(0,773-i0,011)+(0,958+i0,009)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Vologda region	$y_{rt}+iy_{it}=(0,180-i0,017)+(1,075-i0,005)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Kaliningrad region	$y_{rt}+iy_{it}=(-0,265-i0,115)+(1,238+i0,029)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Leningrad region	$y_{rt}+iy_{it}=(0,180-i0,018)+(1,075-i0,005)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Murmansk region	$y_{rt}+iy_{it}=(-0,667+i1,305)+(1,233-i0,609)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Novgorod region	$y_{rt}+iy_{it}=(0,408+i0,179)+(0,836-i0,006)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Pskov region	$y_{rt}+iy_{it}=(0,919+i0,005)+(0,612+i0,003)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$
Saint-Petersburg	$y_{rt}+iy_{it}=(1,182+i0,194)+(0,769-i0,028)(y_{rt-1}+iy_{it-1}),\quad t=2,3,\ldots$

Forecast results of particular regions of Russia (2008 – 2013)

№	Region	The average forecast error of the real part, %	The average forecast error of the imaginary part, %
1	Republic of Karelia	3,6	4,3
2	Komi Republic	1,5	2,7
3	Archangelsk region	4,7	12,5
4	Nenets autonomous district	9,8	12,1
5	Vologda region	8,1	8,5
6	Kaliningrad region	6,5	5,6
7	Leningrad region	5,2	8,9
8	Murmansk region	2,4	5,8
9	Novgorod region	3,9	2,9
10	Pskov region	3,2	2,9
11	Saint-Petersburg	2,6	2,4

References

- Caines, Keng, Sethi (1980). Causality analysis and multivariate Autoregressive modelling with an application to supermarket sales analysis. *Journal of Economic Dynamics and Control*. 3, 267-298. DOI: 10.1016/0165-1889(81)90018-X
- Leeflang, P. and Selva, J. (2012), “Cross-demand effects of price promotions,” Journal of Academy of Marketing Science, 40(4), 572–586.

References

- Gelper, S. E. C., Wilms, I., & Croux, C. (2016). Identifying demand effects in a large network of product categories. *Journal of Retailing*, *92*(1), 25-39. DOI: 10.1016/j.jretai.2015.05.005
- Wilms, Ines and Barbaglia, Luca and Croux, Christophe (2016), Multi-Class Vector Autoregressive Models for Multi-Store Sales Data (May 2016). KU Leuven, Faculty of Economics and Business, KBI_1617. DOI: 10.2139/ssrn.2782370