

Test-Retest Reliability of Subjective Survival Expectations

Jochem de Bresser

Tilburg University and Netspar

October 2016

Motivation – why study subjective survival?

- Individual survival expectations drive behavior
 - Labor supply & saving (French, 2005; De Nardi, French and Jones, 2010)
- Traditional approach: “rational” expectations
- Alternative: subjective expectations
 - Survey respondents report their expectations
 - Multiple probabilities trace distribution (Manski, 2004)
- Previous literature on reliability of subjective survival
 - Predicts mortality/correlates plausibly with risk factors (Hurd and McGarry, 1995; Smith et al., 2001)
 - Updated after death of parent (Hurd and McGarry, 2002)

Motivation – why study subjective survival?

- Individual survival expectations drive behavior
 - Labor supply & saving (French, 2005; De Nardi, French and Jones, 2010)
- Traditional approach: “rational” expectations
- Alternative: subjective expectations
 - Survey respondents report their expectations
 - Multiple probabilities trace distribution (Manski, 2004)
- Previous literature on reliability of subjective survival
 - Predicts mortality/correlates plausibly with risk factors (Hurd and McGarry, 1995; Smith et al., 2001)
 - Updated after death of parent (Hurd and McGarry, 2002)

Motivation – why study subjective survival?

- Individual survival expectations drive behavior
 - Labor supply & saving (French, 2005; De Nardi, French and Jones, 2010)
- Traditional approach: “rational” expectations
- Alternative: subjective expectations
 - Survey respondents report their expectations
 - Multiple probabilities trace distribution (Manski, 2004)
- Previous literature on reliability of subjective survival
 - Predicts mortality/correlates plausibly with risk factors (Hurd and McGarry, 1995; Smith et al., 2001)
 - Updated after death of parent (Hurd and McGarry, 2002)

Motivation – why study subjective survival?

- Individual survival expectations drive behavior
 - Labor supply & saving (French, 2005; De Nardi, French and Jones, 2010)
- Traditional approach: “rational” expectations
- Alternative: subjective expectations
 - Survey respondents report their expectations
 - Multiple probabilities trace distribution (Manski, 2004)
- Previous literature on reliability of subjective survival
 - Predicts mortality/correlates plausibly with risk factors (Hurd and McGarry, 1995; Smith et al., 2001)
 - Updated after death of parent (Hurd and McGarry, 2002)

How reliable are subjective survival expectations?

- Compare responses of same individuals to similar questions in different surveys
- Two levels of aggregation:
 - Individual probabilities
 - Survival curves
- Link to saving and labor supply through life-cycle model

How reliable are subjective survival expectations?

- Compare responses of same individuals to similar questions in different surveys
- Two levels of aggregation:
 - Individual probabilities
 - Survival curves
- Link to saving and labor supply through life-cycle model

How reliable are subjective survival expectations?

- Compare responses of same individuals to similar questions in different surveys
- Two levels of aggregation:
 - Individual probabilities
 - Survival curves
- Link to saving and labor supply through life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

Contributions of the paper

Evaluate test-retest reliability...

- 1 In a nationally representative sample
- 2 With analogous methods used for other subjective data
- 3 At different levels of aggregation
- 4 In different dimensions of the data
- 5 Taking into account measurement error
- 6 In terms of behavior in life-cycle model

One-by-one comparison of probabilities

- Test-retest correlations 0.5-0.7 (similar to wellbeing)
- Rounding matters
 - 20% consistent without rounding
 - Up to 70% consistent with rounding

PH model of life expectancy

- Similar relationships between socio-demographics and hazard
 - Discrepancies cancel out within groups
- Individual effects strongly correlated between surveys (corr. 0.8)
 - Variation between individuals more reliable than longitudinal variation

Sensitivity of behavior in life-cycle model

- Saving more sensitive than labor supply
- Model that accounts for rounding produces sufficiently reliable probabilities to model saving

One-by-one comparison of probabilities

- Test-retest correlations 0.5-0.7 (similar to wellbeing)
- Rounding matters
 - 20% consistent without rounding
 - Up to 70% consistent with rounding

PH model of life expectancy

- Similar relationships between socio-demographics and hazard
 - Discrepancies cancel out within groups
- Individual effects strongly correlated between surveys (corr. 0.8)
 - Variation between individuals more reliable than longitudinal variation

Sensitivity of behavior in life-cycle model

- Saving more sensitive than labor supply
- Model that accounts for rounding produces sufficiently reliable probabilities to model saving

One-by-one comparison of probabilities

- Test-retest correlations 0.5-0.7 (similar to wellbeing)
- Rounding matters
 - 20% consistent without rounding
 - Up to 70% consistent with rounding

PH model of life expectancy

- Similar relationships between socio-demographics and hazard
 - Discrepancies cancel out within groups
- Individual effects strongly correlated between surveys (corr. 0.8)
 - Variation between individuals more reliable than longitudinal variation

Sensitivity of behavior in life-cycle model

- Saving more sensitive than labor supply
- Model that accounts for rounding produces sufficiently reliable probabilities to model saving

Two sets of questions asked in the same panel in the same month

- Pension Barometer (PB):

“Please indicate on a scale from 0 to 100 how likely you think it is that you will live to age 70.”

etc. for 75, 80, 85, 90

- DNB Household Survey (DHS):

“Please indicate your answer on a scale of 0 thru 10, where 0 means ‘no chance at all’ and 10 means ‘absolutely certain’.

How likely is it that you will attain (at least) the age of 65?”

etc. for 75, 80, 85, 90, 100

- Different degrees of rounding imposed by questionnaires
- Average time between surveys: 3.3 weeks

Two sets of questions asked in the same panel in the same month

- Pension Barometer (PB):

“Please indicate on a scale from 0 to 100 how likely you think it is that you will live to age 70.”

etc. for 75, 80, 85, 90

- DNB Household Survey (DHS):

“Please indicate your answer on a scale of 0 thru 10, where 0 means ‘no chance at all’ and 10 means ‘absolutely certain’.

How likely is it that you will attain (at least) the age of 65?”

etc. for 75, 80, 85, 90, 100

- Different degrees of rounding imposed by questionnaires
- Average time between surveys: 3.3 weeks

Two sets of questions asked in the same panel in the same month

- Pension Barometer (PB):

“Please indicate on a scale from 0 to 100 how likely you think it is that you will live to age 70.”

etc. for 75, 80, 85, 90

- DNB Household Survey (DHS):

“Please indicate your answer on a scale of 0 thru 10, where 0 means ‘no chance at all’ and 10 means ‘absolutely certain’.

How likely is it that you will attain (at least) the age of 65?”

etc. for 75, 80, 85, 90, 100

- Different degrees of rounding imposed by questionnaires
- Average time between surveys: 3.3 weeks

Two sets of questions asked in the same panel in the same month

- Pension Barometer (PB):

“Please indicate on a scale from 0 to 100 how likely you think it is that you will live to age 70.”

etc. for 75, 80, 85, 90

- DNB Household Survey (DHS):

“Please indicate your answer on a scale of 0 thru 10, where 0 means ‘no chance at all’ and 10 means ‘absolutely certain’.

How likely is it that you will attain (at least) the age of 65?”

etc. for 75, 80, 85, 90, 100

- Different degrees of rounding imposed by questionnaires
- Average time between surveys: 3.3 weeks

Descriptive statistics

	N	Current age	Mean life table	PB		DHS		Rank corr.
				Mean	S. D.	Mean	S. D.	
a. Men								
Age 75	823	25-63	75.2	65.3	23.0	68.0	19.2	0.66
Age 80	1000	25-68	60.6	52.7	24.9	55.7	22.7	0.68
Age 85	294	65-73	45.7	40.9	25.8	52.5	22.9	0.58
Age 90	188	70-78	25.1	26.4	24.6	38.5	24.6	0.55
b. Women								
Age 75	690	25-63	83.6	65.8	22.5	67.5	19.0	0.56
Age 80	796	25-68	73.7	55.1	24.7	57.0	22.0	0.56
Age 85	168	65-73	61.7	44.5	26.0	54.0	23.0	0.61
Age 90	103	70-78	40.0	29.7	25.0	39.5	24.3	0.53

All probabilities expressed as percentages.

▶ More

Consistent response and rounding

Table: Rates of consistent responses to PB and DHS survival questions

	N	Exactly equal	Minimal rounding	Common rounding	General rounding
Age 75	1513	0.22	0.37	0.46	0.77
Age 80	1796	0.22	0.31	0.40	0.75
Age 85	462	0.18	0.26	0.34	0.68
Age 90	291	0.16	0.24	0.32	0.61
All combined	2,087	0.09	0.18	0.27	0.63

- Minimal rounding: PB to 1 and DHS to 10
- Common rounding: all probabilities reported by individual rounded similarly (Molinari and Manski, 2010)
- General rounding: each reported probability rounded to maximum extent (multiples of 1, 5, 10, 25, 50, 100)

▶ More

Consistent response and rounding

Table: Rates of consistent responses to PB and DHS survival questions

	N	Exactly equal	Minimal rounding	Common rounding	General rounding
Age 75	1513	0.22	0.37	0.46	0.77
Age 80	1796	0.22	0.31	0.40	0.75
Age 85	462	0.18	0.26	0.34	0.68
Age 90	291	0.16	0.24	0.32	0.61
All combined	2,087	0.09	0.18	0.27	0.63

- Minimal rounding: PB to 1 and DHS to 10
- Common rounding: all probabilities reported by individual rounded similarly (Molinari and Manski, 2010)
- General rounding: each reported probability rounded to maximum extent (multiples of 1, 5, 10, 25, 50, 100)

▶ More

Consistent response and rounding

Table: Rates of consistent responses to PB and DHS survival questions

	N	Exactly equal	Minimal rounding	Common rounding	General rounding
Age 75	1513	0.22	0.37	0.46	0.77
Age 80	1796	0.22	0.31	0.40	0.75
Age 85	462	0.18	0.26	0.34	0.68
Age 90	291	0.16	0.24	0.32	0.61
All combined	2,087	0.09	0.18	0.27	0.63

- Minimal rounding: PB to 1 and DHS to 10
- Common rounding: all probabilities reported by individual rounded similarly (Molinari and Manski, 2010)
- General rounding: each reported probability rounded to maximum extent (multiples of 1, 5, 10, 25, 50, 100)

▶ More

Consistent response and rounding

Table: Rates of consistent responses to PB and DHS survival questions

	N	Exactly equal	Minimal rounding	Common rounding	General rounding
Age 75	1513	0.22	0.37	0.46	0.77
Age 80	1796	0.22	0.31	0.40	0.75
Age 85	462	0.18	0.26	0.34	0.68
Age 90	291	0.16	0.24	0.32	0.61
All combined	2,087	0.09	0.18	0.27	0.63

- Minimal rounding: PB to 1 and DHS to 10
- Common rounding: all probabilities reported by individual rounded similarly (Molinari and Manski, 2010)
- General rounding: each reported probability rounded to maximum extent (multiples of 1, 5, 10, 25, 50, 100)

▶ More

PH model for subjective expectations

- True expectations follow Gompertz distribution

$$S_{itk}^q | a_{it} = \Pr(t \geq ta_k | t \geq a_{it}) = g(a_{it}, ta_{it}; \gamma_{it}^q, \alpha^q)$$
$$\gamma_{it}^q = \exp(\mathbf{x}'_{it} \boldsymbol{\beta}^q + \zeta_i^q + \eta_{it}^q)$$

S_{itk}^q for individual i , year t , target age k , questionnaire q

- Two steps from true S_{itk}^q to reported P_{itk}^q :
 1. Add recall error ε_{itk}^q

$$P_{itk}^{*q} = S_{itk}^q + \varepsilon_{itk}^q ; \varepsilon_{itk}^q \sim \mathcal{N}(0, \sigma_{it}^2)$$

- 2a. Model 1: P_{itk}^{*q} censored between 0 and $\min[100, P_{itk-1}^q]$
- 2b. Model 2: P_{itk}^{*q} rounded (1, 5, 10, 25, 50, 100) and censored

PH model for subjective expectations

- True expectations follow Gompertz distribution

$$S_{itk}^q | a_{it} = \Pr(t \geq ta_k | t \geq a_{it}) = g(a_{it}, ta_{it}; \gamma_{it}^q, \alpha^q)$$
$$\gamma_{it}^q = \exp(\mathbf{x}'_{it} \boldsymbol{\beta}^q + \zeta_i^q + \eta_{it}^q)$$

S_{itk}^q for individual i , year t , target age k , questionnaire q

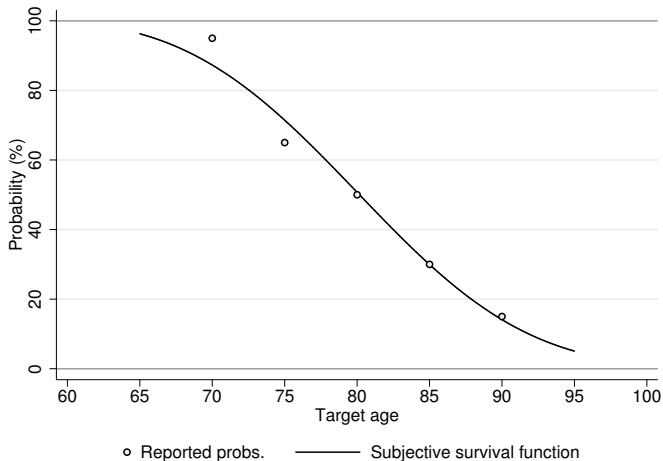
- Two steps from true S_{itk}^q to reported P_{itk}^q :

1. Add recall error ε_{itk}^q

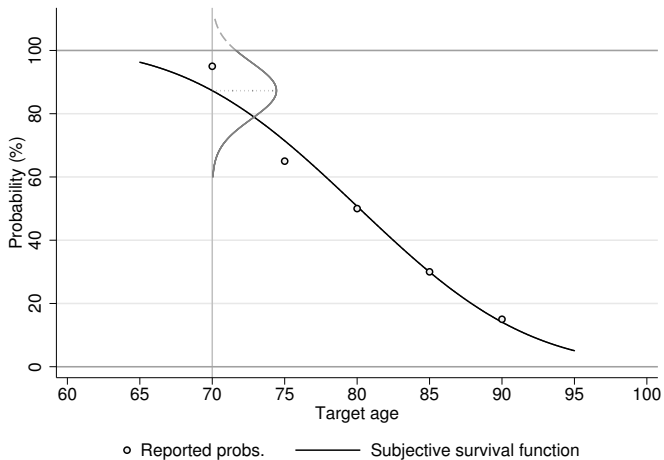
$$P_{itk}^{*q} = S_{itk}^q + \varepsilon_{itk}^q ; \varepsilon_{itk}^q \sim \mathcal{N}(0, \sigma_{it}^2)$$

- 2a. Model 1: P_{itk}^{*q} censored between 0 and $\min[100, P_{itk-1}^q]$
- 2b. Model 2: P_{itk}^{*q} rounded (1, 5, 10, 25, 50, 100) and censored

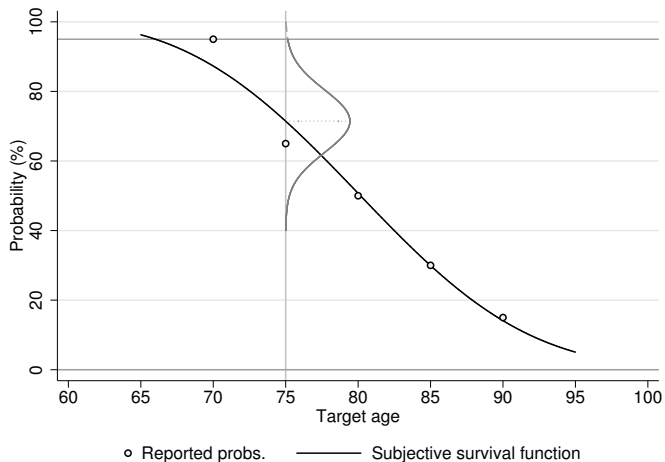
Example 1



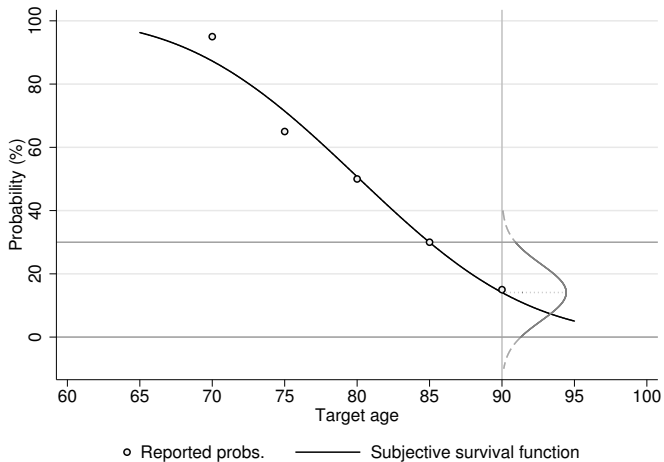
Example II



Example III

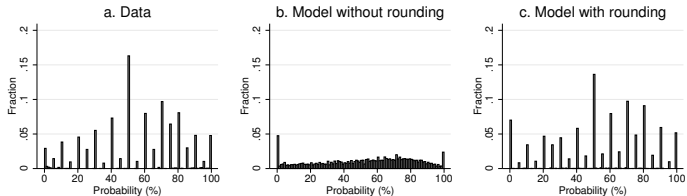


Example IV

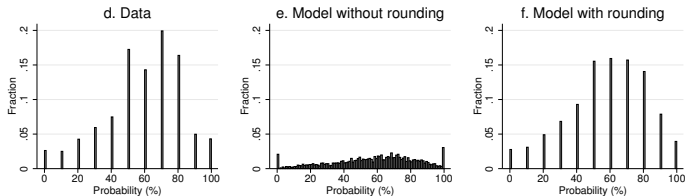


Model fit

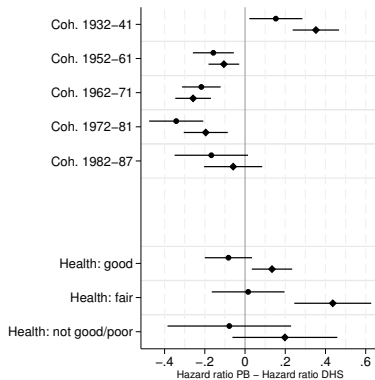
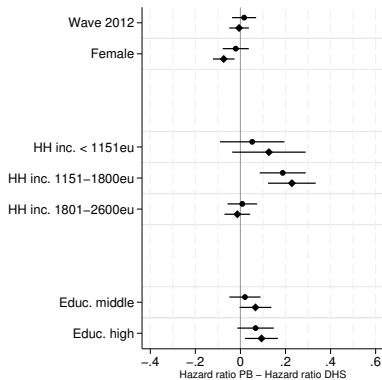
Pension Barometer (PB)



DNB Household Survey (DHS)



Differences between hazard ratios



● Model 1 – no rounding ◆ Model 2 – rounding

- Duration dependence
 - Positive: hazard of death increases with age
 - Significantly different across surveys only if we allow for rounding
- Unobserved heterogeneity
 - Individual effects account for 90% of unexplained variation
 - Individual effects strongly correlated across surveys (corr. 0.78-0.87)
 - Sequence effects smaller and less correlated (corr. 0.03-0.24)
- Rounding: average estimated rounding probabilities

Multiples of...	Pension Barometer (%)	DNB Household Survey (%)
...100	1	2
...50	5	4
...25	11	–
...10	47	95
...5	33	–
...1	4	–

Other estimates

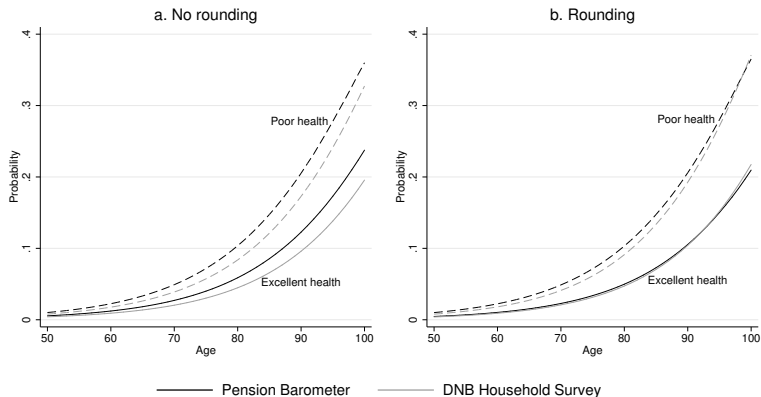
- Duration dependence
 - Positive: hazard of death increases with age
 - Significantly different across surveys only if we allow for rounding
- Unobserved heterogeneity
 - Individual effects account for 90% of unexplained variation
 - Individual effects strongly correlated across surveys (corr. 0.78-0.87)
 - Sequence effects smaller and less correlated (corr. 0.03-0.24)
- Rounding: average estimated rounding probabilities

Multiples of...	Pension Barometer (%)	DNB Household Survey (%)
...100	1	2
...50	5	4
...25	11	–
...10	47	95
...5	33	–
...1	4	–

- Duration dependence
 - Positive: hazard of death increases with age
 - Significantly different across surveys only if we allow for rounding
- Unobserved heterogeneity
 - Individual effects account for 90% of unexplained variation
 - Individual effects strongly correlated across surveys (corr. 0.78-0.87)
 - Sequence effects smaller and less correlated (corr. 0.03-0.24)
- Rounding: average estimated rounding probabilities

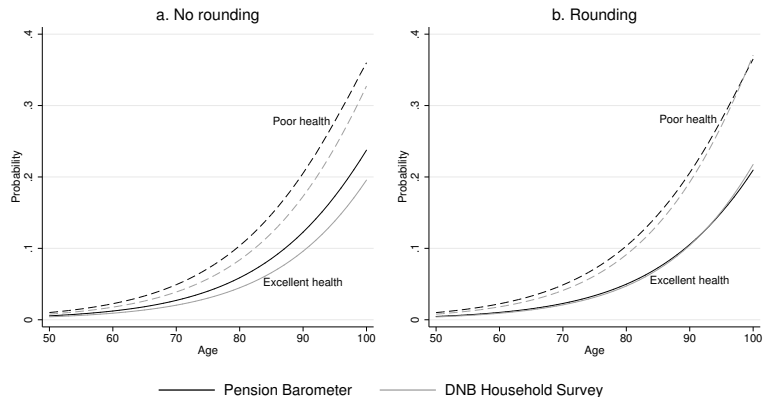
Multiples of...	Pension Barometer (%)	DNB Household Survey (%)
...100	1	2
...50	5	4
...25	11	–
...10	47	95
...5	33	–
...1	4	–

Estimated survival curves and economic behavior



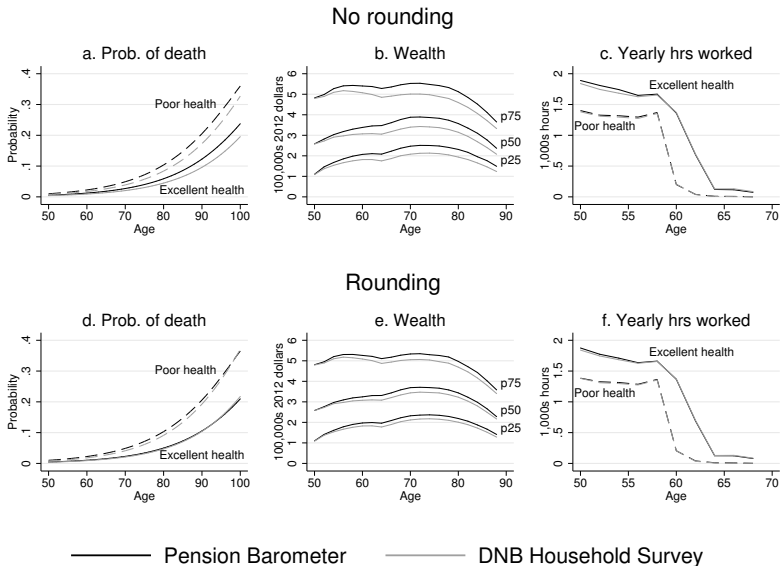
- Life-cycle model maps probabilities into saving and labor supply
- Question: how does reliability of probs. translate into simulated behavior?

Estimated survival curves and economic behavior



- Life-cycle model maps probabilities into saving and labor supply
- Question: how does reliability of probs. translate into simulated behavior?

Estimated survival curves and economic behavior



One-by-one reliability of reported probabilities

- Test-retest correlations 0.5-0.7
- Level differences for older respondents
- Rounding can reconcile many probabilities (40-70% consistent)

Reliability of estimated survival curves

- Similar associations for most covariates (not birth cohorts)
- Similar duration dependence in model without rounding
- Individual effects both important and strongly correlated across surveys
 - Between-individual variation more reliable than longitudinal variation

Conclusion

One-by-one reliability of reported probabilities

- Test-retest correlations 0.5-0.7
- Level differences for older respondents
- Rounding can reconcile many probabilities (40-70% consistent)

Reliability of estimated survival curves

- Similar associations for most covariates (not birth cohorts)
- Similar duration dependence in model without rounding
- Individual effects both important and strongly correlated across surveys
 - Between-individual variation more reliable than longitudinal variation

Conclusion (contd.)

Simulated survival expectations

- More reliable across surveys for model with rounding (differences cancel out)
- Saving is sensitive to expectations
- Model with rounding gives reliable survival curves
 - Difference in simulated wealth below 7%

Consistent response and age

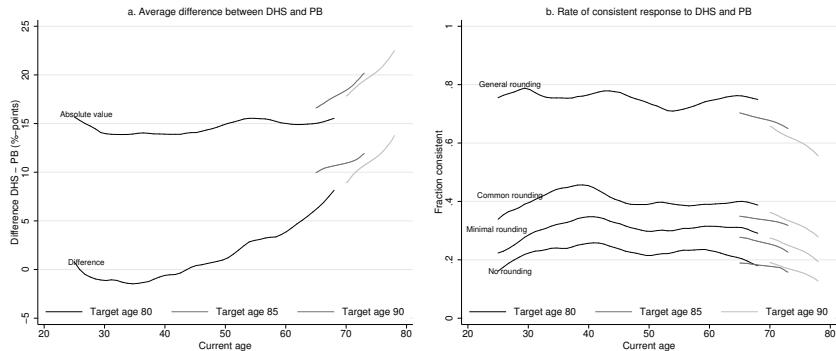
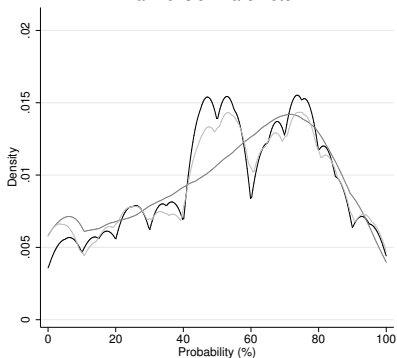


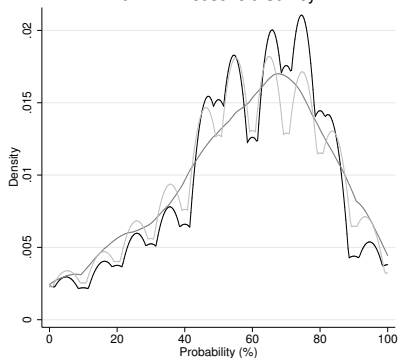
Figure: Kernel regressions of absolute and relative differences between probs. (left) and of rates of consistent responses (right) on current age

Model fit – kernel densities

a. Pension Barometer



b. DNB Household Survey



— Data — Model, no rounding - - - Model, rounding

▶ Back

