Statistiktage 2011 Graz, September 7 – 9, 2011

Female Wage Careers - A Bayesian Analysis Using Markov Chain Clustering

Regina Tüchler, Wirtschaftskammer Österreich

Christoph Pamminger, The Austrian Center for Labor Economics and the Analysis of the Welfare State

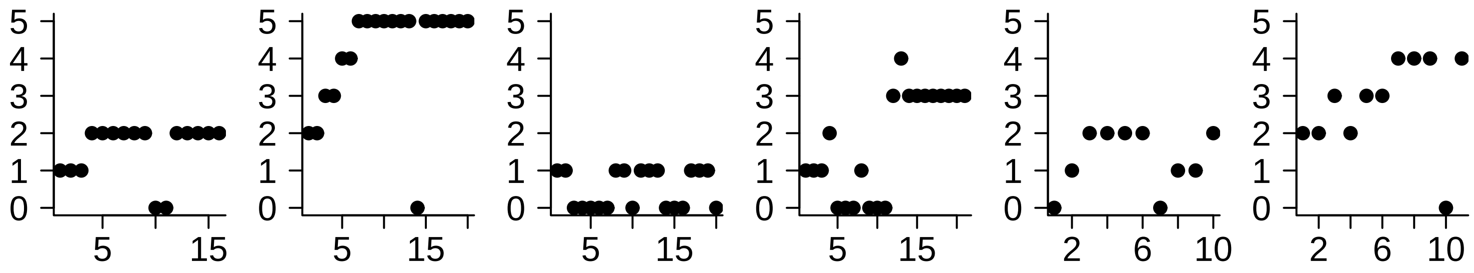
Outline

1. Analyzing wage dynamics
2. The data
3. The method:
   - Markov chain clustering
   - Mixture-of-experts Model
   - MCMC
4. Results
5. Conclusions

Analyzing Wage Dynamics

We analyze female wages **over a time period**.

6 income categories: 0 = no income, 1-5 quintiles of the income distribution



Q1 Are there **groups** of women with similar patterns in their wage dynamic?

Q2 Which **variables** influence the wage dynamic?

The Data

- Austrian Security Data Base with 183 805 female employees
- entry in the labor market between 1980 and 1985
- observation period till 2001 (in 2002 change of qualifying conditions for maternity leave)
- time series length: up to 22 years (median length: 14 years)
- adjusted for long-term unemployed (ts cut after five years of zero-income)

The Data

- age at entry: 14 to 25 years (64.2 % were 17-19 years old)
- start as blue collar worker: 41.1 %; as white collar worker: 58.9 %
- at least once on maternity leave: 72.7 %
- number of children: number of live birth announcements ($\bar{x} = 1.42$, $\tilde{x}_{0.5} = 1$)

see Zweimüller et al. (2009)

Model Based Clustering

$\mathbf{y}_i = \{y_{i1}, \ldots, y_{i,T_i}\}$ ... time series of income states for individual i
$y_{it} \in \{0, \ldots, K\}$ for $i = 1, \ldots, N, t = 1, \ldots, T_i$

Finite mixture model with H components:

$$\sum_{h=1}^{H} \eta_h \, p(\mathbf{y}_i | \boldsymbol{\xi}_h)$$

$\boldsymbol{\xi}_h$ describes the time-series of group h
η_h ... group specific weights

see Frühwirth-Schnatter & Kaufmann (2008)

Markov Chain Model

First-order time-homogeneous Markov chain model:

$$\xi_{jk} = \Pr(y_{it} = k | y_{i,t-1} = j) \qquad \text{and} \qquad \sum_{k=0}^{K} \xi_{jk} = 1$$

$$\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_{0\cdot} \\ \boldsymbol{\xi}_{1\cdot} \\ \vdots \\ \boldsymbol{\xi}_{K\cdot} \end{pmatrix} = \begin{pmatrix} \xi_{00} & \xi_{01} & \cdots & \xi_{0K} \\ \xi_{10} & \xi_{11} & \cdots & \xi_{1K} \\ \vdots & & \ddots & \vdots \\ \xi_{K0} & \xi_{K1} & \cdots & \xi_{KK} \end{pmatrix}$$

Each row represents an unknown discrete probability distribution.

Markov Chain Model

We introduce Markov chain models as clustering kernels. All time series within a cluster are described by the same cluster-specific transition matrix $\boldsymbol{\xi}_h$.

$$p(\mathbf{y}_i|\boldsymbol{\xi}_h) = \prod_{j=0}^{K} \prod_{k=0}^{K} (\xi_{h,jk})^{N_{i,jk}}$$

$N_{i,jk} = \#\{y_{it} = k, y_{i,t-1} = j\}$ is the number of switches of individual i from state j to state k

see Frühwirth-Schnatter & Pamminger (2010)

Modeling Group Membership

We incorporate unit-specific information to assign each individual to one group.

Multinomial Logit Model (MNL):

$$\Pr(S_i = h | \mathbf{x}_i, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_H) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_h)}{1 + \sum_{l=2}^{H} \exp(\mathbf{x}_i \boldsymbol{\beta}_l)}$$

$S_i \in \{1, \ldots, H\}$ ... group indicators, $i = 1, \ldots, N$

$\mathbf{x}_i$ ... row vector of regressors

$\boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_H$ ... group-specific unknown parameters

Priors, Assumptions

Assumptions:

- Set $\boldsymbol{\beta}_1 = \mathbf{0}$, which means that $h = 1$ is the baseline group and $\boldsymbol{\beta}_h$ is the effect on log-odds ratio relative to the baseline.
- Rows of $\boldsymbol{\xi}_h$ are a priori independent.
- Prior independence between $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_H$ and $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_H$.
- Conditional on knowing $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_H$ the observations $\mathbf{y}_1, \ldots, \mathbf{y}_N$ are mutually independent.

Priors:

- $\boldsymbol{\xi}_{h,j\,\cdot} \sim$ Dirichlet distributions with known parameters.
- $\boldsymbol{\beta}_h \sim$ normal distributions with known parameters.

MCMC Estimation

1. **Sample transition matrices $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_H$ given $\mathbf{S}$:**

   draw each $\boldsymbol{\xi}_{h,j\cdot}$ from the Dirichlet distribution $p(\boldsymbol{\xi}_{h,j\cdot}|\mathbf{S}, \mathbf{y})$.

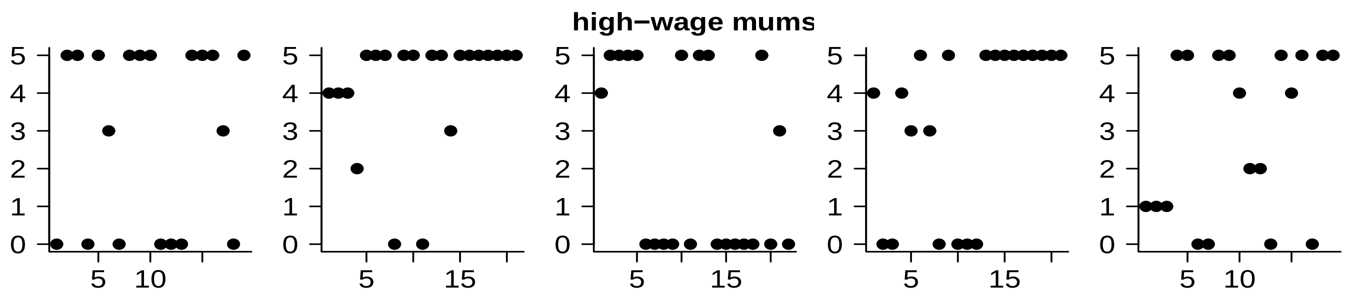
2. **Sample parameters $\boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_H$ given $\mathbf{S}$:**

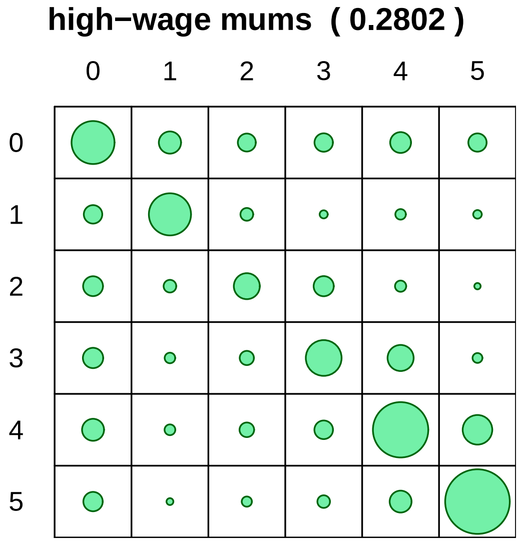
   auxiliary mixture sampling of $\boldsymbol{\beta}_h$ from the MNL involves only standard distributions (Frühwirth-Schnatter and Frühwirth 2010).

3. **Bayes' classification for each individual i:**

$$\Pr(S_i = h|\mathbf{y}_i, \mathbf{x}_i) \propto p(\mathbf{y}_i|\boldsymbol{\xi}_h) \frac{\exp\left(\mathbf{x}_i\boldsymbol{\beta}_h\right)}{1 + \sum_{l=2}^{H} \exp\left(\mathbf{x}_i\boldsymbol{\beta}_l\right)}, \quad h = 1, \ldots, H.$$

Group 1: High-Wage Mums

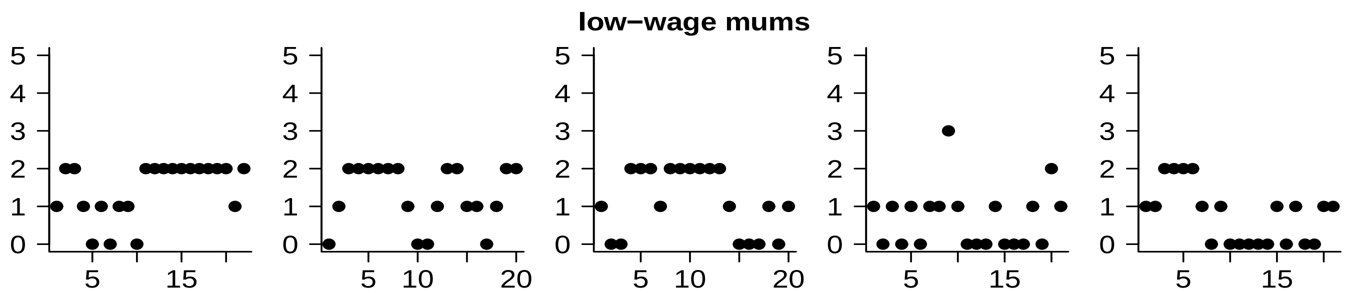


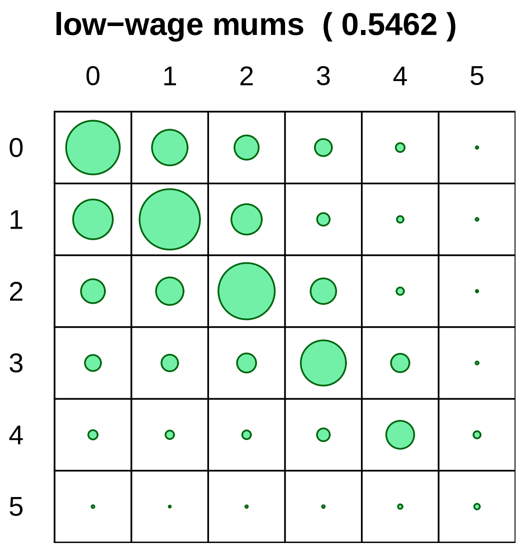


Ex post analysis:

av. age at job entry: 19.2 y.
started as white collar: 83.8 %
at least once on maternity leave: 72.6 %
number of children: $\bar{x} = 1.37$, $\tilde{x}_{0.5} = 1$

Group 2: Low-Wage Mums

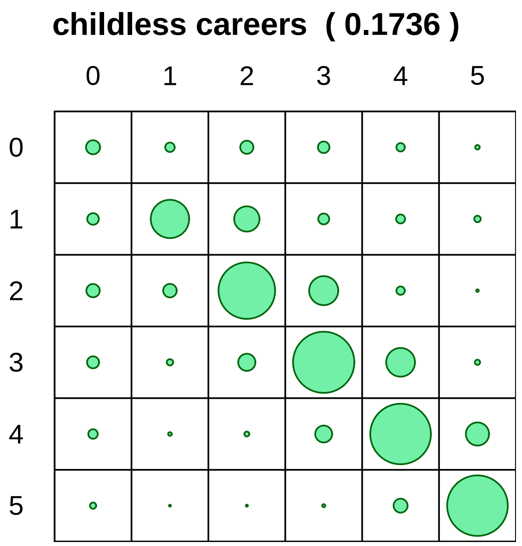




Ex post analysis:

av. age at job entry: 17.7 y.
started as white collar: 44.3 %
at least once on maternity leave: 93.4 %
number of children: $\bar{x} = 1.79$, $\tilde{x}_{0.5} = 2$

Group 3: Childless Careers



Ex post analysis:

av. age at job entry: 18.3 y.
started as white collar: 74.7 %
at least once on maternity leave: 2.6 %
number of children: $\bar{x} = 0.22$, $\tilde{x}_{0.5} = 0$

MNL for Group Membership

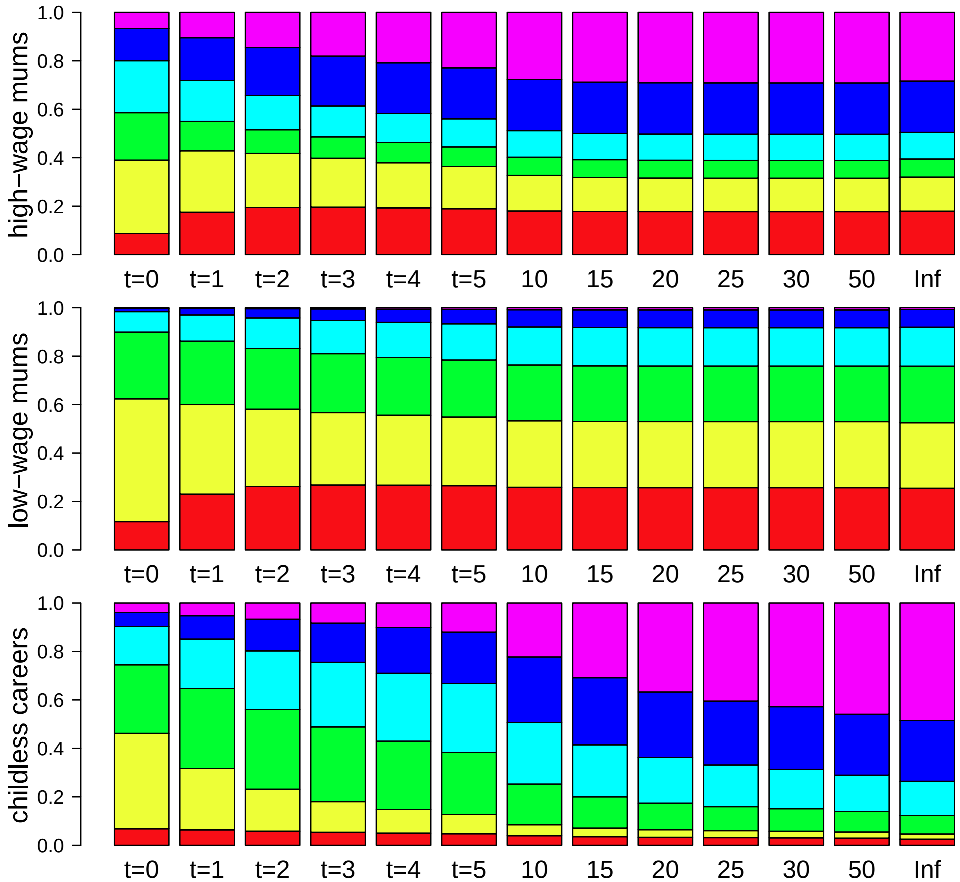
Estimates of Regression coefficients: effect on log odds (baseline: low-wage mums)

	"high-wage"	"childless careers"
Intercept	7.98821	5.05117
blue collar × maternity leave	-1.11074	-6.60941
white collar × no maternity leave	2.38276	2.66115
white collar × maternity leave	0.42969	-1.79956
Number of children	-0.22701	-0.63768
Age at start	-1.10084	-0.53595
Age at start (squared)	0.03443	0.01344

MNL for Group Membership

	"high-wage"	"childless careers"
Start in wage category 1	-0.08364	0.24558
Start in wage category 2	-0.20876	0.59434
Start in wage category 3	0.84706	1.31027
Start in wage category 4	1.80361	1.95150
Start in wage category 5	2.05103	2.58711

Long-Run Distribution



Posterior expectation of the **wage distribution** over the wage categories 0 to 5 after a period of t years in the various clusters.

Conclusions

- 3 groups of women: "high-wage mums", "low-wage-mums", "childless careers"

- variables maternity leave and number of children are very important for finding the 3 groups

- Markov chain model with logit extension allows inclusion of individual attributes

- MCMC samples from standard densities only (Gibbs)

Thank You for Your Attention!

Contact: regina.tuechler@wko.at
christoph.pamminger@jku.at*

*research supported by the Austrian Science Foundation (FWF) under the grant: S 10309-G14 (National Research Network "The Austrian Center for Labor Economics and the Analysis of the Welfare State", Subproject "Bayesian Econometrics").