Risk Patterns and Correlated Brain Activities

Alena Myšičková
Piotr Majer
Song Song
Peter N. C. Mohr
Wolfgang K. Härdle
Hauke R. Heekeren

C.A.S.E. Centre for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Freie Universität Berlin
Max Planck Institute for Molecular Genetics
http://lvb.wiwi.hu-berlin.de
http://www.languages-of-emotion.de
http://www.molgen.mpg.de

Risk Perception

- Which part is activated during *risk related decisions* ?
- Can statistical analysis help to detect this area?
- Response curve (to stimuli)? classify “risky people”?

Risk Perception

- Survey conducted by Max Planck Institute

- 22 young, native German, right-handed and healthy volunteers

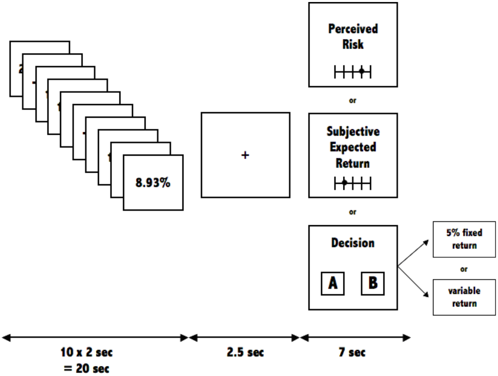
  3 subjects with extensive head movements ($> 5mm$)

  2 subjects with different stimulus frequency

  $n = 22 - (3 + 2) = 17$

- Experiment
  - Risk Perception and Investment Decision (RPID) task ($\times 81$)
  - fMRI images every 2.5 sec.
  - Analysis of the first part ($\times 45$)

Risk Perception



Returns Pause Decision

Risk Perception – Thermodynamics

Theoretical framework

- Risk-return model
  Mohr et al., 2010

Empirical evidence

- fMRI analysis

- Mechanical Equivalent of Heat
  1st law of thermodynamics
  Mayer, 1841

- Experiments "Joule apparatus"
  Joule, 1843

Risk Perception

- functional Magnetic Resonance Imaging

- Measuring Blood Oxygenation Level Dependent (BOLD) effect every 2-3 sec
  High-dimensional, high frequency & large data set

Risk Perception

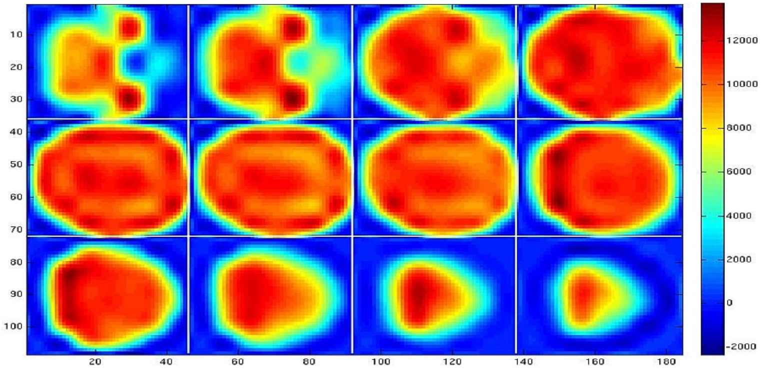
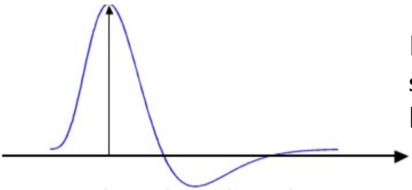


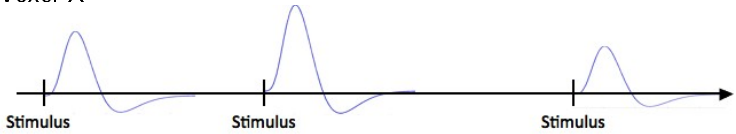
Figure 1: fMRI image observed every 2.5 sec, 12 horizontal slices of the brain's scan, $91 \times 92 \times 71 (x, y, z)$ data points of size 22 MB; scan resolution: $2 \times 2 \times 2mm^3$ fMRI

fMRI



Is there a significant reaction to specific stimuli in the hemodynamic response?

Voxel X



fMRI methods

- Voxel-wise GLM ▸ Voxel-wise GLM
  - linear model for each voxel separately
  - strong a priori hypothesis necessary
- Dynamic Semiparametric Factor Model (DSFM)
  - Use a "time & space" dynamic approach
  - Separate low dim time dynamics from space functions
  - Low dim time series exploratory analysis

Outline

1. Motivation ✓
2. DSFM
3. Results vs. Subject's Behaviour
4. Conclusion
5. Future Perspectives
6. References
7. Appendix

Notation

$$\underbrace{(X_{1,1}, Y_{1,1}), \ldots, (X_{J,1}, Y_{J,1})}_{t=1}, \ldots, \underbrace{(X_{1,T}, Y_{1,T}), \ldots, (X_{J,T}, Y_{J,T})}_{t=T},$$

$X_{j,t} \in \mathbb{R}^d, \; Y_{j,t} \in \mathbb{R}$
T - the number of observed time periods
J - the number of the observations in a period t
$\mathsf{E}(Y_t|X_t) = F_t(X_t)$

Quantify $F_t(X_t)$. How does it move?

Dynamic Semiparametric Factor Model

$$\mathrm{E}(Y_t|X_t) = \sum_{l=0}^{L} Z_{t,l} m_l(X_t) = Z_t^\top m(X_t) = Z_t^\top A^* \Psi$$

$$Z_t = (\mathbf{1}, Z_{t,1}, \ldots, Z_{t,L})^\top \text{low dim (stationary) time series}$$

$$m = (m_0, m_1, \ldots, m_L)^\top, \text{tuple of functions}$$

$$\Psi = \{\psi_1(X_t), \ldots, \psi_K(X_t)\}^\top, \psi_k(x) \text{ space basis}$$

$$A^* : (L+1) \times K \text{ coefficient matrix}$$

DSFM Estimation

$$Y_{t,j} = \sum_{l=0}^{L} Z_{t,l} m_l(X_{t,j}) + \varepsilon_{t,j} = Z_t^\top A^* \psi(X_{t,j}) + \varepsilon_{t,j}$$

- $\psi(x) = \{\psi_1(x), \ldots, \psi_K(x)\}^\top$ tensor B-spline basis

$$(\widehat{Z}_t, \widehat{A^*}) = \arg \min_{Z_t, A^*} \sum_{t=1}^{T} \sum_{j=1}^{J} \left\{ Y_{t,j} - Z_t^\top A^* \psi(X_{t,j}) \right\}^2 \qquad (1)$$

- Minimization by Newton-Raphson algorithm

B-Splines

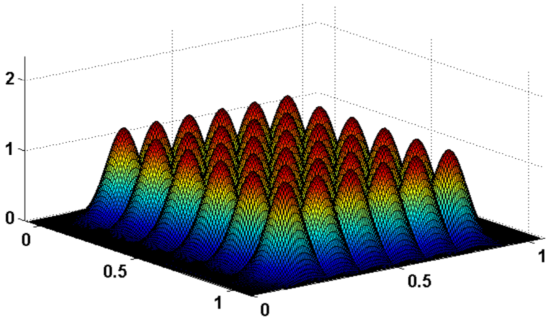


Figure 2: B-splines basis functions; order of B-splines: quadratic; number of knots: $6 \times 6 = 36$ ▸ B-Splines

DSFM Estimation

- Selection of L by explained variance

$$EV(L) = 1 - \frac{\sum_{t=1}^{T}\sum_{j=1}^{J}\left\{Y_{t,j} - \sum_{l=0}^{L} Z_{t,l} m_l(X_{t,j})\right\}^2}{\sum_{t=1}^{T}\sum_{j=1}^{J}\left\{Y_{t,j} - \bar{Y}\right\}^2}$$

number of B-splines (equally spaced) knots: $K = 12 \times 14 \times 14$

$L=2$	$L=4$	$L=5$	$L=10$	$L=20$
92.07	92.25	92.29	93.66	95.19

Table 1: EV in percent of the model with different numbers of factors L, averaged over all 17 analyzed subjects.

Panel DSFM

$$Y^i_{t,j} = \sum_{l=0}^{L} (Z^i_{t,l} + \alpha^i_{t,l}) m_l(X_{t,j}) + \varepsilon_{t,j},^{i}, \;\; 1 \le j \le J, \; 1 \le t \le T,$$

- $n = 17$ weakly/strongly risk-averse subjects
- $Y_{t,j}$ - BOLD signal; X_j voxel's index
  $\alpha^i_{t,l}$ - fixed individual effect; ▸ Residual Analysis
- Identification condition: $\mathrm{E}\left\{ \sum_{i=1}^{n} \sum_{l=0}^{L} \alpha^i_{t,l} m_l(X_{t,j}) | X_{t,j} \right\} = 0$

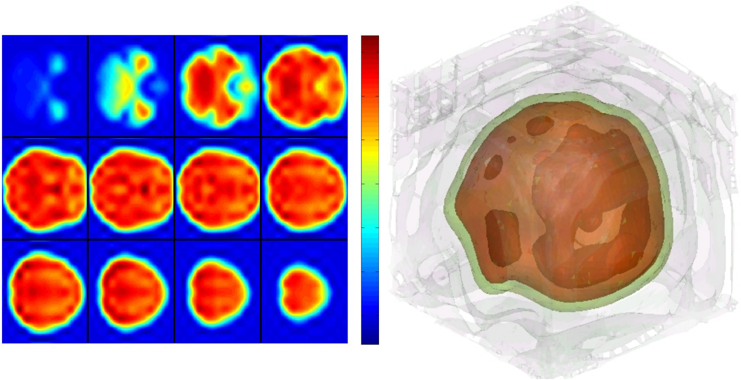
Panel DSFM Estimation

Feasible estimation algorithm:

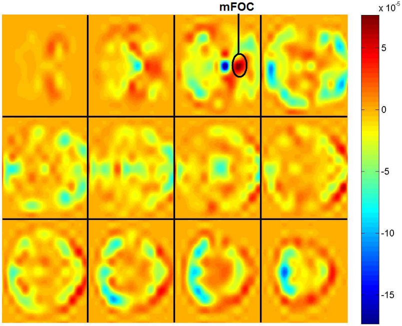
1. Average $Y^i_{t,j}$ over subjects i to obtain $\bar{Y}_{t,j}$
2. Estimate factors m_l for the "average brain" [via (1)]
3. Given $\widehat{m}_l$, for i, estimate $Z^i_{t,l}$

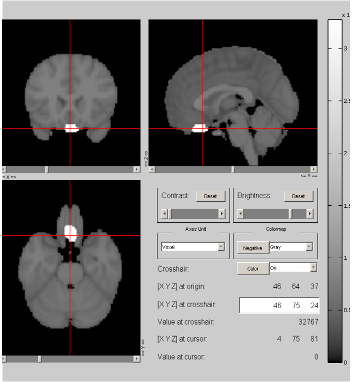
$$Y^i_{t,j} = \sum_{l=0}^{L} Z^i_{t,l} \widehat{m}_l(X_{t,j}) + \varepsilon^i_{t,j}$$

- 26h - computing time; CPU - 2×2.8GHz; data set of size 24.31 GB

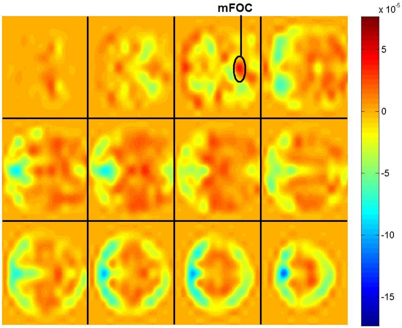


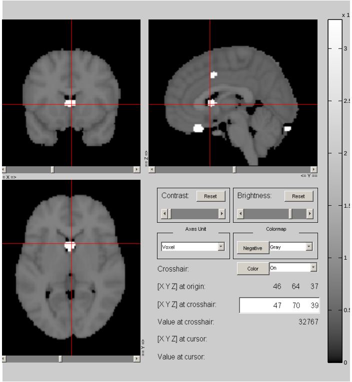
Estimated constant factor $\widehat{m}_0(X) = \sum_{k=1}^{K} \widehat{a}_{0,k} \psi_k(X)$ with $L = 20$



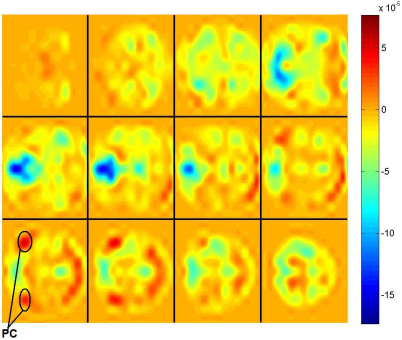


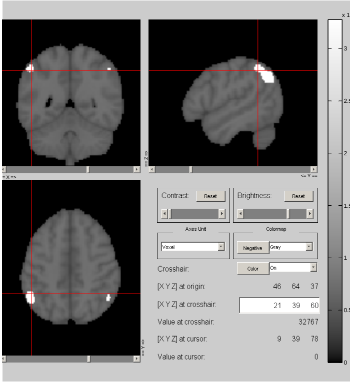
Estimated factor $\widehat{m}_5(X) = \sum_{k=1}^{K} \widehat{a}_{5,k}\psi_k(X)$ with $L = 20$
(MOFC = Medial orbitofrontal cortex)



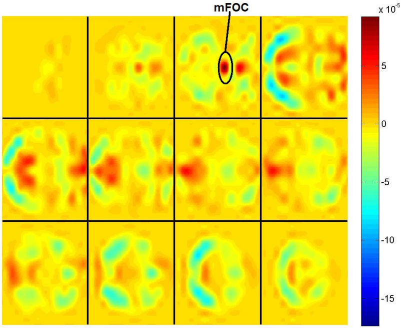


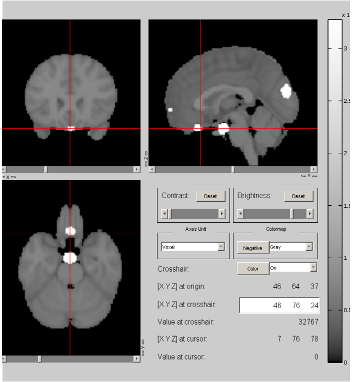
Estimated factor $\widehat{m}_9(X) = \sum_{k=1}^{K} \widehat{a}_{9,k}\psi_k(X)$ with $L = 20$



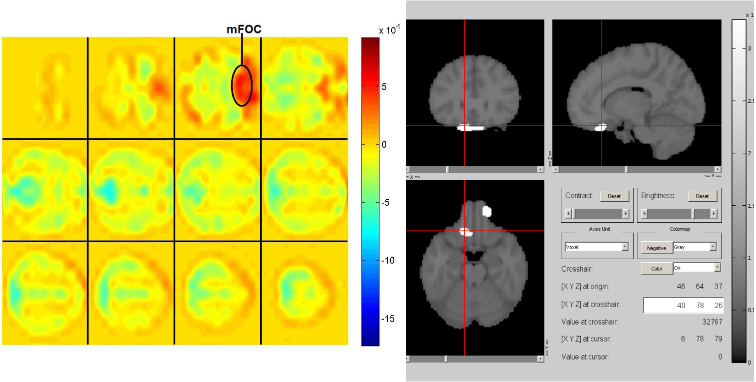


Estimated factor $\widehat{m}_{12}(X) = \sum_{k=1}^{K} \widehat{a}_{12,k}\psi_k(X)$ with $L = 20$
(PC = Paretial Cortex)

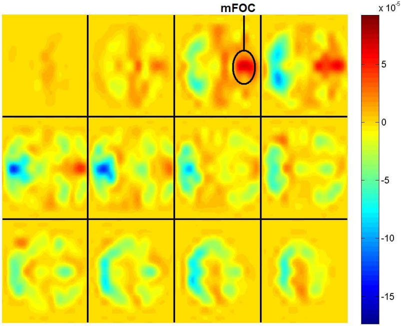


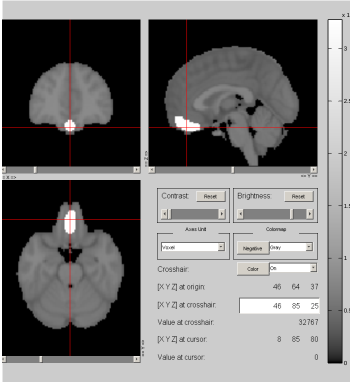


Estimated factor $\widehat{m}_{16}(X) = \sum_{k=1}^{K} \widehat{a}_{16,k}\psi_k(X)$ with $L = 20$



Estimated factor $\widehat{m}_{17}(X) = \sum_{k=1}^{K} \widehat{a}_{17,k}\psi_k(X)$ with $L = 20$





Estimated factor $\widehat{m}_{18}(X) = \sum_{k=1}^{K} \widehat{a}_{18,k}\psi_k(X)$ with $L = 20$

Estimated Factor Loading $\widehat{Z}_5$

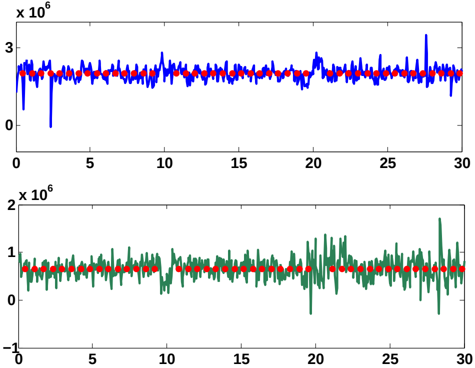


Figure 3: Estimated factor loading $\widehat{Z}_5$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Estimated Factor Loading $\widehat{Z}_9$

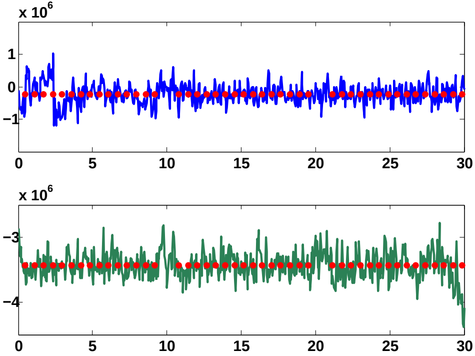


Figure 4: Estimated factor loading $\widehat{Z}_9$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Estimated Factor Loading $\widehat{Z}_{12}$

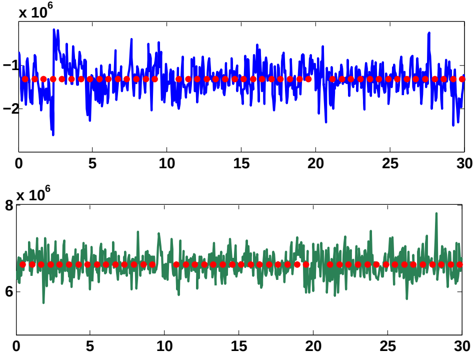


Figure 5: Estimated factor loading $\widehat{Z}_{12}$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Estimated Factor Loading $\widehat{Z}_{16}$

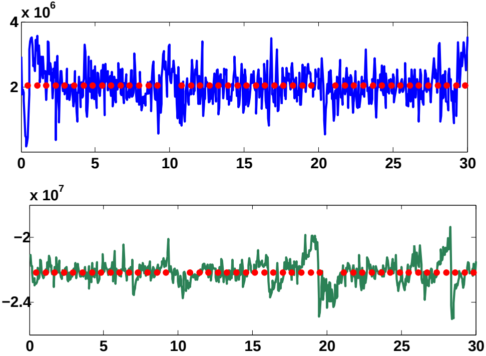


Figure 6: Estimated factor loading $\widehat{Z}_{16}$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Estimated Factor Loading $\widehat{Z}_{17}$

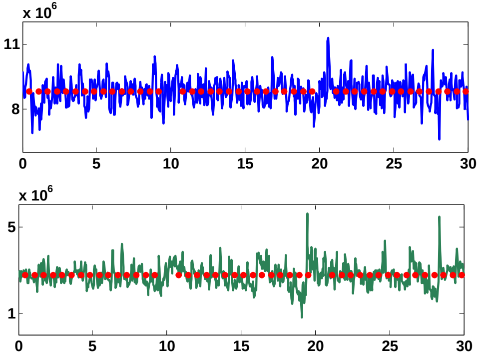


Figure 7: Estimated factor loading $\widehat{Z}_{17}$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Estimated Factor Loading $\widehat{Z}_{18}$

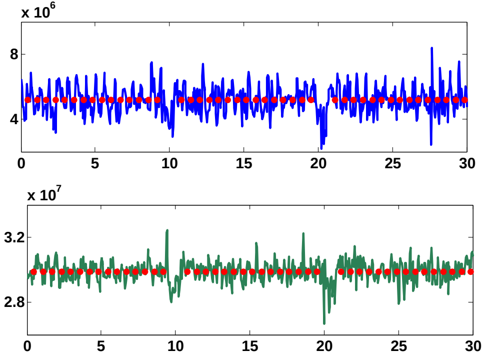


Figure 8: Estimated factor loading $\widehat{Z}_{18}$ for subjects within 30 minutes: 12 (upper panel) and 19 (lower panel) with $L = 20$; red dots denote stimulus

Reaction to the stimulus

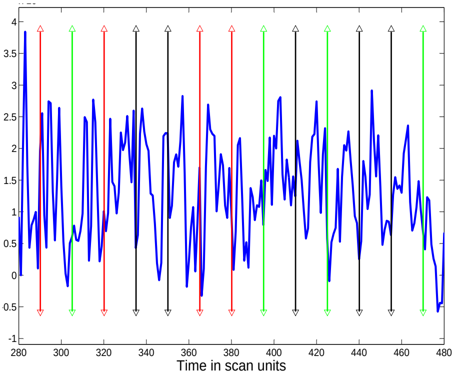


Figure 9: Detailed view of factor loading $\widehat{Z}_1$ for subject 12 with vertical lines in time points of stimuli of 3 different task: decision (red), subjective expected return (green) and perceived risk (black)

Reaction to the stimulus

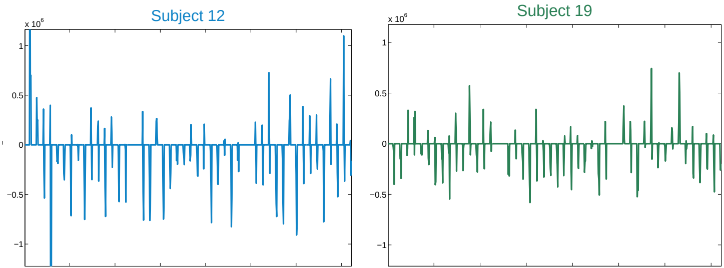


Figure 10: Reaction to stimulus $\overline{\Delta}\widehat{Z}^i_{s,l} = \frac{1}{3}\sum_{\tau=1}^{3}\Delta\widehat{Z}^i_{s+\tau,l}$, where $\Delta\widehat{Z}^i_{t,l} \stackrel{\text{def}}{=} \widehat{Z}^i_{s+t,l} - \widehat{Z}_{s,l}$, $t = 1, 2, 3$, s is the time of stimulus for factors loadings $\widehat{Z}^i_{t,12}$, for subjects 12 (left) and 19 (right) during the experiment (45 stimuli).

Risk attitude

- Subject's risk perception $\widetilde{R}_{i,s}$ - ▸ Risk Metrics
  - standard deviation
  - empirical frequency of loss (negative return)
  - difference between highest an lowest return (range)
  - coefficient of range (range/mean)
  - empirical frequency of ending below 5%
  - coefficient of variation (standard deviation/mean)

- Different subject - different risk perception

  fitted by correlation between risk metrics of return streams and $R_{i,j,s}$ - answers for "perceived risk" task $Q1$, $N = 27$

Risk attitude

- Subjective expected return $\widetilde{m}_{i,s}$ - ▸ Return Ratings
  - recency (higher weights on later returns)
  - primacy (higher weights on earlier returns)
  - below 0% (higher weights on returns below 0%)
  - below 5% (higher weights on returns below 5%)
  - mean

- Selecting return ratings for each subject individually

  best model selected by prediction power of one-leave-out cross validation procedure, $N = 27$

Risk attitude

- Each subject i has (R_i, m_i)
- Risk-return choice model

$$V_i(x_s) = m_i(x_s) - \beta_i R_i(x_s), \quad 1 \le i \le n, 1 \le s \le 27$$

x_s - return stream, m_i-subjective expected return, R_i - perceived risk , V_i - subjective value (unobserved), 5% - risk free return

- β Risk attitude parameter

Risk attitude

- Estimation of individual risk attitude by logistic regression

$$\mathrm{P}\left\{\text{risky choice}|(m,R)\right\} = \frac{1}{1+\exp(m-\beta R-5)}$$
$$\mathrm{P}\left\{\text{sure choice}|(m,R)\right\} = 1-\frac{1}{1+\exp(m-\beta R-5)}$$

  risky choice - unknown return, sure choice - fixed, 5% return

- $\widehat{\beta}$ derived by maximum likelihood method

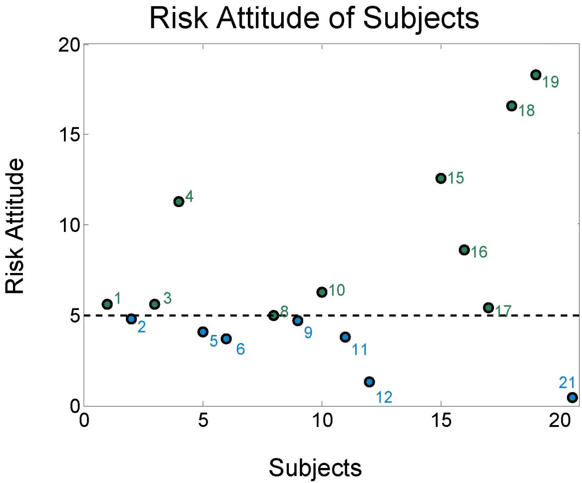


Figure 11: Risk attitude $\widehat{\beta}_i$ for 17 subjects; modeled by the softmax function from individuals' decisions, estimated by ML method ▸ Mohr et. al.

SVM Classification Analysis

- Support Vector Machines (SVM)
  17 subjects, 20 factor loading time series per subject

- Leave-one-out method to train and estimate classification rate
  SVM with Gaussian kernel; (R, C) chosen to maximize classification rate

- Weakly/strongly risk-averse subjects differ in reaction to stimulus $\Delta\widehat{Z}^{i}_{t,l}$ Reaction to Stimulus

SVM Classification Analysis

1. factors attributed to risk patterns: $l = 5, 9, 12, 16, 17, 18$
2. only "Decision under Risk" ($Q3$) stimulus
3. average reaction to s stimulus $\overline{\Delta} \widehat{Z}^{i}_{s,l} = \frac{1}{3} \sum_{\tau=1}^{3} \Delta \widehat{Z}^{i}_{s+\tau,l}$

SVM input data: volatility of $\overline{\Delta} \widehat{Z}^{i}_{s,l}$ over all $Q3$

Std		Estimated	
		Strongly	Weakly
Data	Strongly	1.00	0.00
	Weakly	0.14	0.86

Table 2: Classification rates of the SVM method, **without** knowing the subject's estimated risk attitude ▸ SVM Scores

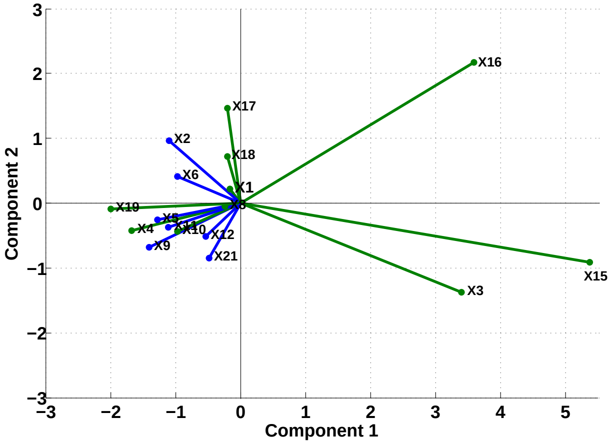
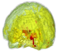


Figure 12: Normalized Principal Component Analysis on volatility of $\overline{\Delta \widehat{Z}}_{s,l}^{i}$ after stimulus for weakly/strongly risk-averse subjects; variance explained by the first and second components: 72%, 85%, respectively

Conclusion

- Factors $\widehat{m}$ identify activated areas, neurological reasonable
- Estimated factor loadings show differences for individuals with different risk attitudes (e.g. 12 vs. 19)
- SVM classification analysis of measurements in $\widehat{Z}_{t,l}$, $l = 5, 9, 12, 16, 17, 18$ after stimulus, can distinguish weakly/strongly risk-averse individuals with high classification rate, **without** knowing the subject's answers

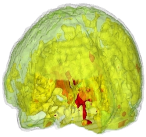


Future Perspectives

- Comparison with the PCA/ICA (PARAFAC) approach
- Analysis of the second part of the experiment (under assumption of independency) to "generate" larger number of subjects
- Improvement of the classification criterion
- Penalized DSFM with seasonal effects

Risk Patterns and Correlated Brain Activities

Alena Myšičková
Piotr Majer
Song Song
Peter N. C. Mohr
Wolfgang K. Härdle
Hauke R. Heekeren



C.A.S.E. Centre for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Freie Universität Berlin
Max Planck Institute for Molecular Genetics
http://lvb.wiwi.hu-berlin.de
http://www.languages-of-emotion.de
http://www.molgen.mpg.de

References

Joule, J. P.
On the Mechanical Equivalent of Heat
British Association Rep., trans. Chemical Section, 35, 1845.

Mayer, R.
Remarks on the Forces of Nature
The Benjamin/Cummings Publishing Company, London, 1841.

Mohr, P., Biele G., Krugel, L., Li S., Heekeren, H.
Neural foundations of risk-return trade-off in investment decisions
NeuroImage, 49: 2556-2563, 2010.

▸ Risk attitudes

References

Park, B., Mammen, E., Härdle, W. and Borak, S.
Time Series Modelling with Semiparametric Factor Dynamics
J. Amer. Stat. Assoc., 104(485): 284-298, 2009.

Ramsay, J. O. and Silverman, B. W.
Functional Data Analysis
New York: Springer, 1997.

Woolrich, M., Ripley, B., Brady, M., Smith, S.
Temporal Autocorrelation in Univariate Linear Modelling of FMRI Data
NeuroImage, 21: 2245-2278, 2010

Voxel-wise GLM

▸ fMRI methods

- FEAT - FMRI Expert Analysis Tool by Department of Clinical Neurology, University of Oxford

- GLM framework

$$Y = XB + \eta,$$

Y - single voxel BOLD time series, X - design matrix (regressors, i.e. visual, auditory)

- Significant, active areas (B) selected by *z-scores* $\equiv \frac{B_i - 0}{\sqrt{\mathrm{Var}(B_i)}}$ and grouping (20 neighbors) scheme

B-Splines

B-Splines

Univariate **B-spline** basis $\Psi = \{\psi_1(X), \ldots, \psi_K(X)\}^\top$ is a series of $\psi_k(X)$ functions defined by $x_0 \leq x_2 \leq \ldots \leq x_{K-1}$, K knots and order p, i.e. for $p = 2$ (quadratic)

$$\psi_j(x) = \begin{cases} \frac{1}{2}(x - x_j)^2 & \text{if } x_j \leq x < x_{j+1} \\ \frac{1}{2} - (x - x_{j+1})^2 + (x - x_{j+1}) & \text{if } x_{j+1} \leq x < x_{j+2} \\ \frac{1}{2}\left\{1 - (x - x_{j+2})^2\right\} & \text{if } x_j \leq x < x_{j+1} \\ x & \text{otherwise} \end{cases}$$

B-Splines

B-Splines

- Knots K and order p has to be specified in advance (EV criterion); K corresponds to bandwidth
- In higher dimensions, for $dim(X) = d > 1$

$$\Psi = \{\psi_1(X_1), \ldots, \psi_{K_1}(X_1)\} \times \ldots \times \{\psi_1(X_d), \ldots, \psi_{K_d}(X_d)\}$$

- Flexible and computationally efficient approach to capture various spatial structures

Residual Analysis

▸ PDSFM

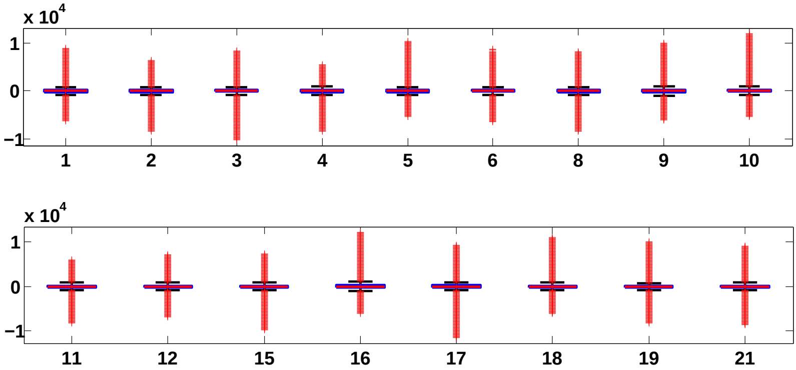


Figure 13: Boxplots of random subsets (size 3×10^7) from $\varepsilon^i_{t,j}$ (4.3×10^9 points) for all 17 analyzed subjects. Kurtosis exceeds 10

Residual Analysis

PDSFM

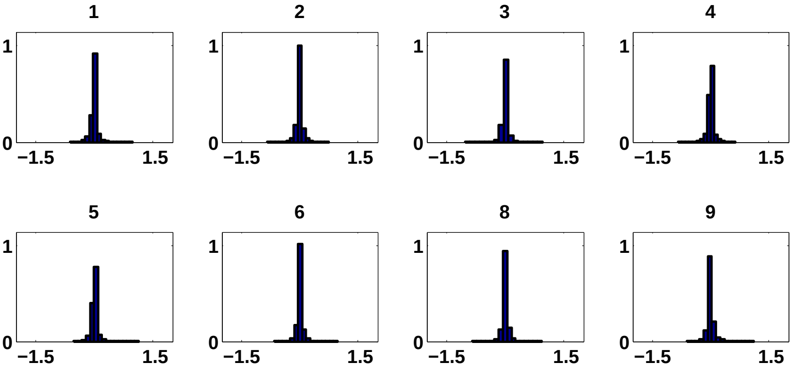


Figure 14: Histograms of random subsets (size 3×10^7) from $\varepsilon^i_{t,j}$ (4.3×10^9 points) for subjects $i = 1, 2, 3, 4, 5, 6, 8, 9$, respectively. Normality hypothesis (**KS test**) for standardized $\varepsilon^i_{t,j}$ rejected for all subjects, $\alpha = 5\%$

Residual Analysis

▸ PDSFM

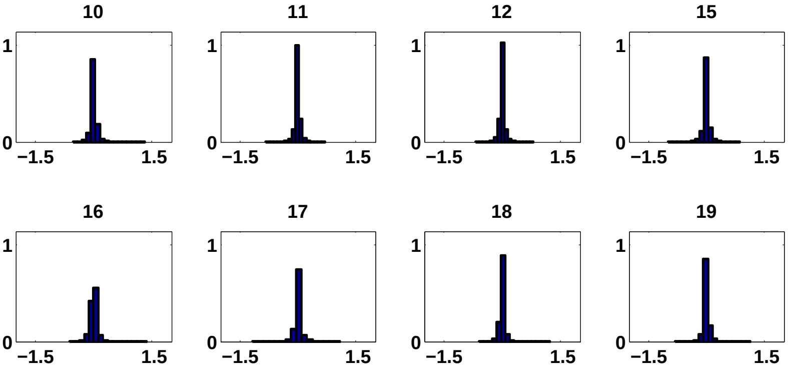


Figure 15: Histograms of random subsets (size 3×10^7) from $\varepsilon_{t,j}^i$ (4.3×10^9 points) for subjects $i = 10, 11, 12, 15, 16, 17, 18, 19$ respectively

Residual Analysis

PDSFM

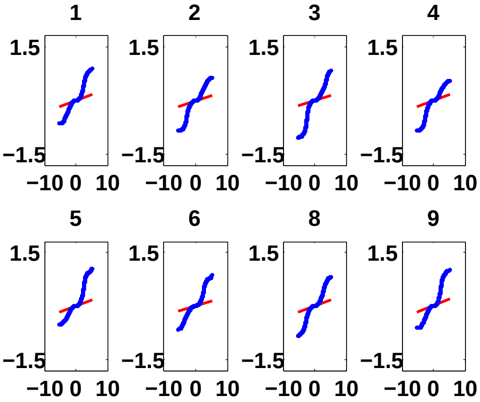


Figure 16: QQplots of random subsets (size 3×10^7) from $\varepsilon^i_{t,j}$ (4.3×10^9 points) for subjects $i = 1, 2, 3, 4, 5, 6, 8, 9$, respectively

Residual Analysis

PDSFM

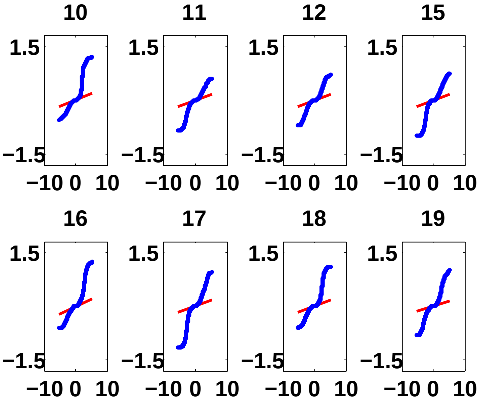


Figure 17: QQplots of random subsets (size 3×10^7) from $\varepsilon^i_{t,j}$ (4.3×10^9 points) for subjects $i = 10, 11, 12, 15, 16, 17, 18, 19$ respectively

Reaction to stimulus

SVM Analysis

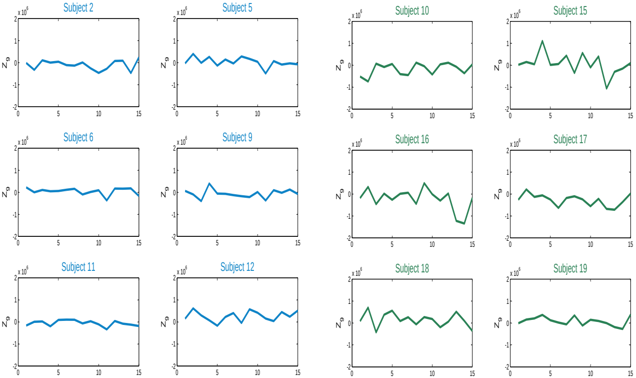


Figure 18: Averaged reaction $\overline{\Delta}\widehat{Z}^{i}_{s,9}$ to stimulus for all 15 $Q3$ questions for weakly/strongly risk-averse individuals

Reaction to stimulus

SVM Analysis

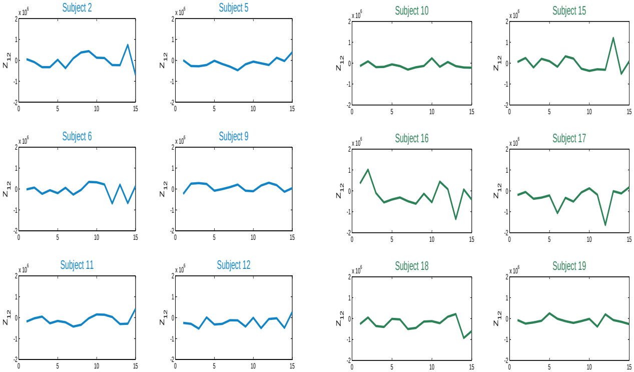


Figure 19: Averaged reaction $\overline{\Delta}\widehat{Z}^{i}_{s,12}$ to stimulus for all 15 $Q3$ questions for weakly/strongly risk-averse individuals

Return Ratings

Risk Attitude

r_i, $i = 1, \ldots, 10$ denotes sequence of random returns in each trial
Subjective Expected Return (**SER**) models:

- Mean

$$SER = \frac{\sum_{i=10-m}^{10} r_i}{m}$$

m-number of returns remembered, $2 \le m \le 10$

- Recency

$$SER = \frac{\sum_{i=10-m}^{10} r_i p}{\sum_{i=10-m}^{10} p}, \quad p = (i - 9 + m)^g$$

g - weighting parameter of returns, $0 < g < 1$

Return Ratings

▸ Risk Attitude

- Primacy

$$SER = \frac{\sum_{i=10-m}^{10} r_i p}{\sum_{i=10-m}^{10} p}, \quad p = (11-i)^g$$

m-number of returns remembered, $2 \leq m \leq 10$

g - weighting parameter of returns, $0 < g < 1$

- Overweight $< 0\%$

$$SER = \frac{\sum_{i=10-m}^{10} r_i p}{\sum_{i=10-m}^{10} p}, p = \begin{cases} 1, & \text{if } r_i \geq 0 \\ 1 + w, & \text{otherwise} \end{cases}$$

w - additional weight of returns , $0 < w < 1$; $1 \leq m \leq 9$

Return Ratings

▸ Risk Attitude

- Overweight $< 5\%$

$$SER = \frac{\sum_{i=10-m}^{10} r_i p}{\sum_{i=10-m}^{10} p}, p = \begin{cases} 1, & \text{if } r_i \geq 5 \\ 1 + w, & \text{otherwise} \end{cases}$$

w - additional weight of returns , $0 < w < 1$; $1 \leq m \leq 9$

- Parameters fitted by Cross Validation over all 27 trials

Return Ratings

▸ Risk Attitude

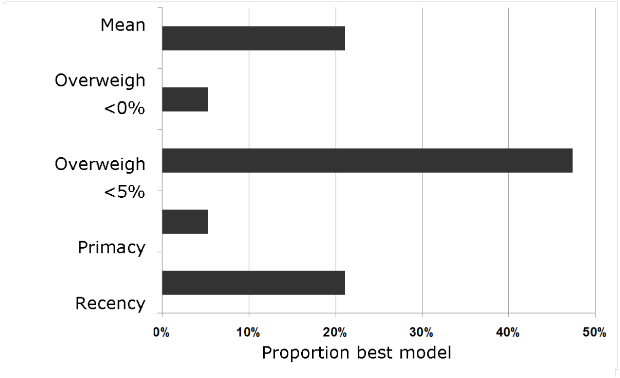


Figure 20: Distribution of return ratings over analyzed subjects

Risk Metrics

▸ Risk Attitude

Risk perception - risk metrics used by individuals

- Standard deviation of a return sequence
- Empirical frequency of loss (negative returns / all returns)
- Range - difference between highest an lowest return in a sequence
- Coefficient of range (range / mean)
- Empirical frequency of ending below 5% (returns $< 5\%$ / all returns)
- Coefficient of variation (standard deviation / mean)

Risk Metrics

Risk Attitude

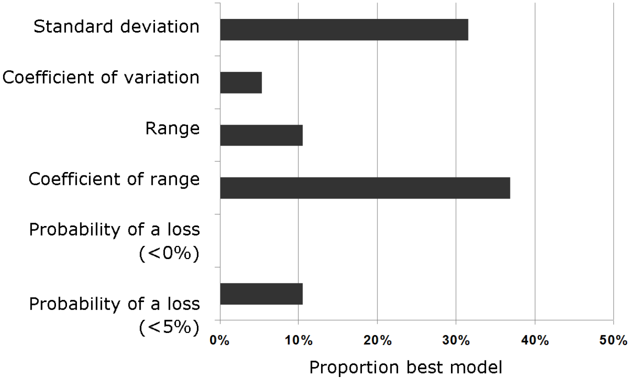


Figure 21: Distribution of risk metrices over analyzed subjects

SVM Scores

▸ SVM Classification

					Strongly					
i	1	3	4	8	10	15	16	17	18	19
β	5.6	5.6	11.3	5.0	6.3	12.6	8.6	5.4	16.6	18.3
Score	0.02	0.43	0.43	0.32	0.58	0.40	0.44	0.23	0.68	0.59
					Weakly					
i	2	5	6	9	11	12	21			
β	4.8	4.1	3.7	4.7	3.8	1.3	1.8			
Score	0.32	−1.03	−0.32	−0.44	−0.79	−0.04	−0.08			

Table 3: Estimated risk attitude and SVM scores (obtained **without** knowing the subject's answers)

SVM Scores

▸ SVM Classification

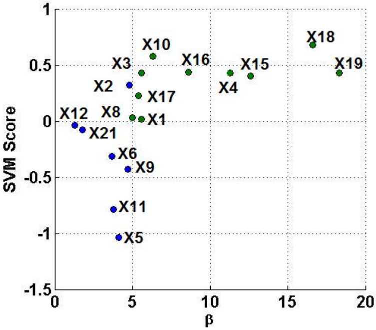


Figure 22: Scatter plot of $\widehat{\beta}_i$ vs SVM scores

Risk Metrics

▸ Risk Attitude

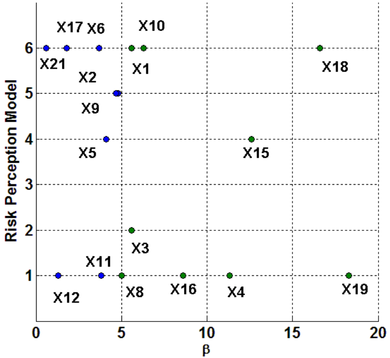


Figure 23: Scatter plot of $\widehat{\beta}_i$ vs risk perception models (vertical line). 1 - Standard deviation, 2 - Coefficient of variation, 3 - Empirical frequency of loss; 4 - Empirical frequency of ending below 5%, 5 - Coefficient of range, 6 - Coefficient of variation.