SUPPLEMENTARY INFORMATION FOR GLYMOUR ET AL.,: DOES CHILDHOOD SCHOOLING AFFECT OLD AGE MEMORY FUNCTION?

JECH/2006/059469

APPENDIX 1: Assumptions for analyzing natural experiments as instrumental variables.

APPENDIX 2: Supplementary sensitivity analyses.

APPENDIX 1: Assumptions for analyzing natural experiments as instrumental variables.

Ideal natural experiments follow a design that mimics randomized trials: exposure to the treatment of interest is influenced by an event or exposure that is otherwise unrelated to the outcome of interest. In randomized trials, this event is the randomized treatment assignment. In the absence of randomized trials, natural experiments, such as policy changes, administrative aspects of health services arrangements, or many other exogenous factors are used as "instrumental variables" that effectively randomize individuals to exposure. The causal structure for a randomized trial and for a natural experiment is shown with a directed acyclic graph (DAG) in appendix figure 1a. To formalize the assumptions for a valid instrument in graphical terms[1]: Z is a valid instrument to estimate the effect of X on Y if in the causal DAG including Z, X, and Y (i) there is an open path from Z to X, and (ii) every open path from Z to Y includes an arrow pointing into X. Assumption (i) simply requires that the instrument and the exposure be associated, typically because the instrument affects exposure. Conceptually, the violations of assumption (ii) arise if either the putative instrument directly affects Y (via a mechanism not mediated by X, as in appendix figure 1b) or if Z and Y share a common cause (as in appendix figure 1c). The power of an instrumental variables approach is that assumption (ii) is not violated even if X and Y share an unmeasured common cause: i.e. instrumental variables are an approach to potentially circumvent unmeasured confounding. These assumptions largely correspond to the assumptions for IV used in econometrics or stated in terms of counterfactuals ([2][3], for an explicit comparison, see [1], page 247).

In a randomized trial, the intent-to-treat estimate is generally attenuated relative to the magnitude of the effect of exposure on outcome. The attenuation is in proportion to the non-adherence of trial participants. The same attenuation applies to instrumental variables derived from natural experiments, but the attenuation may be more severe because the association between the instrument and the exposure is much weaker. IV analyses, such as the two-stage IV approach we use here, scale up the Z-Y association in proportion to the Z-X association, i.e. they de-attenuate the intent-to-treat estimate. [4][5][6][7]. This approach provides a consistent estimate for the effect of X on Y for those individuals among whom the value of X was influenced by the value of Z if Z is a valid instrument and the instrument has a monotonic effect on the exposure (monotonicity defined below). When the exposure is not binary, the IV estimate is a weighted average of the causal effect of each level of exposure among the subpopulation of individuals whose exposure was affected by the instrument. [4]

The monotonicity assumption requires that the effect of the instrument on exposure be in the same direction for everyone in the population: if increases in the CSL influence some people in the population to achieve extra education, there must be nobody in the population for whom increases in CSLs would induce that person to get *less* education than they otherwise would attain. [8][9] Although it is possible that some people in the population dropped out earlier in response to longer requirements, we doubt this is a major source of bias in the current analysis. A much more controversial assumption relates to the validity of the instrument: might CSL changes correlate with differences in memory test scores for any reason other than receipt of extra schooling? The IV effect estimate is scaled up in inverse proportion to the strength of the association

between the instrument and the exposure. In the case of CSLs and education, this is a fairly weak association, so even small violations of the assumptions for valid instruments can introduce a large bias. [10]

APPENDIX 2: Supplementary sensitivity analyses.

The primary analyses we present do not apply the sampling weights to account for the complex, multi-stage sampling design of HRS. This could compromise generalizability of our estimates if education effects are larger for the types of people who are over-represented in the HRS sample. HRS over-sampled people living in predominantly black census blocks, Hispanics, and people living in Florida. The weights provided with HRS cannot be applied directly because we have combined several waves of outcome data (in order to increase the sample size and the reliability of the outcome measures), thus there is no clearly defined population to which we are generalizing. In supplementary analyses, we applied a synthetic weight calculated from the 5 percent sample of the 1980 U.S. Census, based on birth year, sex, state of birth, and educational attainment (5 levels). Weighted estimates for Memory were not statistically distinguishable from the unweighted results. For example, the weighted IV estimate of the effect of education on Memory was 0.21 (0.00, 0.42) using the model 3 covariates and 0.39 (0.09, 0.70) using the model 4 covariates.

To assess whether results might be statistically significant only because we ignore uncertainty in the first stage of our SSIV analyses, we also estimated conventional 2SLS models in which both stages were calculated using the HRS. The 2SLS IV estimate for Memory is 0.15 (CI: 0.01, 0.30) when adjusted for Model 3 covariates and 0.24 (CI: 0.00, 0.47) when adjusted for Model 4 covariates (based on robust variance estimates adjusted for clustering on state of birth).

To assess the sensitivity of our results to missing data, we conducted selected analyses after inverse probability weighting (IPW) on the probability of observing the

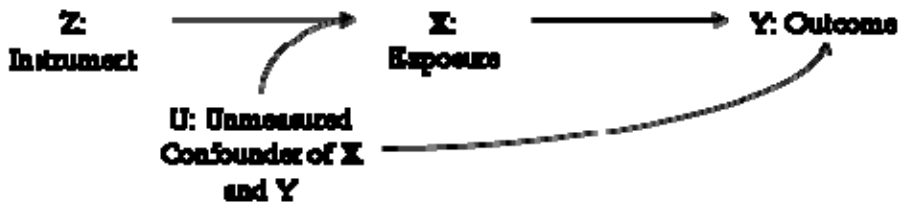
outcome variable.[11] For each model, the IPWs were calculated using the covariates in the model and indicators of parental SES (mother's education greater than 8 years, father's education greater than 8 years, father's occupation), report of heart disease, diabetes, and psychiatric condition at first interview. We used stabilized weights, in which the predicted probability of observing the outcome based on the covariates in the model was used in the numerator of the weights. Applying the IPWs changed the IV effect estimate for Memory in model 3 from 0.18 to 0.19 (95 percent CI: 0.03, 0.36). With IPWs applied, the model 4 estimate for Memory remained 0.34 (0.10, 0.58). IPW corrected IV estimates for Cognition in models 3 and 4 were -0.04 (-0.35, 0.28) and 0.05 (-0.21, 0.30), respectively. We conclude that differential sampling probabilities and missing data in the HRS sample are unlikely to account for the finding that education affects Memory.

APPENDIX Figure Legend.

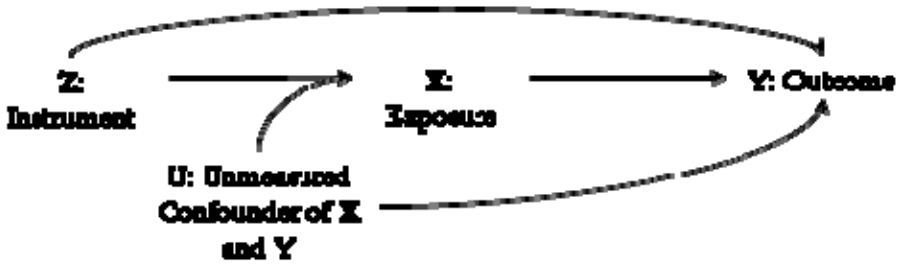
Appendix Figure 1. Causal diagrams for valid and invalid instrumental variables

Appendix Figure 1. Causal diagrams for valid and invalid instrumental variables

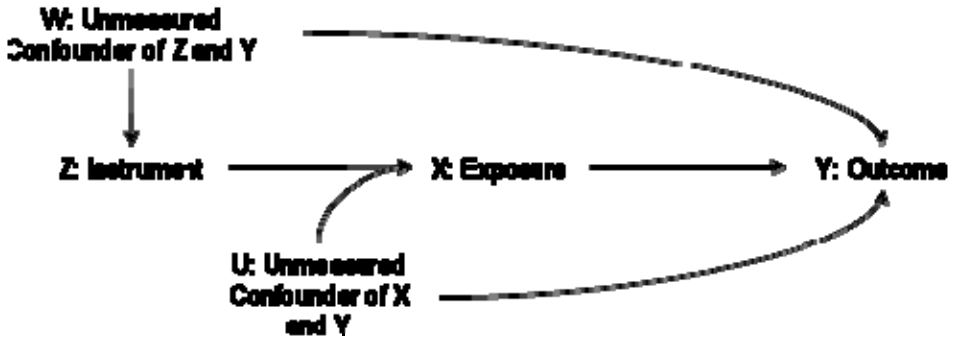
a) Z is a valid instrument for the effect of X on Y.



b) Z is not a valid instrument because it directly affects Y.



c) Z is not a valid instrument because it shares a common cause with Y.



APPENDIX REFERENCES

1 Pearl J. *Causality*. Cambridge, UK: Cambridge University Press 2000.

2 Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996;**91**:444-55.

3 Kennedy P. *A guide to econometrics*. Cambridge, Massachusetts: The MIT Press 1998.

4 Angrist JD, Imbens GW. 2-stage least-squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association* 1995;**90**:431-42.

5 Angrist JD, Krueger AB. Instrumental variables and the search for identification: from supply and demand to natural experiments. *J Econ Perspect* 2001;**15**:69-85.

6 Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;**29**:722-9.

7 Mark SD, Robins JM. A Method for the Analysis of Randomized Trials with Compliance Information - an Application to the Multiple Risk Factor Intervention Trial. *Control Clin Trials* 1993;**14**:79-97.

8 Martens EP, Pestman WR, de Boer A, *et al.* Instrumental variables application and limitations. *Epidemiology* 2006;**17**:260-7.

9 Hernán MA, Robins JM. Instruments for causal inference - An epidemiologist's dream? *Epidemiology* 2006;**17**:360-72.

10 Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica* 1997;**65**:557-86.

11 Robins JM, Gill RD. Non-response models for the analysis of non-monotone ignorable missing data. *Statistics in Medicine* 1997;**16**:39-56.