MRC Medical Research Council

PATHWAYS
biosocial influences on health

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

NCRM National Centre for Research Methods

George B. Ploubidis

The role of sensitivity analysis in the estimation of causal pathways from observational data

Improving health worldwide

www.lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Outline

- Sensitivity analysis
- Causal Mediation -Two examples
- Advantages
- Limitations
- An empirical example
- Summary

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Causal inference

- Causal inference with observational data is a nearly alchemic task
- Estimates depend on the model being correctly specified – no unmeasured confounders – Sequential Ignorability
- Can't be directly tested
- Things become more complicated when mediation is of interest

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

A simple idea

- Sensitivity analysis is an effective method for probing the plausibility of a nonrefutable assumption (sequential ignorability)
- The goal of sensitivity analysis is to quantify the degree to which the key assumption of no unmeasured confounders (sequential ignorability) must be violated for a researcher's original conclusion to be reversed

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- If an inference is sensitive, a slight violation of the assumption may lead to substantively different conclusions
- Given the importance of sequential ignorability, it has been argued that when observational data are employed some kind of sensitivity analysis should always be carried out
- Simply put: What happens to my estimated parameters if I simulate the effect of unmeasured confounders?

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

If only a direct effect is of interest

- Long history for exposure – outcome confounding (as early as the late 50's)
- Many techniques for different outcome types, exposure types, methods of primary analysis have been proposed
- These were generalised and unified by VanderWeele and Arah (Epidemiology, 2011)
- Stata Episens
- However, the difference with causal mediation is that indirect effects need to be estimated

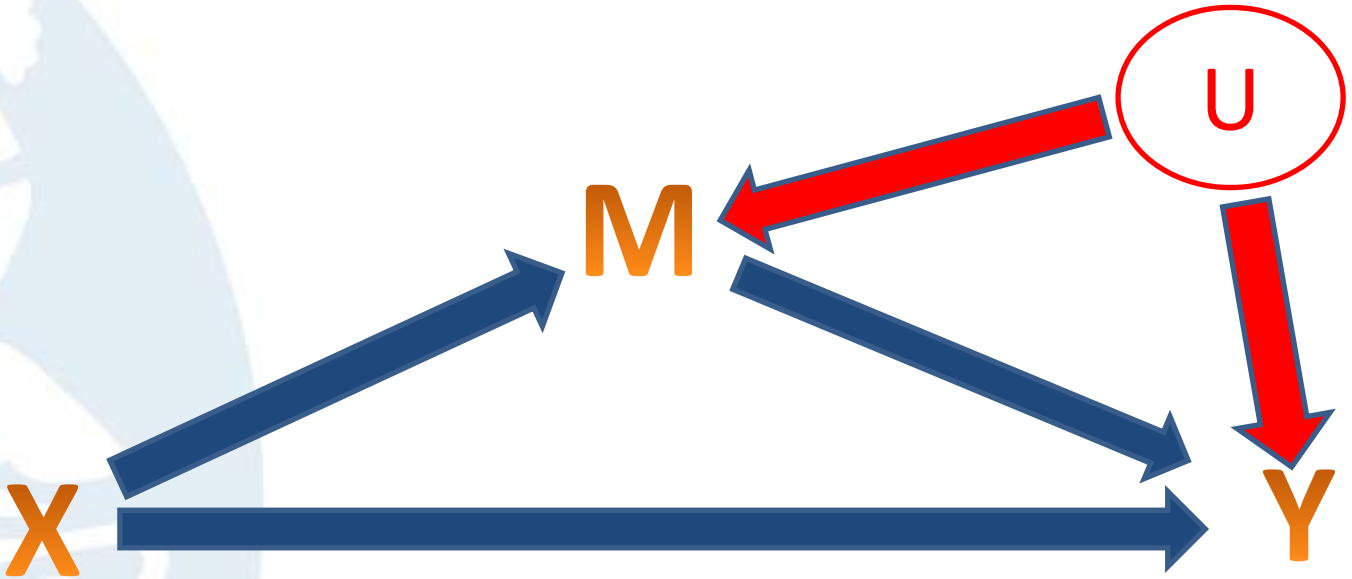
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Three general scenarios

- Mediator – outcome confounders
- Exposure – mediator confounders
- Exposure – mediator – outcome confounders
- Formal approaches available for the first scenario, but model specific approaches available for the remaining two

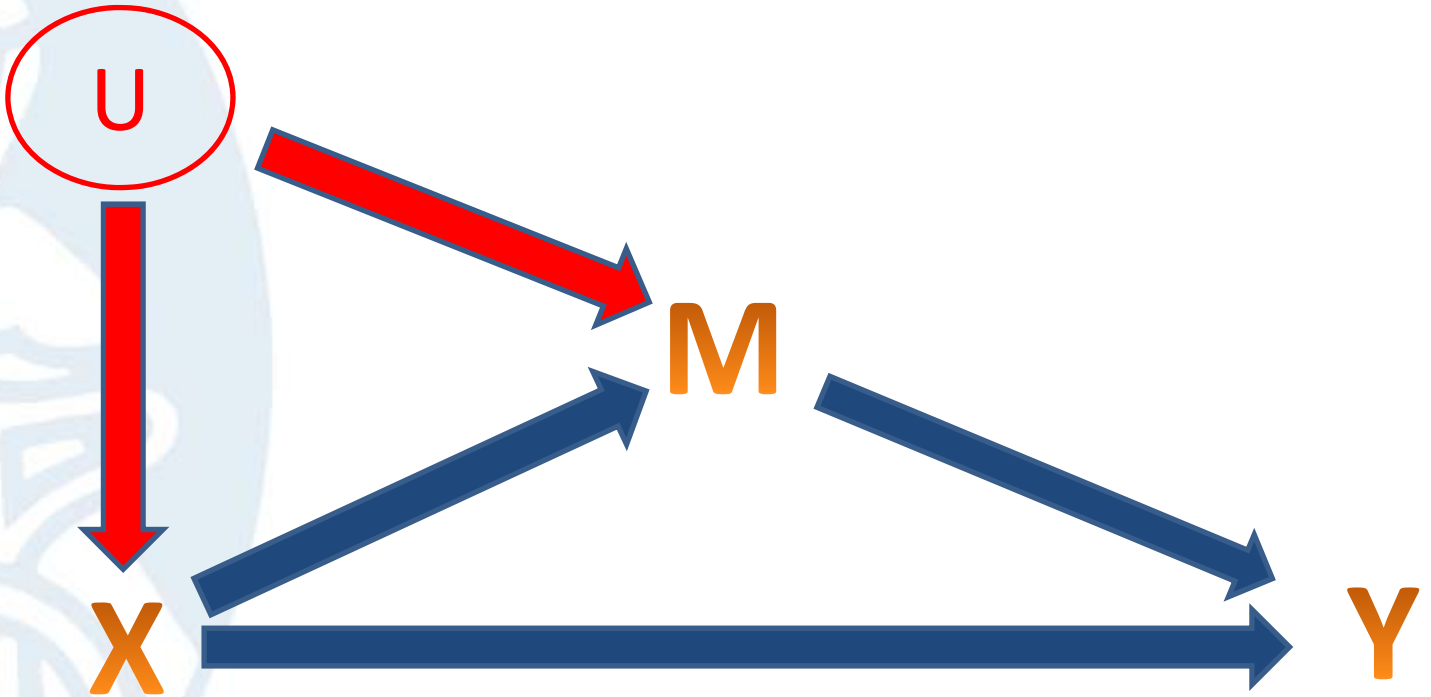
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Mediator - Outcome



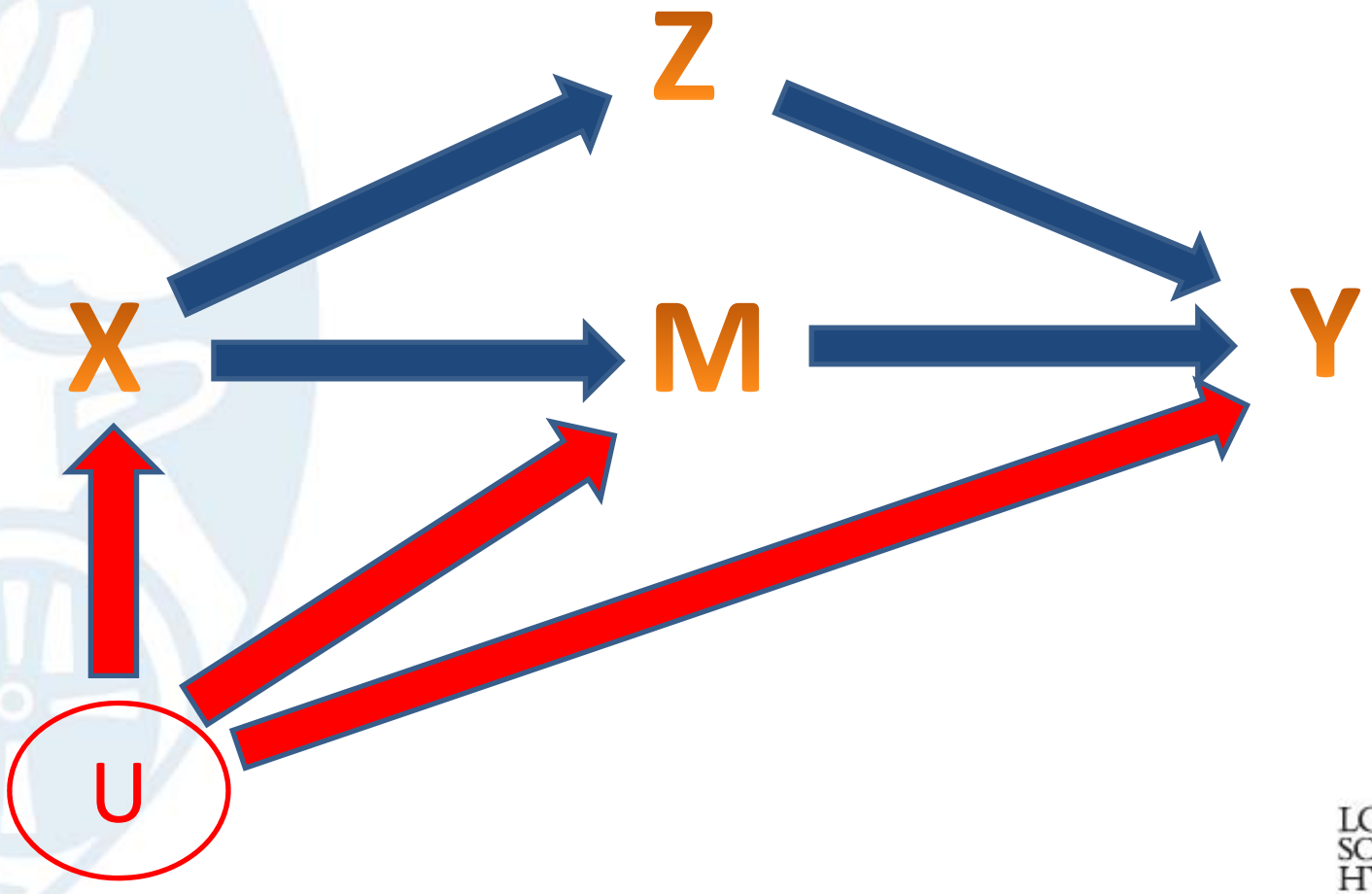
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Exposure – mediator



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Exposure – mediator - outcome



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

When it's about the exposure

- No formal approach thus far
- But under certain assumptions we can "challenge" our parameter estimates
- We can capitalise on the properties of latent variable measurement models
- Latent variables capture unobserved heterogeneity
- Unmeasured confounders can be thought of as sources of unobserved heterogeneity

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

When the exposure is involved

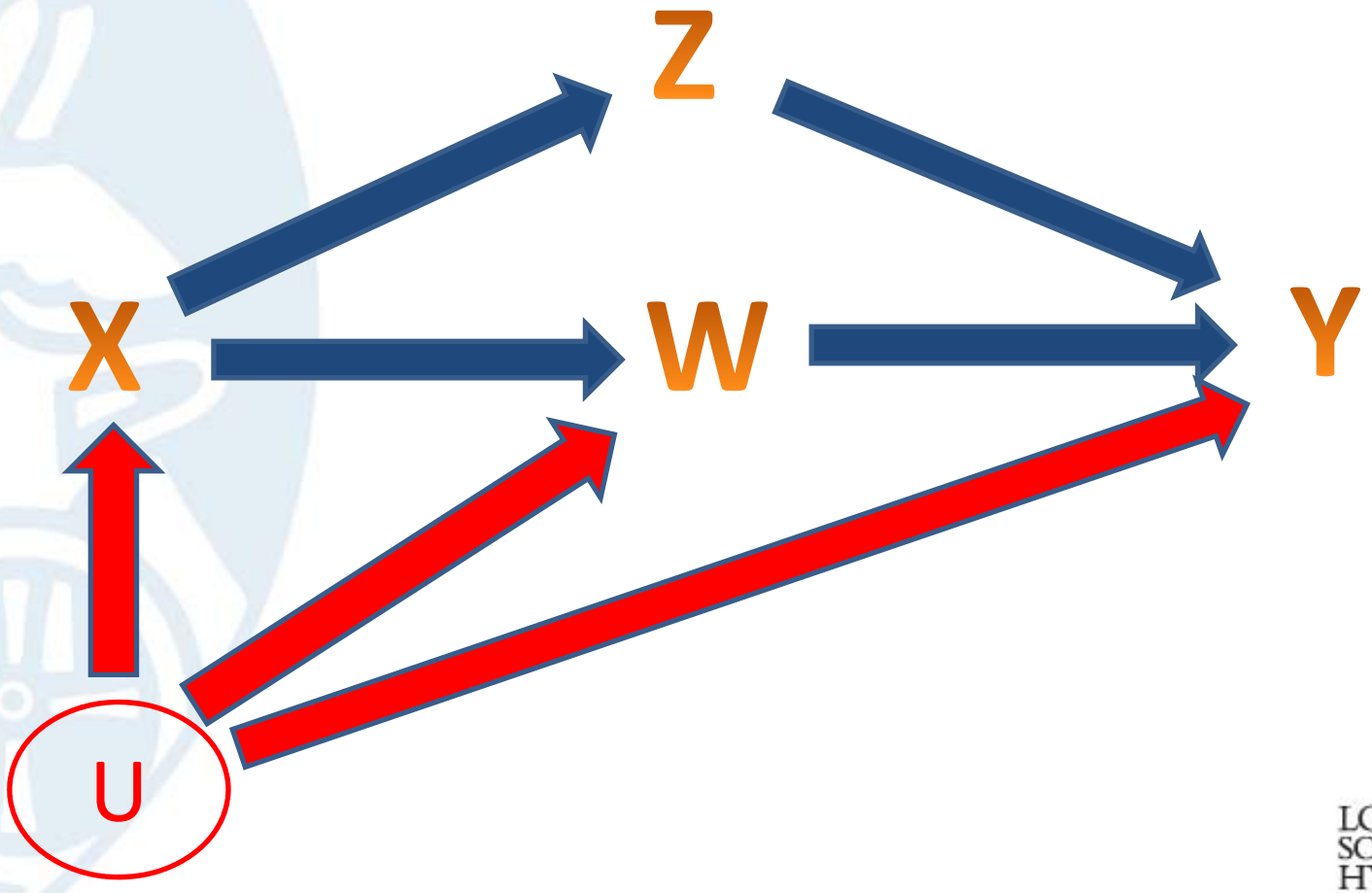
- Include latent variable “U” to represent unmeasured confounder(s)

- U ~ N (0,1)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- The latent variable(s) can represent the effect of one or more confounders
- The goal is to find out what happens to our estimates under several scenarios that involve latent “U”
- It can be shown that under certain assumptions latent variables can “imitate” the effect of observed confounders

A (relatively) simple example



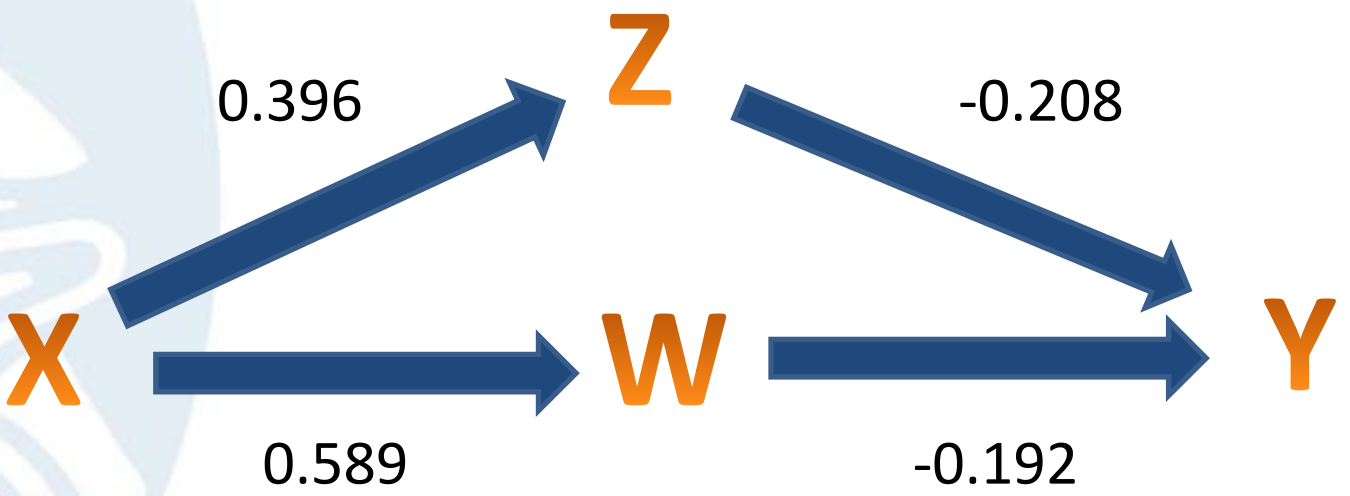
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

A simple LSEM

- All variables continuous and normally distributed
- No other confounders other than “U”
- Linear associations
- No interactions (although they could be accommodated)
- Estimation with MLR

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

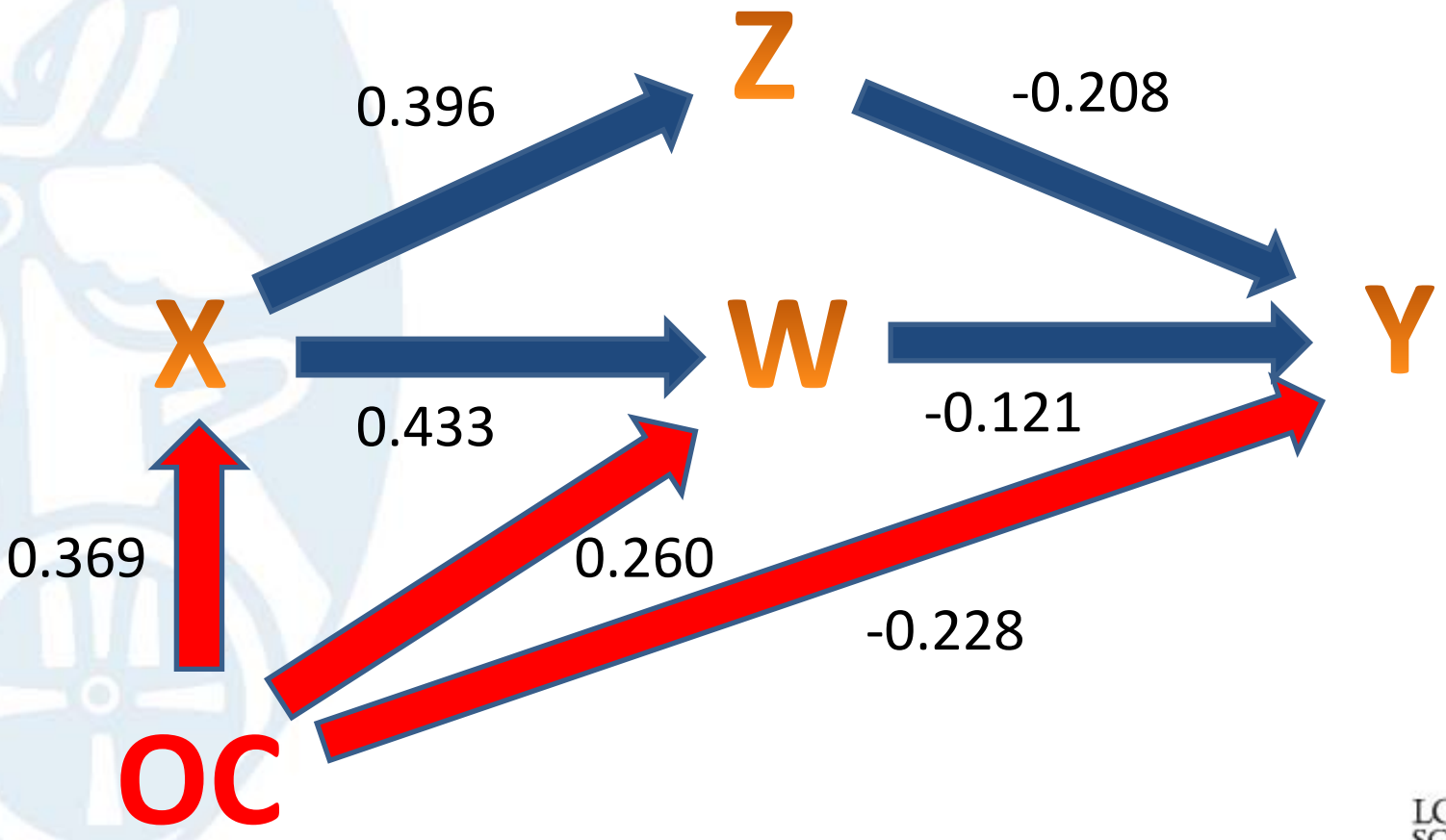
First the parameter estimates without the confounder



Y on X via W = -0.113 (-0.124 - -0.101)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Here comes the (observed) confounder!



Y on X via W = -0.052 (-0.063 - -0.044)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Can a latent variable do the same?

- It can be shown that if we fix the intercepts, slopes (loadings) and variance of the latent variable according to the estimated parameters we can obtain the estimates from the previous model

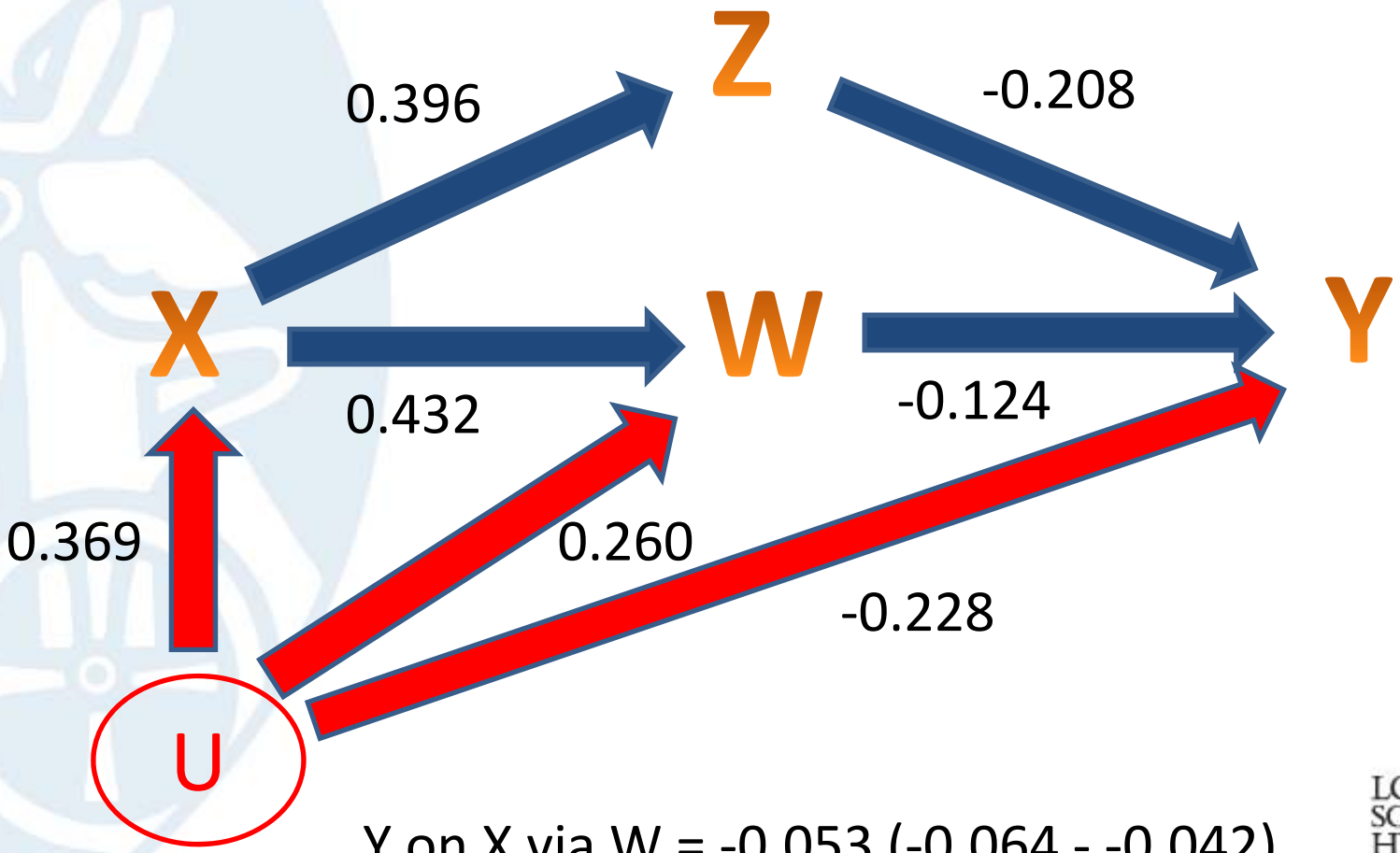
$$X = Ax + \lambda xU + eX$$

$$W = AW + \lambda wU + eW$$

$$Y = AY + \lambda YU + eY$$

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Estimates with the “latent confounder”



Y on X via W = -0.053 (-0.064 - -0.042)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Two possibilities

- a) The researcher suspects a set of unknown confounders

- b) A well known confounder, or a set of well known confounders have not been measured

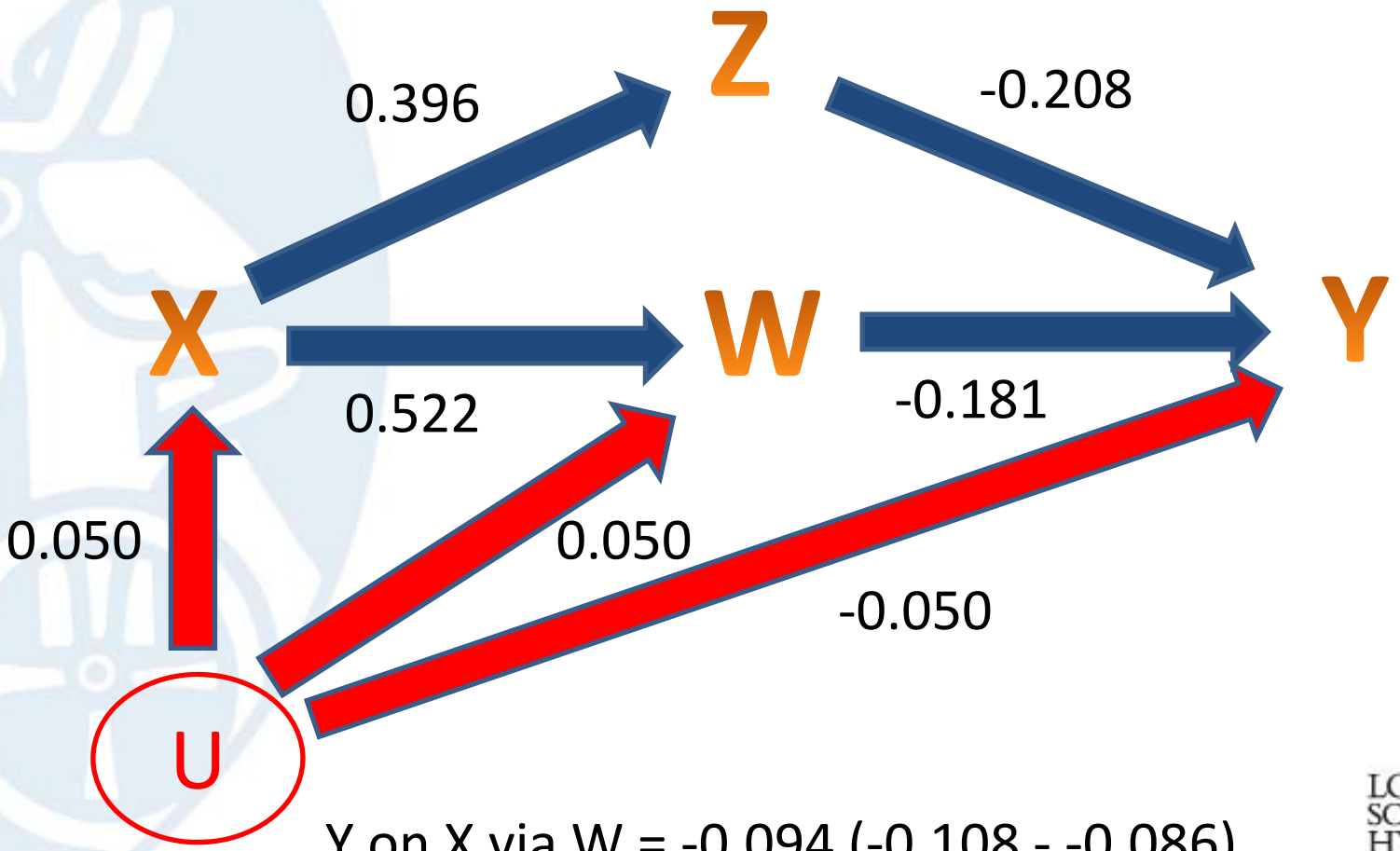
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Frequentist approach

- By specifying values for the effect of the confounder(s), the researcher will be able to test several scenarios of weak/moderate/strong confounding
- An iterative process
- The results of the trials can be quantified

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

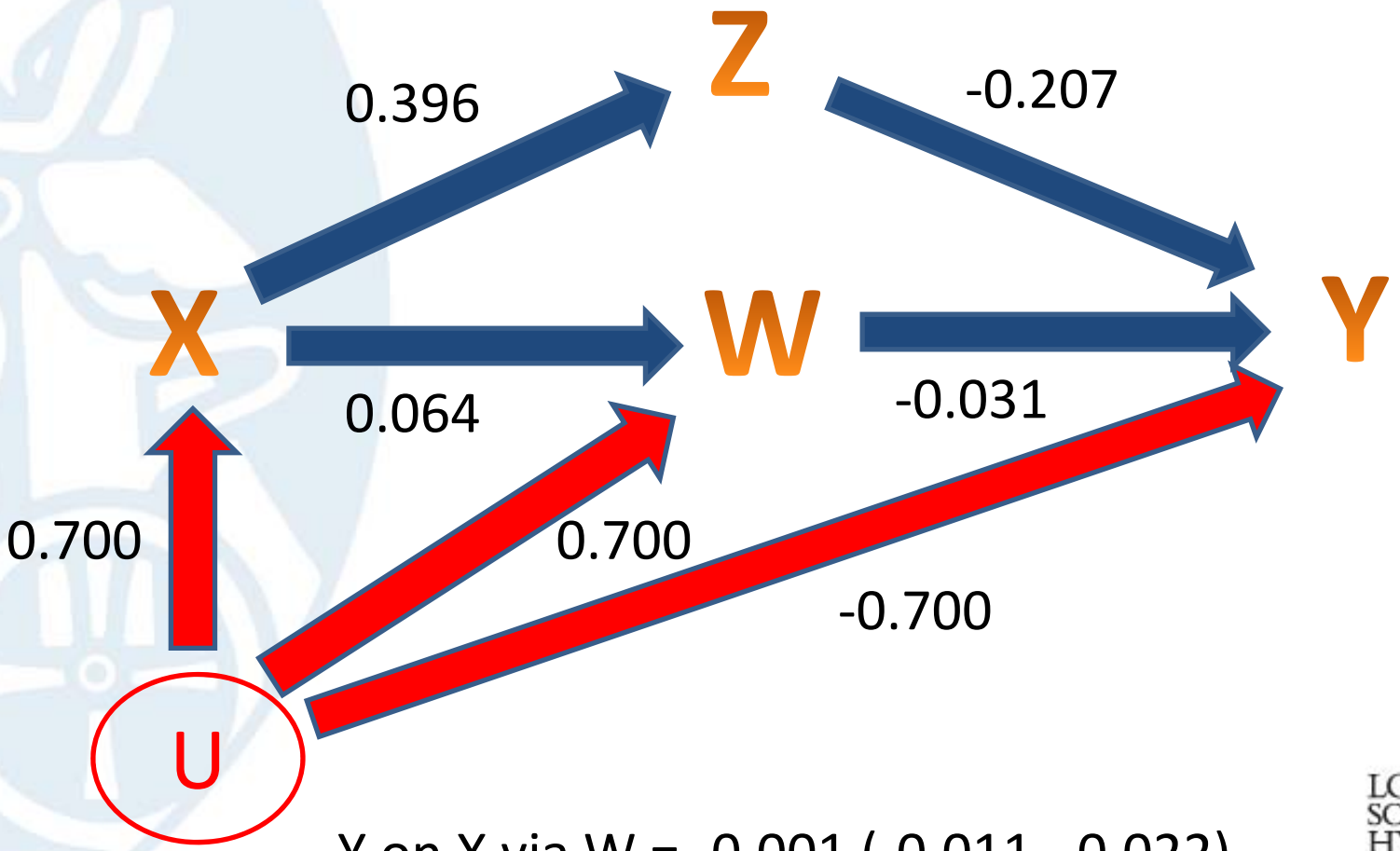
Weak/No confounding



Y on X via W = -0.094 (-0.108 - -0.086)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

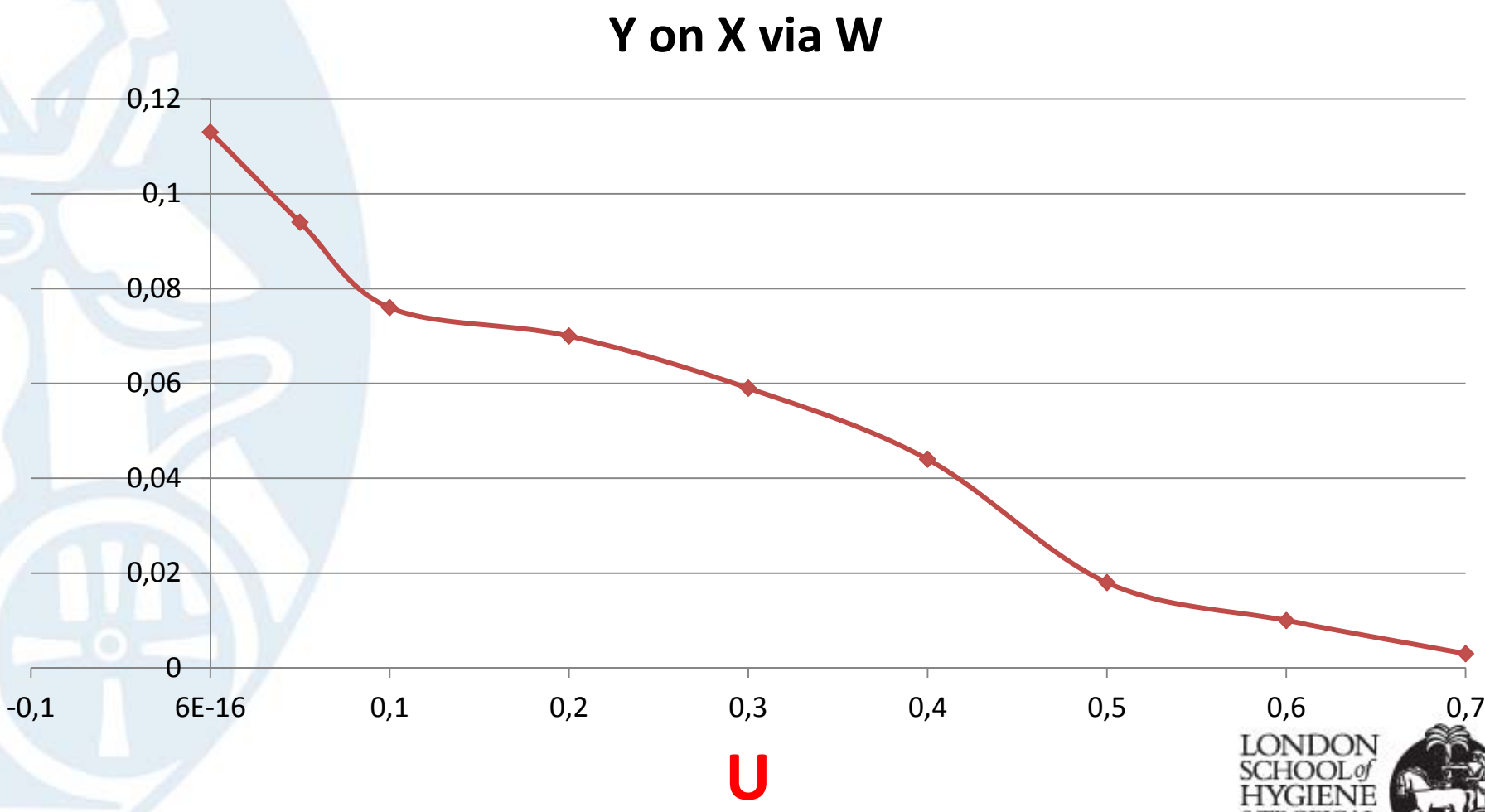
Strong confounding



Y on X via W = -0.001 (-0.011 - 0.022)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

A scree plot
Finding the tipping point



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Let's go Bayesian

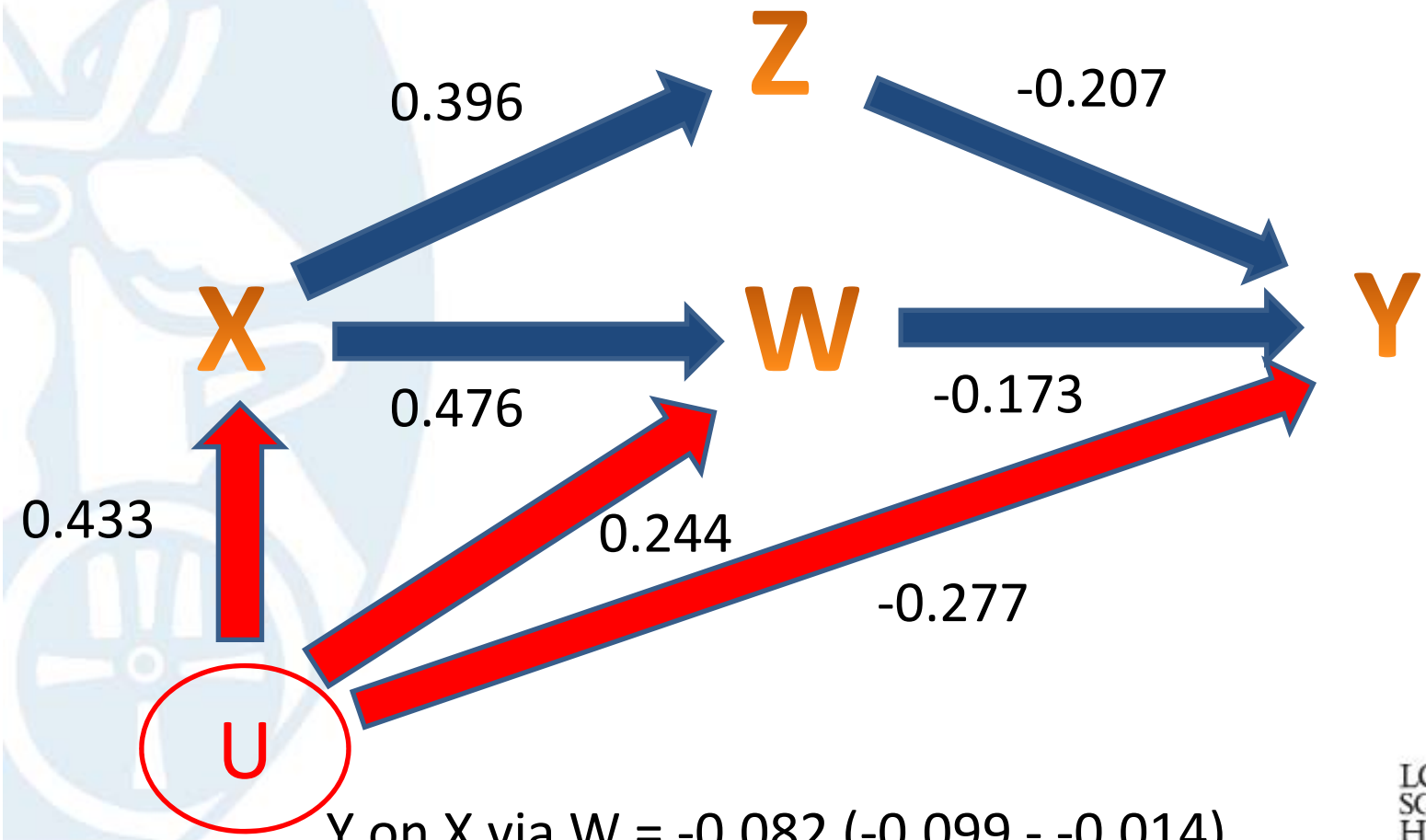
- "U" is a well known confounder
- It's associations with X,W and Y have been quantified in the existing literature
- Hence, we use informative priors for the parameters that link "U" with X,W and Y

$$UX \sim N(0.37, 0.01)$$

$$UW \sim N(0.26, 0.01)$$

$$UY \sim N(-0.23, 0.01)$$

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE



Y on X via W = -0.082 (-0.099 - -0.014)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Limitations

- Not non parametrically identified (i.e. results depend on the distribution of the latent confounder)
- No stopping rule – can't be falsified
- Latent confounder can only be normally distributed
- Discrete latent variables possible, but a lot more work is needed

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Advantages

- Properties of LVMs are well known
- Software availability
- DAG theory can be used to inform the sensitivity analyses

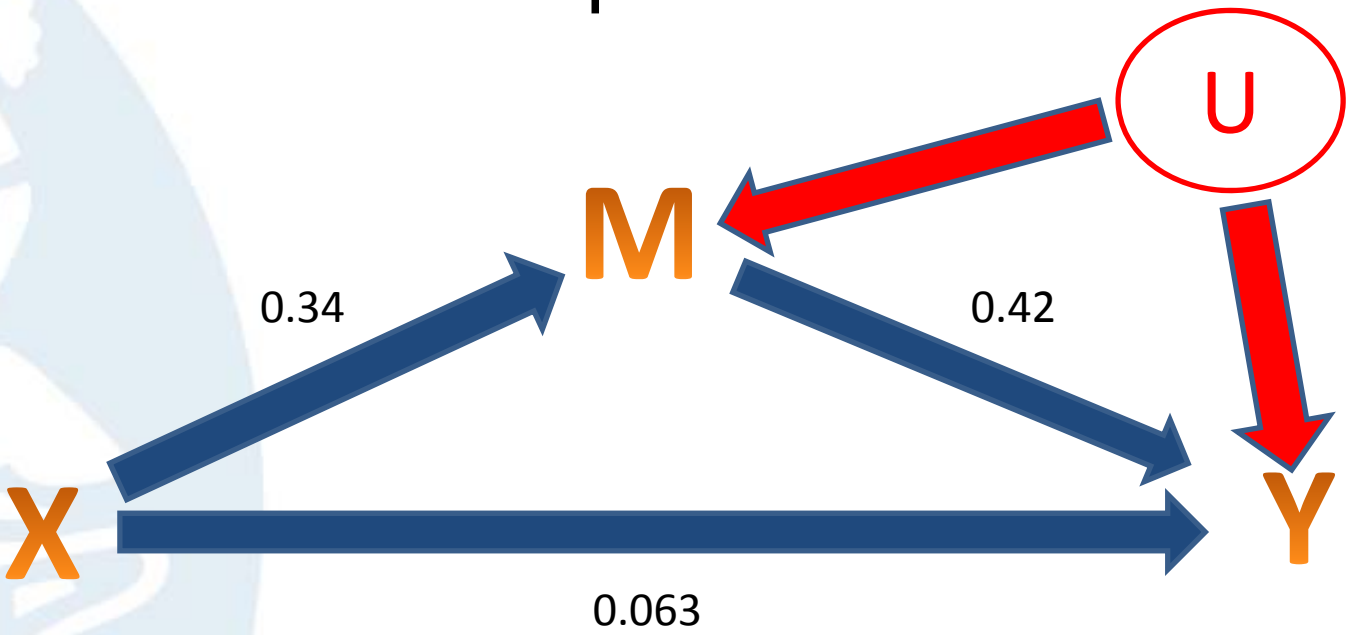
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Mediator – Outcome Confounding

- Medsens (Stata, R, Mplus)
- Employs the correlation (Rho) between the residual variances (errors) of the models for the mediator and outcome
- Effects are computed given different fixed values of the residual covariance.
- The proposed sensitivity analysis asks the question of how large does Rho have to be for the mediation effect (Average Causal Mediation Effect – ACME) to disappear

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

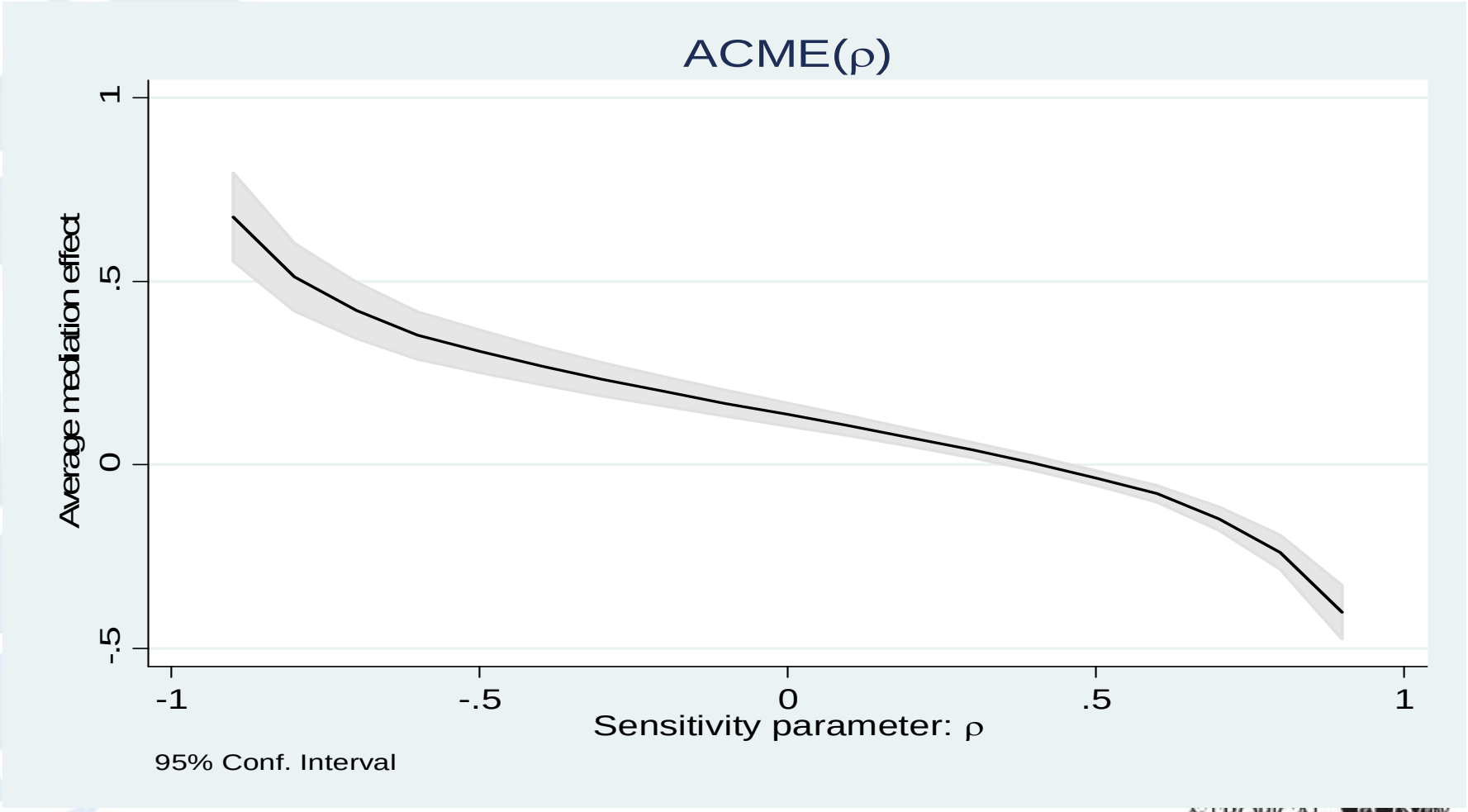
Medsens example



X to Y via M = 0.14 (0.11 – 0.17)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Medsens Results



&TROPICAL MEDICINE

Limitations

- Assumes all confounding due to Rho
- Only available for mediator - outcome associations
- Accommodates continuous mediator and continuous/ binary outcome, and binary mediator and continuous outcome
- Assumes normal distribution of error terms

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Advantages

- No distributional assumption for the unmeasured confounder
- Can accommodate binary, ordinal outcomes
- Quintile regression (for the outcome model) also available (only in R)
- Easy to use software

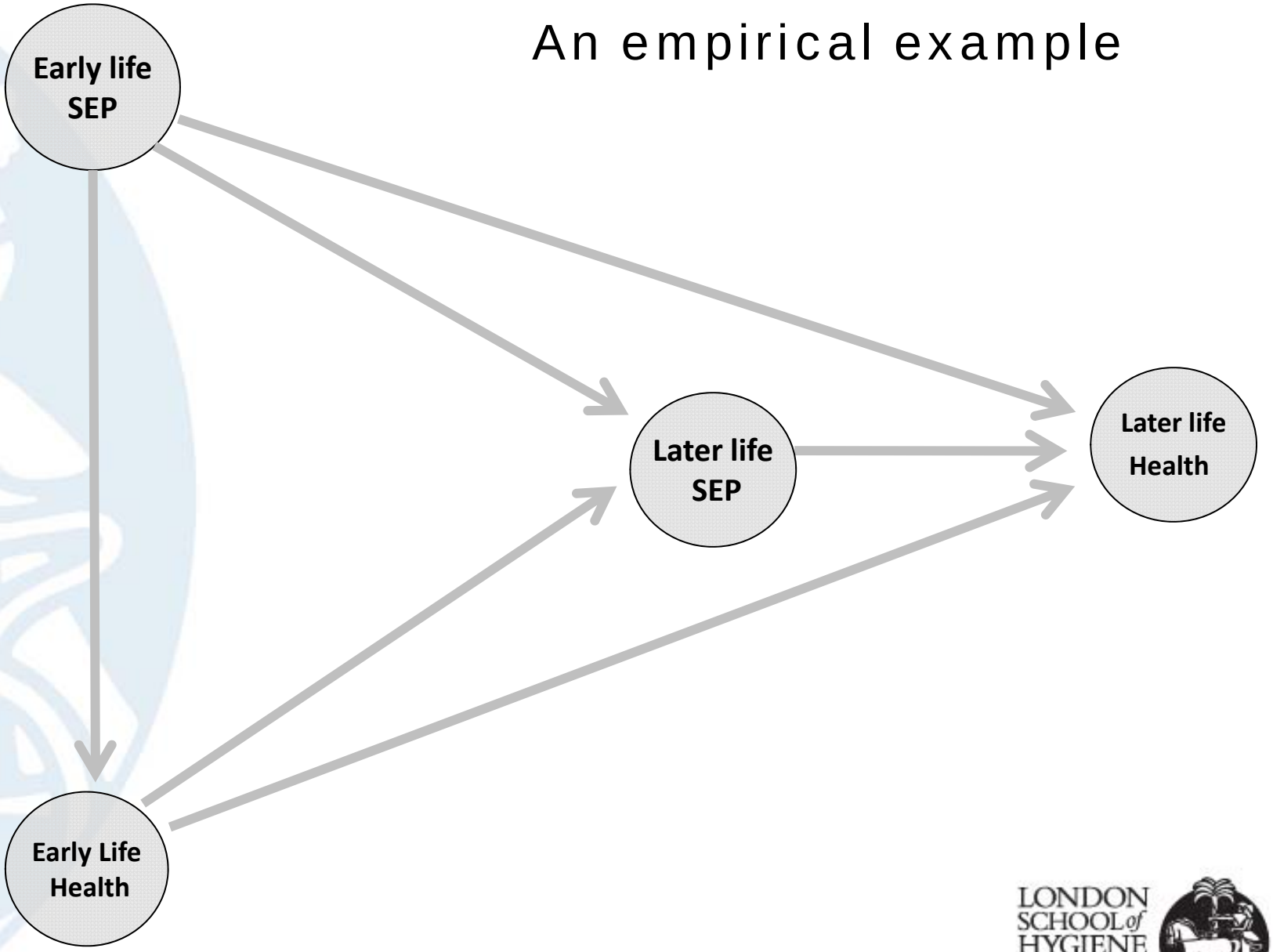
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Summary

- Under certain assumptions latent variables have the potential to “imitate” the effect of unmeasured confounders
- Medsens is a very useful tool to test the effects of mediator – outcome confounders
- Both approaches mostly effective in research areas (like the study of health inequalities) with strong theoretical underpinnings that can inform parameter specification/interpretation

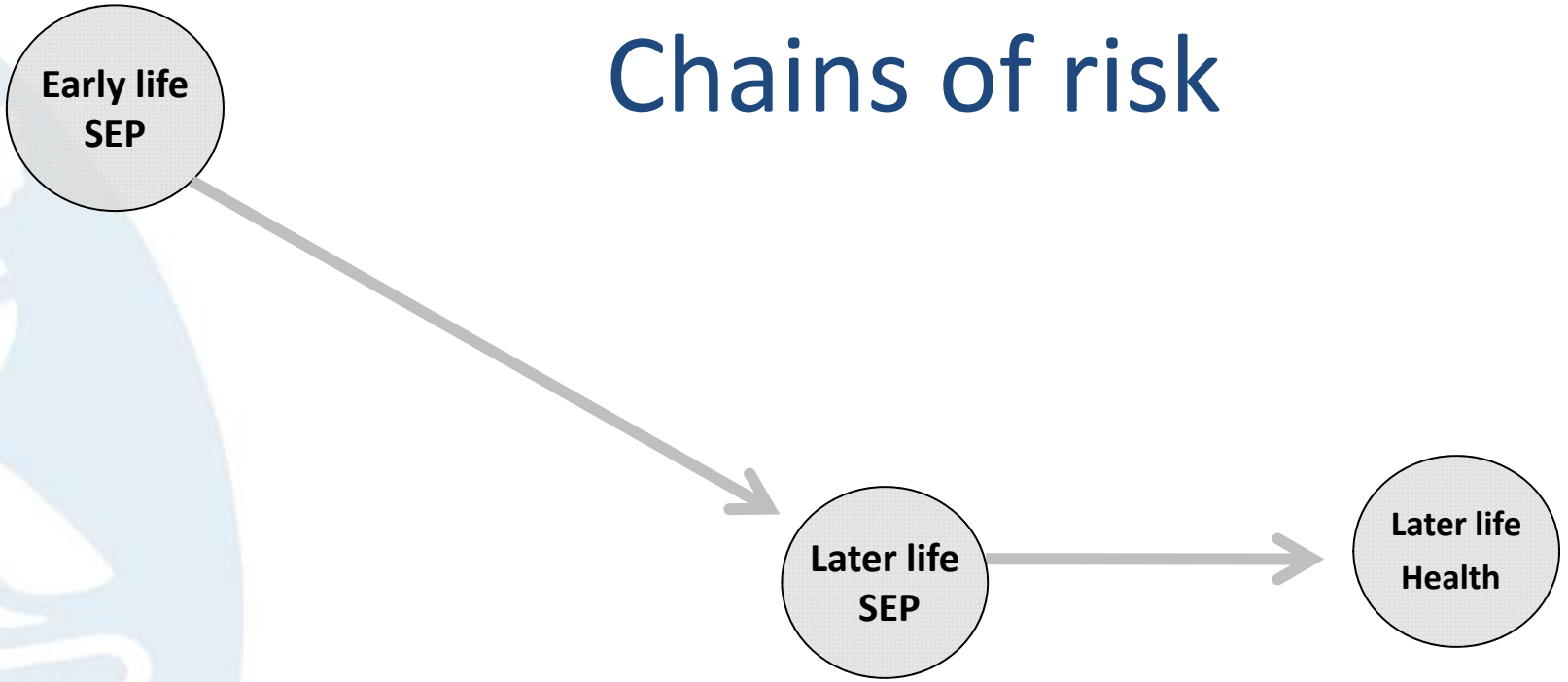
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

An empirical example

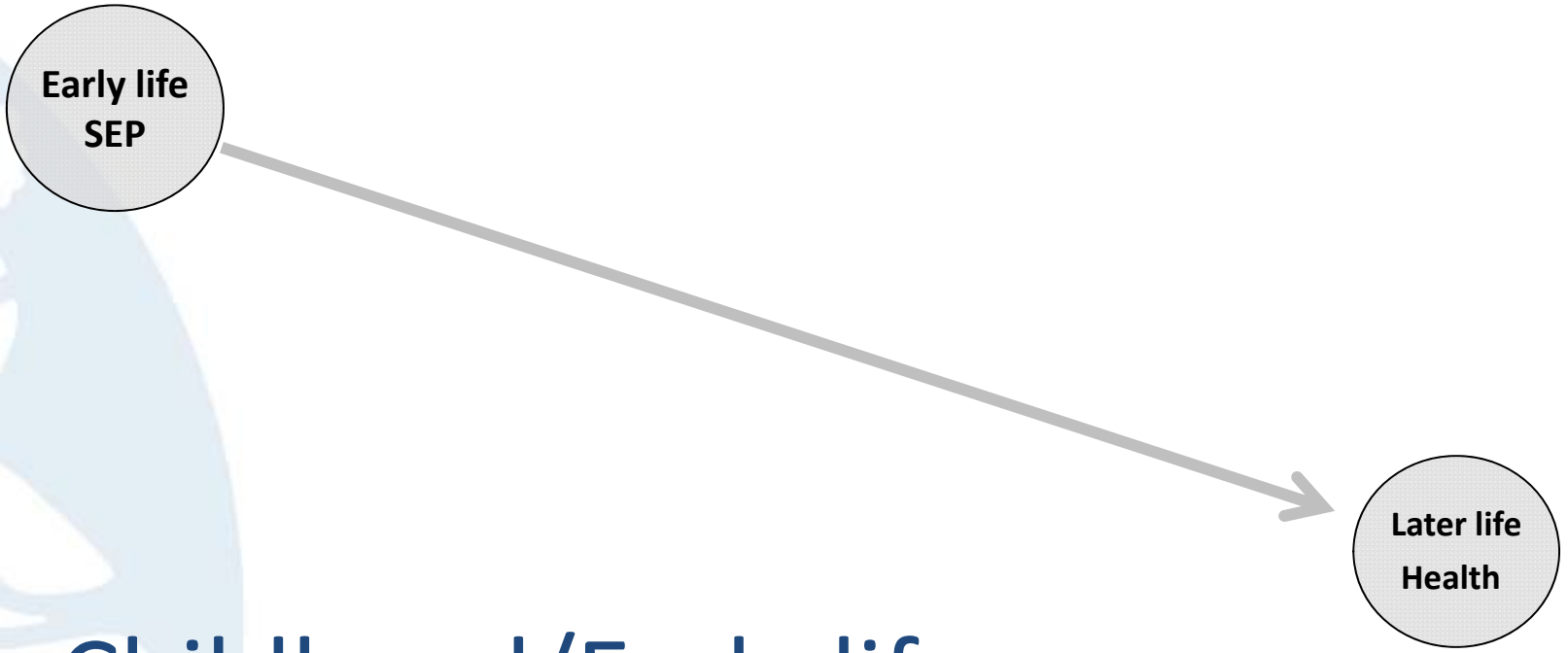


LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Chains of risk



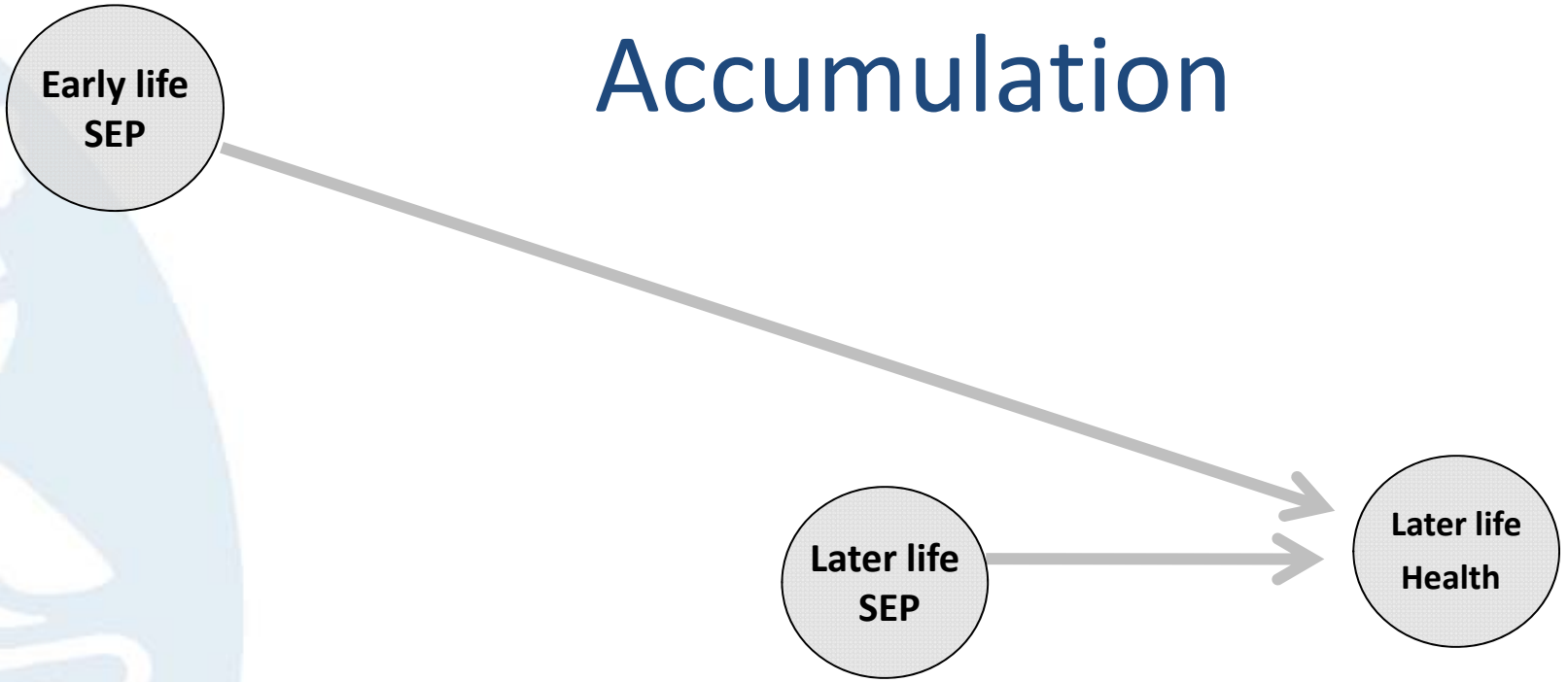
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE



Childhood/Early life

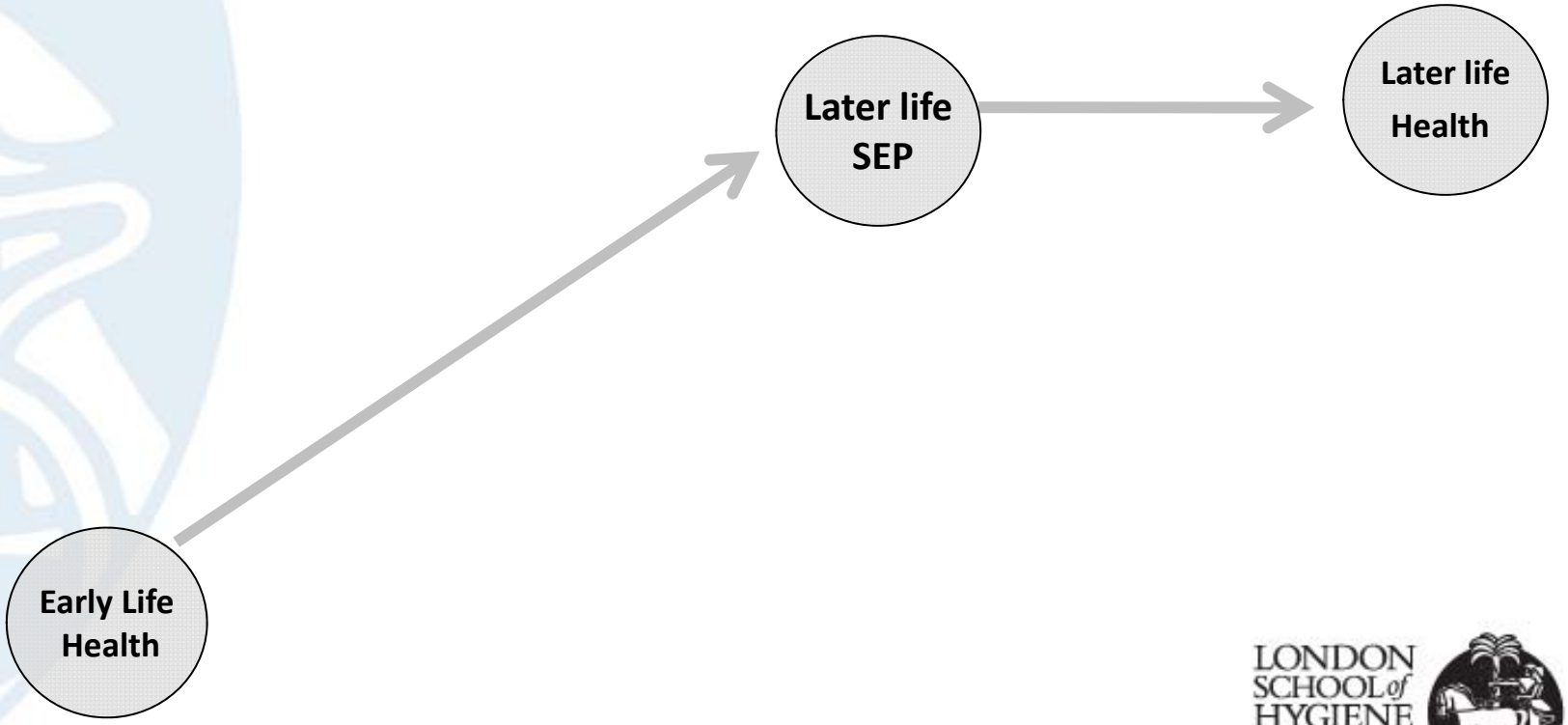
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Accumulation



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Social drift



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Aims

- The major aim of the present study is to test the relative contribution of these hypotheses to later life health inequalities
- A first step in understanding the mechanism that underlies the association between SEP and health
- We need to assign reliable parameters to all these hypotheses – it's all about mediation !!

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Sample

- We used data from the English Longitudinal Study of Ageing (ELSA), a nationally representative multi-purpose sample of the population aged 50 and over living in England
- We analysed a partially incomplete dataset (N = 7758), in which participants were included if they had at least one non missing observation in early life SEP indicators (ELSA Life history interview)
- Stratified by gender and age group (50-64, 65-74, 75+)

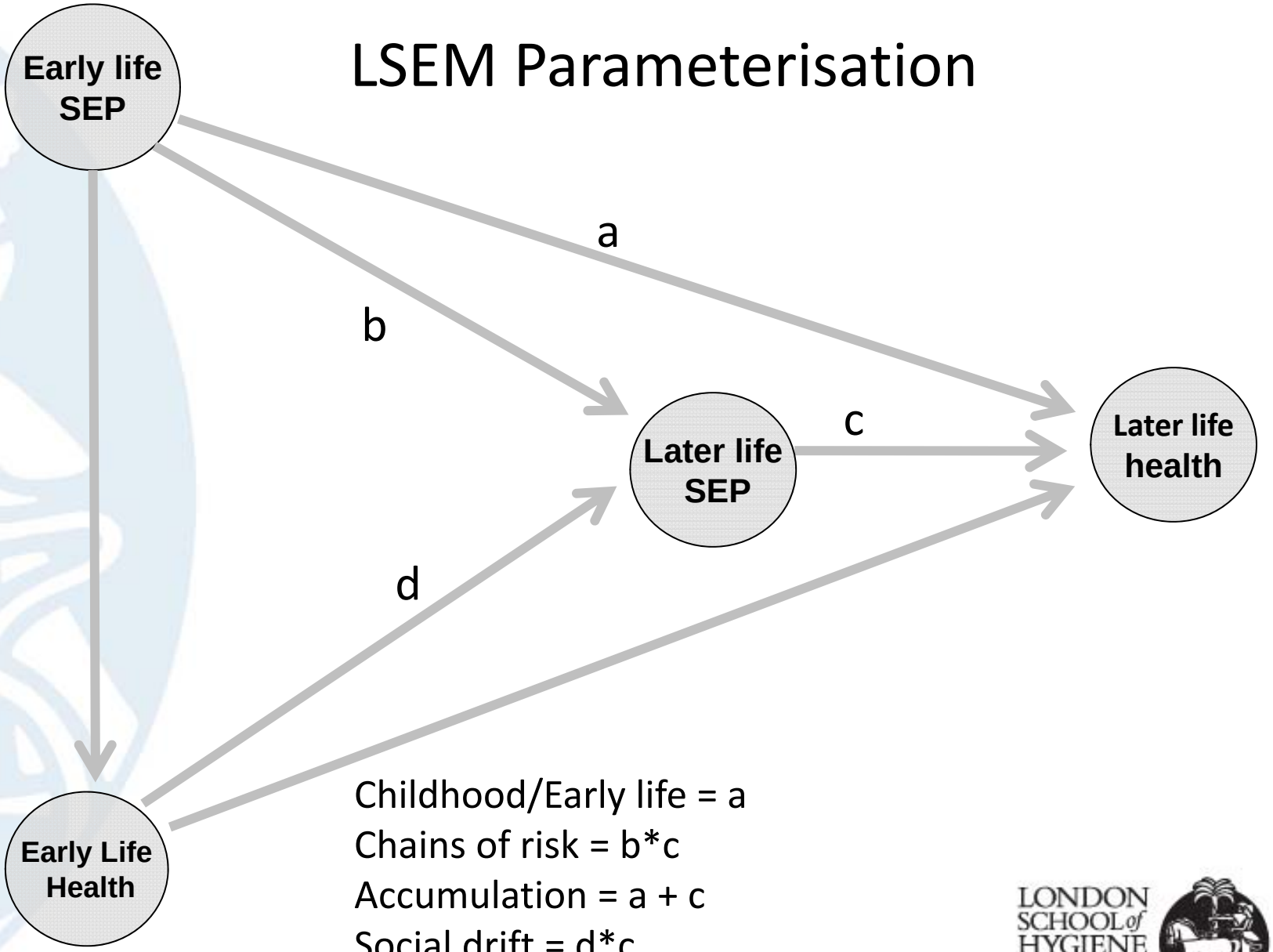
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Statistical modelling

- The specification of each of the latent dimensions was carried out with models appropriate for combinations of binary, ordinal and continuous indicators
- A LSEM was then estimated in order to jointly model the predictors, mediators and health outcomes (adjusted for confounders)
- Preliminary results showed no evidence of interactions
- Missing data on Wave 4 mediators and health outcomes handled with FIML assuming a MAR mechanism
- Estimation with MLR in Mplus 6.12

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

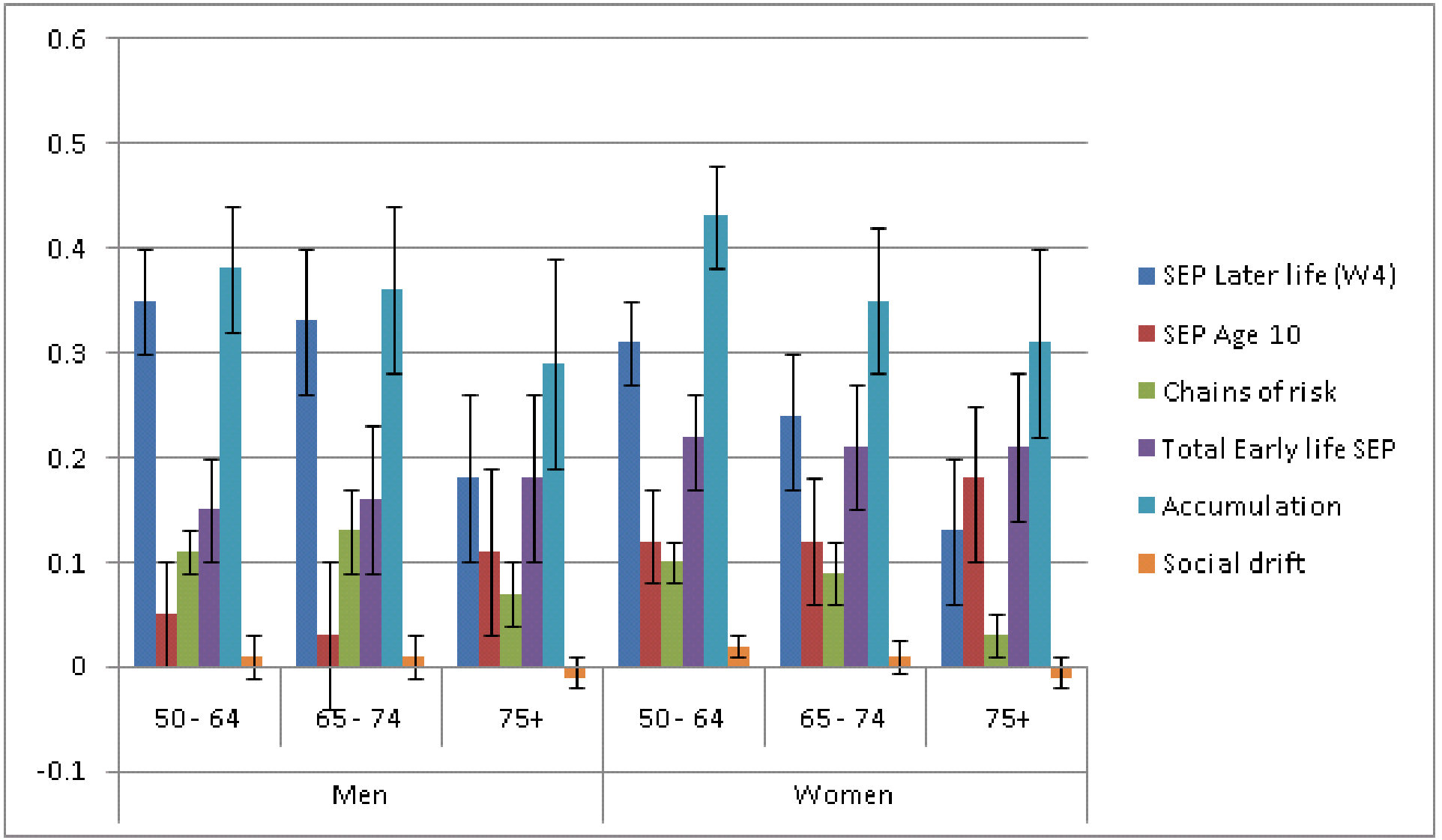
LSEM Parameterisation



Childhood/Early life = a
Chains of risk = b*c
Accumulation = a + c
Social drift = d*c
Total effect of Early life SEP = a + (b*c)

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Results - Physical Health

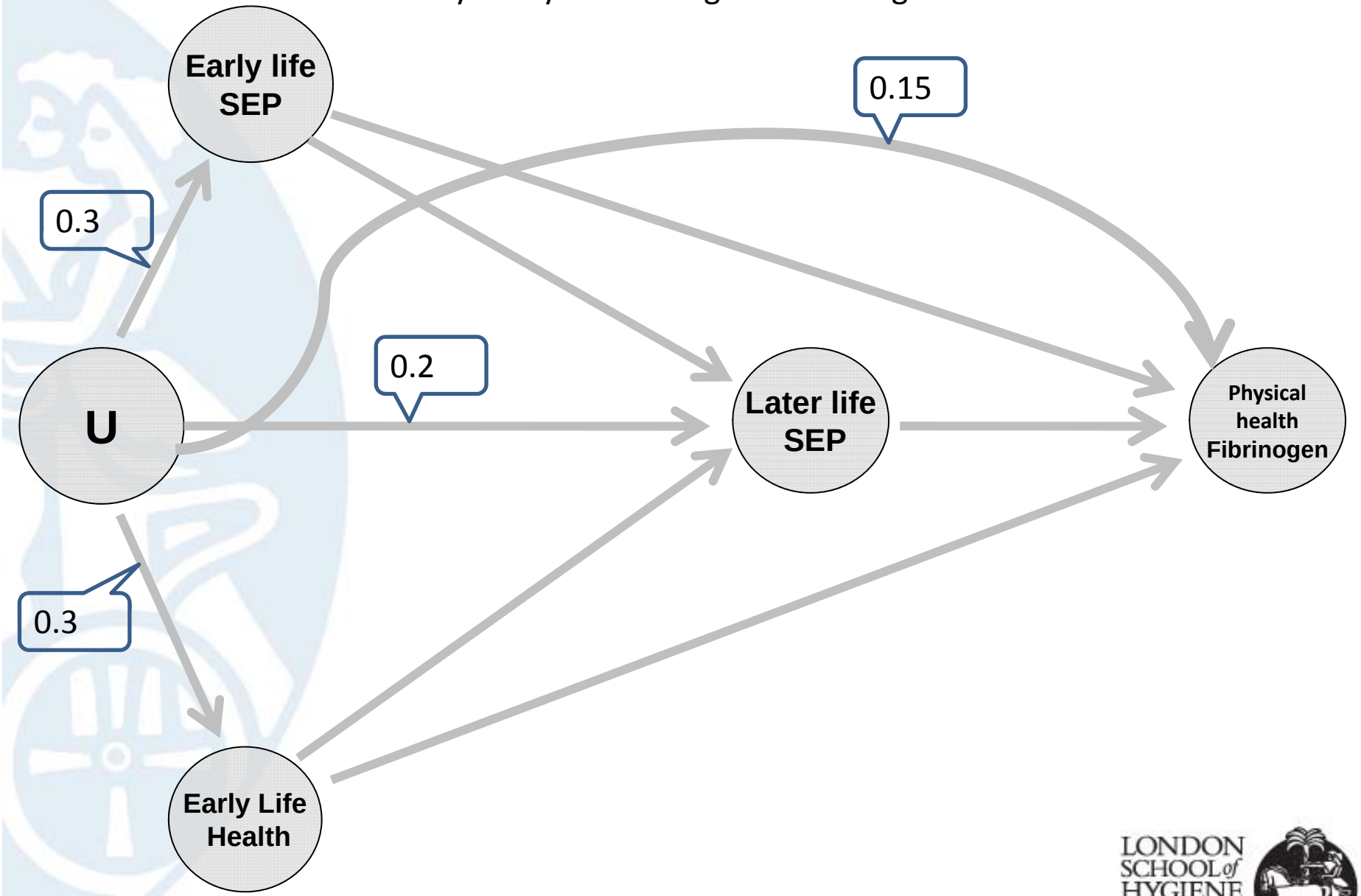


Sensitivity analysis

- Cool parameter estimates, but sequential ignorability implies no unmeasured confounders
- Sufficiently approximated* for the later life SEP – health association, but not for other parts of the diagram
- No data on parental characteristics such as cognitive ability and health status
- Results could reflect the effect of unmeasured parental characteristics

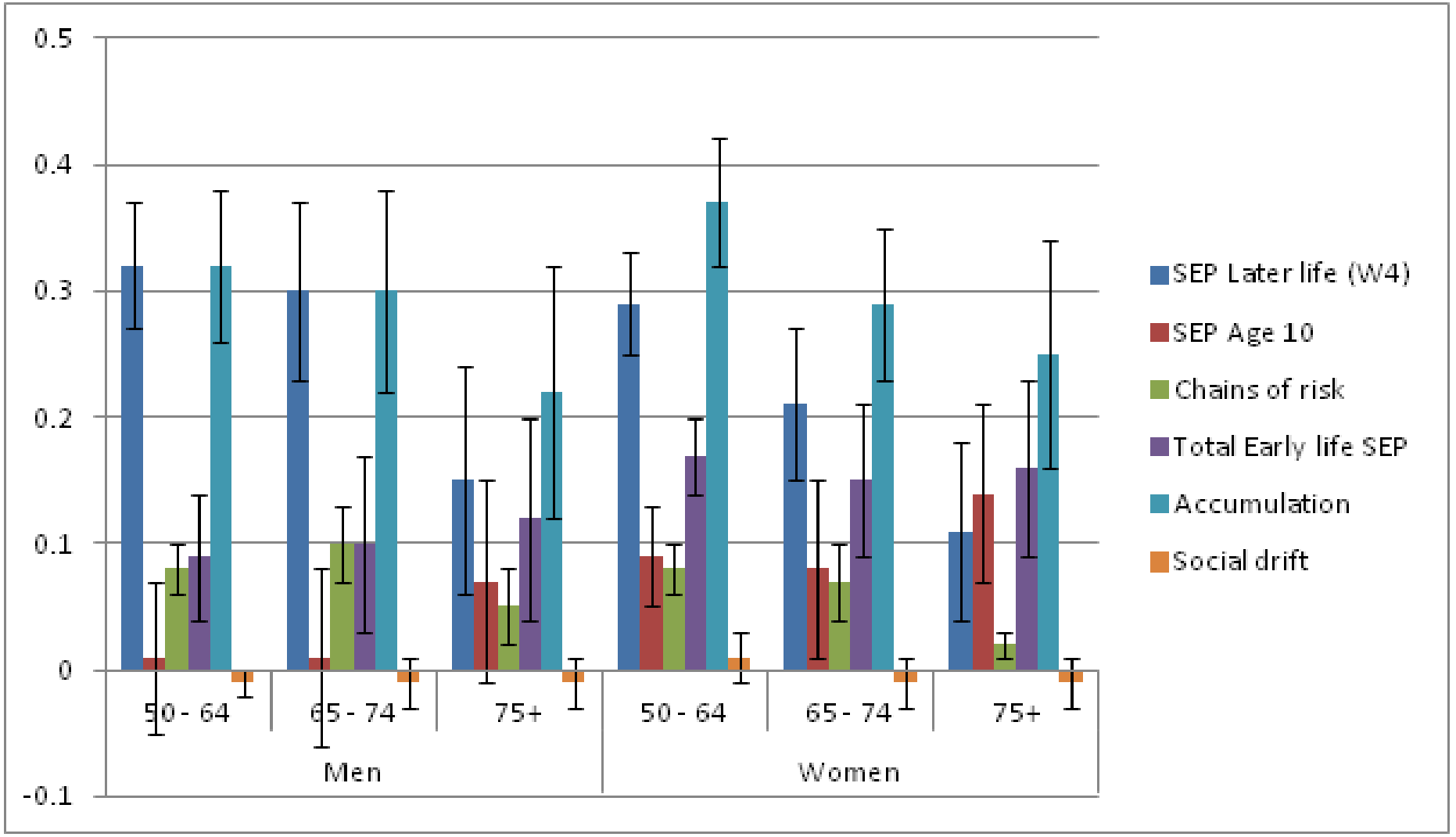
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Sensitivity Analysis – Strong confounding scenario



LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Physical health sensitivity analysis

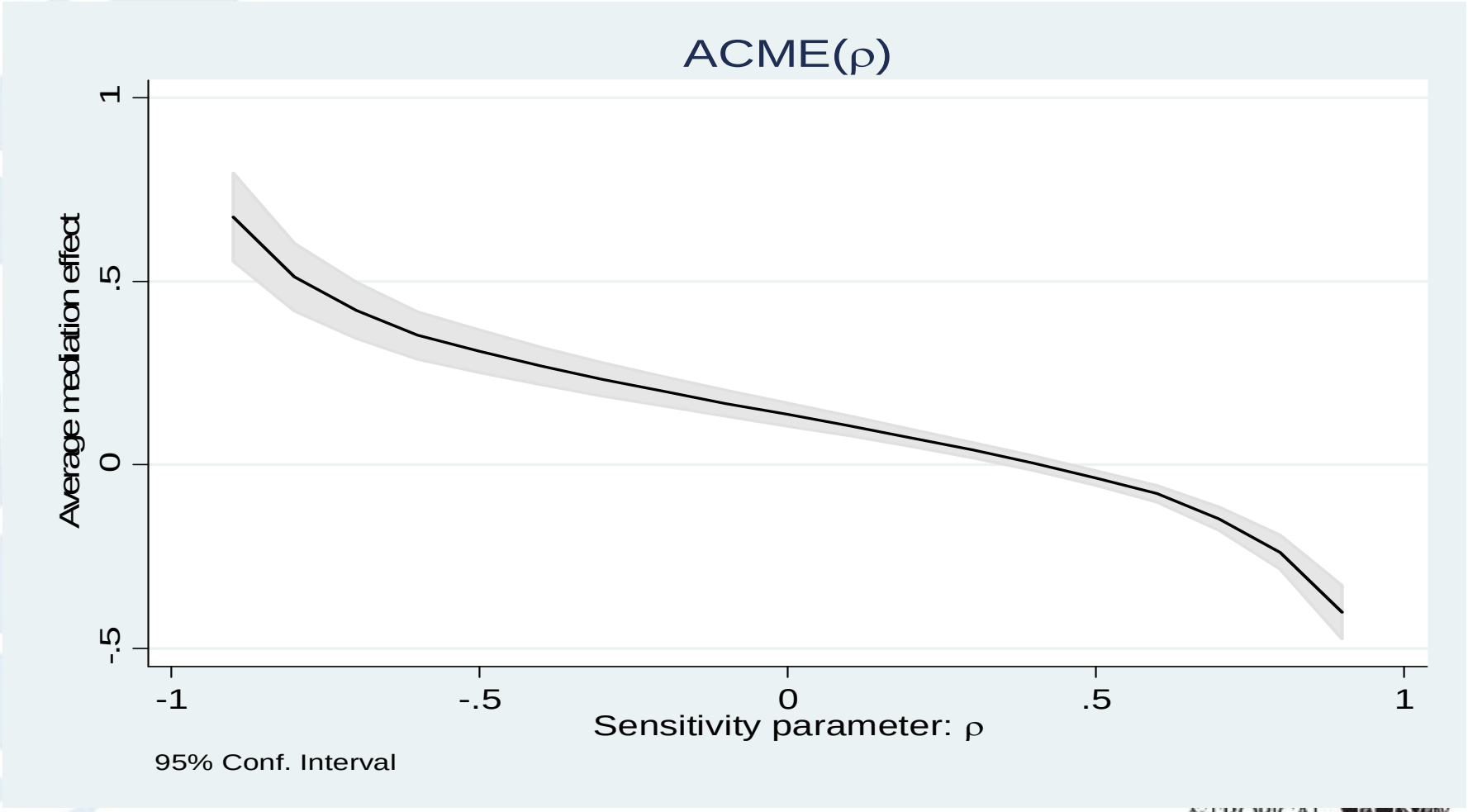


*Sufficiently approximated?

- We have included possible confounders of the later life SEP physical health/fibrinogen associations, but how about running some more sensitivity analyses?

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Medsens Results



&TROPICAL MEDICINE

So?

- Observational data – Causal inference a nearly alchemic task
- However, sensitivity analyses where confounders were simulated supported our results – but bias due to unknown unmeasured confounders cannot be ruled out

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Summary

- Sensitivity analysis not a substitute for randomisation
- Be aware of the assumptions and limitations of all sensitivity analysis approaches
- But especially when estimation of indirect effects (mediation) is required.....
- Always carry out sensitivity analysis!!

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Thank you!