**Description of the files "pdem_withvarnames.dta" and "pdem_withvarnames.sas7bdat"**

These Stata 11 SE and SAS version 9 files contain the predicted probabilities used in the Hurd et al.[1] (2013) study examining the monetary costs of dementia. These datasets contain person-year observations for individuals for whom we could compute predicted probabilities of dementia, and contain the predicted probabilities based on the prediction model described in Hurd et al. (2013). These calculations were done for HRS respondents age 70+ in the 1998-2006 HRS waves. The probabilities of dementia refer to the year following a particular HRS interview (since the prediction model is based on clinical diagnoses of dementia that were made about one year after the preceding HRS interview).[2] For further details on how these probabilities were calculated, please see the online appendix[3] to this paper.

Variables:

- hhidpn: Concatenation of HRS household ID number (hhid) and person number (pn). This variable uniquely identifies HRS respondents
- prob_dementia: This is the predicted probability of dementia for an HRS respondent.
- prediction_year: This is the year for which the prediction was made.
- hhid: Character version of HRS household ID number (hhid). Together with person number (pn), these two variables uniquely identify HRS respondents
- pn: Character version of HRS person number (pn). Together with household ID number (hhid), these two variables uniquely identify HRS respondents

---

[1] Hurd, Michael, Paco Martorell, Kathleen Mullen, Adeline Delavande, and Kenneth Langa. 2013. "The Monetary Costs of Dementia in the United States" (with Michael Hurd, Kenneth Langa, Kathleen Mullen, and Adeline Delavande). *New England Journal of Medicine*. 368:14, 1326-1334.

[2] Note that the cost analyses described in Hurd et al. (2013) use HRS respondents from the 2000-2008 waves. Since the questions on health care utilization and cost generally refer to the two-year period prior to the current interview, the predicted dementia probability associated with the cost responses from a given HRS wave were based on the cognitive scores (and other predictors) from the prior HRS wave. For example, in the Hurd et al. (2013) cost analyses, survey responses on health care spending and utilization from the 2008 HRS interview were related to the predicted probability that this individual had dementia in 2007 (approximately the mid-point of the reference period for the questions on cost and utilization in the 2008 HRS interview). This prediction was generated from information from the 2006 HRS interview.

[3] This can be found at:
http://www.nejm.org/doi/suppl/10.1056/NEJMsa1204629/suppl_file/nejmsa1204629_appendix.pdf