The Rasch Model and its Potential for Empirical Economics Research

Dr. Carolin Strobl

Institut für Statistik, Ludwig-Maximilians-Universität München

Empirical Economics and Econometrics Research Seminar 2011

References
The Rasch Model and its extensions from Item Response Theory (IRT)

Objective

Model specification

Parameter estimation

Model diagnostics

Extended models

Example: Consumer survey

References
Georg Rasch was a Danish mathematician who eventually became professor of statistics at the economics department of the University of Copenhagen, but: “It would be wrong to say that Rasch’s professorship was a indisputable success. [...] Rasch developed the course in statistics. This change was very welcome to a segment of students and scientists, namely the sociologists. But a larger segment of people, namely the economists, found that the course in statistics had become next to useless.” (http://www.rasch.org)
You are in good company...

citations of Rasch’s book from 1960 (reprinted 1980)

(ISI web of knowledge)
You are in good company...

(http://www.sueddeutsche.de)
Reconsidering the problem of data equivalence in international marketing research

Contrasting approaches based on CFA and the Rasch model for measurement

Thomas Salzberger
Vienna University of Economics and Business Administration,
Vienna, Austria, and
Rudolf R. Sinkovics
Manchester Business School, UK

Abstract
Purpose – The paper investigates the suitability of the Rasch model for establishing data equivalence. The results based on a real data set are contrasted with findings from standard procedures based on CFA methods.
METHODOLOGICAL OVERVIEW OF RASCH MODEL AND APPLICATION IN CUSTOMER SATISFACTION SURVEY DATA

FRANCESC A DE BATTIST I  GIOVANNA NICOLINI  SILVIA SALINI

Working Paper n. 2008-04
FEBBRAIO 2008
THE RASCH MODEL

AND ITS USE IN MARKETING

SEMINAR - QUT, MAY 2009

Prof Geoff Soutar
Head of Discipline
Winthrop Professor Marketing
University of Western Australia
Objective

measurement of a latent trait

- intelligence
- math skills
- ...
- attitude
- consumer satisfaction
Data

report for each subject and each item

- was the item answered correctly?
- did the subject agree to the item?
### Data

<table>
<thead>
<tr>
<th>subject</th>
<th>item</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Model specification
whether a subject can answer an item correctly depends on both

- the ability of the subject $\theta_i$ and
- the difficulty of the item $\beta_j$

$$P(u_{ij} = 1 | \theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}}$$
Item-characteristic-curves (ICCs)

\[ P(u_{ij} = 1|\theta_i, \beta_j) \]: probability that subject \( i \) "beats" item \( j \)

for \( \theta_i = \beta_j \): \( P(u_{ij} = 1|\theta_i, \beta_j) = 0.5 \)
Item-characteristic-curves

\[ P(u_{ij} = 1 | \theta_i, \beta_j) \]

items: ← easy hard →
subjects: ← stupid smart →

**The Rasch Model**
Carolin Strobl

**Objective**

**Model specification**
ICCs
Specific objectivity
Local stochastic independence
Sufficient statistics
Derivation of the model

**Parameter estimation**
ML estimation
Joint ML
Conditional ML
Marginal ML
Information of an item

**Model diagnostics**
Graphical test
LR test
Wald tests

**Extended models**
Birnbaum models
Models for ordinal data
The Rasch model as a GLMM

**Example: Consumer survey**

**References**
Specific objectivity

the ordering of the subjects does not depend on which item is used for the comparison

(but the discriminatory power is higher in the center)
Specific objectivity

\[ \frac{P_{ax}}{1 - P_{ax}} \cdot \frac{P_{bx}}{1 - P_{bx}} = \frac{P_{ay}}{1 - P_{ay}} \cdot \frac{P_{by}}{1 - P_{by}} \]

The Rasch Model
Carolin Strobl

Objective
Model specification
ICCs
Specific objectivity
Local stochastic independence
Sufficient statistics
Derivation of the model
Parameter estimation
ML estimation
Joint ML
Conditional ML
Marginal ML
Information of an item
Model diagnostics
Graphical test
LR test
Wald tests
Extended models
Birnbaum models
Models for ordinal data
The Rasch model as a GLMM
Example: Consumer survey
References
Local stochastic independence

for a given ability (i.e. for one subject or several subjects with the same ability) the probability of answering one item does not depend on answering another item and vice versa
Local stochastic independence

for a given ability (i.e. for one subject or several subjects with the same ability) the probability of answering one item does not depend on answering another item and vice versa

⇒ allows us to compute joint probability as product of individual probabilities
Local stochastic independence
could be violated if, e.g.,

- solving one item is crucial for solving another one
- subjects copy each others solutions

or

- the latent trait is not unidimensional
  (subsets of items measure different latent traits
  $\Rightarrow$ scores correlated)
**Sufficient statistics**

row and column sums are sufficient statistics for person and item parameters

<table>
<thead>
<tr>
<th>subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>$r_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

$s_j$ | 1 | 3 | 2 | 3 | 1 | 2 |

sufficient statistics contain all information on parameters
### Sufficient statistics

Row and column sums are sufficient statistics for person and item parameters:

<table>
<thead>
<tr>
<th>subject</th>
<th>item</th>
<th>$r_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 1 0 1 0 1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0 1 1 0 1 1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0 1 1 1 0 0</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
</tr>
</tbody>
</table>

$s_j$ | 1 3 2 3 1 2

Sufficient statistics contain all information on parameters ⇒ allows us to condition on row sums in ML estimation.
models equivalent to the Rasch model can be derived from

▶ continuous and strictly monotone ICCs

and

▶ local stochastic independence and
▶ sufficient statistics

or

▶ specific objectivity

see Fischer and Molenaar (1995, ch. 2) for a summary
models equivalent to the Rasch model can be derived

\[ P(u_{ij} = 1|\theta_i, \beta_j) = \frac{e^{a(\theta_i - \beta_j)}+b}{1 + e^{a(\theta_i - \beta_j)}+b} \]

and have the properties of interval scales of measurement with the same unit \( a \) for person and item parameters

the common form

\[ P(u_{ij} = 1|\theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} \]

with \( a = 1 \) would have the properties of a difference scale, but \( a = 1 \) is not testable (Fischer and Molenaar, 1995, ch. 2)
Parameter estimation
two kinds of parameters: person and item parameters
Joint ML estimation

maximize joint Likelihood $L_u(\theta, \beta)$ w.r.t. $\theta$ and $\beta$ simultaneously
Joint ML estimation

maximize joint Likelihood \( L_u(\theta, \beta) \) w.r.t. \( \theta \) and \( \beta \) simultaneously

- problem: not consistent!
  (\# parameters increases with sample size)
Sufficient statistics

<table>
<thead>
<tr>
<th>subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>ri</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>sj</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

\[
P(u_i | \theta_i, \beta) = P(u_i | r_i, \theta_i, \beta) \cdot P(r_i | \theta_i, \beta)
\]

\[
P(u_i | r_i, \beta) \cdot P(r_i | \theta_i, \beta)
\]
Conditional ML estimation

- step 1: estimate item parameters $\hat{\beta}$ conditional on sufficient statistics $r$ for subject parameters $\theta$

$$L_u(r, \theta, \beta)^{\text{suff. stat.}} = L_u(r, \beta)$$

maximize w.r.t. $\beta$
Conditional ML estimation

- step 1: estimate item parameters $\hat{\beta}$ conditional on sufficient statistics $r$ for subject parameters $\theta$
  
  \[ L_u(r, \theta, \beta)^{\text{suff. stat.}} = L_u(r, \beta) \]

  maximize w.r.t. $\beta$

- step 2: estimate subject parameters $\hat{\theta}$ with estimates for item parameters $\hat{\beta}$ plugged in
  
  \[ L_u(\theta, \hat{\beta}) \]

  maximize w.r.t. $\theta$
Conditional ML estimation

▶ step 1: estimate item parameters $\hat{\beta}$ conditional on sufficient statistics $r$ for subject parameters $\theta$

$$L_u(r, \theta, \beta)^{\text{suff. stat.}} = L_u(r, \beta)$$

maximize w.r.t. $\beta$

▶ step 2: estimate subject parameters $\hat{\theta}$ with estimates for item parameters $\hat{\beta}$ plugged in

$$L_u(\theta, \hat{\beta})$$

maximize w.r.t. $\theta$

▶ problem: uncertainty from estimating $\hat{\beta}$ usually not accounted for (Tsutakawa and Johnson, 1990)
Procedure

The Rasch Model

Carolin Strobl

Objective

Model specification

ICCs
Specific objectivity
Local stochastic independence
Sufficient statistics
Derivation of the model

Parameter estimation

ML estimation
Joint ML
Conditional ML
Marginal ML
Information of an item

Model diagnostics

Graphical test
LR test
Wald tests

Extended models

Birnbaum models
Models for ordinal data
The Rasch model as a GLMM

Example: Consumer survey

References
Procedure

- estimate a few tens of item parameters from a large person sample = test calibration
Procedure

- estimate a few tens of item parameters from a large person sample = test calibration
- estimate the person parameter of one subject from a few tens of items
Procedure

- estimate a few tens of item parameters from a large person sample = test calibration
- estimate the person parameter of one subject from a few tens of items

gives consistent estimates
Marginal ML estimation

different approach to “get rid of” the subject parameters $\theta$ for estimating the item parameters $\hat{\beta}$:

- step 1: assume a distribution (usually the normal) $F(\theta)$
Marginal ML estimation

different approach to “get rid of” the subject parameters $\theta$ for estimating the item parameters $\hat{\beta}$:

- step 1: assume a distribution (usually the normal) $F(\theta)$
- step 2: integrate the $\theta$ out

$$L_u(\beta) = \int_{\Theta} L_u(\theta, \beta) \, \partial F(\theta)$$

maximize w.r.t. $\beta$ + constraints for identifiability
Marginal ML estimation

different approach to “get rid of” the subject parameters $\theta$ for estimating the item parameters $\beta$:

- step 1: assume a distribution (usually the normal) $F(\theta)$
- step 2: integrate the $\theta$ out

$$L_u(\beta) = \int_\Theta L_u(\theta, \beta) \, dF(\theta)$$

maximize w.r.t. $\beta +$ constraints for identifiability

- step 3: estimate subject parameters $\hat{\theta}$ with estimates for item parameters $\hat{\beta}$ plugged in

$$L_u(\theta, \hat{\beta})$$

maximize w.r.t. $\theta$
Marginal ML estimation

different approach to “get rid of” the subject parameters $\theta$ for estimating the item parameters $\hat{\beta}$:

- step 1: assume a distribution (usually the normal) $F(\theta)$
- step 2: integrate the $\theta$ out

$$L_u(\beta) = \int_\Theta L_u(\theta, \beta) \, \partial F(\theta)$$

maximize w.r.t. $\beta$ + constraints for identifiability

- step 3: estimate subject parameters $\hat{\theta}$ with estimates for item parameters $\hat{\beta}$ plugged in

$$L_u(\theta, \hat{\beta})$$

maximize w.r.t. $\theta$

- problems
  - uncertainty from estimating $\hat{\beta}$ not accounted for
  - distribution assumption may be wrong
Alternative estimation approaches

- based on Bayesian MCMC: assume marginal distribution + prior on every parameter

  ... (Fischer and Molenaar, 1995, ch. 3)
Restrictions

\[ P(u_{ij} = 1|\theta_i, \beta_j) = \frac{e^{a(\theta_i - \beta_j) + b}}{1 + e^{a(\theta_i - \beta_j) + b}} \]

for a unique solution

- fix \( a = 1 \) and
- fix \( b \) by means of \( \sum_j \beta_j = 0 \) or set one \( \beta_j = 0 \)

+ for conditional ML: zero and perfect scores must be excluded from the data matrix
in the Rasch model the information (discriminatory power) of an item $j$ is its gradient

$$\mathcal{I}_j(\theta_i) = \frac{\partial}{\partial \theta_i} P(u_{ij} = 1|\theta_i, \beta_j)$$
Information of an item

for all items the information adds up

\[ \mathcal{I}(\theta_i) = \sum_j \mathcal{I}_j(\theta_i) \]

the information is the inverse of the variance, so that the confidence interval for the ML estimator \( \hat{\theta}_i \) is

\[
\left[ \hat{\theta}_i \pm z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{\hat{\mathcal{I}}(\theta_i)}} \right]
\]
Model diagnostics
Graphical model test
idea: item parameter estimates should not depend on the person-sample

- split person sample, e.g., at the median of the raw scores $r_i$
- plot $\hat{\beta}_{\text{group 1}}$ against $\hat{\beta}_{\text{group 2}}$
  $\Rightarrow$ accept model if confidence ellipses cover bisector
Andersen’s likelihood ratio (LR) test

idea: item parameter estimates should not depend on the person-sample

- split person sample into $K$ subsamples based on, e.g., the raw scores
- compare ML-estimates $\hat{\beta}_k$ from $k = 1, \ldots, K$ subsamples and $\hat{\beta}$ from entire sample

$$LR = \frac{L_u(r, \hat{\beta})}{\prod_{k=1}^K L_{u_k}(r_k, \hat{\beta}_k)} = \frac{\prod_{k=1}^K L_{u_k}(r_k, \hat{\beta})}{\prod_{k=1}^K L_{u_k}(r_k, \hat{\beta}_k)}$$

- $T = -2 \log LR \overset{\text{as}}{\sim} \chi^2((K - 1) \cdot (M - 1) - (M - 1))$ for $M$ items
- $H_0 :$ model holds ($LR = 1, T = 0$)
- $H_1 :$ model violated ($LR < 1, T >>> 0$)
  $\Rightarrow$ accept model if p-value is large
The Rasch Model
Carolin Strobl

Objective
Model specification
ICCs
Specific objectivity
Local stochastic independence
Sufficient statistics
Derivation of the model

Parameter estimation
ML estimation
Joint ML
Conditional ML
Marginal ML
Information of an item

Model diagnostics
Graphical test
LR test
Wald tests

Extended models
Birnbaum models
Models for ordinal data
The Rasch model as a GLMM

Example: Consumer survey

References

Wald test

idea: item parameter estimates should not depend on the person-sample

split person sample into $K$ (usually $K = 2$) subsamples based on, e.g., the raw scores

compare ML-estimates $\hat{\beta}_1$ and $\hat{\beta}_2$

$$W = (\hat{\beta}_1 - \hat{\beta}_2)'(\hat{\Sigma}_1 + \hat{\Sigma}_2)^{-1}(\hat{\beta}_1 - \hat{\beta}_2)$$

$W^{\text{as.}} \sim \chi^2$

$H_0 :$ model holds ($W = 0$)

$H_1 :$ model violated ($W >> 0$)

$\Rightarrow$ accept model if p-value is large

note: LR and Wald tests, as well as Lagrange-Multiplier (LM) tests, are asymptotically equivalent
Item specific Wald test

idea: item parameter estimates should not depend on the person-sample

- split person sample into $K$ (usually $K = 2$) subsamples based on, e.g., the raw scores
- compare ML-estimates $\hat{\beta}_{j,1}$ and $\hat{\beta}_{j,2}$

$$W_j = \frac{(\hat{\beta}_{j,1} - \hat{\beta}_{j,2})^2}{\hat{\sigma}^2_{j,1} + \hat{\sigma}^2_{j,2}}$$

- $\text{sign}(\hat{\beta}_{j,1} - \hat{\beta}_{j,2}) \sqrt{W_j} \overset{\text{as.}}{\sim} N(0, 1)$
- $H_0 :$ model holds for item $j$ ($W_j = 0$)
- $H_1 :$ model violated for item $j$ ($|W_j| >> 0$) (note: two-sided test)
  $\Rightarrow$ exclude item $j$ if p-value is small ($< 0.05$)
Sample splitting

for graphical, LR and Wald tests: sample can be split based on

- the raw scores
- any other criterion, including covariates such as gender, age etc.
- usually the median is arbitrarily used for splitting

alternative approaches:

- “mixed” (mixture distribution) Rasch model
- Rasch trees
Rasch trees

Strobl, Kopf, and Zeileis (2010a,b)

⇒ Achim’s Antrittsvorlesung
Extended models
The Birnbaum (two parameter) model

\[
P(u_{ij}=1|\theta_i, \beta_j, \delta_j) = \frac{e^{\delta_j(\theta_i-\beta_j)}}{1 + e^{\delta_j(\theta_i-\beta_j)}}
\]
The Birnbaum (two parameter) model

comparison of items not specifically objective
The (Birnbaum) three-parameter model

\[
P(u_{ij} = 1|\theta_i, \beta_j, \delta_j, \gamma_j) = \gamma_j + (1 - \gamma_j) \cdot \left(\frac{e^{\delta_j(\theta_i - \beta_j)}}{1 + e^{\delta_j(\theta_i - \beta_j)}}\right)
\]

for multiple choice tests set \(\gamma_i = \gamma = \frac{1}{\text{number of options}}\)
Partial credit model

\[ P(u_{ij} = c | \theta_i, \beta_j) = \frac{e^{c \cdot \theta_i - \beta_{jc}}}{\sum_{l=0}^{m_j} e^{l \cdot \theta_i - \beta_{jl}}} \]

with \( c = 0, 1, \ldots, m_j \) and \( \beta_{j0} = 0 \)

(Masters, 1982)
Partial credit model

thresholds $\tau_{j1}, \ldots, \tau_{jm_j}$  
(intersections of ICCs for categories 0 and 1, 1 and 2 etc.)

$\beta_{j0} = 0$, $\beta_{jk} = \sum_{l=1}^{k} \tau_{jl}$

location $\bar{\tau}$
(intersection of ICCs for categories 0 and $m_j$)
Data

- to what degree (for example 0 – 5 credits) was the item answered correctly?
- how strongly (on a scale from 0 – 5) did the subject agree to the item?
Model diagnostics

- if $\beta_{jk}$ and $\tau_{jk}$ are not ordered: item $j$ violates model assumptions
- will be eliminated, e.g., by Wald test
Other “ordinal” models

- Andrich’s rating scale model
  special case of the partial credit model with a fixed number of categories $m_j = m$ for each item

- Samejima’s graded response model
  cumulative probabilities for passing the successive category thresholds

(see also Masters, 1982, for a comparison)
The Rasch model as a generalized linear mixed model (GLMM)

- item parameters are considered as fixed effects
- person parameters are considered as random effects
  \[ \theta \sim N(0, \Sigma), \text{i.e. the random effects } \theta; \text{ are deviations from the average} \]

(Rijmen et al., 2003)
Example: Consumer survey

from Salzberger & Sinkovics (International Marketing Review, 2006)

- consumer study on technophobia (in ATM usage)
- five category Likert items

⇒ Partial Credit model
  category ordering?

- samples from England (N = 278), Mexico (N = 200) and Austria (N = 449)

⇒ Differential Item Functioning?
Example: Consumer survey

from Salzberger & Sinkovics (International Marketing Review, 2006)
References and further reading I


References and further reading II


