Estimating Retail Demand and Lost Sales

Giulio Zotteri – Politecnico di Torino
Ananth Raman – Harvard Business School

Sales Lost due to Stockouts

- Retailers frequently stock out of products;
  - one in three apparel customers does not find the item she/he is looking for in her/his size;
- This leads to lost sales and lost gross margins.
- Lost sales can account for up to 25% of sales (on average 6% in the COER-questionnaire)
- Corsten et al. suggest an 8% stock out rate in the grocery industry, which is likely to lead to substantial lost sales

Estimation of demand and lost sales

- Many retailers expressed interest in estimating demand and lost sales.
- Purpose of estimating demand and lost sales:
  - audit of current performance
  - help forecast future demand
- Yet, most retailers lack systematic procedures to estimate demand and lost sales.
- Current practices are often not completely satisfactory

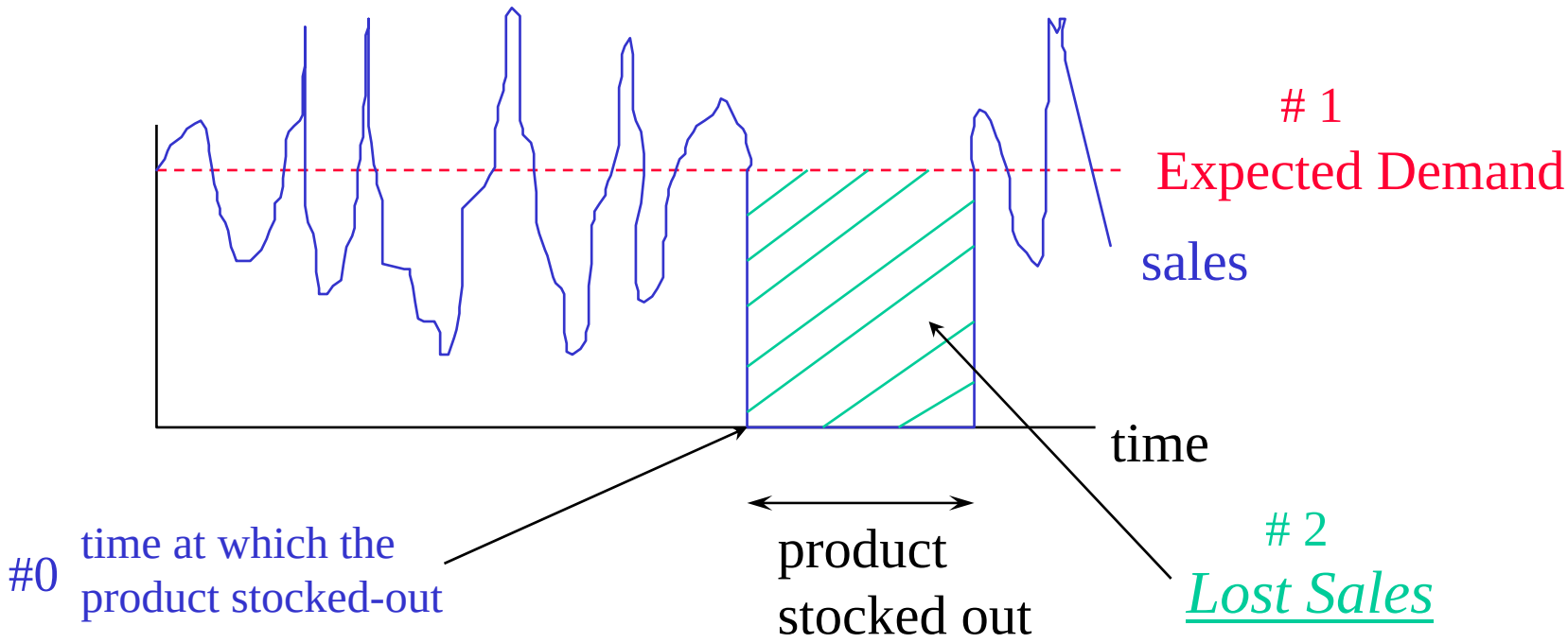
Lost Sales Estimation can be tricky!

- Demand, and lost sales are unobservable at bricks and mortar stores.
- Complex: Analysis needs to be conducted at Store-SKU level.
- Small samples complicate analysis.
  - demand per SKU-store is small
  - many SKUs sell less than 1/week-store
  - tough to draw conclusions based on such small samples.

New Opportunities for Estimating Lost Sales

- Technique should be simple to implement and should work at store-SKU level (I.e. where stock-out occur)
- Combination of POS and inventory data can be used to estimate demand and lost sales.
- The technique is rather simple and can, we believe, be implemented easily but needs to be tailored to the specific set of information available (e.g., in case reliable data about traffic are available we might want to estimate demand per customer, rather than total demand as one of the driver of demand is known)

How does our technique work?



0. Identify time SKU was stocked out at a store, using POS and inventory data.

1. Use #0 and POS data to estimate the expected demand

2. Use # 0 and # 1 to estimate "true" demand and lost sales during "stocked out time".

Step 1: Estimation of expected demand

When estimating demand we face 3 problems:

1 estimation of customer buying rate

(2 estimation of number of unit purchased by each customer)

3 analysis of non-stationarity

Pb1: Customer buying rate

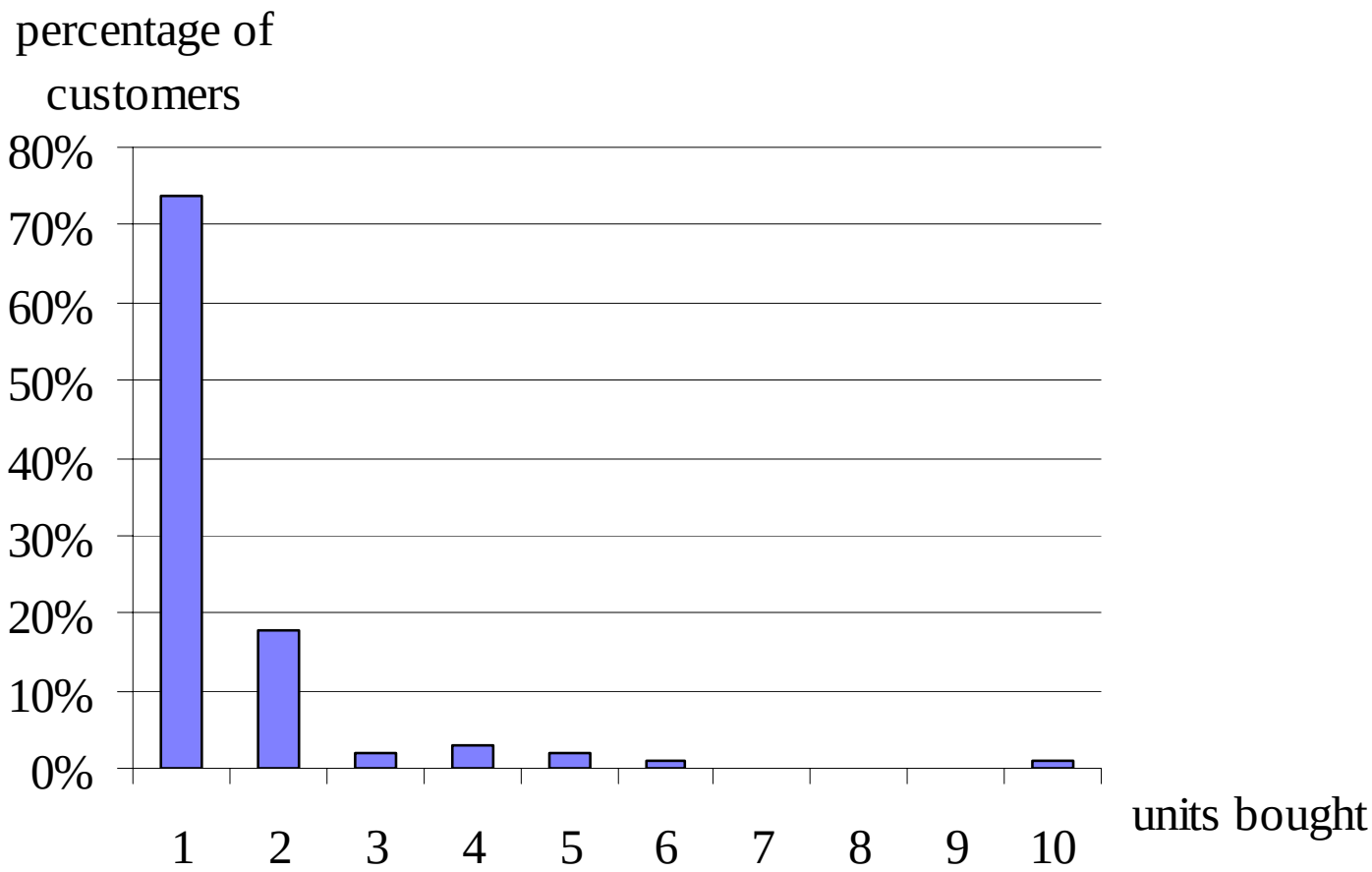
- A retailer opens a new store
- Each customer buys one item
- Demand is stationary during the day
- The store is open from 8 a.m. to 8 p.m.
- At the beginning of the first day he stocks 3 units
- He stocks-out at 2 p.m.
- He has to estimate demand for the following day.

- What is your estimate for average demand rate?

Pb2: Order size distribution

- Lost sales are due to:
  - the units that the last customer would have bought, had they been available.
  - the requests of customers that walked into the store during the stock-out.

Pb 2: Order size distribution: an example



Order size distribution: an example.

- Customers buy either 1, or 2 or 3 units of a given product at the same time.
- The retailer stocks three units.
- We observed the following situations:

	1	*2*	*3*	*Tot*
1	50	30	30	*110*
2		10	15	*25*
3			15	*15*

Number of units available

Number of units bought

Order size distribution

Intuition might suggest that the expecte order size is

$$\frac{110\cdot 1+25\cdot 2+15\cdot 3}{150}=1.37$$

However, to estimate the expected order size we have to estimate the probability of each possible order size:

$$P(o=1)=\frac{30+30}{30+30+10+15+15}=0.6$$

$$P(o=2)=(1-0.6)\cdot\frac{15}{15+15}=0.2$$

$$P(o=3)=(1-0.6)\cdot\frac{15}{15+15}=0.2$$

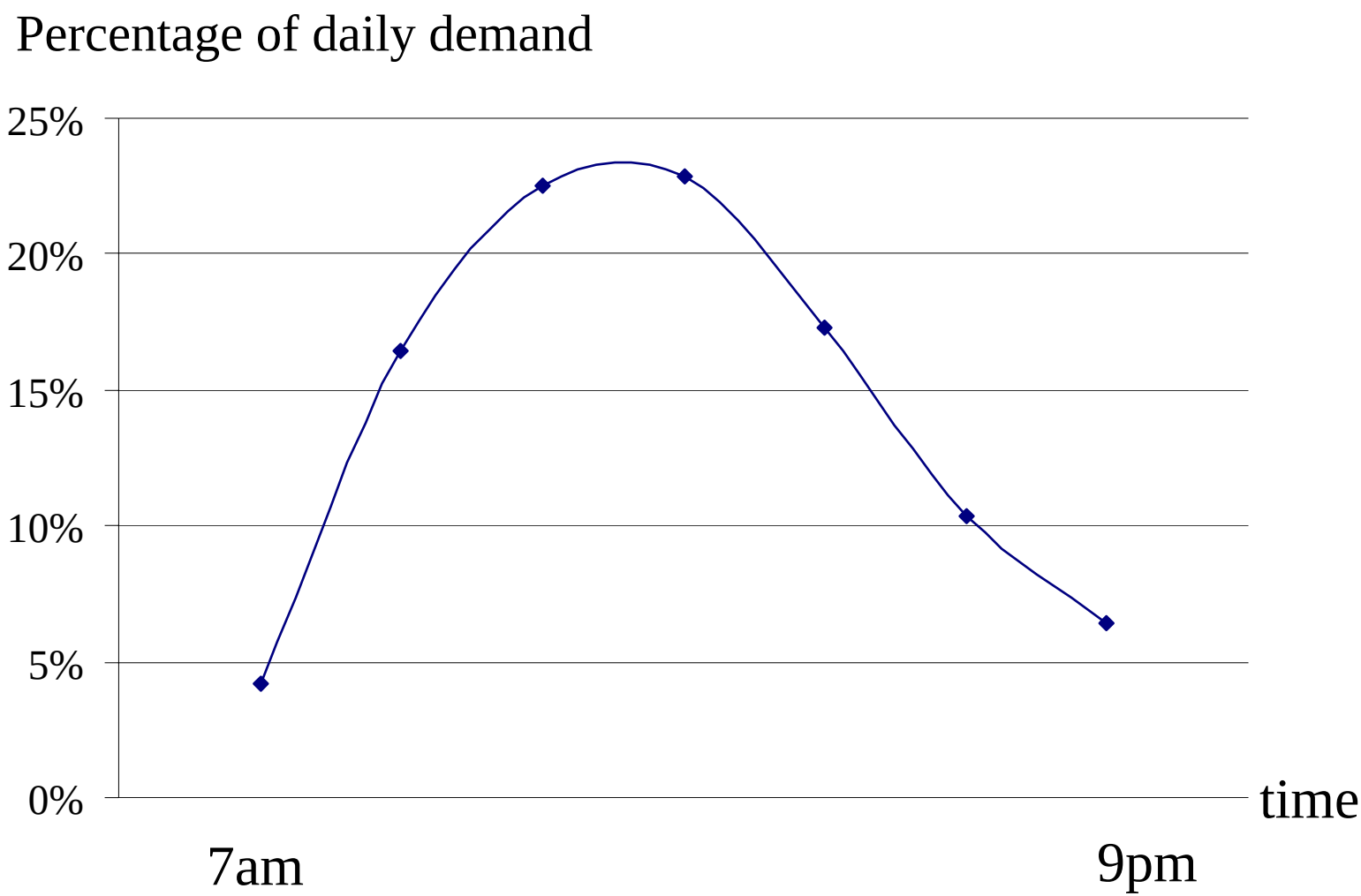
$$P(o=1)+P(o=2)+P(o=3)=1$$

$$Expected\ \ order\ \ size=1\cdot 0.6+2\cdot 0.2+3\cdot 0.2=1.6$$

Pb 3. Non stationarity

- Retail demand is non-stationary:
  - short term non-stationarity (e.g., variations within the day);
  - long term seasonality (e.g., variations within the year).
- We can deal with non stationarity in two different ways:
  - split the period into sub-periods where demand can be assumed to be stationary;
  - use traffic or other products' sales data as proxies

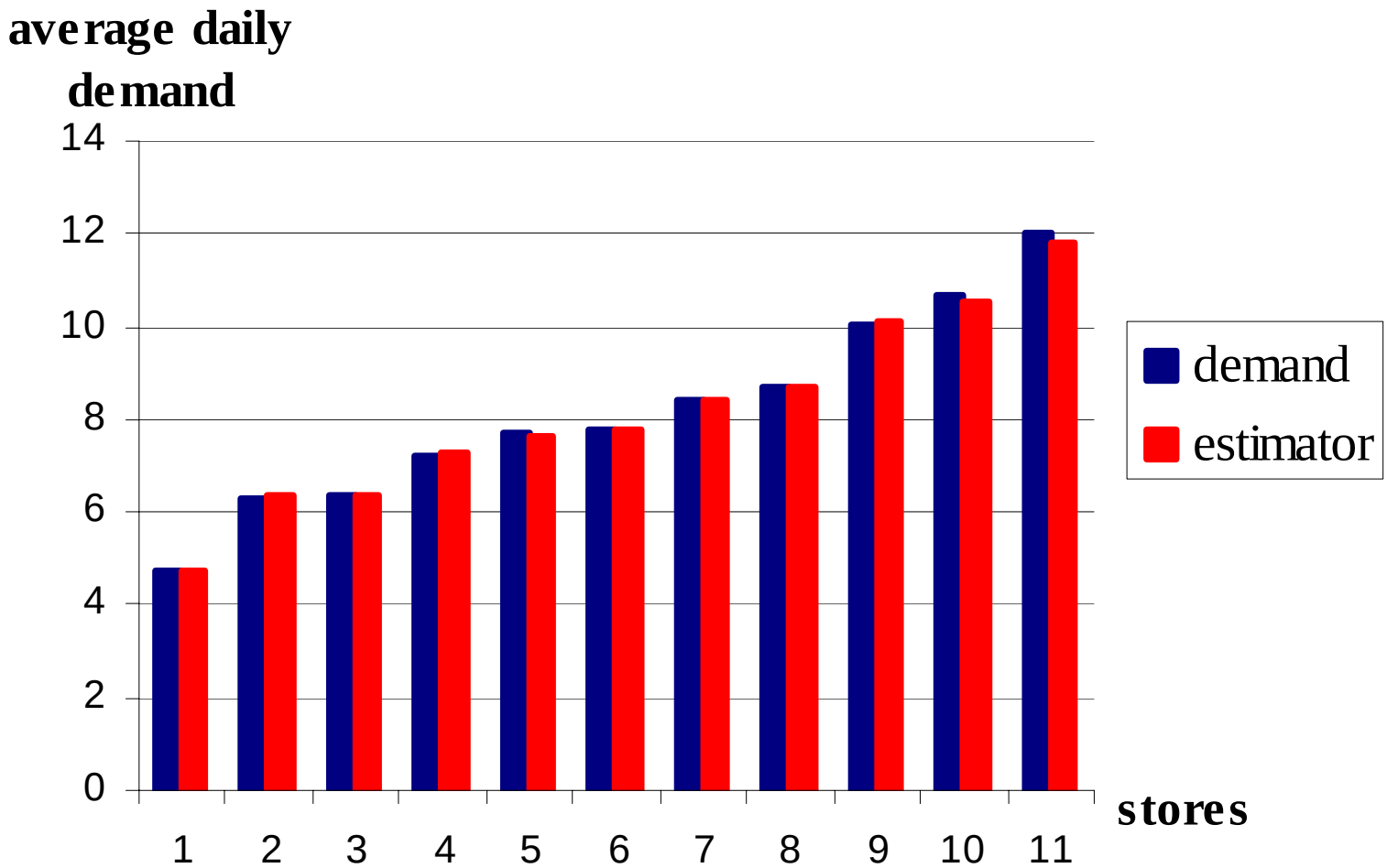
Demand non stationarity: an example



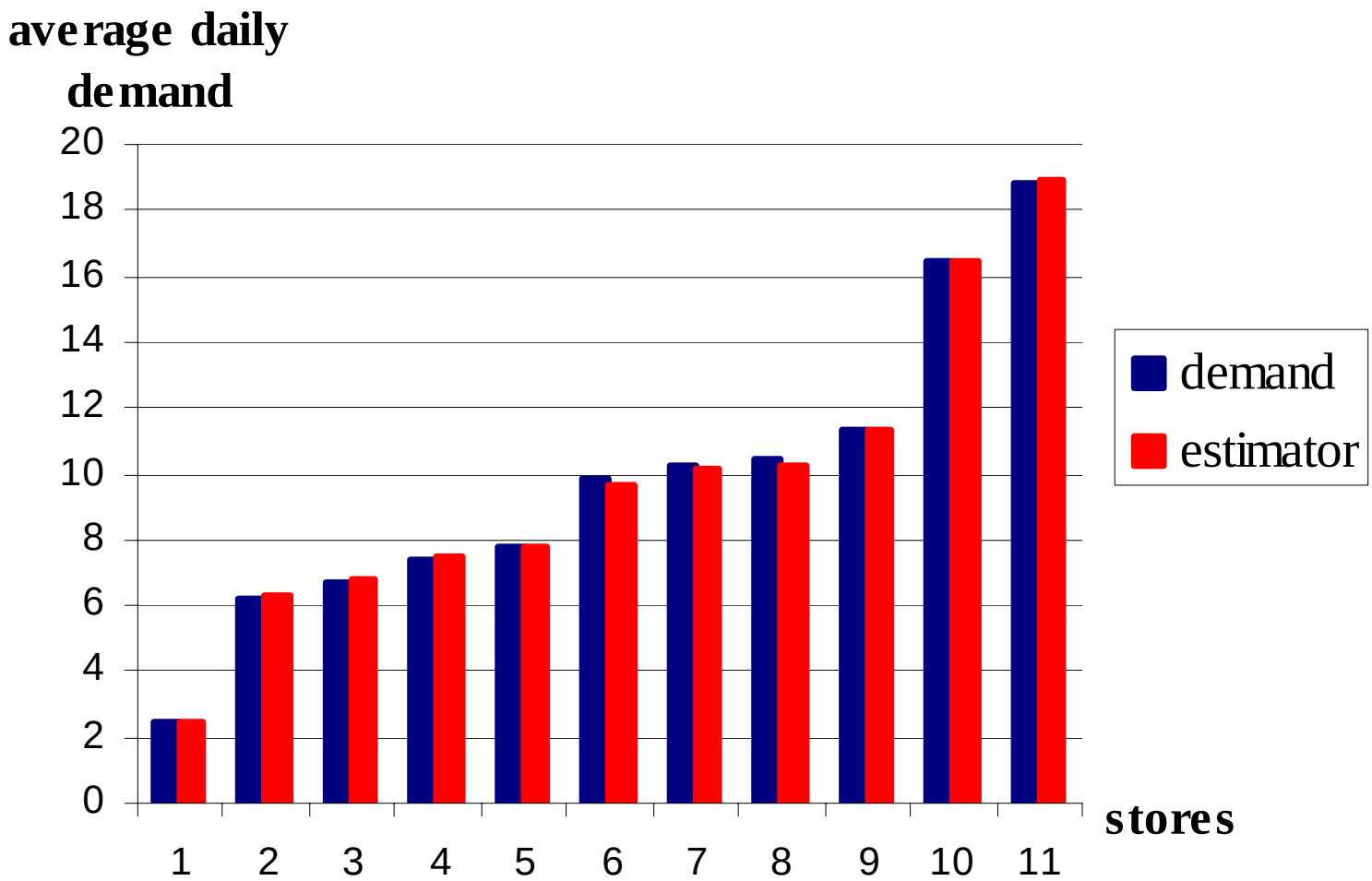
Validating the technique

- We received data from two major retailers in the COER panel (93 & 100 days - 11 stores).
- Two similar items that never stocked out $\Rightarrow$ we know real demand.
- We can pretend that the retailers had stocked less than they actually did.
- We can simulate sales and lost sales.
- We test whether we are able to estimate actual demand

Demand Estimate: Retail Chain A



Demand Estimate: Retail Chain B



Demand estimation: Retail Chain A

	Demand	**Estimated**	**Percentage Error**
store			
A	10.69	10.52	-1.53%
B	12.04	11.82	-1.85%
C	6.35	6.36	0.13%
D	8.70	8.70	0.04%
E	4.75	4.75	-0.01%
F	8.43	8.44	0.17%
G	6.27	6.35	1.22%
H	7.76	7.75	-0.22%
I	10.02	10.09	0.73%
J	7.69	7.64	-0.68%
K	7.22	7.25	0.52%
District	78.51	78.09	-0.20%

Demand estimation: Retail Chain B

	Demand	**Estimator**	**Percentage Error**
store			
A	10.47	10.24	-2.13%
B	7.36	7.45	1.22%
C	16.47	16.50	0.21%
D	6.24	6.28	0.69%
E	18.83	18.96	0.66%
F	7.83	7.77	-2.16%
G	10.22	10.12	-0.96%
H	9.87	9.66	-2.12%
I	11.35	11.32	-0.22%
J	2.45	2.44	-0.16%
K	6.74	6.90	2.31%
District	97.61	97.42	-0.38%

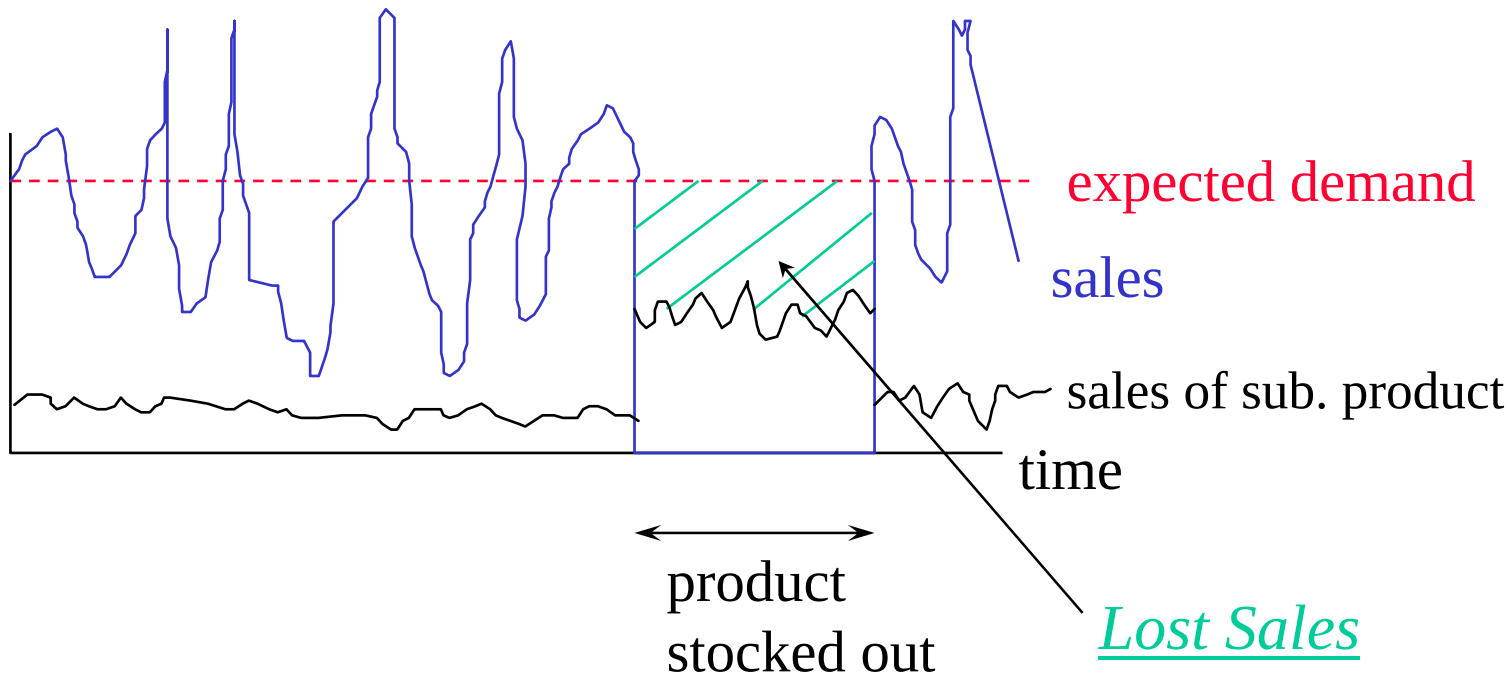
Lost sales estimation: Retailer A

Store	Actual Lost Sales	Estimated Lost Sales	Error (%demand)
A	1.61	1.44	-1.61
B	3.71	3.51	-1.70
C	0.90	0.85	-0.83
D	1.46	1.58	1.28
E	0.88	0.88	-0.01
F	1.29	1.31	0.22
G	0.80	0.91	1.71
H	0.00	0.01	0.11
I	1.02	1.09	0.72
J	1.41	1.36	-0.69
K	0.99	0.93	-0.76
District	14.08	13.87	-0.23

Lost sales estimation: Retailer B

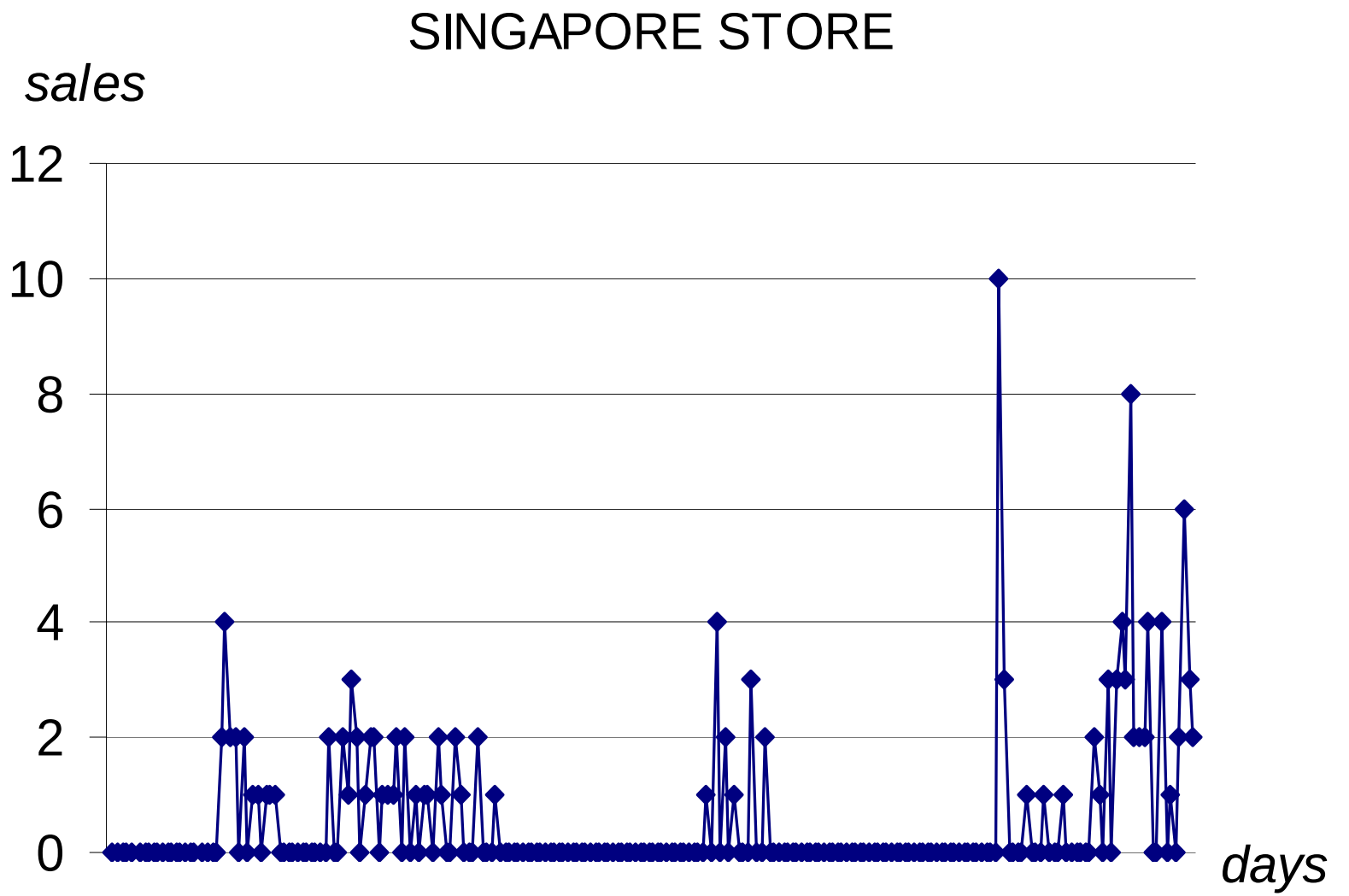
Store	Actual Lost Sales	Estimated Lost Sales	Error (%demand)
A	2.14	2.21	0.62
B	1.23	1.36	1.77
C	2.25	2.48	1.37
D	0.77	0.78	0.15
E	2.98	2.84	-0.72
F	1.18	1.14	-0.56
G	1.39	1.23	-1.61
H	1.6	1.39	-2.17
I	1.53	1.35	-1.56
J	0.49	0.50	0.42
K	1.19	1.36	2.50
District	16.75	16.62	-0.12

Extentions:Substitution



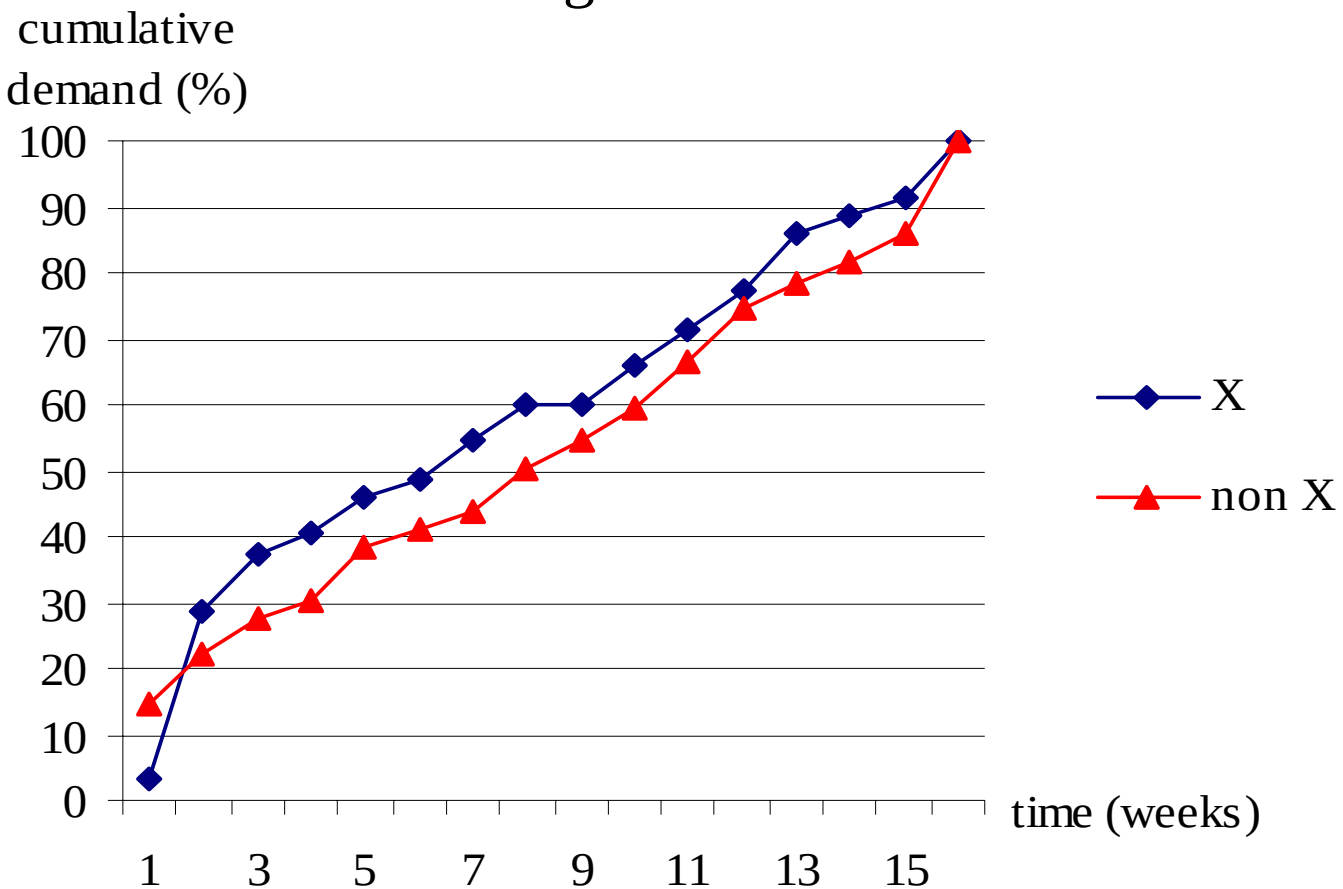
Some customers who did not find product A switched to another product at the same store.

An implementation in the luxury industry



How can we estimate potential demand?

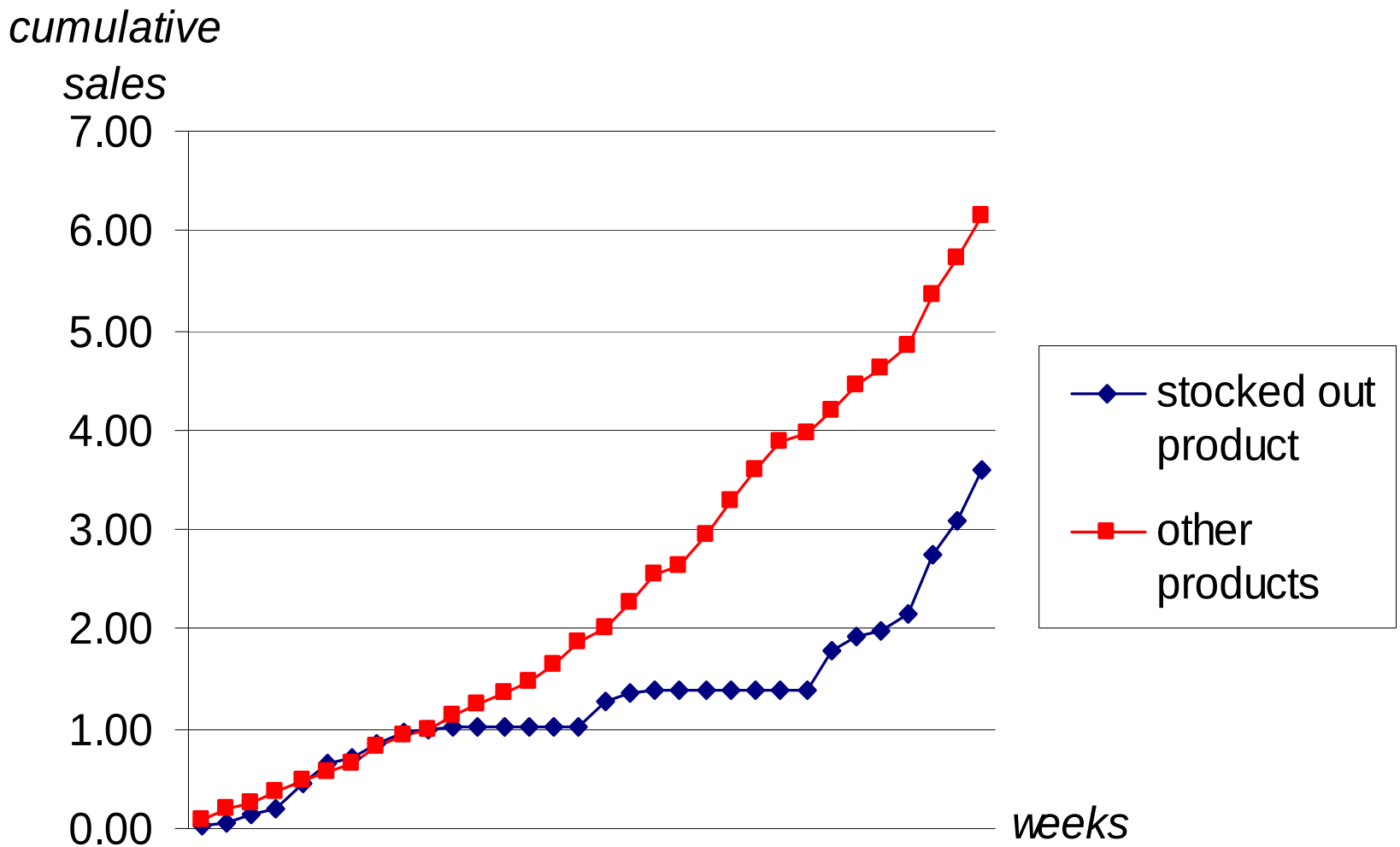
We can identify a variable which is correlated to the demand: sales of other items or traffic are good candidates



Implementation of the technique

- Given the technique we implemented it at some stores in the reatil chain.
- We considered a red-hot product.
- It stocked out early in the season.
- We provide estimates of demand and lost sales.
- Pb: non stationarity. Average demand in March is different from average demand in December.

The analysis



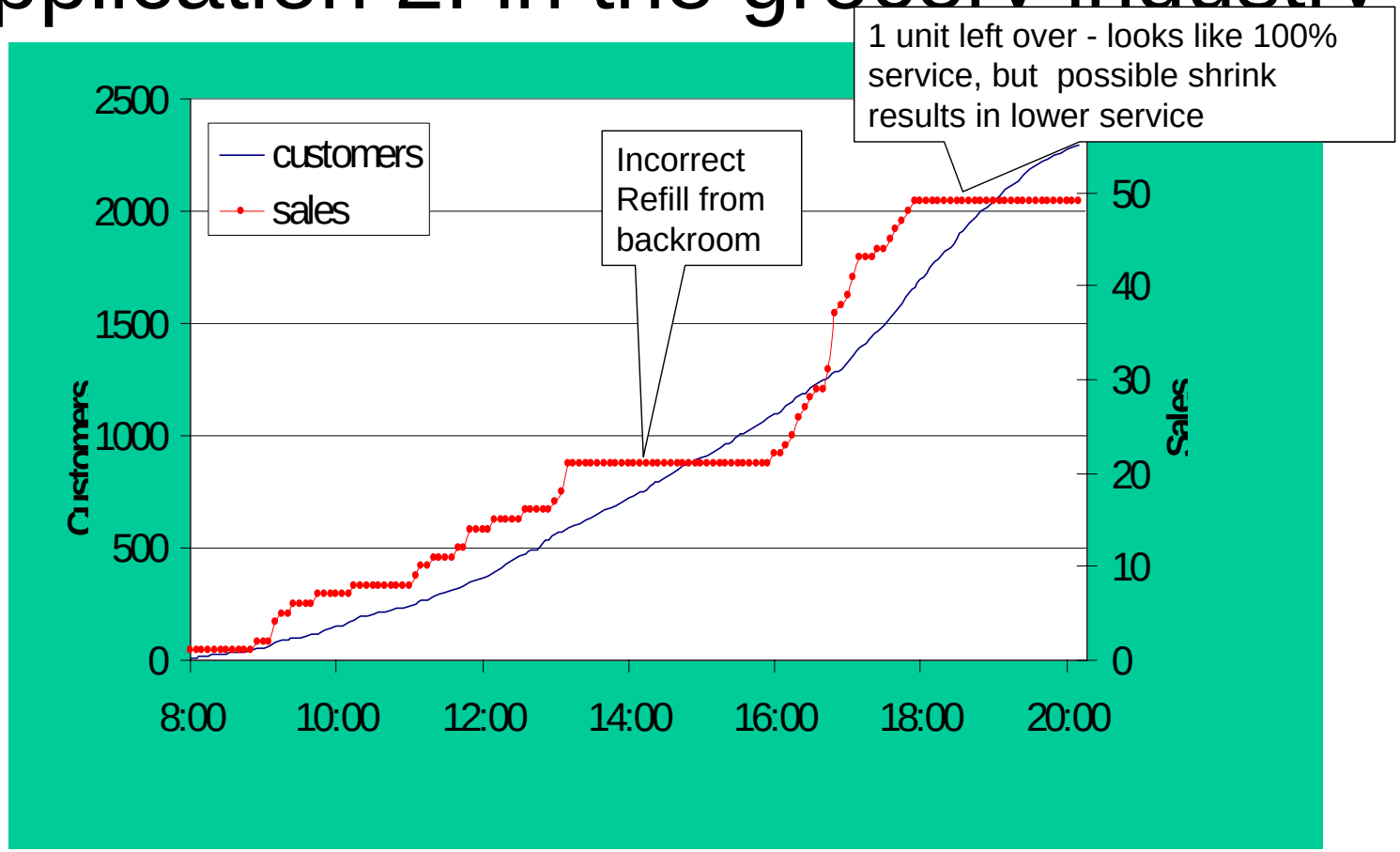
Extensive Analysis

- We extended the previous analysis to Italy and USA for new products
- USA
  - on average 29.7% of sales of new products
- Italy
  - on average 36% of sales of new products

Application 2: Grocery industry

- A top European retailer has already implemented these ideas and creates a stock out report.
- One problem they face, as most retailers do, is inventory records incaccuracy, i.e., it is hard to figure out when a product actually stocks out.
- The company on the one hand works hard to improve inventory record accuracy (it is a pre-condition to automate many processes including unattended replenishment)
- On the other hand figure out how to spot stock-out in the presence of inventory record inaccuracy.

Application 2: in the grocery industry



Demand can level-off either

-because none of the customers that entered the store after 5.45 were interested or

-Because the unit left was not there

Conclusions & Managerial Impact

- Lost sales are a relevant problem for the company.
- We can track this performance and try to improve forecasting by using more reliable demand data.
- The techniques are not trivial but easy to implement

$\Rightarrow$ managers can believe what the techniques say

- Problems:
  - they require some detailed data (store/sku level);
  - they might require some customisation

  $\Rightarrow$ a fruitful application requires managers' support.