Using SAS to Assess the Fit of Survey Response Data to the Rasch Measurement Model

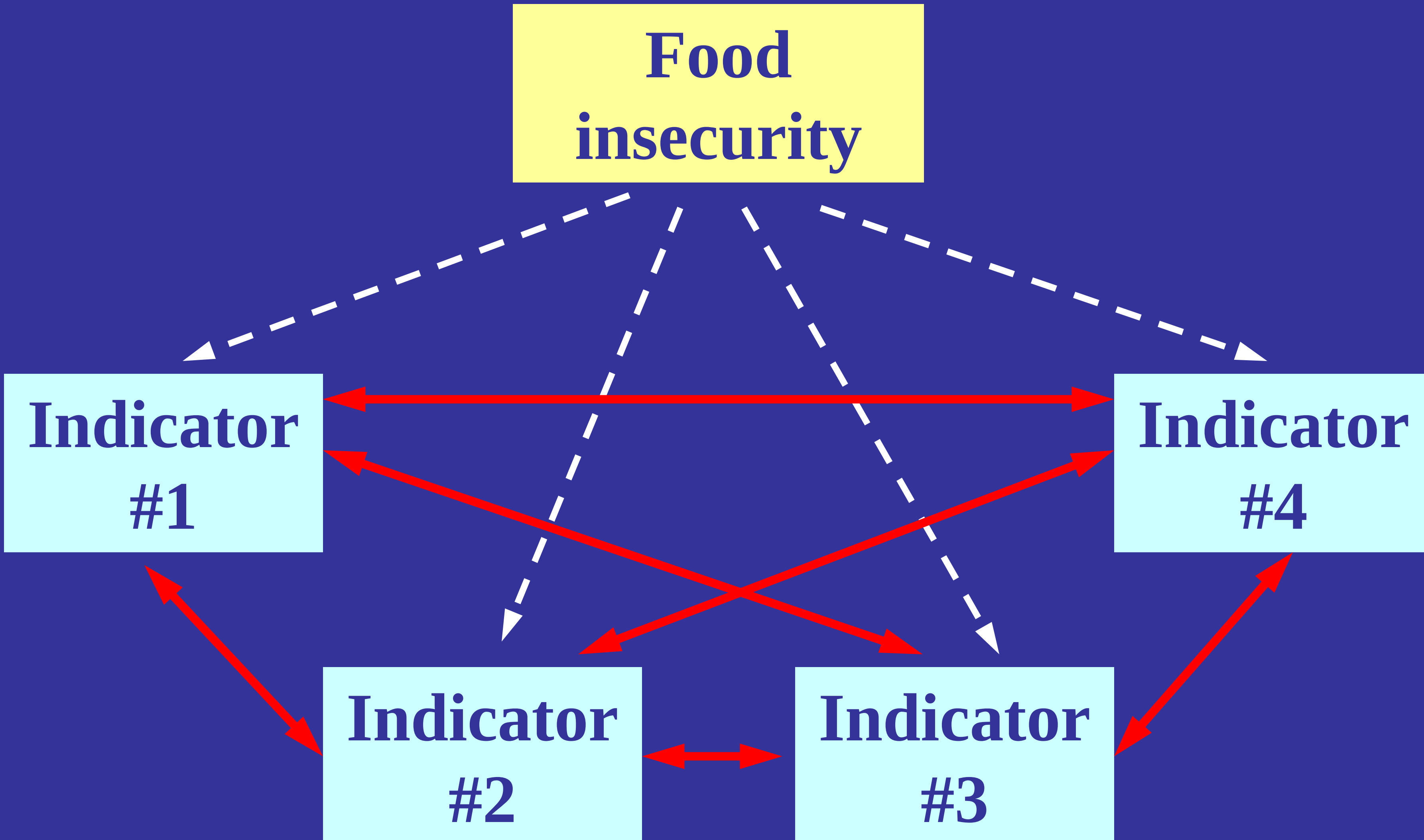
Presentation at SIGSTAT

Mark Nord
Economic Research Service
USDA

Preview

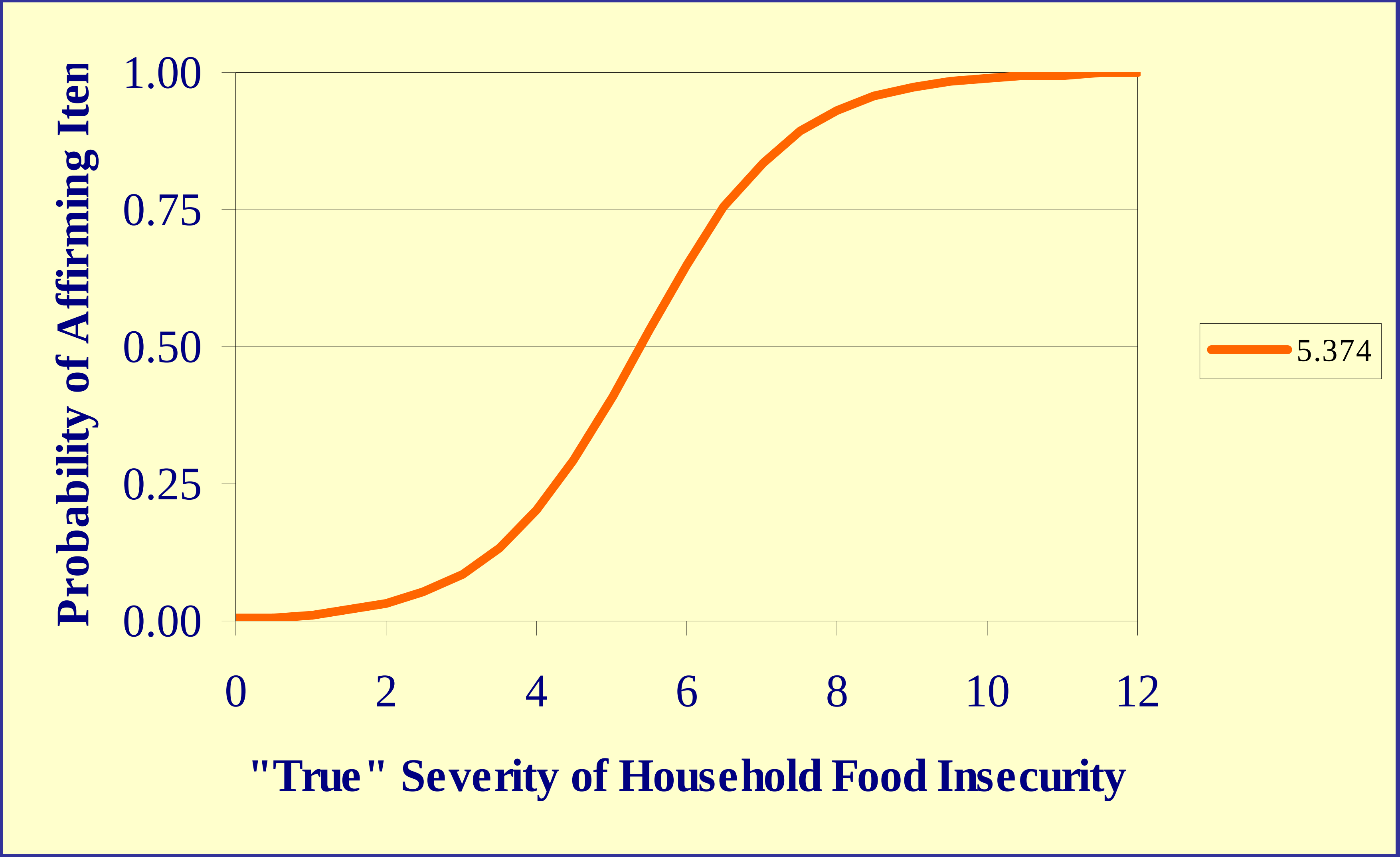
- Overview of latent trait measurement models
- Mathematics of the Rasch measurement model
- Why use this model?
- Assessment statistics, item parameters, item-fit statistics
- Estimation methods
- JML estimation using PROC LOGISTIC
- CML estimation using PROC LOGISTIC with the STRATA command
- MML estimation using PROC NLMIXED

Theory of multiple-indicator measures

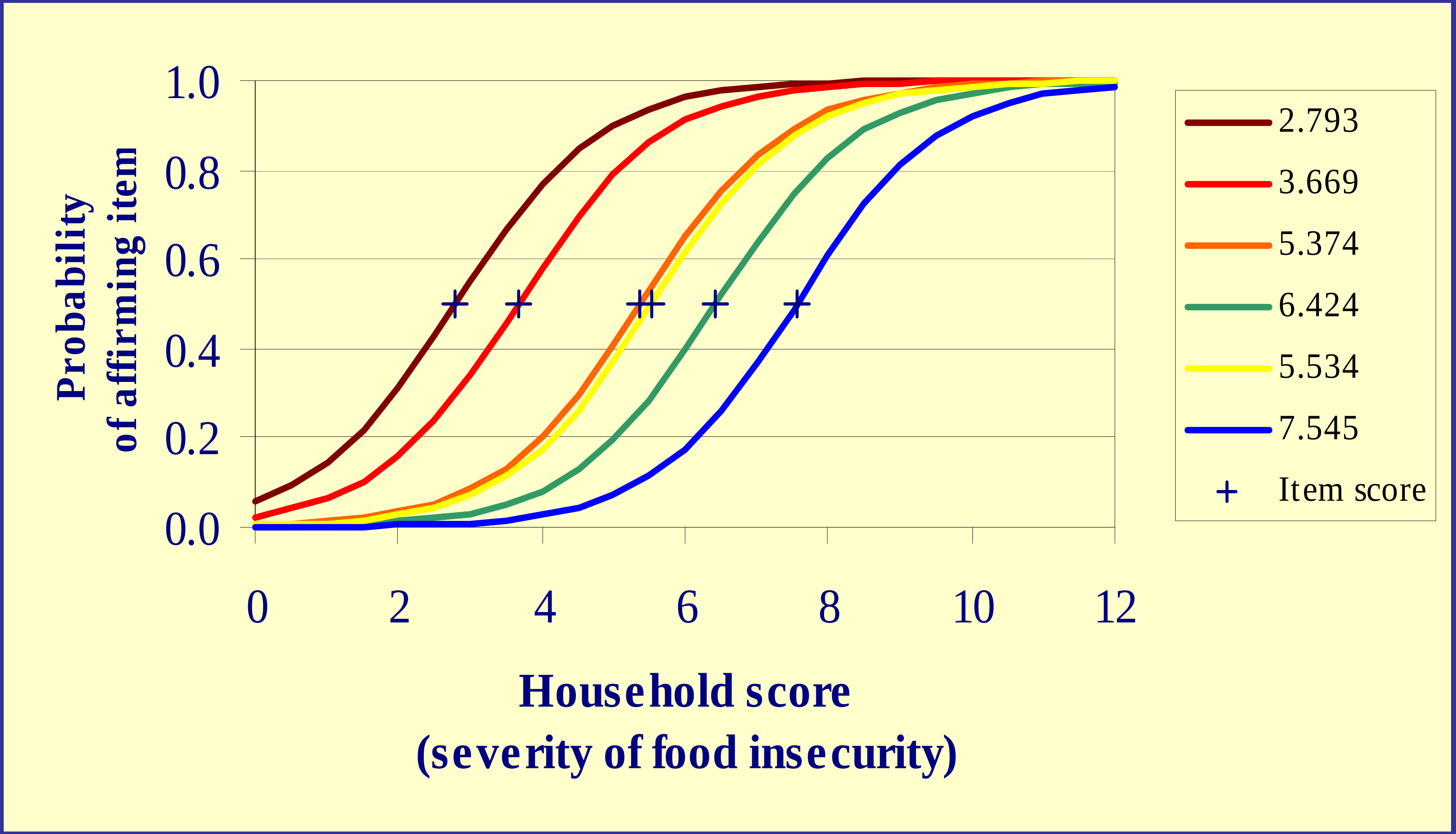


Probability That a Respondent Affirms Item as Function of Severity of Food Insecurity Experienced

log[p/(1-p)] = householdseverity - itemseverity



Response Functions of 6 Items with Different Severity but Equal Discrimination



Why use the Rasch model?

- Simplicity: Raw score is ordinal
- Specific objectivity
- Statistically defensible measurement

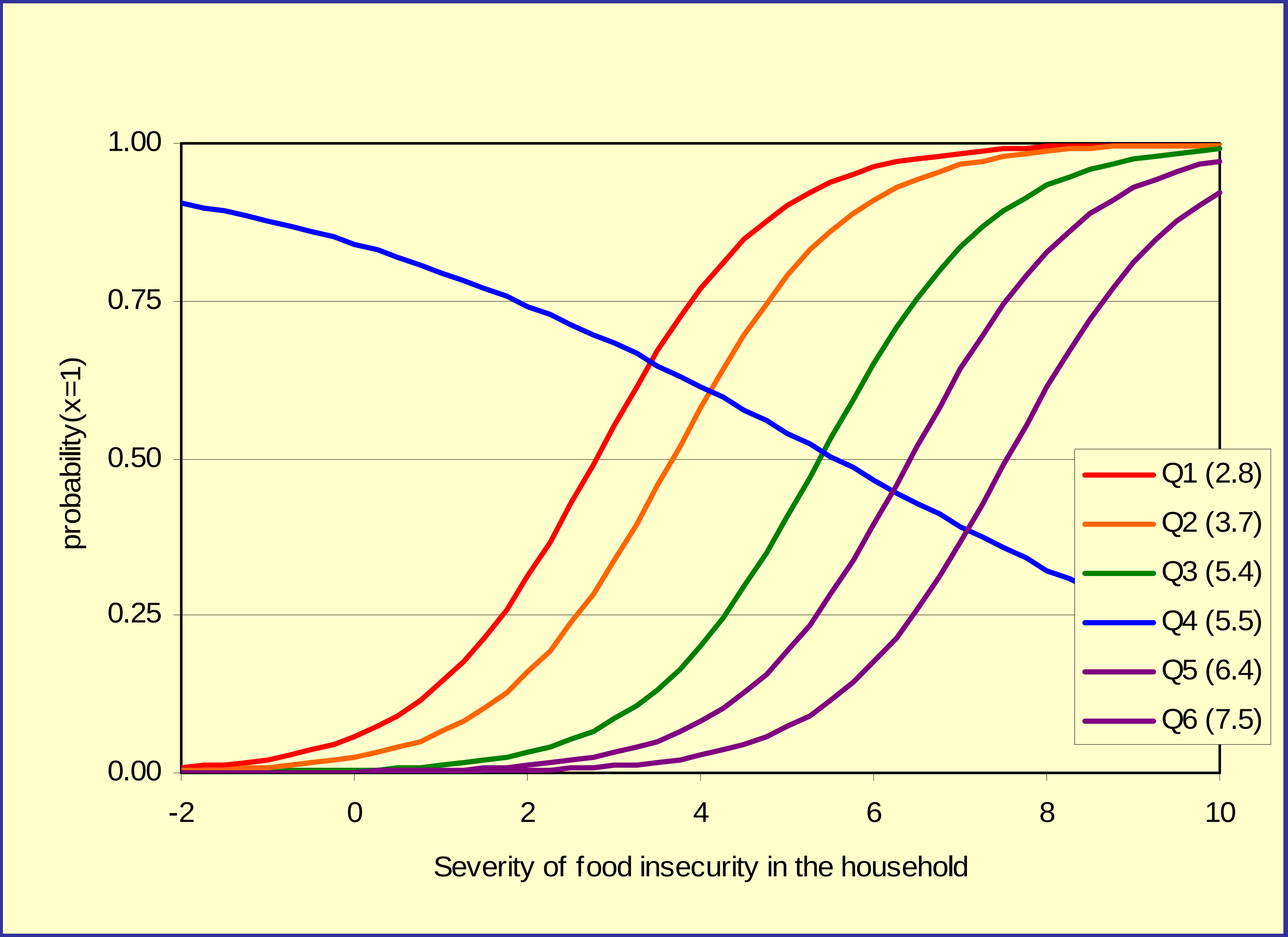
Key Assumptions of the Rasch model

- Items discriminate equally and are related to the underlying trait by the logistic function
    - Item-infit and Item-outfit statistics

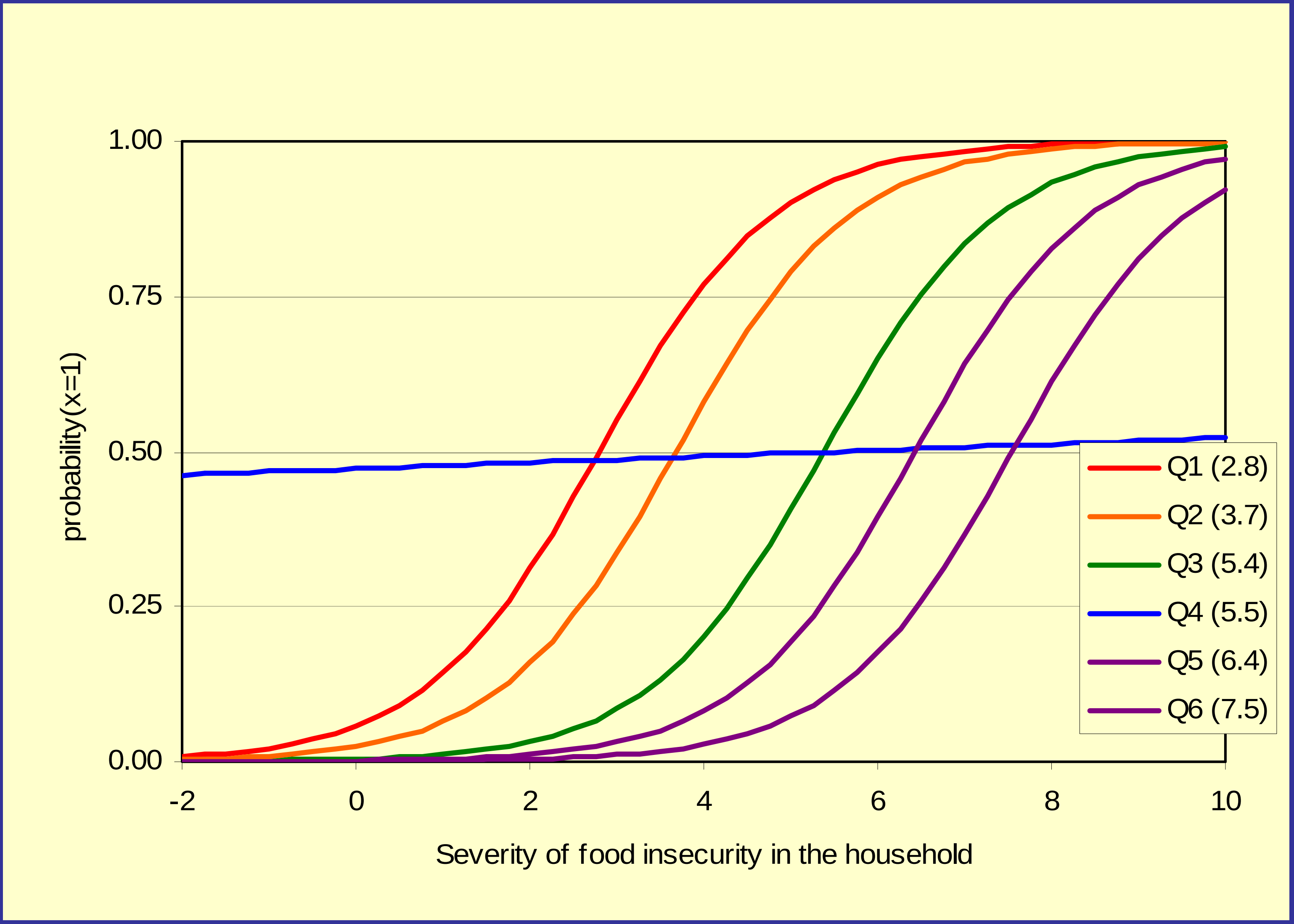
- Items are conditionally independent
    - Factor analysis of standardized residuals

- Items function the same in all subpopulations
    - Differential Item Function (DIF) or comparison of item parameters between subpopulations

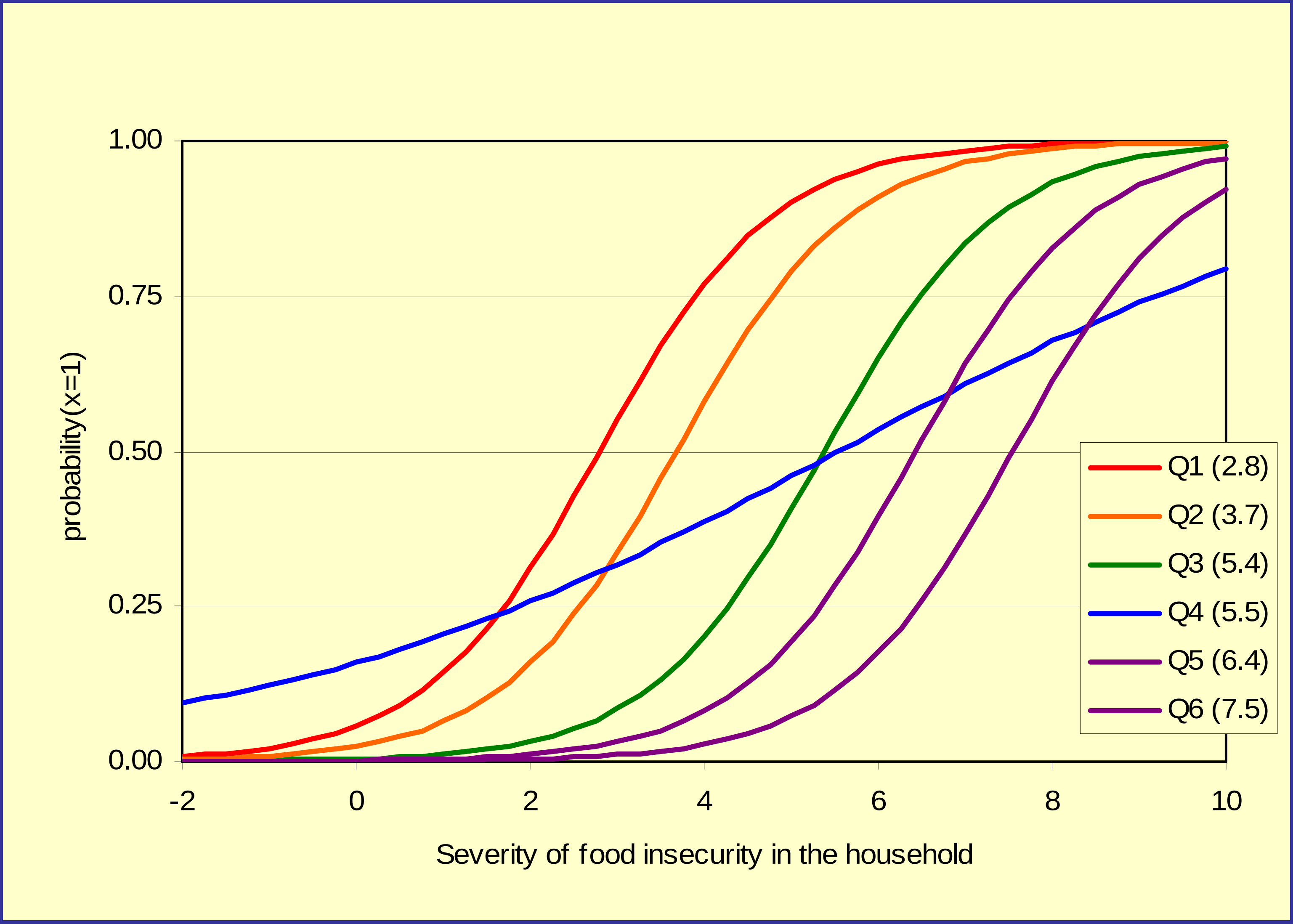
Would you want Q4 in your scale?



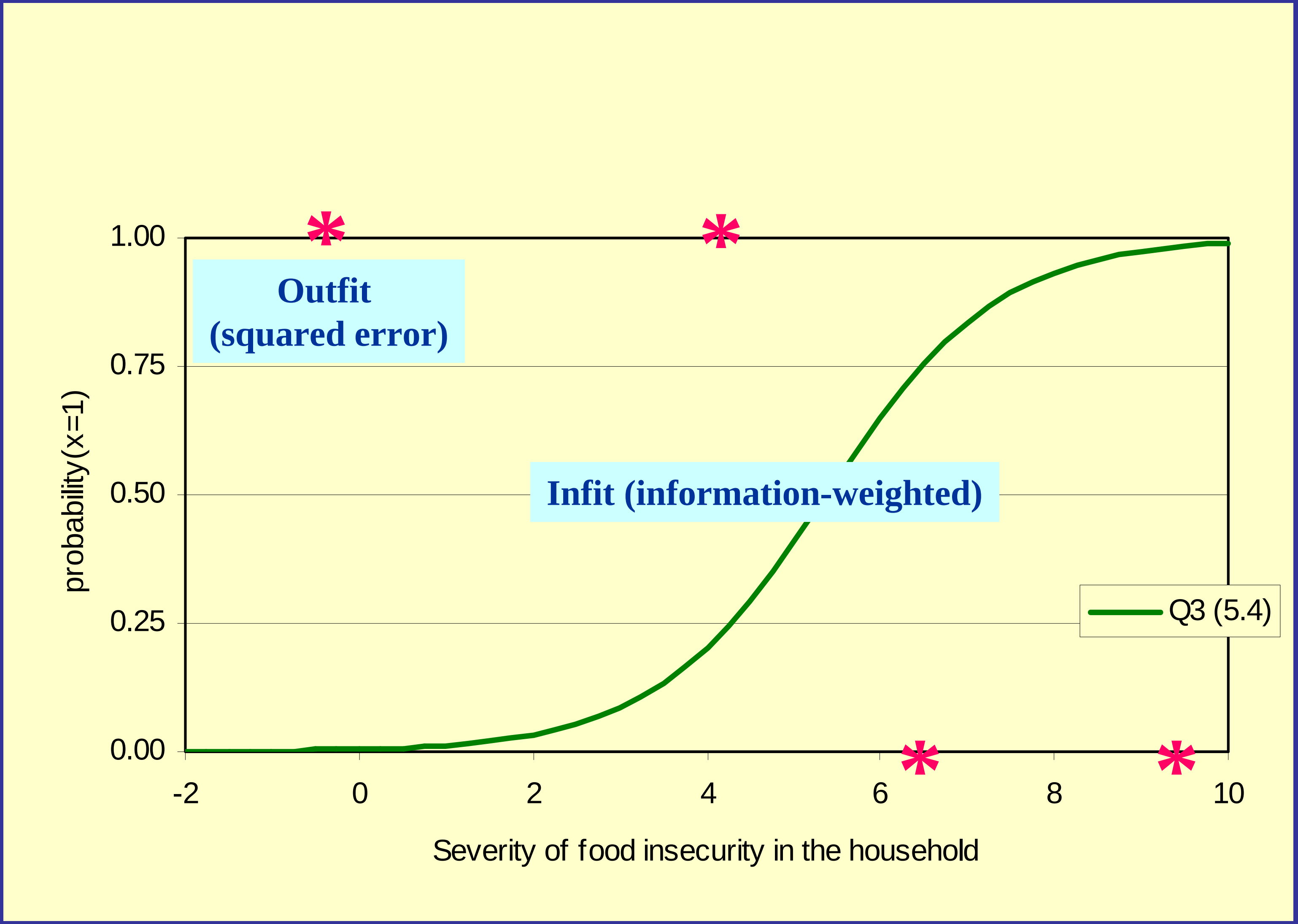
Would you want Q4 in your scale?



Which item has low discrimination?

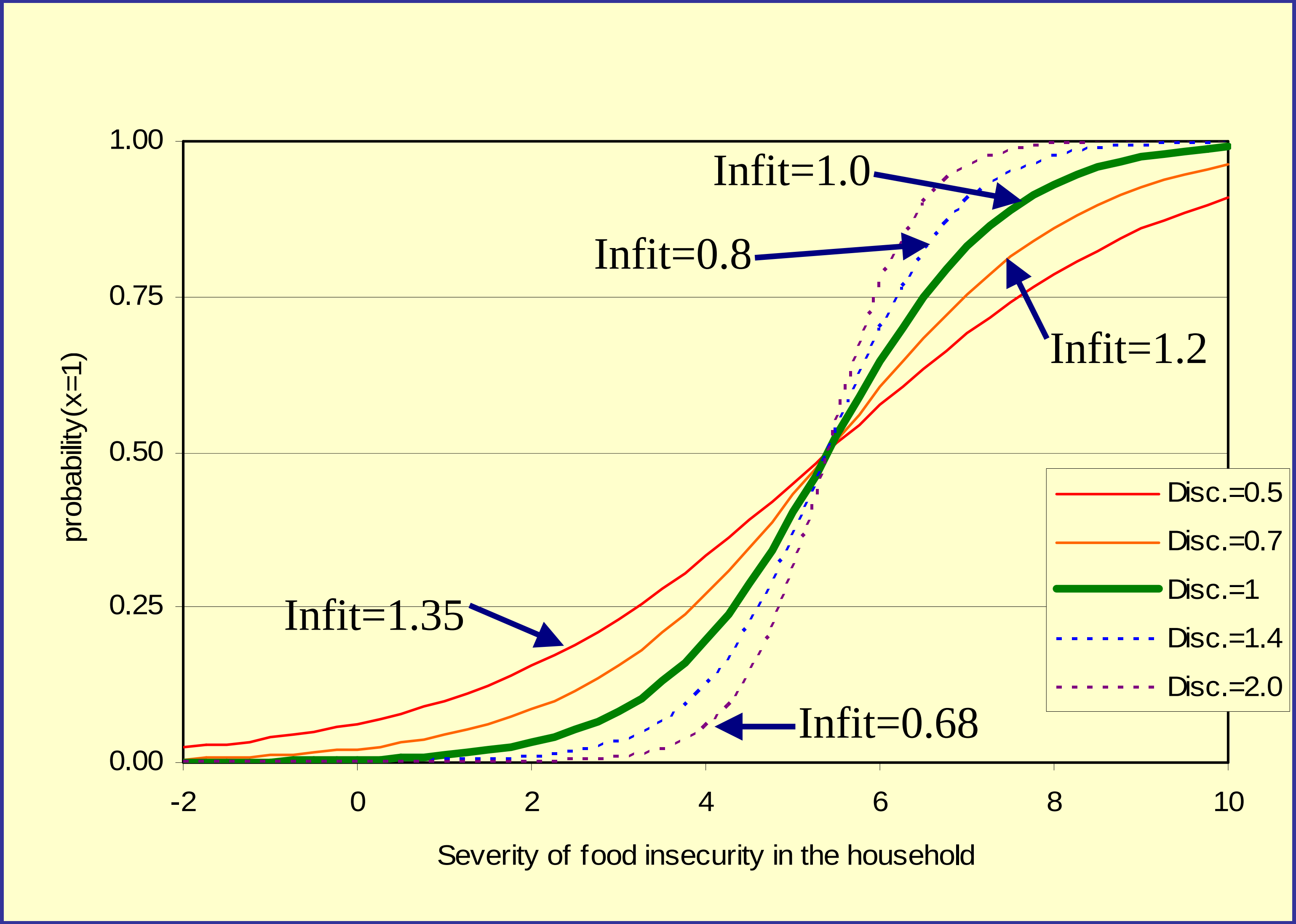


Item-(mis)fit statistics



Items with the same severity but different discrimination

log[p/(1-p)] = (householdseverity – itemseverity)*k



Estimation Methods

- Conditional maximum likelihood (CML)
  - Conditioned on raw score
  - Sample independent
  - Estimates are statistically consistent
  - SAS procedure cannot use case weights
- Joint (or unconditional) maximum likelihood (JML)
  - Respondent and item parameters estimated simultaneously
  - Sample independent
  - Estimates are not statistically consistent*
  - SAS procedure can use case weights
- Marginal maximum likelihood (MML)
  - Somewhat sample dependent
  - SAS procedure cannot use case weights

2006 Food Security Data

Data Set Name	WORK.TEMP06	Observations	12442
Member Type	DATA	Variables	10
Engine	V9	Indexes	0
Created	Monday, December 31, 2007 10:47:34 AM	Observation Length	80
Last Modified	Monday, December 31, 2007 10:47:34 AM	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Variables in Creation Order

#	Variable	Type	Len
1	PESEX	Num	8
2	worried	Num	8
3	fnotlast	Num	8
4	balmeal	Num	8
5	cutskip	Num	8
6	eatless	Num	8
7	hungry	Num	8
8	losewt	Num	8
9	whlday	Num	8
10	rawsc	Num	8

2006 Food Security Data

The FREQ Procedure

rawsc	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	10182	81.84	10182	81.84
1	517	4.16	10699	85.99
2	384	3.09	11083	89.08
3	368	2.96	11451	92.04
4	244	1.96	11695	94.00
5	238	1.91	11933	95.91
6	198	1.59	12131	97.50
7	184	1.48	12315	98.98
8	127	1.02	12442	100.00

2006 Food Security Data

```
Data temp06;
set temp06 (where=(rawsc gt 0 and rawsc lt 8));

proc means data=temp06;
var worried--whlday;
run;
```

```
                    The MEANS Procedure

Variable      N           Mean        Std Dev     Minimum     Maximum
---------------------------------------------------------------------
worried    2133      0.7834037      0.4120218           0   1.0000000
fnotlast   2133      0.6516643      0.4765547           0   1.0000000
balmeal    2133      0.6793249      0.4668455           0   1.0000000
cutskip    2133      0.4078762      0.4915552           0   1.0000000
eatless    2133      0.3886545      0.4875588           0   1.0000000
hungry     2133      0.2062822      0.4047304           0   1.0000000
losewt     2133      0.1157993      0.3200592           0   1.0000000
whlday     2133      0.0632911      0.2435430           0   1.0000000
---------------------------------------------------------------------
```

Restructure the data

(One record per respondent)

Rawsc	Q1	Q2	Q3	Q4
2	1	0	1	0
3	1	1	1	0

(One record for each item for each respondent)

Draw1	Draw2	Draw3	DQ1	DQ2	DQ3	DQ4	RX
0	1	0	1	0	0	0	1
0	1	0	0	1	0	0	0
0	1	0	0	0	1	0	1
0	1	0	0	0	0	1	0
0	0	1	1	0	0	0	1
0	0	1	0	1	0	0	1
0	0	1	0	0	1	0	1
0	0	1	0	0	0	1	0

Restructure data

```
*restructure data, calculate dummy sets for raw score and item;
data tempx; set temp06;
array dq{8} dq1-dq8; *dummy indicating which response var is operative;
array qd{8} worried fnotlast balmeal cutskip eatless hungry losewt whlday;
                               *original response variables;
array draw{7} draw1-draw7; *dummy set for raw score;
if _n_ eq 1 then hhid=0;
retain hhid;
hhid=hhid+1; *creates unique hhid to be used as strata;
do q=1 to 8; *create 1 record per variable;
  do qq=1 to 8; dq{qq}=(qq eq q); end; *sets dummies for operating item;
  do r=1 to 7; draw{r}=(r eq rawsc); end; *sets dummies for raw score;
  rx=qd{q};
  output;
  end;
run;


proc means n mean sum min max data=tempx;
var rx dq1--dq8 hhid draw1--draw7;
title1 'Run 117: HH x 8, 2006 w single person, nonextreme on 8 items';
run;
```

Means omitting extreme responses

```
* The MEANS Procedure

 Variable          N           Mean            Sum        Minimum        Maximum
 --------------------------------------------------------------------------------
 rx            17064      0.4120370        7031.00              0      1.0000000
 dq1           17064      0.1250000        2133.00              0      1.0000000
 dq2           17064      0.1250000        2133.00              0      1.0000000
 dq3           17064      0.1250000        2133.00              0      1.0000000
 dq4           17064      0.1250000        2133.00              0      1.0000000
 dq5           17064      0.1250000        2133.00              0      1.0000000
 dq6           17064      0.1250000        2133.00              0      1.0000000
 dq7           17064      0.1250000        2133.00              0      1.0000000
 dq8           17064      0.1250000        2133.00              0      1.0000000
 hhid          17064        1067.00    18207288.00      1.0000000        2133.00
 draw1         17064      0.2423816        4136.00              0      1.0000000
 draw2         17064      0.1800281        3072.00              0      1.0000000
 draw3         17064      0.1725270        2944.00              0      1.0000000
 draw4         17064      0.1143929        1952.00              0      1.0000000
 draw5         17064      0.1115799        1904.00              0      1.0000000
 draw6         17064      0.0928270        1584.00              0      1.0000000
 draw7         17064      0.0862635        1472.00              0      1.0000000
 --------------------------------------------------------------------------------
```

Estimate JML model

```
*estimate JML model;
proc logistic data=tempx order=internal descending;
model rx=dq2--dq8 draw2--draw7;
output out=tempxp pred=pyes;
run;
```

```
NOTES:
Omitted one dummy variable from each set.
Output file will be used to calculate fit statistics.
```

SAS output from proc logistic

```
Response Variable             rx
Number of Response Levels     2

   Number of Observations Read       17064
   Number of Observations Used       17064

   Response Profile

      Ordered                        Total
        Value            rx      Frequency

            1             1           7031
            2             0          10033

         Probability modeled is rx=1.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

                                 Intercept
                  Intercept            and
    Criterion          Only     Covariates

    AIC           23126.839      11006.734
    SC            23134.584      11115.160
    -2 Log L      23124.839      10978.734
```

SAS output from proc logistic (continued)

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12146.1047	13	<.0001
Score	8794.1408	13	<.0001
Wald	3889.0428	13	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.2642	0.0707	13.9772	0.0002
dq2	1	-0.9985	0.0861	134.4834	<.0001
dq3	1	-0.7973	0.0859	86.0791	<.0001
dq4	1	-2.8373	0.0953	886.2781	<.0001
dq5	1	-3.0000	0.0966	963.5800	<.0001
dq6	1	-4.8453	0.1164	1734.1118	<.0001
dq7	1	-6.1155	0.1347	2060.6192	<.0001
dq8	1	-7.1284	0.1543	2132.9405	<.0001
draw2	1	1.1926	0.0754	249.9695	<.0001
draw3	1	2.2340	0.0792	795.5265	<.0001
draw4	1	3.2979	0.0945	1217.0346	<.0001
draw5	1	4.4620	0.1054	1793.7841	<.0001
draw6	1	5.7537	0.1241	2148.0271	<.0001
draw7	1	7.2083	0.1452	2465.0919	<.0001

NOTE: Standard errors are not correct, especially true of respondent score errors.

Analize in spreadsheet

		Item	Item param	Raw sco	Respondent param
Intercept	-0.26	worried	0.00	1	-0.26
dq2	-1.00	fnotlast	1.00	2	0.93
dq3	-0.80	balmeal	0.80	3	1.97
dq4	-2.84	cutskip	2.84	4	3.03
dq5	-3.00	eatless	3.00	5	4.20
dq6	-4.85	hungry	4.85	6	5.49
dq7	-6.12	losewt	6.12	7	6.94
dq8	-7.13	whlday	7.13		
draw2	1.19				
draw3	2.23	How likely is it that a respondent with raw score 6			
draw4	3.30	will report that their food did not last?			
draw5	4.46	odds=exp(5.49-1.00)			89.21
draw6	5.75				
draw7	7.21	prob=odds/(1+odds)			0.989

Calculate item-fit statistics and item parameter estimation error

```
*calculate item-fit statistics and item estimation errors;
data tempxp; set tempxp;
errsq=(rx-pyes)**2;
exerrsq=pyes*(1-pyes);
outfit=errsq/exerrsq;
ncases=1;
proc sort data=tempxp; by q;
proc summary data=tempxp; by q;
var ncases rx errsq exerrsq outfit;
output out=tempfit sum=;

data tempfit; set tempfit;
itemerr=sqrt(1/exerrsq);
infit=errsq/exerrsq;
outfit=outfit/ncases;

proc print data=tempfit;
Var q ncases rx itemerr infit outfit;
run;
```

SAS output: item-fit statistics and item parameter estimation error

Obs	q	ncases	rx	itemerr	infit	outfit
1	1	2133	1671	0.06206	1.12109	3.90637
2	2	2133	1390	0.05826	0.93695	2.13751
3	3	2133	1449	0.05856	1.25699	2.85469
4	4	2133	870	0.06258	0.75635	0.60714
5	5	2133	829	0.06342	0.78921	0.64631
6	6	2133	440	0.07565	0.81701	0.56705
7	7	2133	247	0.08793	0.98813	0.70588
8	8	2133	135	0.10475	1.06535	1.12632

Note: Outfit statistics are doubly distorted, first by the inconsistency of the JML estimates and second by the screening inherent in module as administered.

Infit statistics of the least severe and most severe items are biased slightly upward by the inconsistency of the JML estimates. Module screening has minimal effect on infits.

Estimate CML model

```
*estimate CML model;
proc logistic data=tempx order=internal descending;
model rx=dq2--dq8;
strata hhid; * /nosummary;
output out=tempxp2 pred=pyes2;
run;
```

Note: Unlike JML, no dummy variables are included for raw score. Instead, the strata statement creates individual fixed effects.

The nosummary option could be included to suppress the information about strata in the output.

This took 16 minutes of computer time!!

SAS Proc Logistic output, CML

Data Set	WORK.TEMPX
Response Variable	rx
Number of Response Levels	2
Number of Strata	2133
Model	binary logit
Optimization Technique	Newton-Raphson ridge

Number of Observations Read	17064
Number of Observations Used	17064

Response Profile

Ordered Value	rx	Total Frequency
1	1	7031
2	0	10033

Probability modeled is rx=1.

SAS Proc Logistic output, CML (continued)

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
AIC	13746.055	7267.344
SC	13746.055	7321.558
-2 Log L	13746.055	7253.344

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6492.7106	7	<.0001
Score	5314.3108	7	<.0001
Wald	3095.3462	7	<.0001

SAS Proc Logistic output, CML (continued)

The LOGISTIC Procedure

Conditional Analysis

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
dq2	1	-0.7574	0.0747	102.8364	<.0001
dq3	1	-0.5972	0.0742	64.8060	<.0001
dq4	1	-2.2815	0.0837	743.2962	<.0001
dq5	1	-2.4202	0.0849	812.7233	<.0001
dq6	1	-3.9722	0.1014	1534.0085	<.0001
dq7	1	-4.9982	0.1155	1871.7910	<.0001
dq8	1	-5.7330	0.1281	2002.1155	<.0001

Note: The CML model does not directly provide estimates of the respondent parameters. They can be calculated using ML methods.

Standard errors are slightly high because they are all relative to the omitted dq1 and effectively incorporate estimation error for that item as well.

Calculate item-fit statistics from CML output dataset

```
*calculate item-fit statistics and item estimation errors;
data tempxp2; set tempxp2;
errsq=(rx-pyes2)**2;
exerrsq=pyes2*(1-pyes2);
outfit=errsq/exerrsq;
ncases=1;
proc sort data=tempxp2; by q;
proc summary data=tempxp2; by q;
var ncases rx errsq exerrsq outfit;
output out=tempfit2 sum=;

data tempfit2; set tempfit2;
itemerr=sqrt(1/exerrsq);
infit=errsq/exerrsq;
outfit=outfit/ncases;

proc print data=tempfit2;
var q ncases rx itemerr infit outfit;
run;
```

SAS output: item-fit statistics and item parameter estimation error, CML

Obs	q	ncases	rx	itemerr	infit	outfit
1	1	2133	1671	0.06137	1.09362	2.56558
2	2	2133	1390	0.05892	0.95130	1.73292
3	3	2133	1449	0.05907	1.28279	2.33355
4	4	2133	870	0.06361	0.77097	0.61476
5	5	2133	829	0.06447	0.80709	0.65683
6	6	2133	440	0.07691	0.82862	0.56194
7	7	2133	247	0.08860	1.00123	0.63647
8	8	2133	135	0.10209	1.00918	0.79386

Note: Outfit statistics are still distorted by module screening, but are otherwise unbiased.

Infit statistics are essentially unbiased (bias due to module screening is minimal.

Highest infit (q3) is “balanced meals” question, which is known to be somewhat problematic.

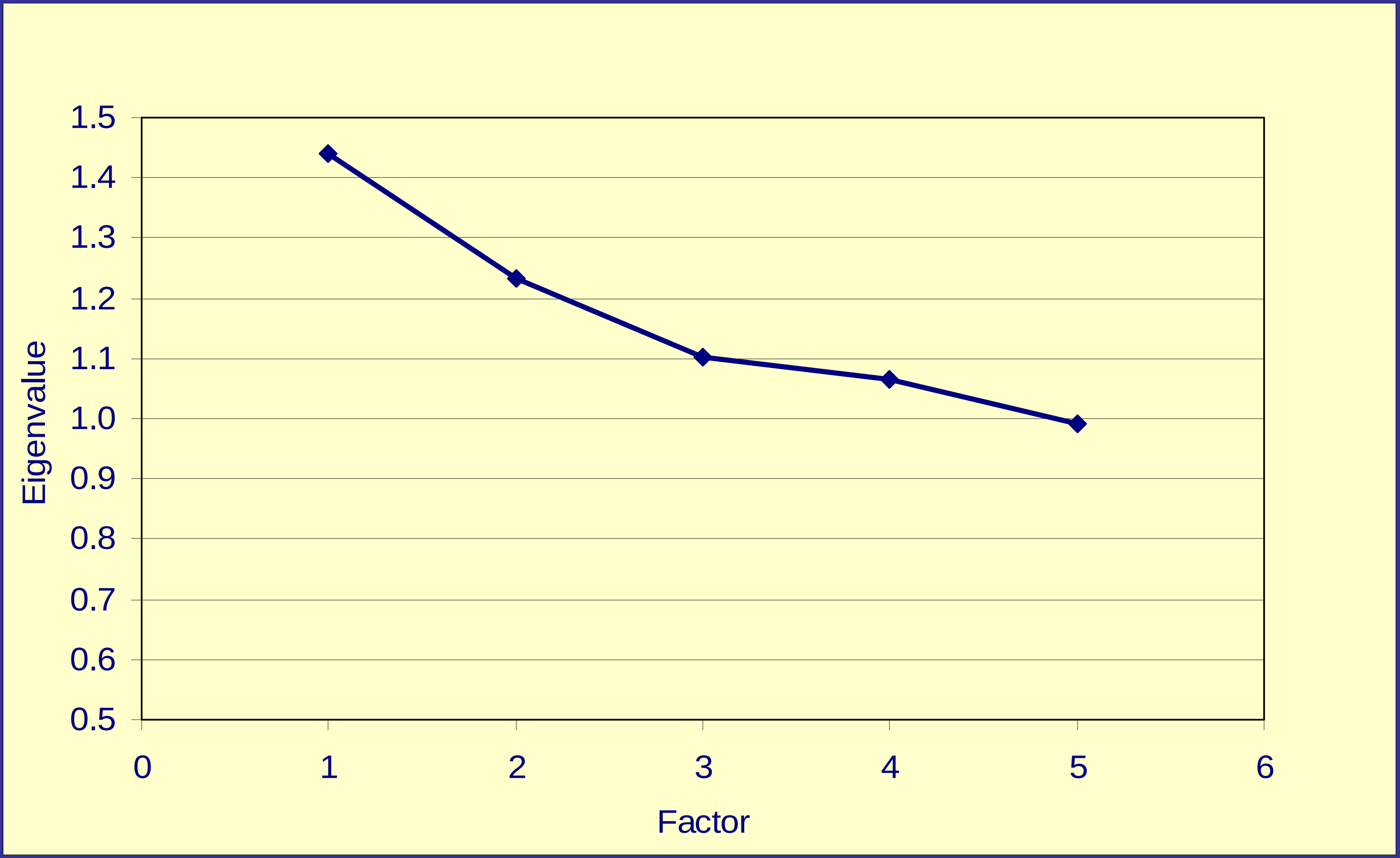
Assess conditional independence by factor analysis of standardized residuals from CML output dataset

```
*factor analysis of residuals from cml;
proc sort data=tempxp2; by hhid;
data tempfact; set tempxp2; by hhid;
array qresid {8} qresid1-qresid8;
retain qresid1--qresid8;
qresid{q}=(rx-pyes2)/sqrt(pyes2*(1-pyes2));
*std residual is error divided by espected std dev of error;
if last.hhid then output; *create record in tempfact;

proc factor data=tempfact mineigen=1 method=principal;
var qresid1--qresid8;
run;
```

SAS output: factor analysis of standardized residuals, CML

```
The FACTOR Procedure
             Initial Factor Method: Principal Components

                 Prior Communality Estimates: ONE



     Eigenvalues of the Correlation Matrix: Total = 8  Average = 1

             Eigenvalue    Difference    Proportion    Cumulative

        1    1.44044115    0.20814007        0.1801        0.1801
        2    1.23230108    0.12993192        0.1540        0.3341
        3    1.10236916    0.03830276        0.1378        0.4719
        4    1.06406641    0.07322238        0.1330        0.6049
        5    0.99084403    0.02434823        0.1239        0.7288
        6    0.96649580    0.05607504        0.1208        0.8496
        7    0.91042075    0.61735913        0.1138        0.9634
        8    0.29306162                      0.0366        1.0000

        4 factors will be retained by the MINEIGEN criterion.
```

Scree plot: factor analysis of standardized residuals, CML

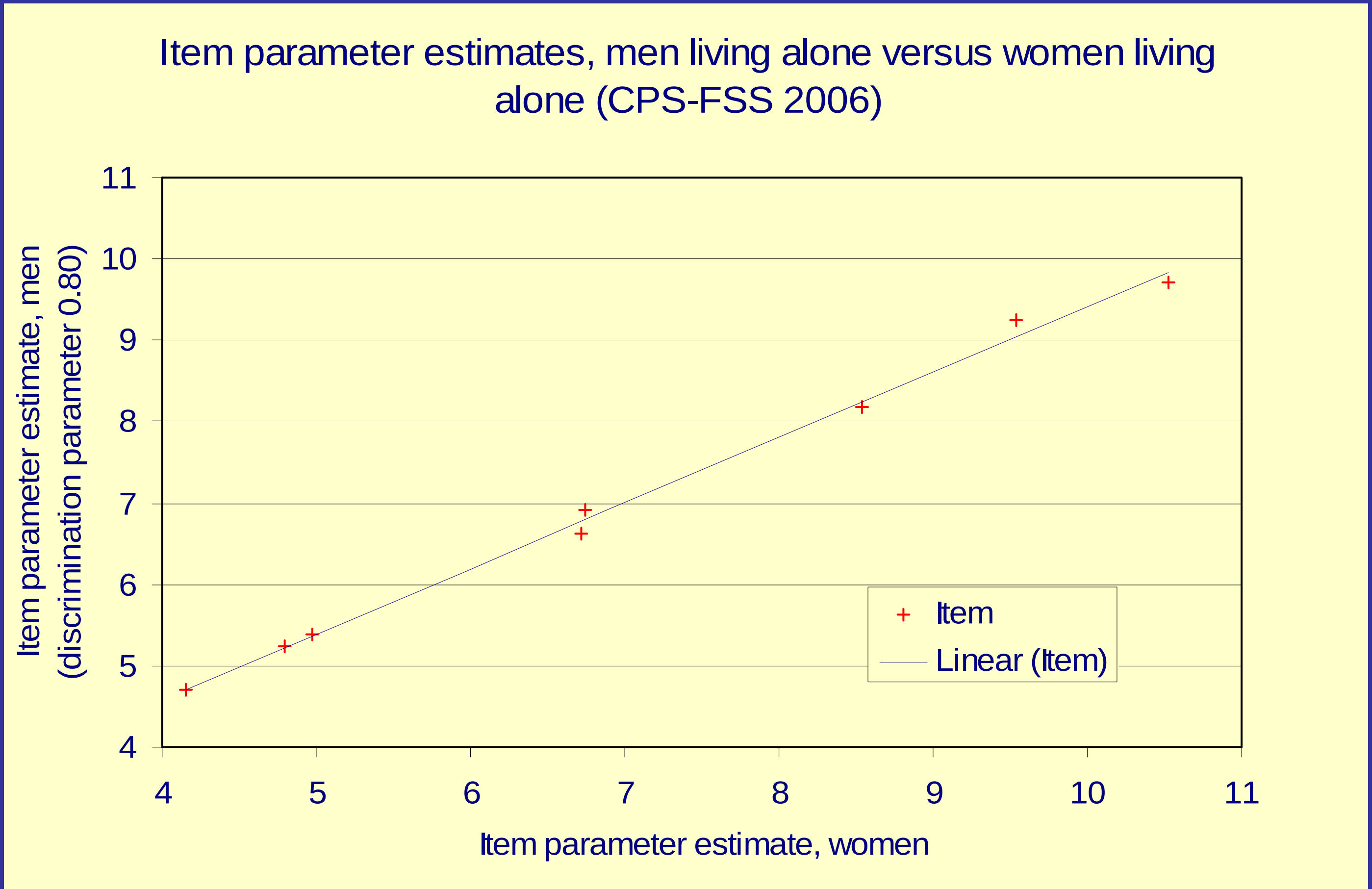


SAS output: factor analysis of standardized residuals, CML (continued)

Factor Pattern

	Factor1	Factor2	Factor3	Factor4
qresid1	-0.46543	0.22196	0.65892	0.03233
qresid2	-0.59490	0.21655	0.03361	-0.02564
qresid3	-0.44894	-0.40414	-0.67575	-0.04587
qresid4	0.36383	-0.25949	0.08386	0.65929
qresid5	0.28421	-0.47978	0.24182	0.18461
qresid6	0.28291	0.60982	-0.34030	0.15209
qresid7	0.38606	-0.29263	0.15085	-0.73904
qresid8	0.47553	0.46667	-0.07968	-0.14899

Comparing Measurement between subgroups



Differences can be compared with item estimation errors to assess statistical significance of differences.

Here it is apparent that substantive importance of differences are small.

Restructure data for CML estimation when some responses are missing

```
*restructure data, calculate dummy set for items;
data tempx; set temp06m;
array dq{8} dq1-dq8; *dummy indicating which response variable is operative;
array qd{8} worried fnotlast balmeal cutskip eatless hungry losewt whlday;
                                *original response variables;
if _n_ eq 1 then hhid=0;
retain hhid;
hhid=hhid+1; *creates unique hhid to be used as strata;
do q=1 to 8; *each item;
  if qd{q} in (0,1) then do; *valid response, create record;
    do qq=1 to 8; dq{qq}=(qq eq q); end; *sets dummies for operating item;
    rx=qd{q};
    output;
    end; *of valid response, create record;
  end; *of each item;
run;
```

Notes:
Extreme high cases are now those with all valid items ‘yes.’
Sorry, I haven’t worked out the math of this, but it seems to work. . .

Estimate MML Model

```
*estimate MML model;
proc nlmixed data=tempx;
parms b1=-3 b2=-2 b3=-2 b4=-1 b5=0 b6=1 b7=2 b8=3 s2u=1;
*specifying starting values for parameters speeds estimation;
*   but is not required;
eta=u-(b1*dq1+b2*dq2+b3*dq3+b4*dq4+b5*dq5+b6*dq6+b7*dq7+b8*dq8);
expeta=exp(eta);
pyes=expeta/(1+expeta);
model rx ~ binomial (1,pyes);
random u ~ normal(0,s2u) subject=hhid;
id pyes; *includes pyes in the output dataset;
predict u out=tempx3;

proc freq data=tempx3;
tables pred;
run;
```

Notes:
The random statement creates individual random effects.
This took 48 seconds of computer time.
Item-fit stats can be calculated using pyes as in JML and CML.
The proc freq provides household parameter estimates.

Questions?

marknord@ers.usda.gov

202-694-5433