The cognitive and economic impact of social activities in older age: evidence from 17 European countries

by

Dimitris Christelis
CSEF, CFS, and CEPAR

Loretti Isabella Dobrescu
University of New South Wales and CEPAR

Outline of the talk

- Economic importance of cognition in older age
- Reasons why being socially active can affect cognition
- Estimation methodology
- Data
- Results
- Discussion

Economic importance of cognition in older age

- Can lead to positive economic and other outcomes
- Strong positive association with:
  - financial literacy (Delavande, Rohwedder and Willis 2008)
  - wealth and risky portfolio choice (Lee and Willis, 2001; McArdle, Smith and Willis, 2009 Smith, McArdle and Willis, 2010; Christelis, Jappelli and Padula, 2010)
  - consumption smoothing and life-satisfaction in retirement (Banks, O'Dea and Oldfield, 2010)
  - health outcomes and behavior (Cutler and Lleras-Muney, 2010; Plassman et al. (2010) and Deary et al., 2009)

Stylized fact: strong positive association between cognition and social activities

- Lack of social networks in which social activities are performed can lead to loneliness, which can lead to depression, which can lead to cognitive decline

- Social activities can provide a meaningful social role and a sense of purpose in old age, and thus could have direct neurohormonal influences on the brain, including the reduction of the stress response

Stylized fact: strong positive association between cognition and social activities (cont.)

- As social activities often involve the challenge of effective communication and participation in complex interpersonal exchanges, they are likely to inhibit cognitive decline in the elderly

- Social activity might also require a degree of physical activity above and beyond regular exercise and walking, which could enhance physical health, which can also delay cognitive decline

Challenges for identification

- Unobserved individual characteristics (e.g. intellectual curiosity, zest for life, ease of relating, listening to and learning from others) affect both social activities and cognition

- Time varying circumstances like the death of a partner, divorce, a new partner or health problems could affect social activities and cognition in the same direction

- Simultaneity: cognition could also affect social activities

Observable and unobservable magnitudes

- Outcome Y, observed treatment w, generic value d. Two possible values in our case: 0, 1.

- Average potential outcome for treatment t

$$E[Y(d)] = E[Y(d)|w = d]\, P(w = d) + E[Y(d)|w \neq d]P(\mathrm{w} \neq d)$$

- We do not observe $E[Y(d)|w \neq d]$, i.e. the counterfactual

- We observe everything else

Observable and unobservable magnitudes (cont.)

- $E[Y(d)|w = d]$=$E[Y(d)|w \neq d] \rightarrow$
  the treatment is exogenous $\rightarrow$
  $E[Y(d)] = E[Y(d)|w = d]$, which is observed in the data

- Average treatment effect for w= t, u with $u > t$
  $$E[Y(u)] - E[Y(t)]$$

- Interpretation: give everybody treatment u and treatment u and take the average difference, taking all other observables and unobservables as given

Observable and unobservable magnitudes (cont.)

- Under exogenous treatment selection,

$$ATE(u,t) = E[Y(u)] - E[Y(t)] = E[Y(u)|w = u] - E[Y(t)|w = t]$$

- In other words the ATE is equal to the difference in observed means (or dummy variable coefficient in a OLS regression)

Observable and unobservable magnitudes (cont.)

- In most cases, however,

$$E[Y(d)|w = d] \neq E[Y(d)|w \neq d]$$

- Treatment is not randomly assigned, units choosing different treatments are not homogeneous → endogeneity

- Need to replace $E[Y(d)|w \neq d]$ with something observable

Bounds – Partial Identification

- We replace $E[Y(d)|w \neq d]$with observable magnitudes $M_L$and M_U such that

$$M_L \leq E[Y(d)|w \neq d] \leq M_U$$

- Then we calculate lower and upper bounds $LB[d]$ and $UB[d]$ for $E[Y(d)]$. We thus obtain partial identification (Manski 1989, 1997; Manski and Pepper, 2000)

$$LB[t] = E[Y(d)|w = d]\, P(w = d) + M_L P(w \neq d) \leq E[Y(d)] \leq E[Y(d)|w = d]P(w = d) + M_U P(w \neq d) = UB[d]$$

Observable and unobservable magnitudes (cont.)

- Suppose we have two treatments u and t with $u > t$. We then have

$$LB[t] \leq E[Y(t)] \leq UB[t]$$
$$LB[u] \leq E[Y(u)] \leq UB[u]$$

- The average treatment effect $ATE(u,t) = E[Y(u)] - E[Y(t)]$ can then be bounded.

Observable and unobservable magnitudes (cont.)

- Suppose we have 2 treatments u and t with $u > t$. We then have

$$LB[u] - UB[t]$$
$$\leq$$
$$ATE(u,t) = E[Y(u)] - E[Y(t)]$$
$$\leq$$
$$UB[u] - LB[t]$$

Numerical example

- Suppose

$$2 \leq E[Y(t)] \leq 5$$
$$6 \leq E[Y(u)] \leq 10$$

- Then

$$1 \leq ATE(u,t) = E[Y(u)] - E[Y(t)] \leq 8$$

No assumptions bounds

- If in $M_L \leq E[Y(d)|w \neq d] \leq M_U$ we put $M_L = Y_{min}$ and $M_U = Y_{max}$ Then we have

$$E[Y(d)|w = d]P(w = d) + Y_{min}P(w \neq d) \leq E[Y(d)] \leq E[Y(d)|w = d]P(w = d) + Y_{max}P(w \neq d)$$

- These bounds tend to be very wide, which is the price for making no assumptions. Narrower bounds require additional assumptions.

Monotone Treatment Response (MTR)

- If treatment $u > t \rightarrow E[Y(u)|w = d] \geq E[Y(t)|w = d], \forall d$
- In our context, being socially active has a weakly positive effect on cognition on average
- Use the assumption to narrow identification range

Monotone Treatment Response (cont.)

- $E[y(1)] = E[y(1)|w = 0]P(w = 0) + E[y(1)|w = 1]P(w = 1)$
- $MTR \rightarrow E[y(1)|w = 0] \geq E[y(0)|w = 0]$
- Get an lower bound on $E[y(1)]$ by replacing the unobservable $E[y(1)|w = 0]$ with the observable $E[y(0)|w = 0]$

Monotone Treatment Response (cont.)

- Since $E[Y(0)|w = 0] \geq Y_{min}$ the lower bound becomes larger and the width of the identification range of $E[Y(0)]$ becomes smaller

Monotone Treatment Selection (MTS)

- If treatment $u > t \rightarrow E[y(d)|w = u] \geq E[y(d)|w = t], \forall d$

- More socially active individuals have characteristics (intelligence, curiosity, resourcefulness) which lead to higher cognition for any level of social activity actually observed

MTR + MTS are testable

- Manski and Pepper (2000): $u > t \rightarrow E[y(u)|w = u] \geq E[y(t)|w = t]$
- Hence MMTR + MTS implies weak monotonicity in the observed mean outcomes. This can be tested
- In our data MTR+MTS are not refuted

Instruments

- One can use instruments to further narrow the bounds

- Instruments are useful if one wants the lower bound of the ATE to become larger than zero. With only MTR+MTS this is not possible (Manski, 2000).

Exogenous instruments

- An instrument z is exogenous if

$$E[y(d)|z = m] = E[y(d)], \forall m$$

- Hence one can increase the lower bound by maximizing over all values of the instrument, and decrease the upper bound by minimizing over all values of the instrument, i.e.

$$\max_m LB[d|z = m] \leq E[y(d)] \leq \min_m UB[d|z = m]$$

Monotone instruments

- Exogenous instrument are hard to find
- A weaker instrument is a monotone one. An instrument z is monotone if

$$l \leq o \rightarrow E[y(d)|z = l] \leq E[y(d)|z = o]$$

$$\sum_m P(z = m) \left(\max_{m_1 \leq m} LB[d|z = m_1] \right)$$

$$\leq \sum_m P(z = m) E[y(d)|z = m] = E[y(d)] \leq$$

$$\sum_m P(z = m) \left(\min_{m_2 \geq m} UB[d|z = m_2] \right)$$

Role of instruments

- They help identify the ATE and not the local average treatment effect (LATE), which is what conventional IV estimation identifies under heterogeneity of the treatment

- They can, however, lead to the lower bound being greater than the upper bound, either in the point estimate or the bootstrap

Role of instruments (cont.)

- If the lower bound greater than the upper bound then not all assumptions are compatible with the data
- Up to the researcher to judge which combination of assumptions is more plausible
- Manski (2003) suggests the possibility to apply the instrument to part of the estimation, i.e. to only some $E[y(d)]$

Role of instruments (cont.)

- In our case, bounds cross for $E[y(0)]$
- The identification region for $E[y(0)]$ is about 4 times narrower than those of $E[y(1)]$
- Narrow region due to small $P(w = 1) \cong 21\%$
- We use MTR+MTS for $E[y(0)]$ and MTR+MTS+MIV for $E[y(1)]$

Data – Survey of Health, Ageing and Retirement in Europe (SHARE)

- European study of the 50+. 1st wave in 2004-2005, 2nd wave in 2006-7, 4th wave in 2010-11. Modeled after US and English surveys
- 15 countries in all: Sweden, Denmark, Germany, the Netherlands, Belgium, France, Switzerland, Austria, Italy, Spain, Greece, Czech Republic, Poland, Slovenia, Hungary, Portugal, Estonia
- About 94,000 obs (numeracy), 113,000 (other scores)
- Modules: demographics, children, physical and mental health, cognition, income, assets, social activities, expectations

SHARE (cont.)

- Numeracy: 4 questions asked with possible answers suggested

- How many people out of 1,000 would be expected to get the disease if the chance of getting a disease is 10%
- What is the sale cost of a sofa, given the initial price and a 50 % discount
- What is the initial price of a car if two-thirds of what it costs new is 6,000 euro
- What is the final balance of a savings account that initially hold 2,000 euro, at 10% interest after 2 years

SHARE (cont.)

- Fluency: number of animals respondents can enumerate in one minute
- Immediate recall: respondents are read ten words and are asked to repeat them immediately after. The score is the number of correctly recalled words
- Delayed recall: respondents are asked to recall again the ten words after the fluency and numeracy tests are conducted. The score is again the number of correctly recalled words

SHARE (cont.)

- Respondents are asked whether in the last month (waves 1 and 2) or year (wave 4) they
  - engaged in voluntary or charity work
  - participated in an educational or training course
  - went to a sport, social or other kind of club
  - participated in a in a political or community-related organization
- Respondents are asked whether they engaged in these activities daily, almost every week, or less often
- We use as our main measure of social activities whether they engaged in any of the four activities at least once a week

Instrument

- Monotone instrument used: whether one smokes
- Negative association between smoking and cognition well established empirically. Smoking associated with cerebral lesions, decreased oxygenation, increased cardiovascular risk, Alzheimer's disease
- Hypothesis behind monotone instrument untestable, as it refers to the unobserved $E[y(d)]$. However, monotonicity weaker than exogeneity, which is also untestable

Descriptive Statistics

Numeracy	2.19	1.18	0	4	93,756
Fluency	18.13	7.65	0	100	113,015
Immediate Recall	4.84	1.91	0	10	113,616
Delayed Recall	3.39	2.09	0	10	113,627
Social activities	0.21	0.41	0	1	117,819

Results

- Results will be presented by adding assumptions progressively
- Identifying power of each assumption becomes clear
- 95% confidence intervals for the ATE computed as in Imbens and Manski (2004)

Results – Numeracy, Fluency

Method	Numeracy				Fluency			
	Lower Bound	Upper Bound	Low 95% CI	High 95% CI	Lower Bound	Upper Bound	Low 95% CI	High 95% CI
Panel A. Whole Sample								
Exogenous Treatment Selection	0.564		0.533	0.595	4.191		3.990	4.393
No Assumptions	-1.939	2.061	-1.952	2.073	-30.325	69.675	-30.629	69.978
MTR	0.000	2.061	0.000	2.073	0.000	69.675	0.000	69.978
MTR + MTS	0.000	0.564	0.000	0.593	0.000	4.191	0.000	4.375
MTR + MTS + MIV	0.164	0.566	0.140	0.596	0.898	4.200	0.717	4.386
Number of observations	92,582				112,164			

Other Results

- Similar results for immediate and delayed recall (lower bound about 30% of 1 std
- Results stronger for females than for males (esp. for numeracy and fluency)
- Results consistent across European regions (Northern, Middle, Southern and Eastern Europe)
- Results robust to alternative definition of cognition (binary, above/below median)

Economic impact of social activity

- Does social activity impact, through increased cognition, economic welfare?
- Examine the effect of cognition on household net financial assets, household net worth
- Use binary indicator for cognition being above/below the median

Economic impact of social activity - results

Method	Net Worth				Net Financial Assets			
	Lower Bound	Upper Bound	Low 95% CI	High 95% CI	Lower Bound	Upper Bound	Low 95% CI	High 95% CI
Panel A. Numeracy								
Exogenous Treatment Selection	1.402		1.281	1.524	2.144		1.965	2.322
No Assumptions	-17.087	17.655	-17.232	17.800	-15.538	17.159	-15.651	17.272
MTR	0.000	17.655	0.000	17.800	0.000	17.159	0.000	17.272
MTR + MTS	0.000	1.402	0.000	1.508	0.000	2.144	0.000	2.298
MTR + MTS + MIV	0.000	0.970	0.000	1.080	0.002	1.827	0.000	2.005
Number of observations	61,275				61,275			
Panel B. Fluency								
Exogenous Treatment Selection	1.188		1.057	1.319	1.882		1.707	2.058
No Assumptions	-14.613	20.129	-14.748	20.264	-14.065	18.632	-14.176	18.744
MTR	0.000	20.129	0.000	20.264	0.000	18.632	0.000	18.744
MTR + MTS	0.000	1.188	0.000	1.305	0.000	1.882	0.000	2.035
MTR + MTS + MIV	0.000	0.654	0.000	0.763	0.000	1.461	0.000	1.624
Number of observations	73,652				73,652			

Economic impact of social activity – cont.

- Combining
  - the lower bound estimates of the effect of social activity on cognition
  - A 5% elasticity of household net worth and net financial assets with respect to cognition
  - One gets a 0.195% increase due to numeracy, corresponding effects of fluency and immediate and delayed recall are equal to 0.26%, 0.34%, and 0.31%. Total effect is about 1.11%.

Discussion

- Advantages of the method
  - Non-parametric: no assumptions about functional form
  - Does not involve any other control variables. No need to worry about their endogeneity, interactions or intermediate outcomes
  - Allows for complete heterogeneity of the treatment effect
  - No assumptions about the error
  - Deals with all forms of endogeneity of the treatment
  - Allows the use of instruments if available, and they need not be exogenous – being monotonic can suffice
  - Estimates the average treatment effect and not the local average treatment effect as in most IV estimation contexts

Discussion (cont.)

- Advantages of the method
  - Uses relatively mild assumptions, some of which are testable (MTR + MTS)
  - Is completely transparent about how each assumption affects the outcome. It's up to the modeler/reader to assess the plausibility of the assumptions
  - Makes no assumption about behavior or expectations
  - Can be used in difficult estimation circumstances, e.g. cross-sectional data with no exogenous instruments
  - The ATE can be computed for quantiles as well
  - Analysis can be done within subsamples of interest
  - No need to use complicated empirical models (no need to maximize likelihoods, worry about numerical issues like convergence, local maxima etc.)

Discussion (cont.)

- Disadvantages of the method
  - Sometimes the width of the range defined by the bounds can be large – but one should be wary of unwarranted certainty
  - MTR and/or MTS need not hold. Judgment needed.
  - Not easy to bound the ATE away from zero – importance depends on the context. Instruments are needed
  - As in conventional estimation, assumptions about instrument are untestable. Judgment needed.
  - More complicated than OLS, IV, panel data methods, although not terribly so
  - Estimation is done in subsamples defined by the treatment/instrument combination – large samples might be needed
  - Continuous monotone instruments need to be discretized

Summary

- Used non-parametric methods based on bounds in order to assess the effect of being socially active on the cognition of older individuals
- The lower bound of the effect is about 15-30% of a standard deviation – smaller than the one assuming exogeneity. Upper bound about 50% of a standard deviation
- Effect is much and more uncertain than the one estimated under exogeneity

Summary (cont.)

- Results are much stronger for females than for males
- Through increased cognition, social activities have an economically significant impact on economic welfare