Test-Retest Reliability of Subjective Survival Expectations

Jochem de Bresser
Tilburg University and Netspar
January 2017
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)
  - Affected by framing (Payne et al., 2012)
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)
  - Affected by framing (Payne et al., 2012)
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)
  - Affected by framing (Payne et al., 2012)
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)
  - Affected by framing (Payne et al., 2012)
Research question

**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
Research question

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
Research question

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. With analogous methods used for other subjective data
2. At different levels of aggregation
3. In terms of behavior in life-cycle model
Contribution of the paper

Evaluate test-retest reliability...

1. With analogous methods used for other subjective data
2. At different levels of aggregation
3. In terms of behavior in life-cycle model
Contributions of the paper

Evaluate test-retest reliability...

1. With analogous methods used for other subjective data
2. At different levels of aggregation
3. In terms of behavior in life-cycle model
Findings

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
Findings

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
Findings (contd.)

Sensitivity of behavior in life-cycle model

- Saving more sensitive than labor supply
- Model that accounts for rounding halves wealth difference
- Heterogeneity in expectations larger than average difference between surveys
Two sets of questions in the same month (avg. 3.3 weeks apart)

- **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
Data

Two sets of questions in the **same month** (avg. 3.3 weeks apart)

- **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
## Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Current age</th>
<th>Mean life table</th>
<th>PB</th>
<th>DHS</th>
<th>Rank corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S. D.</td>
<td>Mean</td>
<td>S. D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Men</td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 75</td>
<td>823</td>
<td>25-63</td>
<td>75.2</td>
<td>65.3</td>
<td>23.0</td>
<td>68.0</td>
</tr>
<tr>
<td>Age 80</td>
<td>1000</td>
<td>25-68</td>
<td>60.6</td>
<td>52.7</td>
<td>24.9</td>
<td>55.7</td>
</tr>
<tr>
<td>Age 85</td>
<td>294</td>
<td>65-73</td>
<td>45.7</td>
<td>40.9</td>
<td>25.8</td>
<td>52.5</td>
</tr>
<tr>
<td>Age 90</td>
<td>188</td>
<td>70-78</td>
<td>25.1</td>
<td>26.4</td>
<td>24.6</td>
<td>38.5</td>
</tr>
<tr>
<td>b. Women</td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 75</td>
<td>690</td>
<td>25-63</td>
<td>83.6</td>
<td>65.8</td>
<td>22.5</td>
<td>67.5</td>
</tr>
<tr>
<td>Age 80</td>
<td>796</td>
<td>25-68</td>
<td>73.7</td>
<td>55.1</td>
<td>24.7</td>
<td>57.0</td>
</tr>
<tr>
<td>Age 85</td>
<td>168</td>
<td>65-73</td>
<td>61.7</td>
<td>44.5</td>
<td>26.0</td>
<td>54.0</td>
</tr>
<tr>
<td>Age 90</td>
<td>103</td>
<td>70-78</td>
<td>40.0</td>
<td>29.7</td>
<td>25.0</td>
<td>39.5</td>
</tr>
</tbody>
</table>

All probabilities expressed as percentages.
### Table: Rates of consistent responses to PB and DHS survival questions

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

- **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)
Consistent response and rounding

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

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>Exactly equal</th>
<th>Minimal rounding</th>
<th>Common rounding</th>
<th>General rounding</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>1513</td>
<td>0.22</td>
<td>0.37</td>
<td>0.46</td>
<td>0.77</td>
</tr>
<tr>
<td>80</td>
<td>1796</td>
<td>0.22</td>
<td>0.31</td>
<td>0.40</td>
<td>0.75</td>
</tr>
<tr>
<td>85</td>
<td>462</td>
<td>0.18</td>
<td>0.26</td>
<td>0.34</td>
<td>0.68</td>
</tr>
<tr>
<td>90</td>
<td>291</td>
<td>0.16</td>
<td>0.24</td>
<td>0.32</td>
<td>0.61</td>
</tr>
<tr>
<td>All</td>
<td>2,087</td>
<td>0.09</td>
<td>0.18</td>
<td>0.27</td>
<td>0.63</td>
</tr>
</tbody>
</table>

- **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)
### Table: Rates of consistent responses to PB and DHS survival questions

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

- **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)
## Consistent response and rounding

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

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Exactly equal</th>
<th>Minimal rounding</th>
<th>Common rounding</th>
<th>General rounding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age 75</strong></td>
<td>1513</td>
<td>0.22</td>
<td>0.37</td>
<td>0.46</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>Age 80</strong></td>
<td>1796</td>
<td>0.22</td>
<td>0.31</td>
<td>0.40</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>Age 85</strong></td>
<td>462</td>
<td>0.18</td>
<td>0.26</td>
<td>0.34</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Age 90</strong></td>
<td>291</td>
<td>0.16</td>
<td>0.24</td>
<td>0.32</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>All combined</strong></td>
<td>2,087</td>
<td>0.09</td>
<td>0.18</td>
<td>0.27</td>
<td>0.63</td>
</tr>
</tbody>
</table>

- **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)
PH model for subjective expectations

- True expectations follow Gompertz distribution

\[
S_{itk}^{q} | a_{it} = \Pr (t \geq t_{ak} | t \geq a_{it}) = g(a_{it}, t_{ait}, \gamma_{it}^{q}, \alpha^{q})
\]

\[
\gamma_{it}^{q} = \exp (x_{it}^{'} \beta^{q} + \xi_{i} + \eta_{it}^{q})
\]

\(S_{itk}^{q}\) 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 \(\varepsilon_{itk}^{q}\)

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

  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

\[ S^q_{itk} \mid a_{it} = \Pr (t \geq ta_k \mid t \geq a_{it}) = g (a_{it}, ta_{it}; \gamma^q_{it}, \alpha^q) \]

\[ \gamma^q_{it} = \exp (x_{it}' \beta^q + \zeta^q_i + \eta^q_{it}) \]

\( S^q_{itk} \) for individual \( i \), year \( t \), target age \( k \), questionnaire \( q \)

- Two steps from true \( S^q_{itk} \) to reported \( P^q_{itk} \):
  1. Add recall error \( \varepsilon^q_{itk} \)

\[ P^*_{itk} = S^q_{itk} + \varepsilon^q_{itk} ; \varepsilon^q_{itk} \sim \mathcal{N} \left( 0, \sigma^2_{it} \right) \]

  2a. Model 1: \( P^*_{itk} \) censored between 0 and min \([100, P^q_{itk-1}]\)
  2b. Model 2: \( P^*_{itk} \) rounded \((1, 5, 10, 25, 50, 100)\) and censored
Illustration of the likelihood

a. Model without rounding

b. Model with rounding

Reliability Subjective Survival
Differences between hazard ratios

Wave 2012
Female
HH inc. < 1151eu
HH inc. 1151−1800eu
HH inc. 1801−2600eu
Educ. middle
Educ. high

Health: good
Health: fair
Health: not good/poor

-0.4 -0.2 0 0.2 0.4 0.6
Hazard ratio PB − Hazard ratio DHS

Differences between hazard ratios
Model 1 − no rounding Model 2 − rounding

Model 1 – no rounding ○ Model 2 – rounding
One-by-one vs. aggregate reliability

Shaded areas are 90% confidence bands.

95% CIs indicated as vertical bars.

Based on estimates from model without rounding.
Estimated survival curves and economic behavior

- Life-cycle model maps probabilities into saving and labor supply
- 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

How does reliability of probs. translate into simulated behavior?
Estimated survival curves and economic behavior

No rounding

- a. Prob. of death
- b. Wealth
- c. Yearly hrs worked

Rounding

- d. Prob. of death
- e. Wealth
- f. Yearly hrs worked

---

Pension Barometer  DNB Household Survey

Reliability Subjective Survival
Variation in survival curves and economic behavior

Pension Barometer (PB)

a. Prob. of death

b. Wealth

c. Yearly hrs worked

DNB Household Survey (DHS)

d. Prob. of death

e. Wealth

f. Yearly hrs worked

All simulations are based on estimates from model without rounding.
Conclusion

One-by-one reliability of reported probabilities

- Test-retest correlations 0.5-0.7
- Rounding can reconcile many probabilities (40-70% consistent)

Reliability of estimated survival curves

- Similar associations for most covariates (not birth cohorts)
- 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
- Rounding can reconcile many probabilities (40-70% consistent)

Reliability of estimated survival curves
- Similar associations for most covariates (not birth cohorts)
- Individual effects both important and strongly correlated across surveys
  - Between-individual variation more reliable than longitudinal variation
Simulated survival expectations

- Accounting for rounding improves reliability
- Saving is sensitive to expectations
- Model with rounding gives reliable survival curves
  - Difference in simulated wealth below 7%
- Unobserved heterogeneity larger than average difference between surveys
Distribution of PB conditional on DHS

Age 75

Age 80

Age 85

Age 90

- Median
- IQR

Reliability Subjective Survival
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

Pension Barometer (PB)

a. Data

b. Model without rounding
c. Model with rounding

DNB Household Survey (DHS)
d. Data
e. Model without rounding
f. Model with rounding
Model fit – kernel densities

- a. Pension Barometer
- b. DNB Household Survey

Density vs Probability (%)

- Data
- Model, no rounding
- Model, rounding

Reliability Subjective Survival
Estimates – hazard ratios from PB and DHS (I)

Model 1 – no rounding

Wave 2012
Female
HH inc. < 1151eu
HH inc. 1151–1800eu
HH inc. 1801–2600eu
Educ. middle
Educ. high

Hazard ratio

Model 2 – rounding

PB  ■  DHS  □
Estimates – hazard ratios from PB and DHS (II)

Model 1 – no rounding

Model 2 – rounding


Health: good  |  Health: fair  |  Health: not good/poor

PB  |  DHS

Reliability Subjective Survival