OFFICE OF FINANCIAL RESEARCH

A Simple Direct Estimate of Rule-of-Thumb Consumption Using the Method of Simulated Quantiles & Cross Validation

Nathan M. Palmer
The Office of Financial Research

George Mason University Department of Computational Social Science
November 3, 2017

Motivation: Want to Build Agent-Based Macro Models

- Assume you want to solve a "very complicated" dynamic stochastic general equilibrium problem
  - Eg. you think that the complex structure of mortgage or repo markets mattered for the crisis
- Problem: rational expectations solutions are intractable – what to do?

Motivation: Selecting Household Behavior

- Agent-based literature: "rules of thumb."
- Which rules to use?
-
    -

Motivation: Selecting Household Behavior

- Agent-based literature: "rules of thumb."
- Which rules to use?
- Growing literature: use optimization framework, approximate / learning solutions:
  - Howitt & Ozak (2014), Evans & McGough (2015), Lettau and Uhlig (1999), Allen and Carroll (2001) [and cottage literature], Arifovic (many), Gabaix QJE (2014), ...

Technical Difficulties in Selecting Household Behavior

- Note: this only side-steps the question.
- We'd still like to select between those. They are fairly different.
- Difficulties:
  - No closed form solution for behavior
  - Behavioral models are not nested
  - Likelihood surface: "very hard" to compute, if it exists (semiparametric estimation)
  - Easily available data on life cycle choices is not great

Technical Difficulties in Selecting Household Behavior

- Note: this only side-steps the question.
- We'd still like to select between those. They are fairly different.
- Difficulties:
  - No closed form solution for behavior
  - Behavioral models are not nested
  - Likelihood surface: "very hard" to compute, if it exists (semiparametric estimation)
  - Easily available data on life cycle choices is not great

Possible Solution: Simulation-Based Estimation, Selection

- Possible solution: semi-parametric, simulation-based estimation. Specifically:
  - "Method of Simulated Quantiles," Dominicy & Veredas (2013)
- ...coupled with k-fold cross-validation
- Highly general approach to both estimation and selection

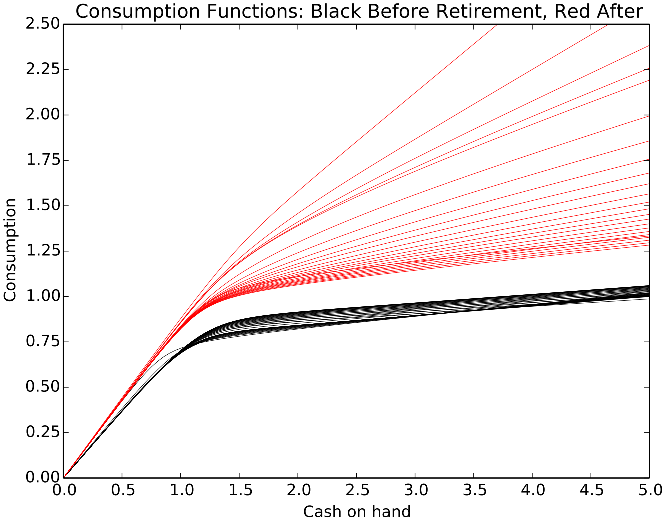
A Simple Implementation 0^{th}-order Example

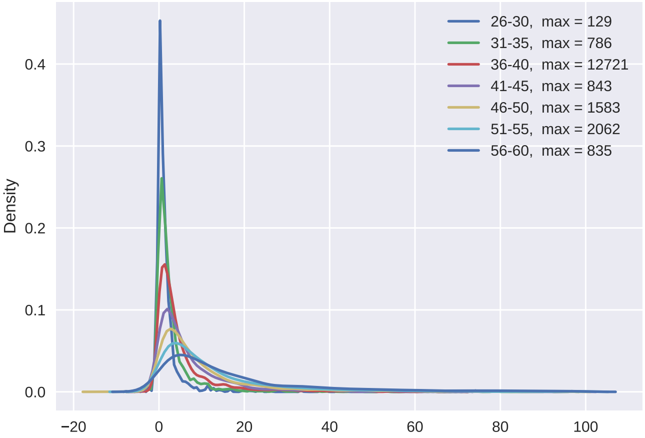
- This presentation: a specific, simple implementation
- Specifically:
  - Estimate a "textbook" structural, semi-parametric household life-cycle consumption-savings problem
  - Then **jointly** re-estimate the model with additional parameter(s): number of agents with different beta
  - Then use k-fold cross-validation to formally select between models
- Roughly a micro version of Campbell and Mankiw (1989, 1990)*
- Work is **preliminary** and very much **in progress**

Outline

Household Consumption Functions Example

An example solution:





Survey of Consumer Finance, Federal Reserve Board

"Textbook" Household Problem

A household solves the T -horizon problem described by sequence of Bellmans:

$$v_t^*(m_t) = \max_{c_t} u(c_t) + \beta \mathrm{B}_t \mathbb{E}_t \left[\Gamma_{t+1}^{1-\rho} \ v_{t+1}^*(m_{t+1}) \right]$$
$$s.t.$$
$$m_{t+1} = R_{t+1}(m_t - c_t) + \xi_{t+1}$$
$$m_0 \text{ given}$$

where

- β, ρ are discount factor and risk aversion
- m_t is total "cash on hand:" total assets + total income
- R_t is risk-free return on assets
- ξ_t are mean-1 temporary shocks to income
- Entire problem is normalized by permanent income process, not shown[2]

[2] See Carroll (2012a, b) for extensive discussion.

Outline

Estimation Method

Denote behavioral parameters:

$$\phi \equiv \{\beta, \rho\}$$

Denote structural parameters:

$$\begin{aligned} \rho &= \{\rho_t\}_{t=0}^{T} \\ &= \{\Gamma_t, \xi_t, R_t, \mathrm{B}_\mathrm{t}, \, etc\}_{t=0}^{T} \end{aligned}$$

Estimation Method: Simulation Step

- Given ρ, choose ϕ, solve problem for optimal policy functions $\{c_t^*\}_{t=1}^{T}$
- Simulate large panel of artificial wealth data under ϕ
- From SCF data, create weath distributions for 7 age groups:
  - 21-30, 31-35, 31-40, 41-45, 41-50, 51-55, 51-60
- Pool simulated data to match age groups of SCF data

Functions of Quantiles and Loss Function

- Construct functions of quantiles $\forall\ \tau$:

$$\text{Empirical:} \quad \hat{\varphi}_{\tau} = \begin{pmatrix} \hat{q}_{\tau,75} - \hat{q}_{\tau,25} \\ \hat{q}_{\tau,50} \end{pmatrix}$$

$$\text{Theoretical:} \quad \varphi^{N}_{\phi,\tau} = \begin{pmatrix} q^{\phi}_{\tau,75} - q^{\phi}_{\tau,25} \\ q^{\phi}_{\tau,50} \end{pmatrix}$$

Functions of Quantiles

- Define vectors of functions of quantiles:

$$\hat{\varphi} = \left(\hat{\varphi}_1^\top, \hat{\varphi}_2^\top, ..., \hat{\varphi}_7^\top\right)^\top$$

$$\varphi_\phi^{\mathbf{N}} = \left(\varphi_{\phi,1}^\top, \varphi_{\phi,2}^\top, ..., \varphi_{\phi,7}^\top\right)^\top$$

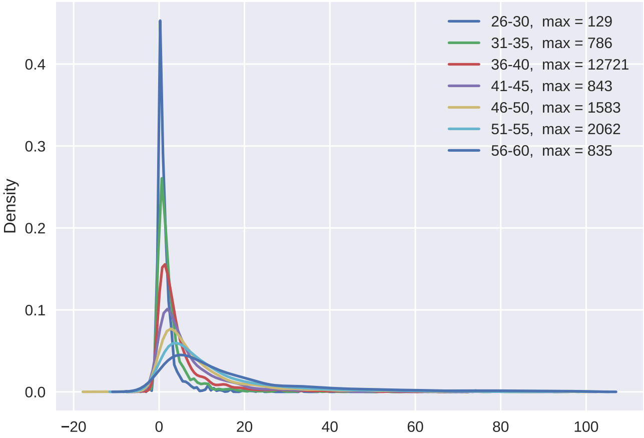
Loss Function

- Define the loss function:

$$\varpi_\rho(\phi) = \left(\hat{\varphi} - \varphi_\phi^N\right) \mathbf{W} \left(\hat{\varphi} - \varphi_\phi^N\right)$$

  where **W** is a positive definite matrix of weights
- Minimize the loss function, obtain ϕ
  - Bootstrap for variance

Motivation: Why Consider N Types?

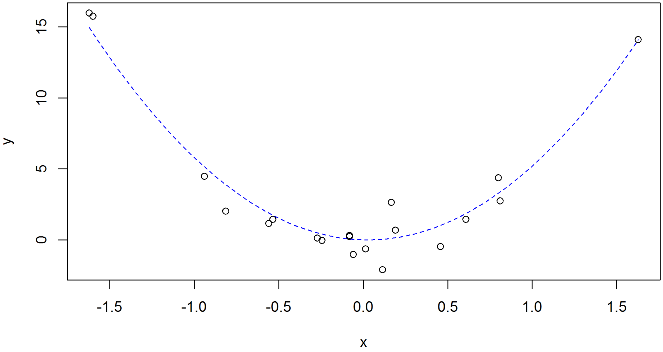


Survey of Consumer Finance, Federal Reserve Board

Outline

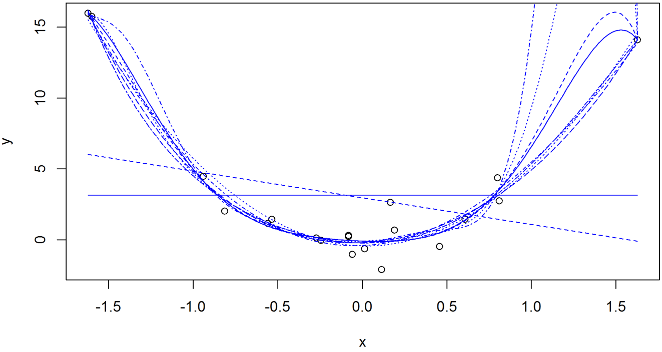
Quick Illustration

Consider the following artificial data:[3]

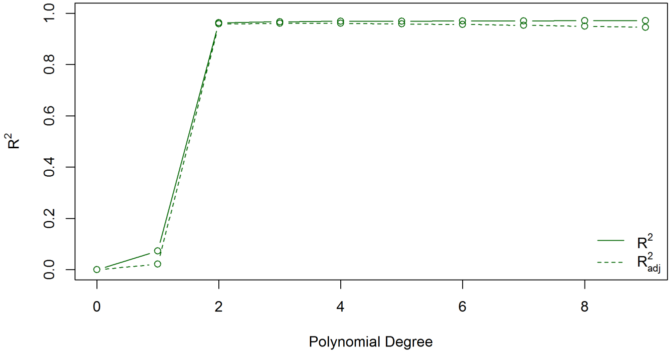


[3] Discussion from Cosma Shalizi's "Advanced Data Analysis from an Elementary Point of View"

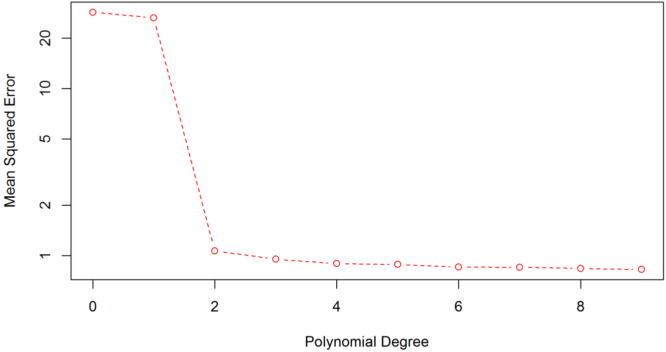
Polynomial Overfitting



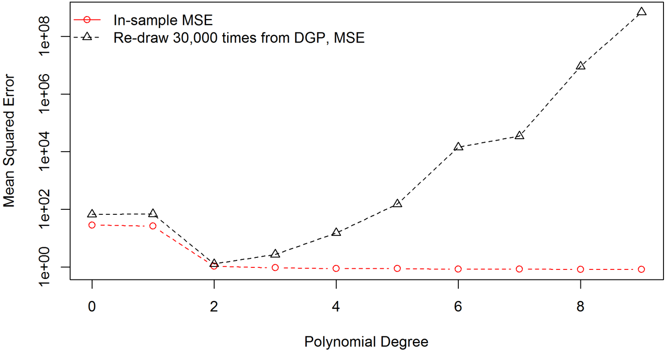
R-squared Looks Great



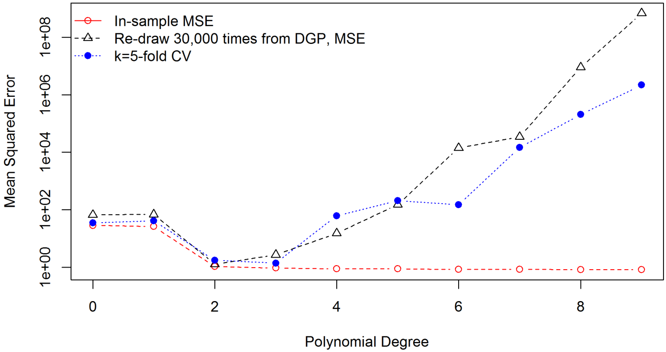
Loss Function (SSE) Looks Great



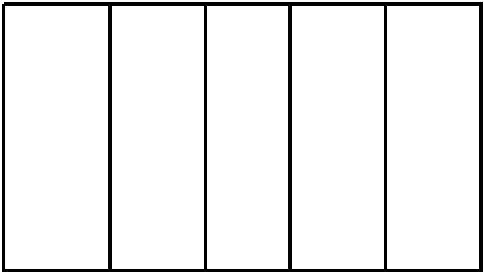
However, Very Poor Fit



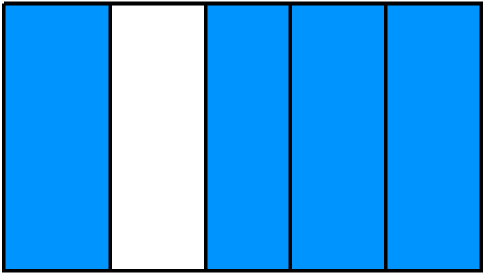
However, Very Poor Fit



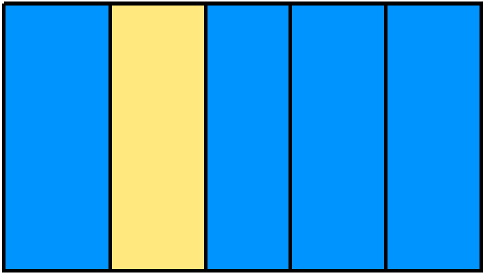
Data Selection, $K = 5$



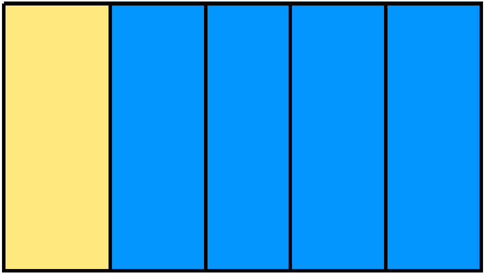
Data Selection, $K = 5$



Data Selection, $K = 5$



Data Selection, $K = 5$



Outline

Very Preliminary Results

1. Original, median-only[4], $N_{type} = 1$: $\beta = 1.007, \quad \rho = 4.4$
2. Original, median+IQR, $N_{type} = 1$: $\beta = 1.01, \quad \rho = 1.6$
3. With $N_{type} = 2 - 6$:
   - $\beta_{lo} \approx 0.25 - .45, \ \beta_{hi} \approx 1.04; \quad \rho \approx 4.7$
   - ...with "low" fraction ≈ 0.46
4. Cross-validation: evidence for selecting $N \geq 2$ types

[4] Due to: p_{death}, β_{τ}.

Very Preliminary Results

1. Original, median-only[4], $N_{type} = 1$: $\beta = 1.007, \quad \rho = 4.4$
2. Original, median+IQR, $N_{type} = 1$: $\beta = 1.01, \quad \rho = 1.6$
3. With $N_{type} = 2 - 6$:
   - $\beta_{lo} \approx 0.25 - .45, \ \beta_{hi} \approx 1.04; \quad \rho \approx 4.7$
   - ...with "low" fraction ≈ 0.46
4. Cross-validation: evidence for selecting $N \geq 2$ types

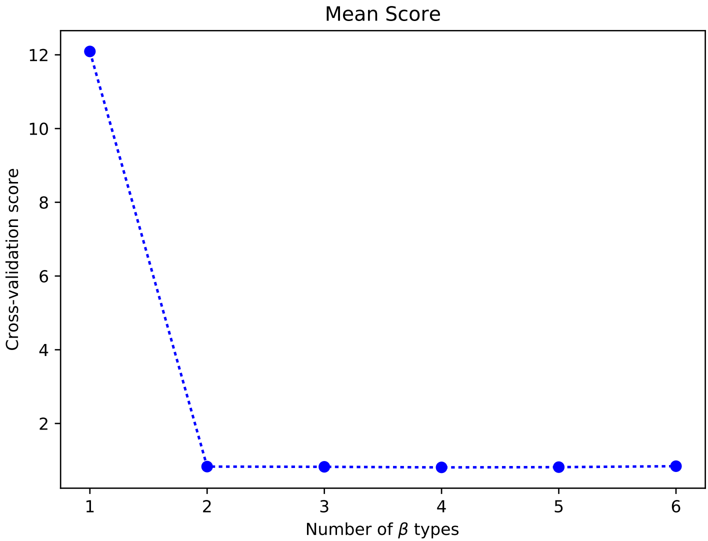
[4] Due to: p_{death}, β_{τ}.

Very Preliminary Results

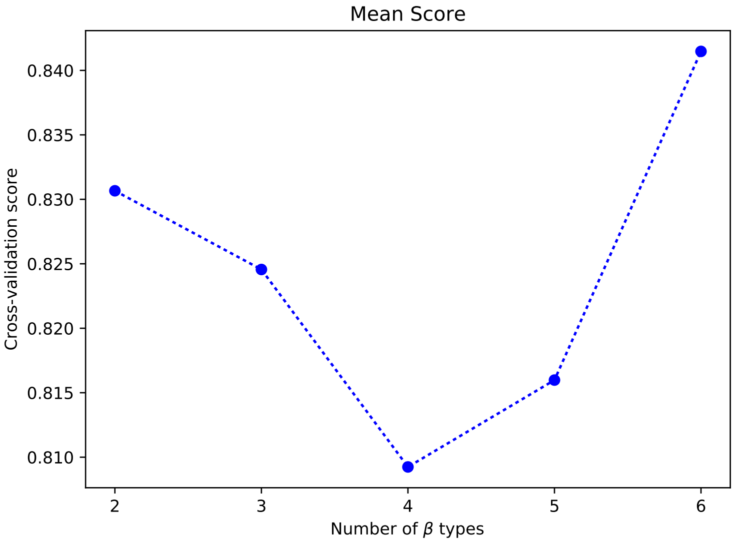
1. Original, median-only[4], $N_{type} = 1$: $\beta = 1.007, \quad \rho = 4.4$
2. Original, median+IQR, $N_{type} = 1$: $\beta = 1.01, \quad \rho = 1.6$
3. With $N_{type} = 2 - 6$:
    - $\beta_{lo} \approx 0.25 - .45, \; \beta_{hi} \approx 1.04; \quad \rho \approx 4.7$
    - ...with "low" fraction ≈ 0.46
4. Cross-validation: evidence for selecting $N \geq 2$ types

[4]Due to: p_{death}, β_{τ}.

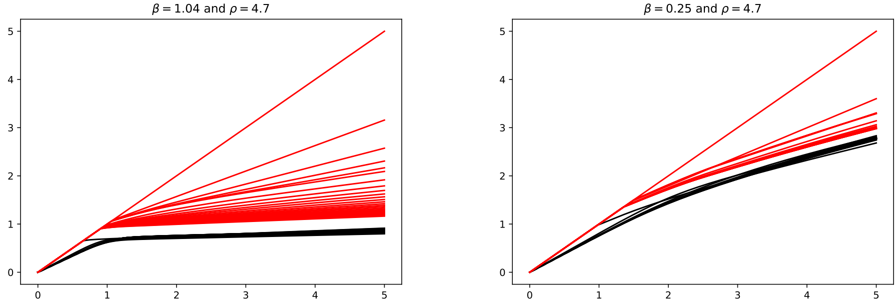
K-Folds CV on N Types



“Zoom In:” K-Folds CV: 2-6 Types



N=2 Consumption Functions



Ugly Tables: Full Estimation Results, $N \in (1, 2, 3, 4)$

N_β	1	2	3	4
ρ	1.65	4.65	4.94	4.24
β	1.01	0.25, 1.04	0.29, 0.99, 1.05	0.01, 0.41, 0.81, 1.04
$frac$	n.a.	0.46, 0.54	0.26, 0.35, 0.38	0.17, 0.19, 0.09, 0.54
β_{lo}	1.01	0.25	0.69	0.34
$frac_{lo}$	n.a.	0.46	0.62	0.46

Ugly Tables: Full Estimation Results, $N \in (5, 6)$

N_β	5	6
ρ	4.70	4.74
β	0.00, 0.01, 0.52, 0.79, 1.04	0.11, 0.16, 0.24, 0.28, 0.45, 1.04
$frac$	0.09, 0.07, 0.15, 0.17, 0.53	0.03, 0.17, 0.09, 0.08, 0.1, 0.54
β_{lo}	0.44	0.25
$frac_{lo}$	0.47	0.46

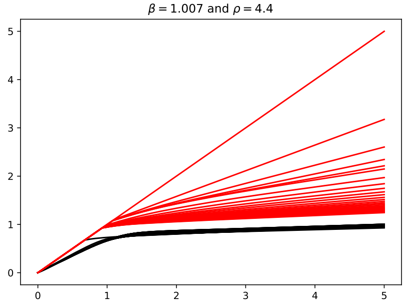
Estimation Results, $N \in (1, 2, 3, 4, 5, 6)$

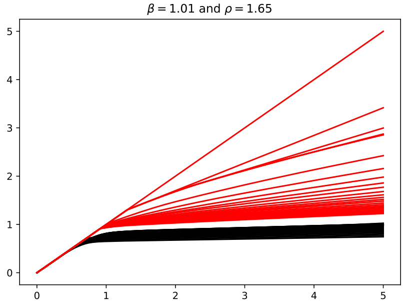
N_β	2	3	4	5	6
ρ	4.65	4.94	4.24	4.70	4.74
$\beta_{\{lo,hi\}}$	0.25, 1.04	0.69, 1.05	0.34, 1.04	0.44, 1.04	0.25, 1.04
$frac_{\{lo,hi\}}$	0.46, 0.54	0.62, 0.38	0.46, 0.54	0.47, 0.53	0.46, 0.54

Summary, Next Steps

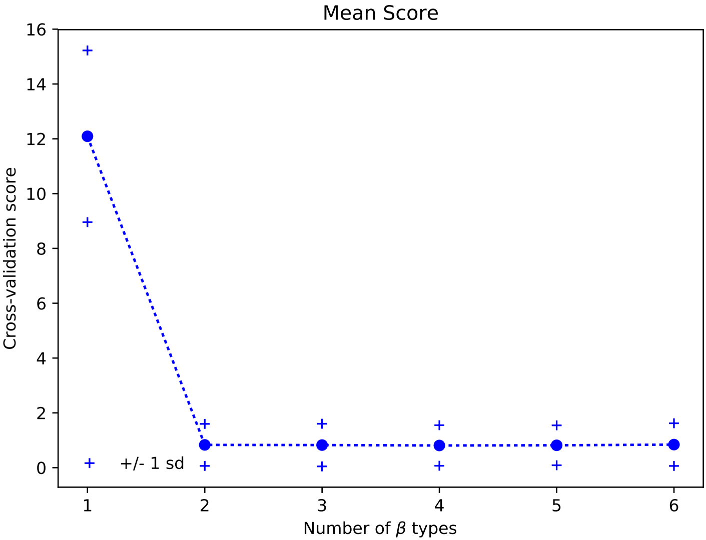
- Alternative models of consumption-savings behavior have no closed form solution, extremely hard to calculate likelihood surface, and are not nested. None-the-less we would like to select between possible candidates.
- This project jointly estimates a basic structural life cycle consumption-savings problem multiple types and selects between number of types via k-fold CV.
  - Fraction of "low beta" consumers is estimated at ˜0.46
- Next steps: many
  - Data update
  - Robustness checks
  - Selection with simple learning

Appendix: "Regular" Consumption Functions

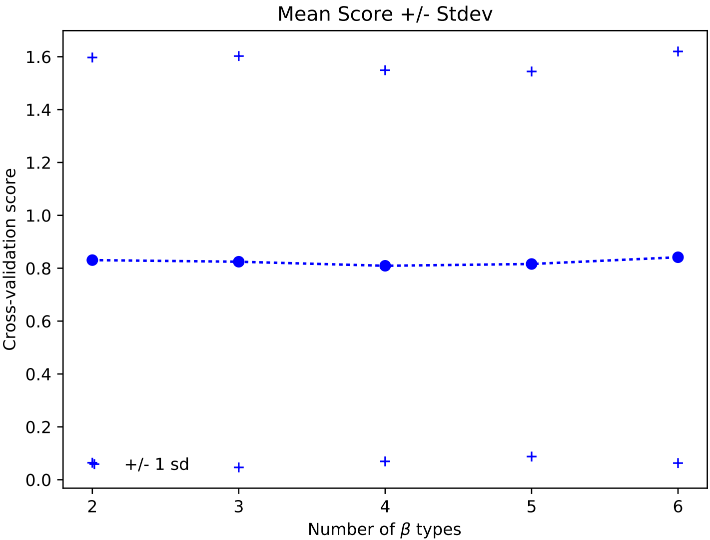




Appendix: K-Folds CV on N Types



Appendix: K-Folds CV on N Types



Appendix: K-Folds CV on N Types