“Unveiling Hidden Markov Models in Marketing”

Oded Netzer

Is This Model Dynamic?

$$U_{ijt} = \alpha_j - \beta_1 price_{ijt} + \beta_2 adv_{ijt} + \beta_3 display_{ijt} + \varepsilon_{ijt}$$

Individual i
Brand j
Time t

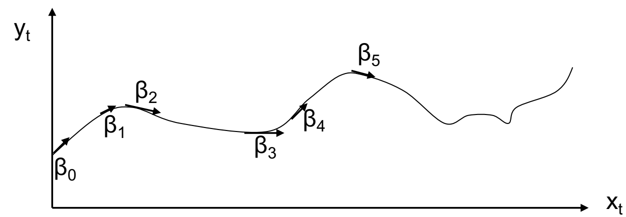
$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}$$

$$P(choice_{it} = j) = \frac{\exp(V_{ijt})}{\sum_j \exp(V_{ijt})}$$

Dynamic Parameters – State Space Models

$$U_{ijt} = \alpha_{ijt} - \beta_{i1t} price_{ijt} + \beta_{i2t} adv_{ijt} + \beta_{i3t} display_{ijt} + \varepsilon_{ijt}$$

Individual i
Brand j
Time t



Dynamic Parameters - Kalman Filter

$$Y_{ijt} = \boldsymbol{\beta}_{\mathbf{it}} \mathbf{X}_{\mathbf{ijt}} + \varepsilon_{ijt} \quad \text{Observation Equation}$$

$$\boldsymbol{\beta}_{\mathbf{it}} = \mathbf{T}_{\mathbf{t}} \boldsymbol{\beta}_{\mathbf{it-1}} + \mathbf{v}_{\mathbf{it}} \quad \text{Transition Equation}$$

Where
β_{it} is a state vector (MX1)
X_{ijt} is covariates vector (1XM)
T_t is a transition matrix (MXM) (subscript t is optional)

$$\varepsilon_{itj} \sim N(0, \sigma_\varepsilon^2);\ \ v_{it} \sim N(0, \Omega_v^2);\ \mathbf{v}_{i0} \sim N(\mathbf{m_0}, \mathbf{c_0})$$

References:

Meinhold and Singpurwalla (1983, TAS); Naik, Mantrala and Sawyer (1998, MS); Laachab, Ansari, Jedidi, and Trabelsi (2006, QME).

Kalman Filter – Special Cases

$Y_{ijt} = \boldsymbol{\beta}_{\mathbf{it}} \mathbf{X}_{\mathbf{ijt}} + \varepsilon_{ijt}$ Observation Equation

$\boldsymbol{\beta}_{\mathbf{it}} = \mathbf{T}_{\mathbf{t}} \boldsymbol{\beta}_{\mathbf{it+1}} + \mathbf{v}_{\mathbf{it}}$ Transition Equation

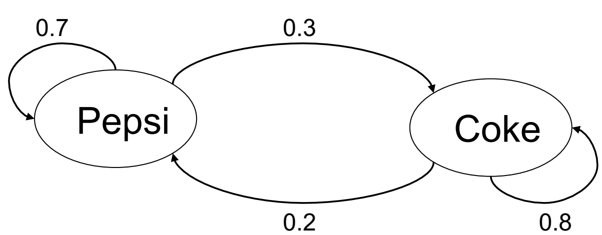
For $T_t = I$ and $v_{it} = 0$ we get…

For $T_t = I$ we get…

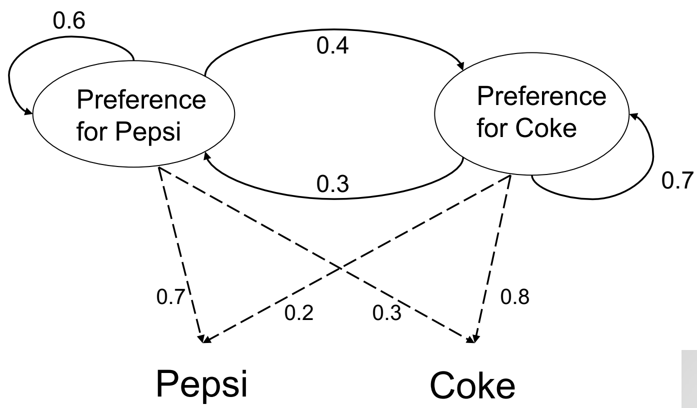
Other Special Cases…

Kalman Filtering on other latent states

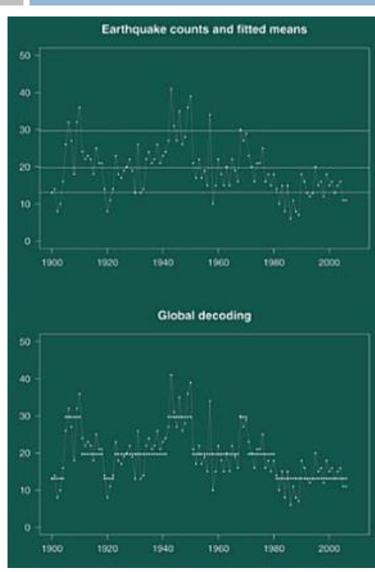
Markov Models

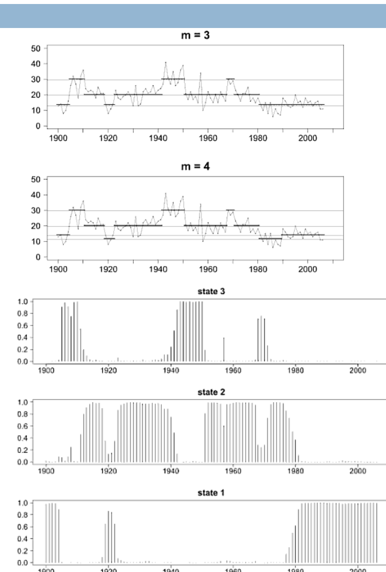


Hidden Markov Models (HMMs)



Example – Earthquakes

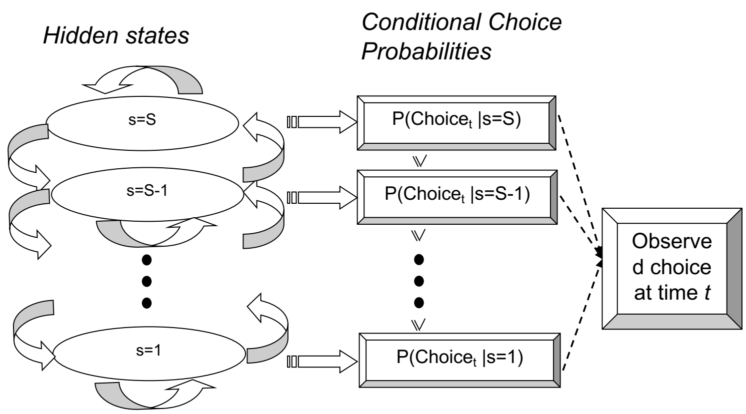




Source: Zucchini and MacDonald 2009

Hidden Markov Models

$$u_{it|s} = \beta_{is} x_{it} + \varepsilon_{ist}, \quad \varepsilon_{ist} \sim N(0, \sigma_s^2)$$



HMMs- Applications

- Engineering
  - Speech/image recognition (e.g., Jiang and Rabinar 1991)
  - Robot Navigation (e.g., Hannford and Lee 1991)
- Biology
  - DNA sequence information (e.g., Krogh 1998)
  - Ultrasound movement data (e.g., Leroux and Puterman 1992)
- Education
  - Students interest dynamics (e.g., Vermunt, Langeheine, Böckenholt 1999)
- Meteorology
  - Precipitation prediction (e.g., Hughes and Guttorp 1994)
- Economics and Finance
  - Predict stock market behavior (e.g., Ryden, Terasvirta and Asbrink 1998)
  - Identifying economic downturn (e.g., Hamilton 1989, 1990; Albert and Chib 1993)

…More examples in Zucchini and MacDonald 2009…

HMMs – Marketing Applications (incomplete list)

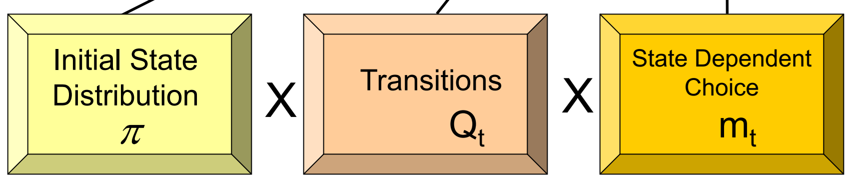
- **Dynamic segmentation-** Poulsen (1990); Brangule-Vlagsma, Pieters and Wedel (2002); Lemmens, Croux and Stremersch (2011)
- **Internet browsing** - Montgomery, Li, Srinivasan and Liechty (2004)
- **Family lifecycle** – Du and Kamakura (2006)
- **Visual attention states** – Liechty, Pieteres and Wedel (2003, 2008); van der Lans, Pieters and Wedel (2008)
- **Augmenting competitive response** – Moon, Kamakura and Ledolter (2007)
- **CRM** – Netzer, Lattin and Srinivasan (2008); Ascarza, Netzer and Hardie (2018); Zhang, Watson and Palmatier (2016)
- **Marketing mix allocation** – Montoya, Netzer and Jedidi (2010)
- **Social media interventions** - Ma, Sun and Kekre (2015)
- **B2B relationships** - Sriram et al. (2011); Zhang, Netzer and Ansari (2014); Lue and Kumar (2013)
- **Competitive dynamics** – Ebbes, Grewal and DeSarbo (2010)
- **Portfolio choice** – Paas, Vermunt and Bijmolt (2007); Schweidel, Bradlow and Fader (2011)
- **Cyclical buying** – Park and Gupta (2011)
- **Search** - Stuttgen, Boatwright and Monroe (2012)
- **Multi-channel** – Mark et al. (2013); Mark, Lemon and Vandenbosch (2014)
- **Dynamics learning in behavioral games** – Ansari, Montoya and Netzer (2012); Shachat and Wei (2012)
- **Predicting churn** – Ascarza and Hardie (2013)
- **Buy till you die models** - Schmittlein, Morrison and Colombo (1987); Fader, Hardie and Huang (2004); Schwartz, Bradlow and Fader (2014)

Today's plan

- Definition
- Components and Likelihood
- Estimation and model inference
- HMMs in Marketing
- Hands-on experience with R

The HMM Components

$$P_i(Y_{i1}=y_{i1}\ldots Y_{iT}=y_{iT}) = \sum_{s1\ldots sT} P(s_1)\prod_{\tau=2}^{T} P(s_\tau \mid s_{\tau-1})\prod_{\tau=1}^{T} P(Y_{i\tau}=y_{i\tau} \mid s_\tau)$$



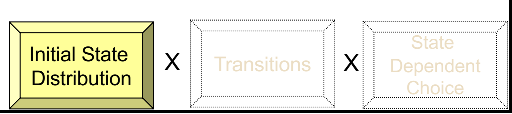
HMM - Initial State Membership

State membership at time 1

$$\boldsymbol{\pi}' = [\pi_1, \pi_2 \ldots, \pi_S] \quad \sum_{s=1}^{S} \pi_s = 1 \quad 0 \le \pi_s \le 1$$

Several options

1. Determine the distribution a-priori (e.g.,$\boldsymbol{\pi}' = [1,0,0,\ldots,0]$)
2. Assume the stationary distribution $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$
3. Estimate $\pi_1, \pi_2 \ldots, \pi_{S-1}$



HMM - Transition Matrix

State at t

$$\mathbf{Q}_{\mathbf{it}} = \begin{array}{c|ccccc|} & 1 & 2 & 3 & \cdots & S \\ \hline 1 & q_{it11} & q_{it12} & q_{it13} & \cdots & q_{it1S} \\ 2 & q_{it21} & q_{it22} & q_{it23} & \cdots & q_{it2S} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ S & q_{itS1} & q_{itS2} & q_{itS3} & \cdots & q_{itSS} \\ \hline \end{array}$$

State at t-1

$0 \le q_{itss'} \le 1 \; \forall$ s and s'

$\sum_{s'=1}^{S} q_{itss'} = 1 \; \forall$ s

- Homogenous HMM – $Q_{it} = Q_{ir} \;\; \forall \, t, r$
- Non-Homogenous HMM (Netzer et al. 2008) – $Q_{itss'} = f(Z_{it})$
  Could be modeled as logit or ordered logit
- Non-stationary HMM - $Q_{itss'} = f(\tau_s), \tau_s$ - is state duration in state s
- Restrictions on the transition matrix (e.g., absorbing state)

Initial State Distribution X **Transitions** X State Dependent Choice

HMM - State Dependent Behavior

The modular aspect of the HMM

- Binary choice - binary logit/probit
- Multinomial choice – multinomial logit/probit
- Count data – Poisson
- Continuous – Normal, Exponential, Gamma
- Multinomial – Multinomial distribution
- Multiple observations - Any combination of the above…

Can be a function of variables and covariates

Initial State Distribution X Transitions X **State Dependent Choice**

HMM - State Dependent Behavior - Logit

- The dichotomous choice given the state

$$m_{it|s} = \frac{\exp(\beta_{0s} + \mathbf{x_{it}}'\boldsymbol{\beta_s})}{1 + \exp(\beta_{0s} + \mathbf{x_{it}}'\boldsymbol{\beta_s})} \qquad s=1,\ldots,S$$

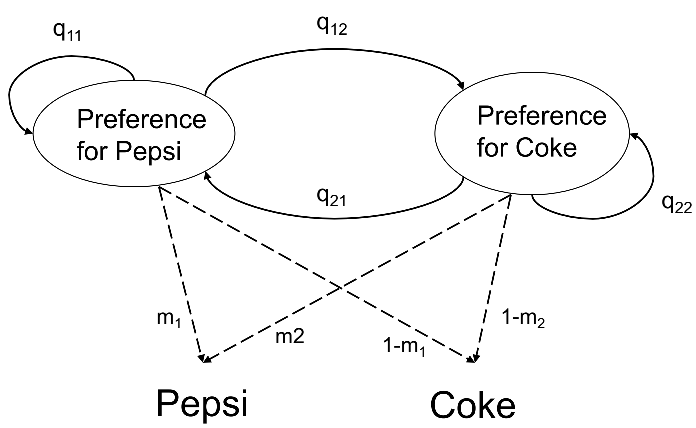
$$\mathbf{m_{it}}' = \left[m_{it|s=1}, m_{it|s=2}, \ldots, m_{it|s=N}\right]$$

  x_{it} are immediate influence covariates (e.g., price)

- Label Switching problem
  - Could put restrictions $\beta_{01} \le \beta_{02} \le \ldots \le \beta_{0S}$
  - Constrained permutation sampler – Fruhwirth-Schnatter (2001)

Initial State Distribution X Transitions X State Dependent Choice

Generating the HMM Likelihood



Generating the HMM Likelihood

Time →

1	2	3	4	5	…
PEPSI	PEPSI	Coca-Cola	Coca-Cola	Coca-Cola	
$\pi_1*m_1 + \pi_2*m_2$	$\pi_1*(q_{11}*m_1 + q_{12}*m_2) + \pi_2*(q_{21}*m_1 + q_{22}*m_2)$	$\pi_1*[q_{11}*q_{11}*(1-m_1) + q_{11}*q_{12}*(1-m_2) + q_{12}*q_{21}*(1-m_1) + q_{12}*q_{22}*(1-m_2)] + \pi_2*[q_{21}*q_{11}*(1-m_1) + q_{21}*q_{12}*(1-m_2) + q_{22}*q_{21}*(1-m_1) + q_{22}*q_{22}*(1-m_2)]$	…		

Generating the HMM Likelihood

1. Draw the Initial state membership π
2. Draw observation from the state dependent behavior m_t
3. Generate a transition from the current state s_t to s_{t+1} following the transition matrix Q_t
4. Repeat steps 2-3 until the last observation

HMM Likelihood Function

$$\begin{aligned} L_{iT} &= P(Y_{i1},...,Y_{it},...,Y_{iT}) \\ &= \sum_{X_{i1}=1}^{S} ... \sum_{X_{iT}=1}^{S} P(Y_{i1},...,Y_{iT} \mid X_{i1},...,X_{iT}) P(X_{i1},...,X_{iT}) \\ &= \sum_{X_{i1}=1}^{S} ... \sum_{X_{iT}=1}^{S} P(Y_{i1} \mid X_{i1}) \times P(Y_{iT} \mid X_{iT}) P(X_{i1}) P(X_{i2} \mid X_{i1}) ... P(X_{iT} \mid X_{iT-1}) \\ &= \sum_{X_{i1}=1}^{S} ... \sum_{X_{iT}=1}^{S} P(X_{i1}) P(Y_{i1} \mid X_{i1}) P(X_{i2} \mid X_{i1}) P(Y_{i2} \mid X_{i2}) ... P(X_{iT} \mid X_{iT-1}) P(Y_{iT} \mid X_{iT}) \end{aligned}$$

- "Simplified" likelihood function (S^T elements!)

$$L_{iT} = P(Y_{i1}, Y_{i2}, ..., Y_{iT}) = \boldsymbol{\pi}_{\mathbf{i}} (\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S) \mathbf{Q}_{\mathbf{i12}} (\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S) ... \mathbf{Q}_{\mathbf{iT\text{-}1T}} (\tilde{\mathbf{m}}_{\mathbf{iT}} \otimes \mathbf{I}_S) \mathbf{1}'$$

Today's plan

- Definition
- Components and Likelihood
- Estimation and model inference
- HMMs in Marketing
- Hands-on experience with R

HMM Estimation

- Baum-Welch EM Algorithm
  - Based on the forward/backward probabilities
  - References: Baum et al. (1970); Baum (1972); Zucchini and MacDonald 2009 (book)
- Maximum Likelihood
  - Maximizing the likelihood function from before
  - References: Zucchini and MacDonald 2009 (book); Netzer, Lattin and Srinivasan (2008) – Bayesian version
- Bayesian Estimation – Augmenting the states
  - Latent states are treated as missing data
  - References: Frühwirth-Schnatter (2006) – book; Albert and Chib (JBES 1993); Scott S. (JASA 2002)

HMM Estimation - MLE

- Advantages
  - Easy to estimate with standard maximum likelihood optimizer
  - Easy to handle missing data and constraints
  - Can be extended to HB framework to account for heterogeneity
- Difficulties
  - Numerical underflow of the L_{it} (see solution in ZM p. 46-47)
  - Local maxima
  - Initial state distribution
  - MCMC - Adaptive tuning constant - Atchade (2006)

MLE code provided at the workshop

HMM State Recovery

- State transition

$$P(S_{it} \mid S_{it-1}) = \pi_{it-1} Q_{it-1t}$$

- Filtering

$$P(S_{it} \mid Y_{i1} \ldots Y_{it}) = L_{it}(s^{th} \text{ element}) / L_{it} = L_{it}(s^{th} \text{ element}) / \sum_{j=1}^{S} L_{it}(j^{th} \text{ element}) =$$

$$\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t.s}\, m_{it|s} / L_{it}$$

$Q_{it\text{-}1t.s}$ is the s^{th} column of Q

- Smoothing

$$P(S_{it} \mid Y_{i1} \ldots Y_{iT}) =$$

$$\underbrace{\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t.s}}_{\alpha(s)} \boxed{m_{it|s}} \underbrace{\mathbf{Q}_{\mathrm{i}tt+1s.}(\tilde{\mathbf{m}}_{\mathrm{i}t+1} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathrm{i}1t+2}(\tilde{\mathbf{m}}_{\mathrm{i}t+1} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}1T}(\tilde{\mathbf{m}}_{\mathrm{i}T} \otimes \mathbf{I}_S)\mathbf{1}'}_{\beta(s)} / L_{iT}$$

HMM Prediction

- Prediction

$$P(Y_{it+1} \mid Y_{i1},\ldots,Y_{it}) = L_{it+1} / L_{it} = \frac{\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t}(\tilde{\mathbf{m}}_{\mathrm{i}t} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathrm{i}tt+1}(p(Y_{it+1j|s} \otimes \mathbf{I}_S)\mathbf{1}'}{\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t}(\tilde{\mathbf{m}}_{\mathrm{i}t} \otimes \mathbf{I}_S)\mathbf{1}'}$$

$$P(Y_{it+h} \mid Y_{i1},\ldots,Y_{it}) = \frac{\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t}(\tilde{\mathbf{m}}_{\mathrm{i}t} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathrm{i}tt+1}\mathbf{Q}_{\mathrm{i}t+1t+2}\cdots\mathbf{Q}_{\mathrm{i}t+h-1t+h}(p(Y_{it+hj|s} \otimes \mathbf{I}_S)\mathbf{1}'}{\boldsymbol{\pi}_{\mathbf{i}}(\tilde{\mathbf{m}}_{\mathbf{i1}} \otimes \mathbf{I}_S)\mathbf{Q}_{\mathbf{i12}}(\tilde{\mathbf{m}}_{\mathbf{i2}} \otimes \mathbf{I}_S)\ldots\mathbf{Q}_{\mathrm{i}t\text{-}1t}(\tilde{\mathbf{m}}_{\mathrm{i}t} \otimes \mathbf{I}_S)\mathbf{1}'}$$

- Can predict several periods a head

Selecting the Number of States

- Similar to latent class models – model selection criteria
- Classic estimation
  - Penalized fit measures: e.g., BIC, AIC.
  - Markov Switching Criterion (MSC)- Smith, Naik and Tsai (2006)
  - Predictive measures
- Bayesian Estimation
  - Log Marginal Density – Bayes factor
  - DIC
  - Posterior predictive distributions
- Interpretation and parsimony considerations

Incorporating Heterogeneity

- Why is it important in dynamic models?
- Heterogeneity in:
  - Initial state distribution
  - Transitions
  - State dependent choice
- Relating the distribution of heterogeneity to observed individual characteristics

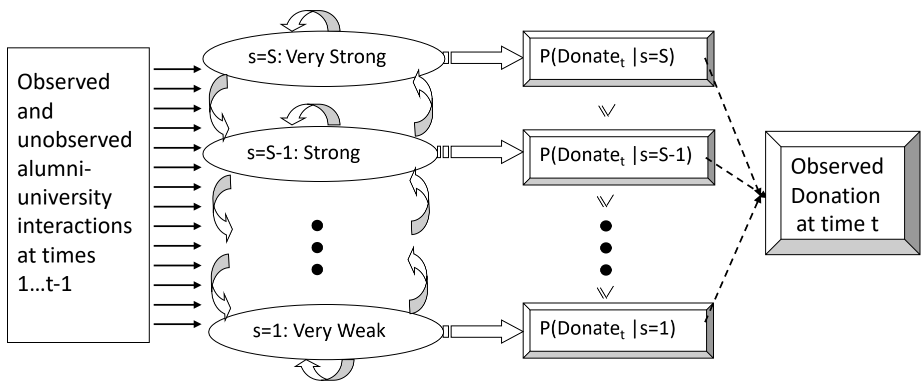
Discrete vs. Continuous States
HMM vs. Kalman Filter

- Semi-parametric (HMM) vs. parametric (KF) form of dynamics
- Regime shift dynamics
- HMM with a large number of states should approximate KF (but it can become very expensive)
- HMM is often easier to interpret and communicate

Today's plan

- Definition
- Components and Likelihood
- Estimation and model inference
- HMMs in Marketing
- Hands-on experience with R

A Hidden Markov Model of Customer Relationship Dynamics (Netzer, Lattin and Srinivasan 2008)



The Conditional Choice Behavior

We identified three states of relationships with the following donation rates:

State	Donation Probability	Name
State 1	**0%**	**"Dormant"**
State 2	**32%**	**"Occasional"**
State 3	**99%**	**"Active"**

Netzer, Lattin and Srinivasan 2008

The Transition Matrices

No Alumni-University Interactions			
t-1 \ t	Dormant	Occasional	Active
Dormant	97%	3%	0%
Occasional	18%	73%	9%
Active	0%	22%	78%

Reunion Attendance			
t-1 \ t	Dormant	Occasional	Active
Dormant	67%	33%	0%
Occasional	4%	62%	34%
Active	0%	21%	79%

- The states are relatively "sticky"
- Twice as likely to fall down than go up from occasional

- Reunion attendance
  - Strong effect on dormant and occasional
  - Minimal effect on active

Netzer, Lattin and Srinivasan 2008

Crossing the States with Survey Data

Question	Scale	Dormant	Occasional	Active
Satisfaction with your experience at Stanford	1-5	4.51	4.75	4.80
Strong feeling about Stanford	1-5	4.31	4.50	4.60
Pride in your Degree	1-4	3.47	3.62	3.69
University experience helped *shape your life*	1-4	**2.89**	**3.24**	**3.43**
Emotional connection	1-4	**2.72**	**3.14**	**3.22**
Responsibility	1-4	**2.28**	**2.66**	**3.03**
Affinity with graduating class	1-4	**1.94**	**2.52**	**2.26**
Recommend to prospective students	1-4	**3.35**	**3.67**	**3.74**
University serves your needs as an alum?	1-4	2.74	2.93	2.96
University *values* its alumni	1-3	**2.25**	**2.41**	**2.55**
Parents have a degree from Stanford	Yes/No	19%	18%	12%
Received *financial aid*	Yes/No	40%	40%	39%
Median Lifetime donation		$100	$475	$1382
Sample Size (N)		64	29	35

* Bolded means are significantly different across the states at the 0.05 level

- Active alumni show favorable ratings

Netzer, Lattin and Srinivasan 2008

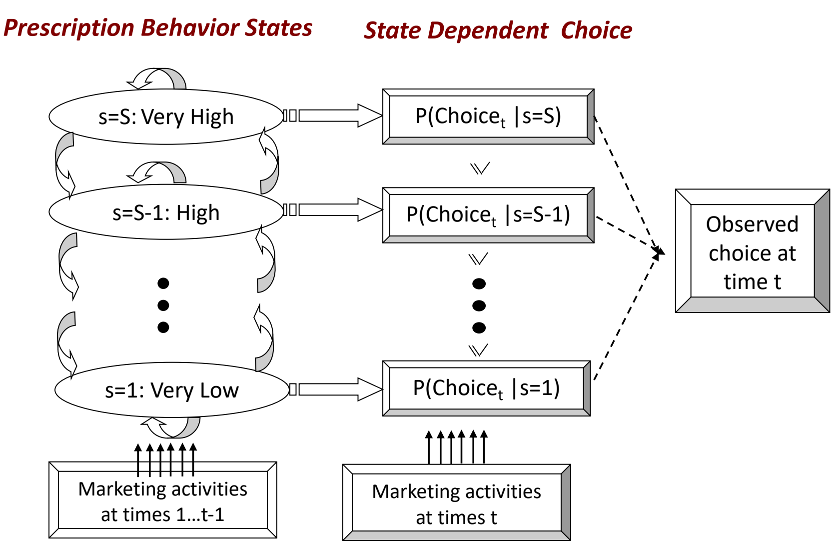
Dynamic Allocation of Pharmaceutical Detailing and Sampling for Long-Term Profitability (Montoya, Netzer and Jedidi 2010)

- A drug for treating a medical condition in postmenopausal women
- New drug introduction
- Several competing brands
- Two years of monthly-level data
- Sample of 300 physicians
- Monthly data:
  - Prescriptions of the drug
  - Prescriptions in the category
  - Sampling and detailing at the physician-level

	Mean	95% C. I.	
Prescriptions	1.62	0.54	3.21
Details	2.18	1.22	3.71
Samples	9.07	4.17	16.33
Category prescriptions	22.5	10.10	37.79
Share-of-prescriptions	0.079	0.026	0.143

- Average monthly statistics across physicians

HMM of prescription behavior



Montoya, Netzer and Jedidi 2010

Short and Long-Term Effects of Marketing Actions

Probability of prescribing - Share

State	Intercept	Detailing	Sampling
Inactive	0.4 %	0.7%	0.4%
Infrequent	6.2 %	6.5%	6.7%
Frequent	19.6%	18.7%	20.1%

Transition Matrix Detailing

I	0.62	0.38	0.00
IF	0.16	0.79	0.05
F	0.15	0.45	0.40
	I	IF	F

Transition Matrix No Marketing

I	0.75	0.25	0.00
IF	0.17	0.78	0.05
F	0.15	0.46	0.39
	I	IF	F

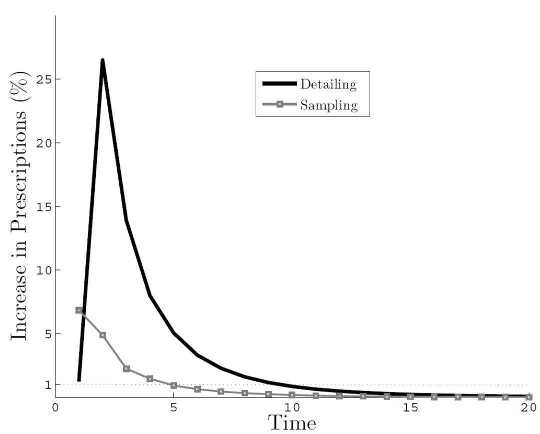
Transition Matrix Sampling

I	0.70	0.30	0.00
IF	0.13	0.81	0.06
F	0.10	0.41	0.49
	I	IF	F

- Detailing mainly affect physicians in the Inactive states
- Sampling mainly affect physicians only in the Frequent state

Montoya, Netzer and Jedidi 2010

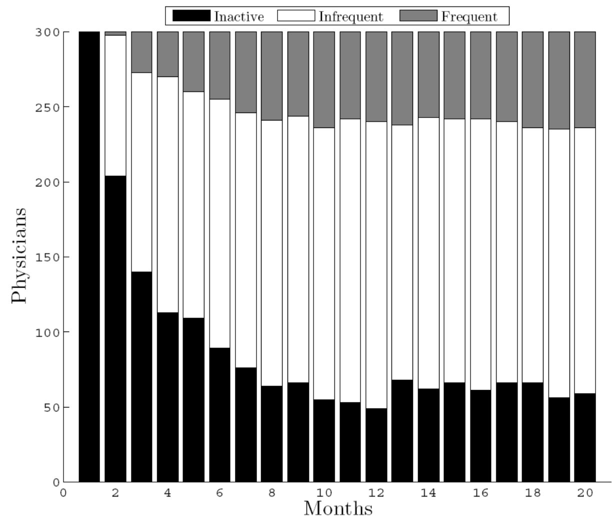
Duration of Marketing Actions



- 25% of the total effect of *detailing* occurs the first month
- 35% of the total effect of *sampling* occurs the first month

Montoya, Netzer and Jedidi 2010

Recovering the State Membership



Montoya, Netzer and Jedidi 2010

From Estimation to Control

	Model Estimation
Observable States	Markov Chain
Unobservable States	HMM Hidden Markov Model

Montoya, Netzer and Jedidi 2010

Dynamic Targeted Pricing in B2B Settings (Zhang, Netzer and Ansari 2014)

Two Buying Behavior States

	"Relaxed" State	"Vigilant" State
Quote request prob.	23%	86%
Bid accept prob.	65%	52%
Average quantity	432 lb	502 lb
Inter-purchase time	5.5 weeks	8.1 weeks
Average price elasticity	1.3	3.4
Average loss aversion parameter	0.92	3.11
Average sensitivity to market characteristics	0.8	6.7

Zhang, Netzer and Ansari 2014

Transition Matrices

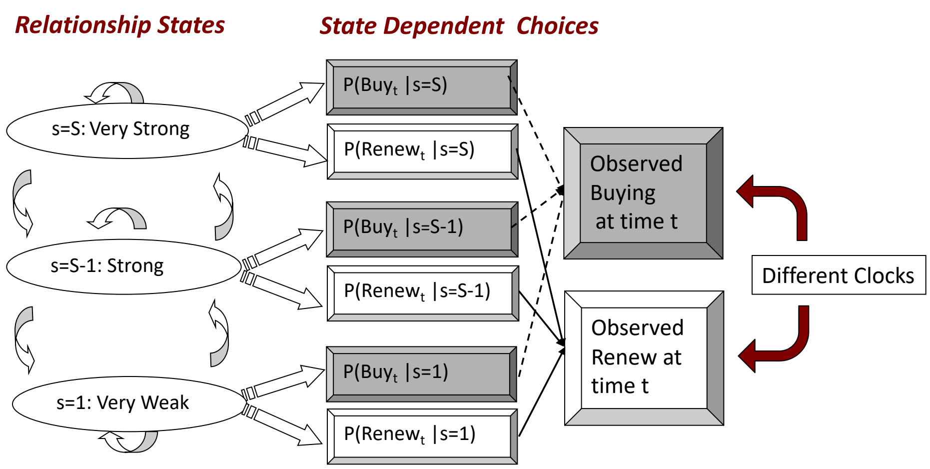
Average price

Relaxed (t+1)	Vigilant (t+1)
86.2%	13.8%
7.1%	92.9%

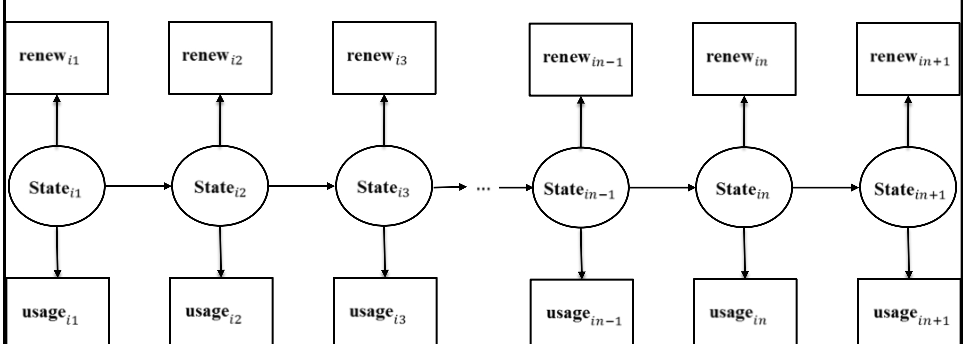
- Both states are sticky
- 10% price increase -> 50% increase of transition from relaxed state to vigilant state
- Loss aversion
- Capturing long-term effect of reference prices

Zhang, Netzer and Ansari 2014

A Joint Model of Usage and Churn in Contractual Settings (Ascarza and Hardie 2013)

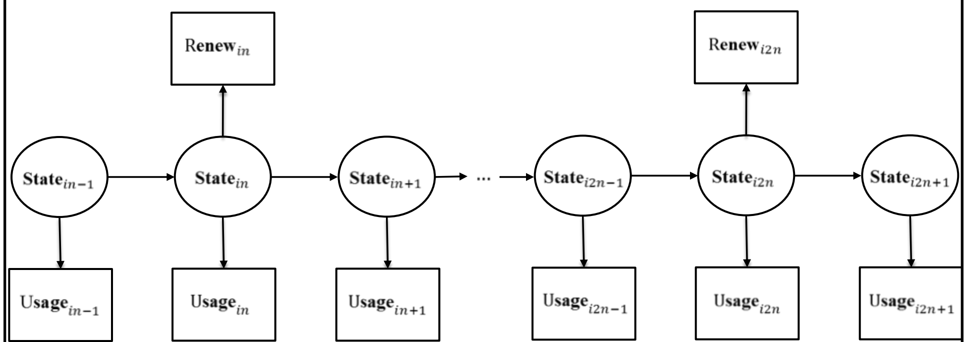


A Joint Model of Usage and Churn



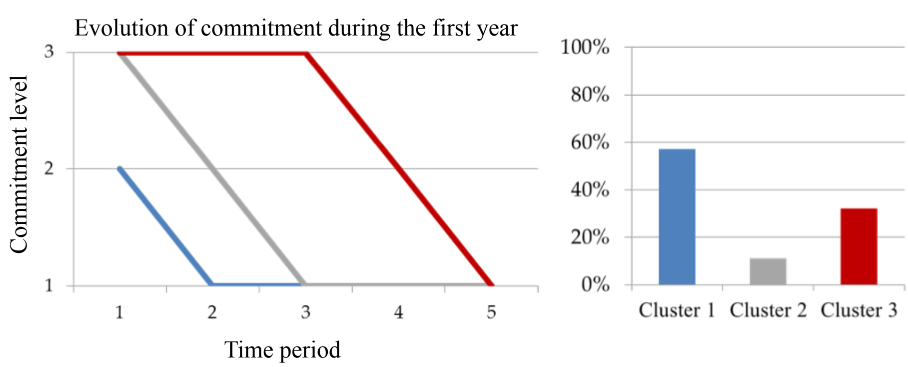
Ascarza and Hardie 2013

Different "Clocks" in Contractual Settings



Ascarza and Hardie 2013

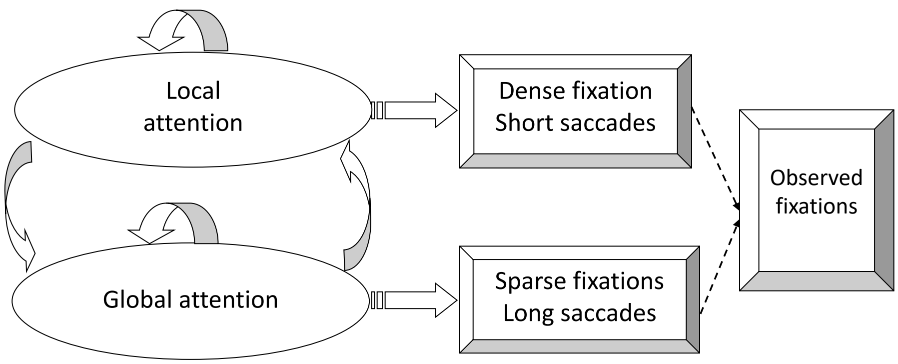
Identifying "Paths to Death"



Ascarza and Hardie 2013

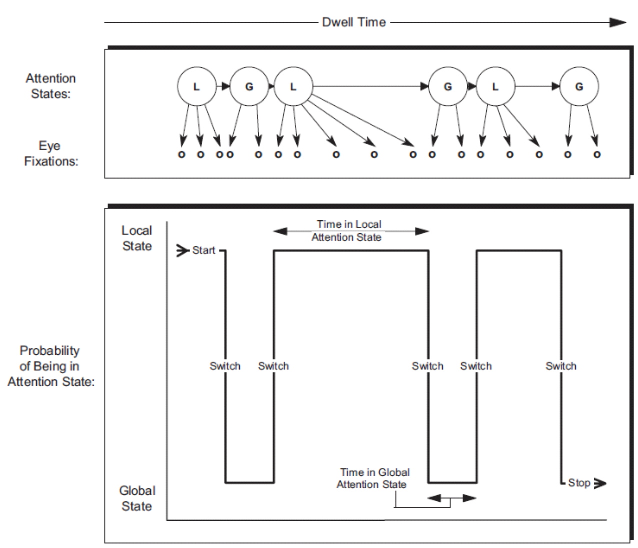
HMM applied to Eye Movement Data (Psychology) (Wedel, Pieters and Liechty 2008)

States of Attention **Attention behavior**



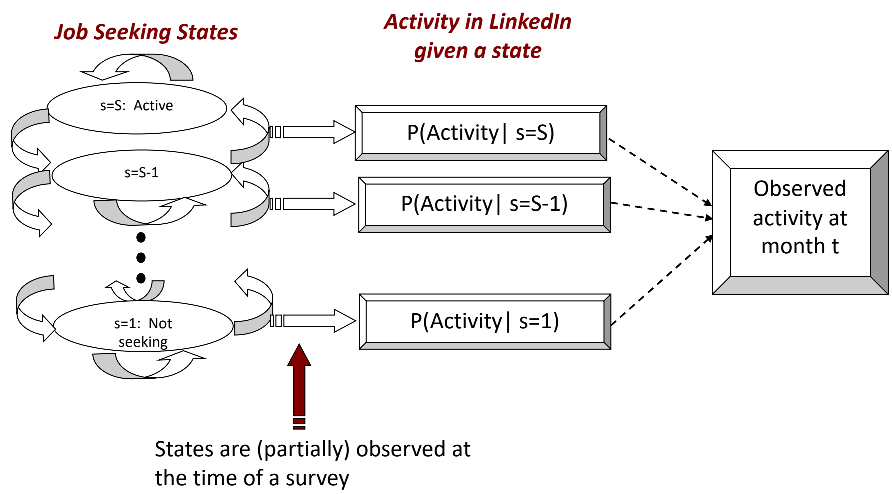
Wedel, Pieters and Liechty 2008

Transitions between local and global attention states



Wedel, Pieters and Liechty 2008

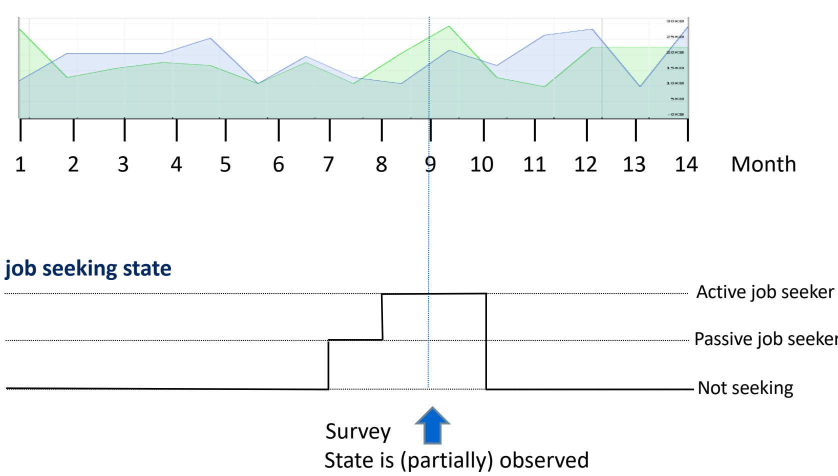
Using Social Media Data to Identify and Target Job Seekers (Ebbes and Netzer 2016)



Fusing the Two Sources of Data

Multivariate site activity

Page views, invitations to connect, job searches, etc…



Ebbes and Netzer 2016

Site Activity

	State 1 - Non Job Seeker	State 2 - Passive Job Seeker	State 3 - Active Job Seeker
Profile update	3%	17%	23%
Job search	1%	12%	40%
Total searches	1.7	1.6	5.0
Page views	7.7	37.9	112.2

(calibrated on monthly data)

Job seekers leverage the LinkedIn website while being an active searcher

Ebbes and Netzer 2016

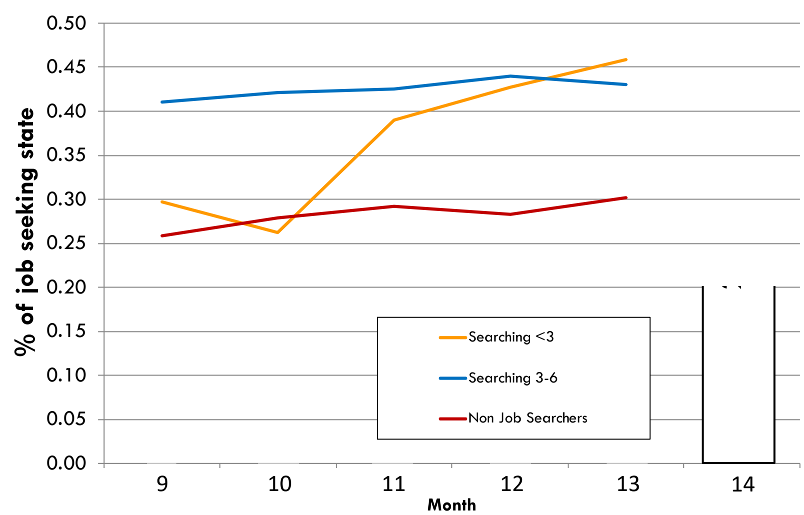
Social Network Activity

	State 1 - Non Job Seeker	State 2 - Passive Job Seeker	State 3 - Active Job Seeker
Invitations outside company > inside company	2%	8%	37%
Invitations sent	1.7	1.3	2.4
Invitations received / sent	0.84	0.74	0.49
# connections formed	1.3	1.8	3.1
# connections invitees	19.8	20.6	26.2

Active job seekers (try to) grow their network faster, in a strategic way, but are also a bit the “Homers”

Ebbes and Netzer 2016

Predicting Length of Job Seeking

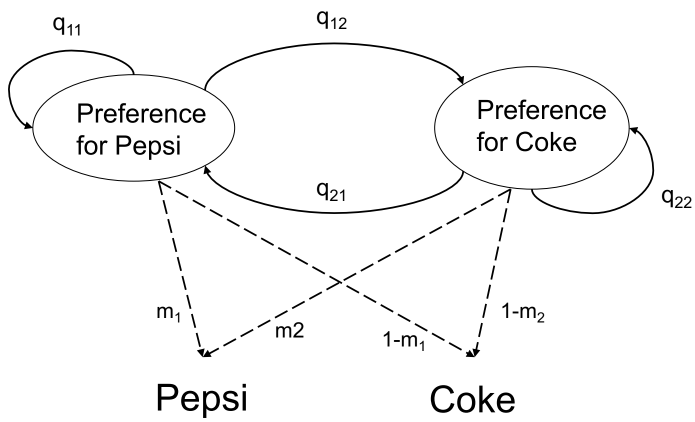


Ebbes and Netzer 2016

Today's plan

- Definition
- Components and Likelihood
- Estimation and model inference
- HMMs in Marketing
- Hands-on experience with R

Model to Estimate



R code – To recap…

- Simulation: Data generating process
- Likelihood building
- Maximize likelihood
  - Local optima
  - Fit measures
- Model inferences
  - Parameters
  - States
- Model variants
  - # states
  - Continuous vs. Discrete (e.g., #purchases)

Conclusions

- HMMs in marketing can be used to:
  - dynamically segment the firm's customer base
  - understand how customers transition among segments over time (possible due to touch points with the firm)
  - capture the long and short-term effect of marketing actions and pricing decisions
  - capture transitions between distinct behaviors (e.g. learning rules)
  - augment unobserved behaviors
  - identify behavioral states in behavioral research
  - fuse different sources of data

References – General

- Books
  - Zucchini and MacDonalnd – 'Hidden Markov Models for Time Series"
  - Cappe, Moulines and Ryden - "Inference in Hidden Markov Models"
  - Elliot, Aggoun and Moore - "Hidden Markov Models"
  - Kim and Nelson – "State-space Models with Regime Switching"
  - Fruhwirth-schnatter - "Finite Mixture And Markov Switching Models"
- Tutorials
  - Rabiner, L. (1989) "A Tutorial in Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, 77(2): 257-286.
  - Visser I. (2011) "Seven Things to Remember about Hidden Markov Models: A Tutorial on Markovian Models for Time Series," *Journal of Mathematical Psychology, 55,* 403-415
  - Netzer Oded, Peter Ebbes and Tammo Bijmolt (2017), "Hidden Markov Models in Marketing." Advanced Methods for Modeling Markets, edited by Peter Leeflang, Jaap Wieringa, Koen Pauwels, Springer