On the identification of sales forecasting models in the presence of promotions

The 32nd Annual International Symposium on Forecasting

J. R. Trapero

Universidad de Castilla-La Mancha

N. Kourentzes, R. Fildes

Lancaster University

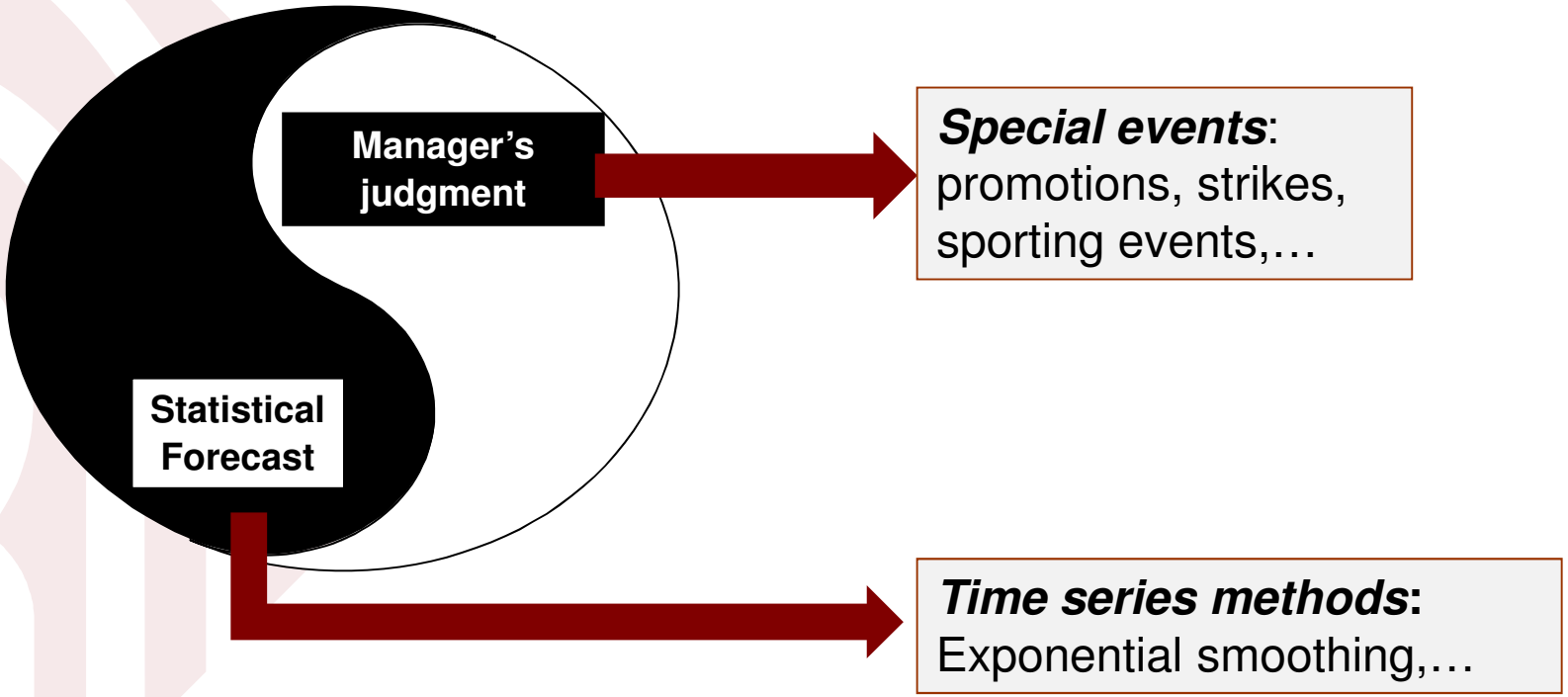
This work is supported by the SAS/IIF research grant

Agenda

1. Judgmental Forecasting and promotions
2. Case Study
3. Models
4. Conclusions

Judgmental forecasting

- Sales forecasting relies on Forecasting Support Systems:



- Promotional and advertising activity is one of the main reasons for adjusting statistical forecasts (Fildes and Goodwin, 2007)

Promotional modelling

- It's not a new topic:
  - **PromoCastTM** (Cooper *et al.,* 1999)
  - **Scan*PRO** evolutionary model building (Leeflang *et al.,* 2002)
  - **CHAN4CAST** (Divakar *et al.,* 2005)
- Common features:
  - PSS are based on regression models
  - Sales=f(regular prices, price cuts, feature advertising, special displays, …)

Promotional modelling

- Issues?:
  - Important data needs. 67 variables CHAN4CAST or more.
  - Cost of acquiring information
  - Selection methods are required. (Stepwise method)
  - Multicollinearity
  - Managers have to change their forecasting process:
    - Before: judgmental forecasting.
    - Now: Econometric models….

Can we build PSS with limited data (=reality) to support operations?

2. Case study

- A manufacturing company specialized in household detergent products:
  - Shipments
  - One-step-ahead system forecasts (SF)
  - One-step-ahead adjusted or final forecasts (FF)
  - Promotional information:
    - Price cuts
    - Shelf display
    - Feature advertising
    - Days promoted in each week

- The data contains 60 SKUs
  - In total, 8800 observations
  - Weekly sampled between October 2008 and July 2011.

Let's try some models

- Experimental setup:
  - A predictive validation experiment is carried out.
  - Last 30 weeks are reserved for validation purposes.
  - A rolling origin experiment is designed
  - The forecasting horizon is one week ahead.
- Promotional models:
  - Benchmarks: SES, Naïve, Last like promotion.
  - Proposed model …

Let's try some models

- Dynamic regression
  - Identify univariate structure automatically by minimizing the Schwartz Information Criterion
- 4 types of promotions
  - Multicollinear → Some types of promotion take occur simultaneously for almost all SKUs
  - Use **principal components** of promotional inputs → simplify estimation & smaller models
- Build model for promotional and non-promotional periods separately → Aid in resolving bias and estimation issues

Experimental results

- Table of results:
  - MAE on the hold-out sample
  - Data has been normalized.

	SF	FF	Naïve	SES	DR
No Promo	0.607	0.652	0.762	0.629	**0.592**
Promo	1.069	1.327	1.092	0.904	**0.861**
Overall	0.671	0.745	0.808	0.667	**0.624**

- Results:
  - Dynamic regression outperforms SF & FF and benchmarks
  - PCA performs well → results without PCA worse (not shown)
  - **Catch 1: No promotion ramp-up & -down effects → Build single model?**
  - **Catch 2: How do you forecast promotion on a time series with no or limited promotional history?**

More models…

- PCA on pooled promotions across all time series → More estimation sample

- Dynamic regression on each time series:

  – Identify univariate structure automatically by minimizing the Schwartz Information Criterion

  – Add pooled PCA with cross-SKU estimated coefficients → able to model promotional effects even with no history

  – Add lags of PCA → Model ramp-up, ramp-down effects

  – Estimate univariate coefficients simultaneously with pooled PCA coefficients

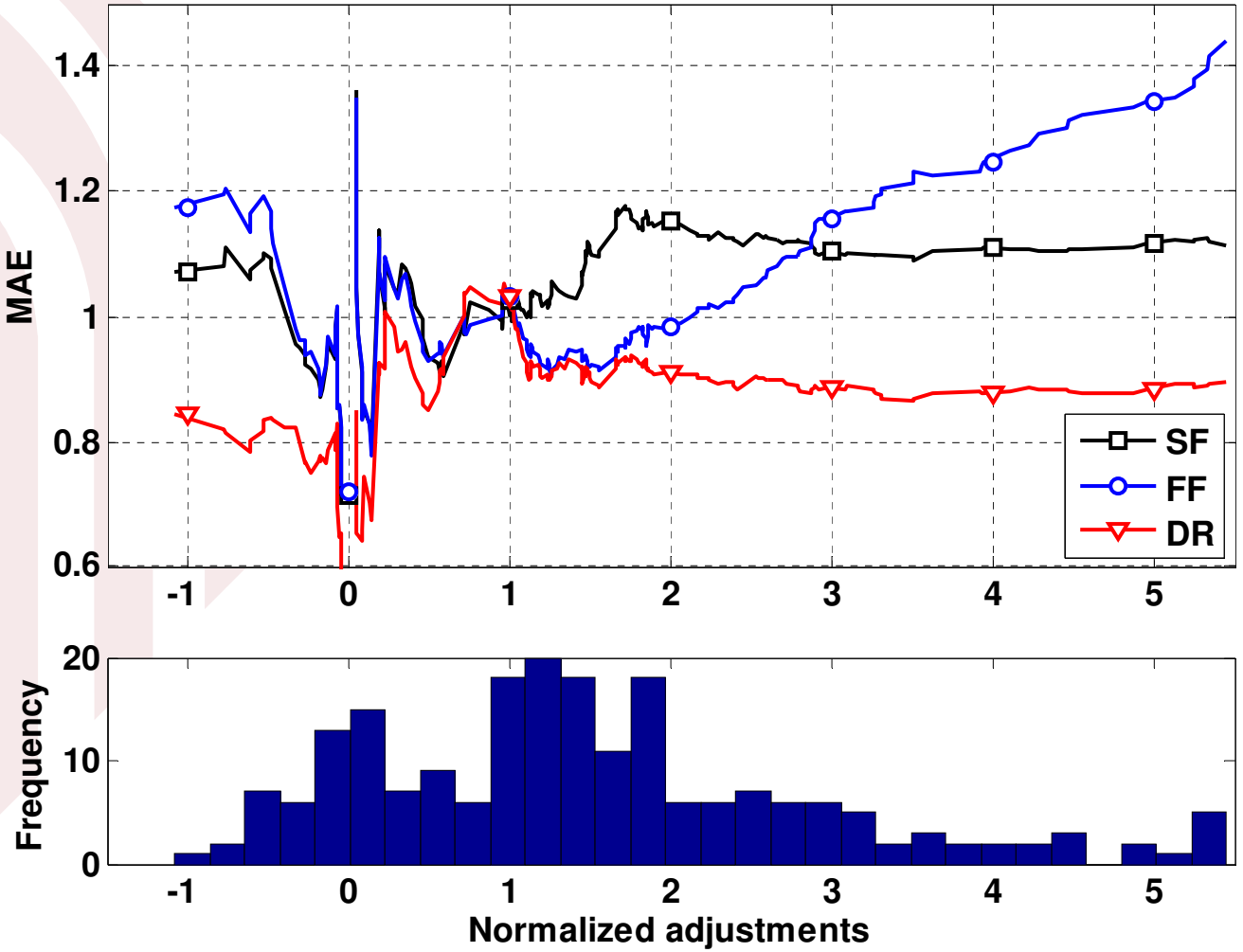
Experimental results

	SF	FF	Naïve	SES	DR	DR-p
No Promo	0.607	0.652	0.762	0.629	**0.592**	0.609
Promo	1.069	1.327	1.092	0.904	0.861	**0.853**
Overall	0.671	0.745	0.808	0.667	0.624	**0.624**

- Results:
  - Superior promotional & overall forecasting performance
  - **Provide promotional forecasts for SKUs with no promotional history**
  - PCA dynamics allow forecasting well non-promotional periods
  - Estimation problems for univariate (on each SKU) coefficients?

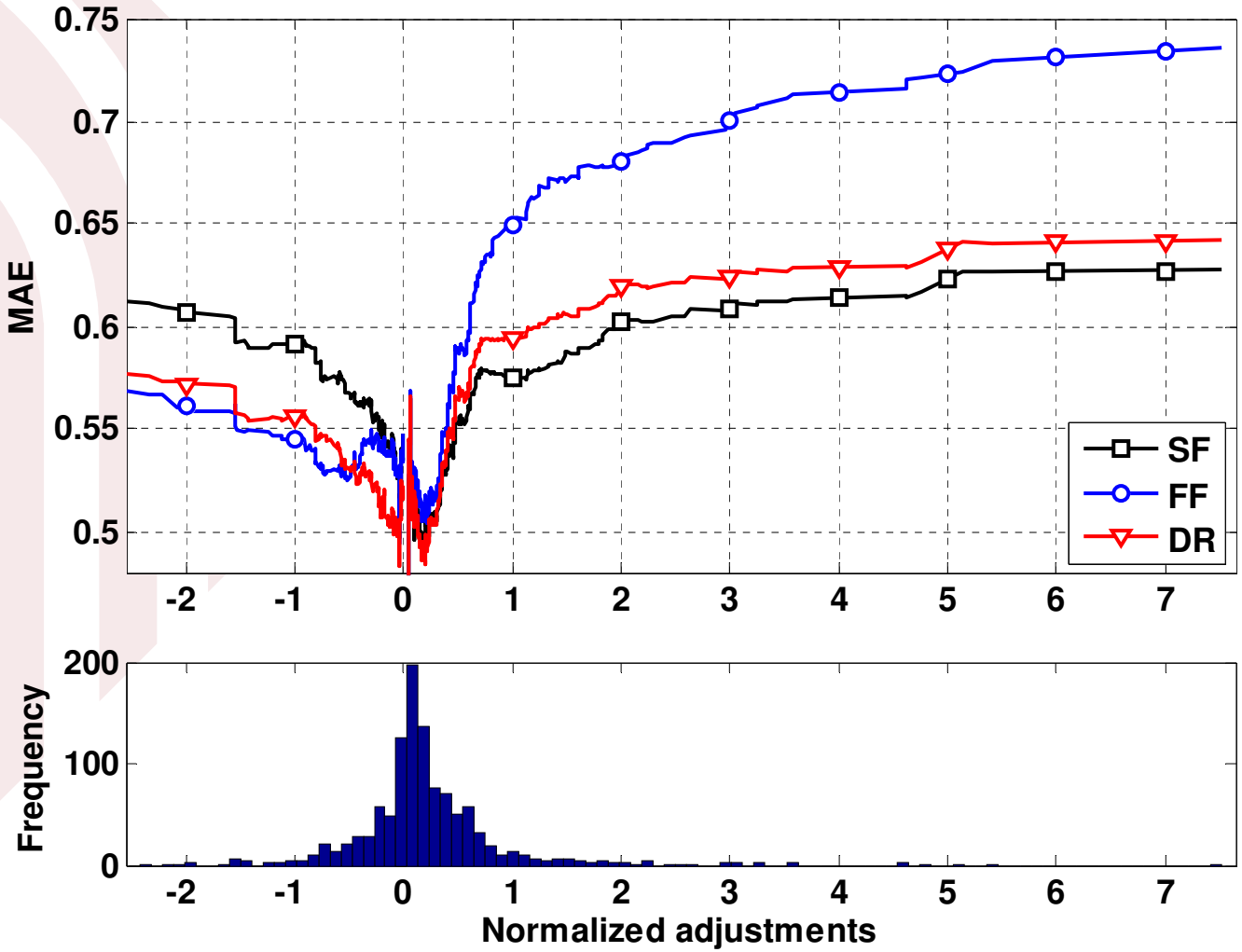
Experimental results

- **Dynamic Regression under promotions**



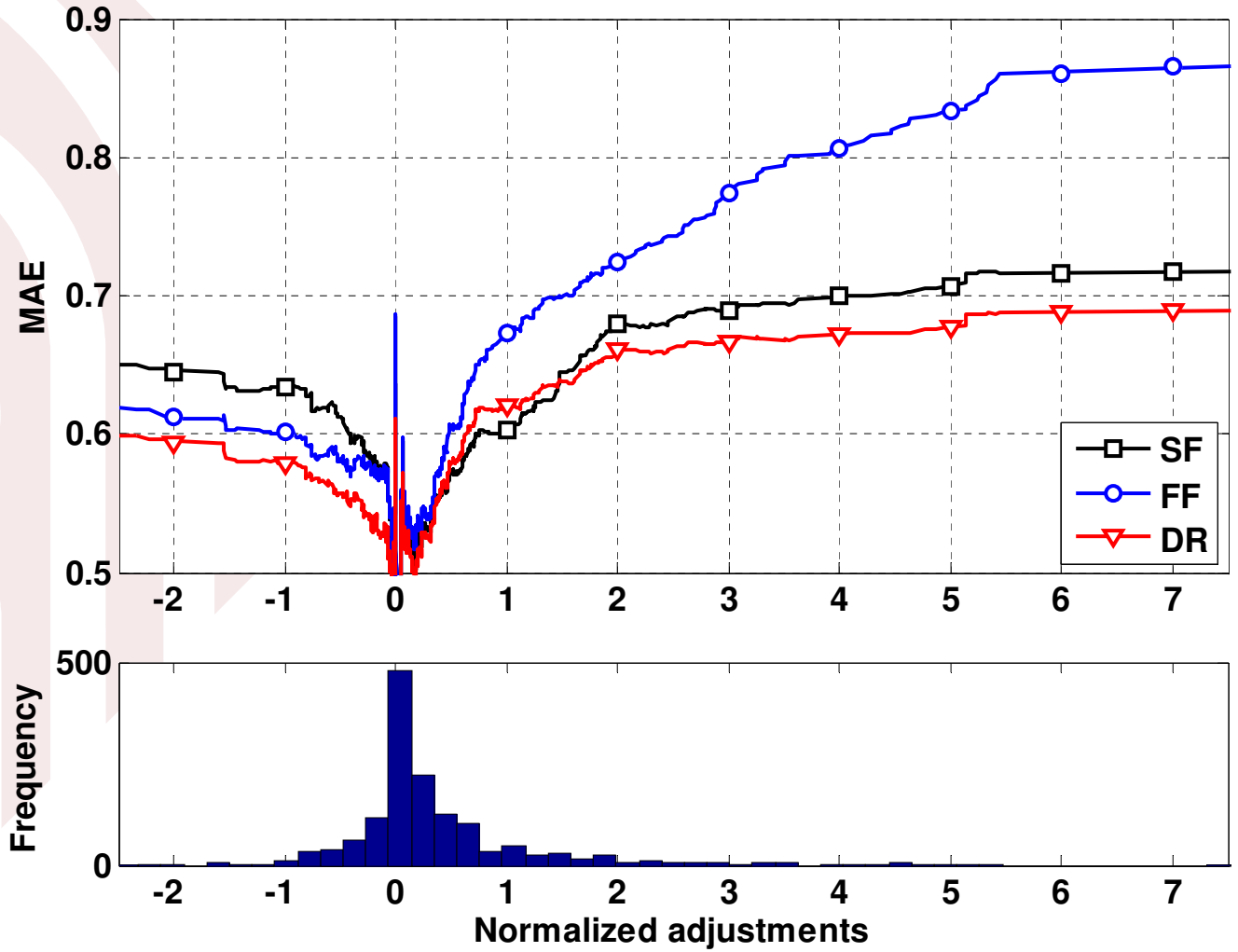
Experimental results

- **Dynamic Regression under no promotions**



Experimental results

- **Dynamic Regression → Overall**



4. Conclusions

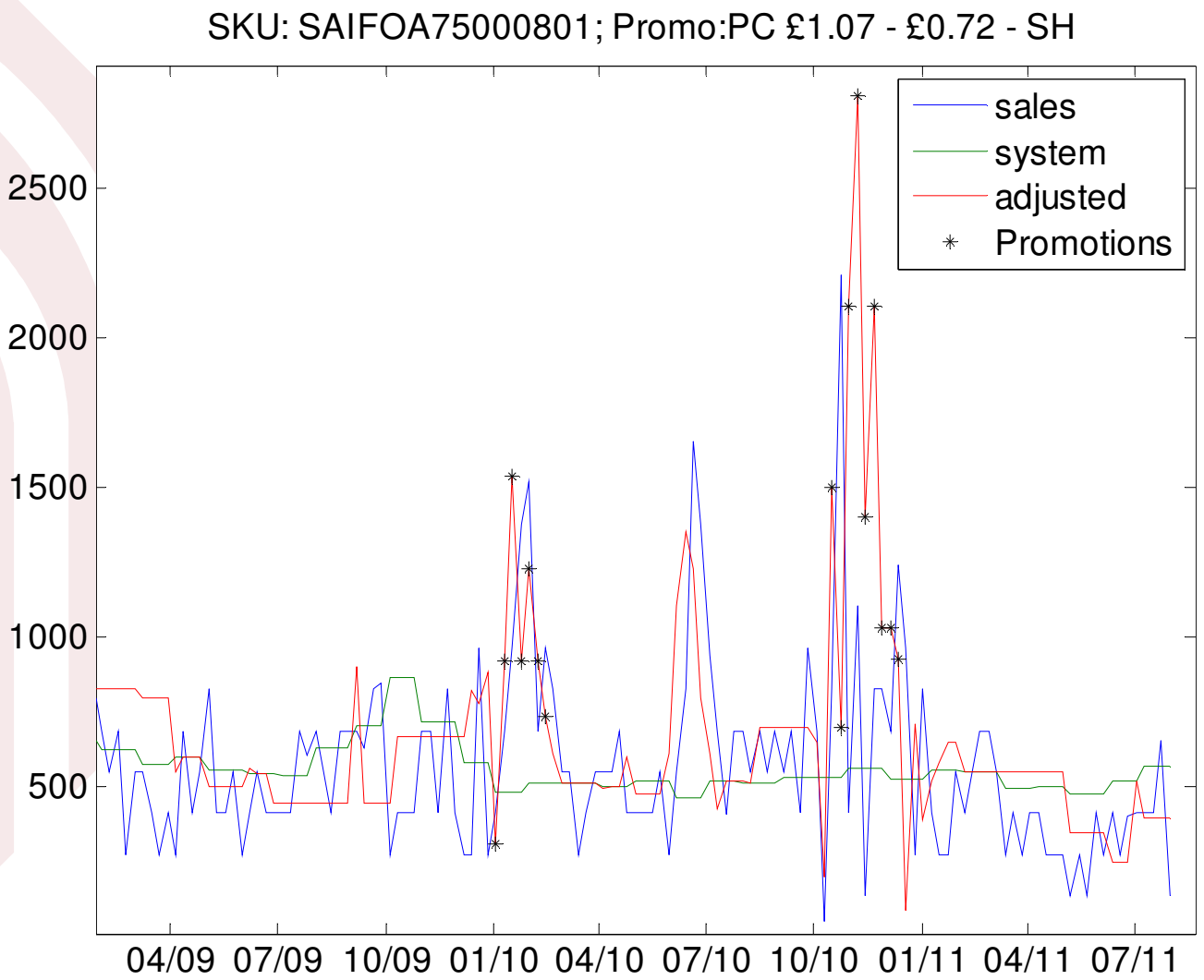
- Promotional modelling might substitute judgmental forecasts with promising results.
- Dynamic Regression based on Principal components and noise term modeling have reduced the forecasting error on promotional periods.
- Pooled PCA on promotions allows for better estimation of promotional effects and ability to produce forecasts with limited or no promotional history
- PCA → transparent calculation → track promotional effects
- Some estimation issues remain

Thank you for your attention!

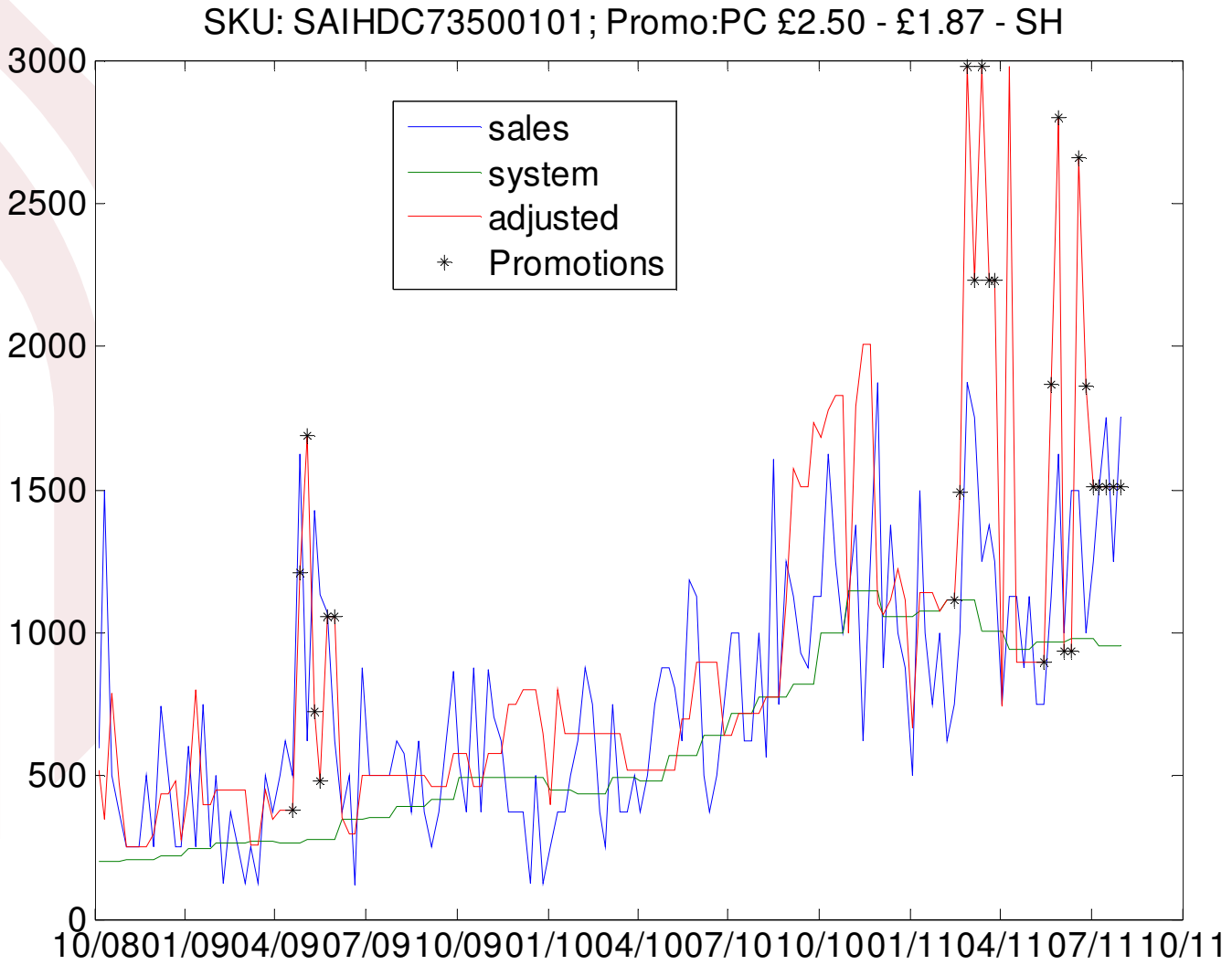
E-mail:

Juan Trapero Arenas
juanramon.trapero@uclm.es
Nikolaos Kourentzes
n.kourentzes@lancaster.ac.uk

2. Case study


SKU: SAIFOA75000801; Promo:PC £1.07 - £0.72 - SH
sales
system
adjusted
Promotions
2500
2000
1500
1000
500
04/09
07/09
10/09
01/10
04/10
07/10
10/10
01/11
04/11
07/11

2. Case study


SKU: SAIHDC73500101; Promo:PC £2.50 - £1.87 - SH
sales
system
adjusted
Promotions
3000
2500
2000
1500
1000
500
0
10/08
01/09
04/09
07/09
10/09
01/10
04/10
07/10
10/10
01/11
04/11
07/11
10/11