Causal Models for Work and Health Disparities

Bryan Dowd
Division of Health Policy and Management
University of Minnesota

National Institute on Minority Health and Health Disparities (NIMHD)
September 28-29, 2020

The Basic Relationship of Interest

WORK
Paid and Unpaid → HEALTH

The Basic Question: The Role of Disparities

- Age
- Race / ethnicity
- Gender

(all predetermined)

- Discrimination
- Biological, Behavioral differences
- Physical/built and socioeconomic environments (including health care)

WORK
Paid and Unpaid

HEALTH

Disparities (differences) by:

- Age
- Race / ethnicity
- Gender

A Word About Disparities

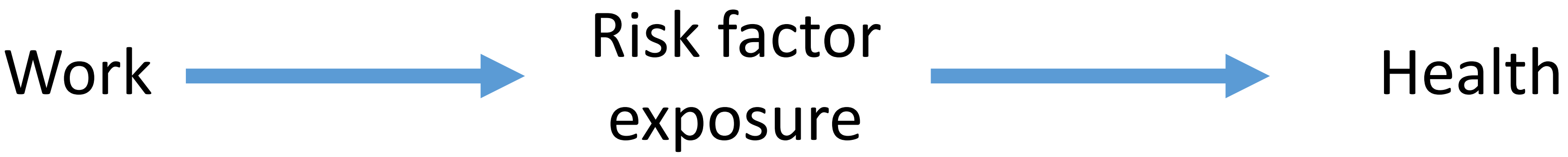
Three important questions about disparities (and possible responses):

1. Why do we care about disparities?
   - We want everyone to be able to have an equal opportunity to achieve the best possible state of health.
   - At an individual level, we want individuals to be able to invest \$1 in their own health if the resultant effect is worth more than \$1 to them.
   - We want society to be able to invest \$1 to improve the health of individuals if that investment results in more that \$1 of societal well being (somehow measured).
2. What do disparities represent?
   - Health differences that adversely affect defined disadvantage population (racial and ethnic minorities, low SES, sexual/gender minorities, rural)
   - Discrimination (and racism) at multiple levels such as structural, interpersonal and internalized
   - Possible biological, behavioral, physical/built environment, sociocultural environment, and/or health care system determinants
3. What, if anything, can or should be done about them?

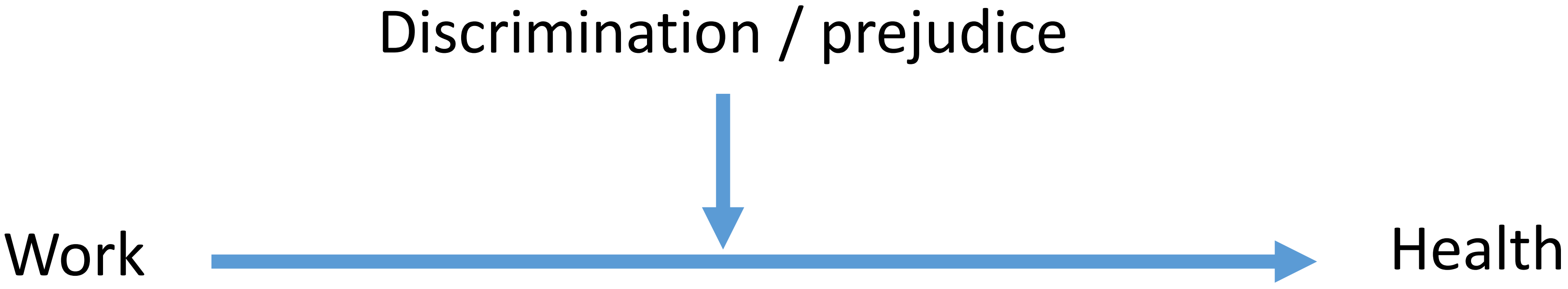
To answer that question, we need a theory of the role of disparities in causal models.

Terminology and Directed Acyclic Graphs (Pearl, 2000)

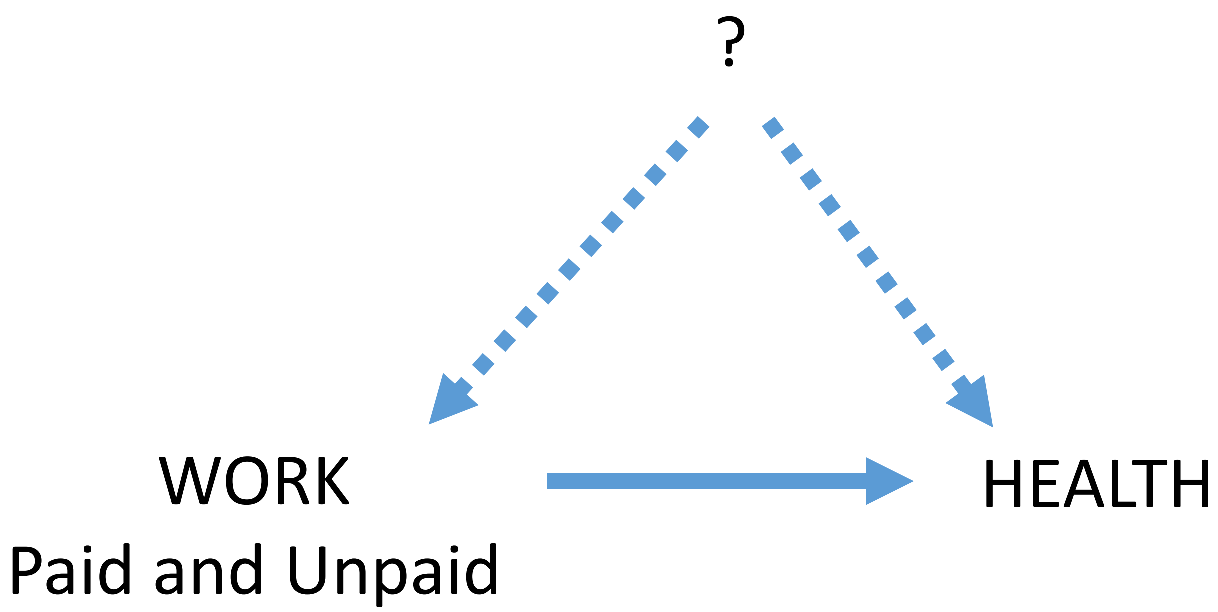
- **Risk factor exposure** is a mediator. It is one of the ways that **Work** affects **Health**.



- **Discrimination/prejudice** is a moderator. It modifies the effect of **Work** on **Health** perhaps through greater exposure to risk factors.



Estimation problem #1: Omitted Variable Bias



Estimation problem #2: Reverse Causality

Satisfaction /
Risk-factor Exposure

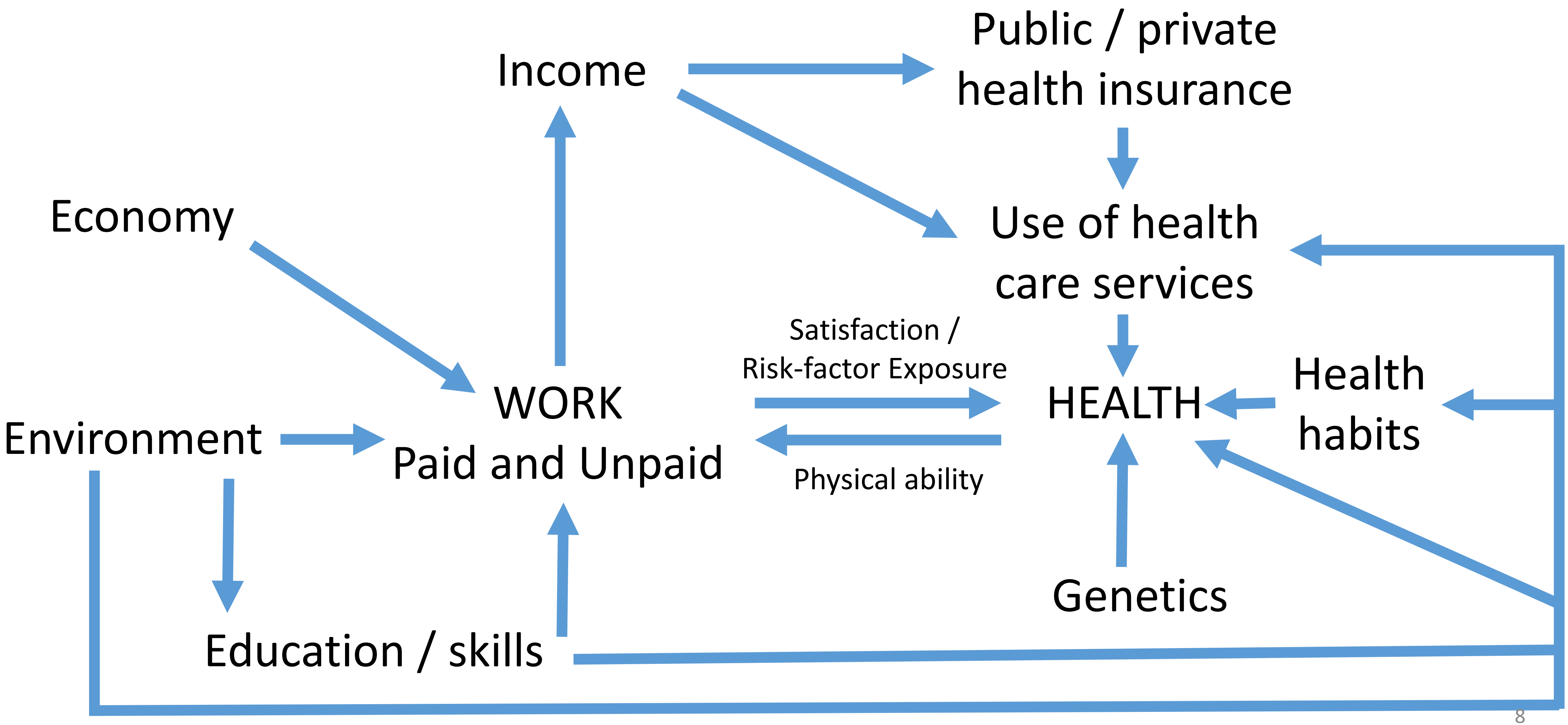
WORK
Paid and Unpaid

HEALTH

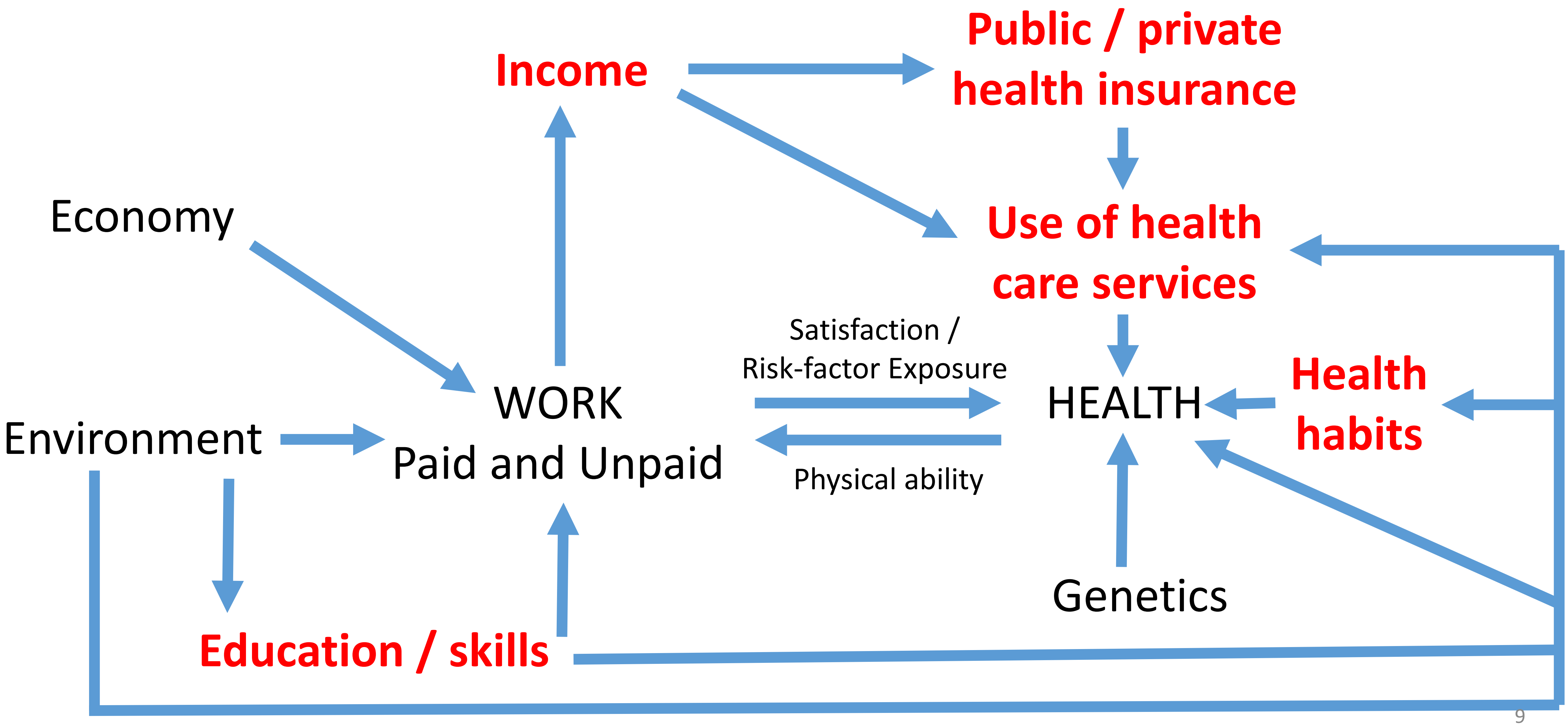
Physical ability

A More Complete Model

Structural Form

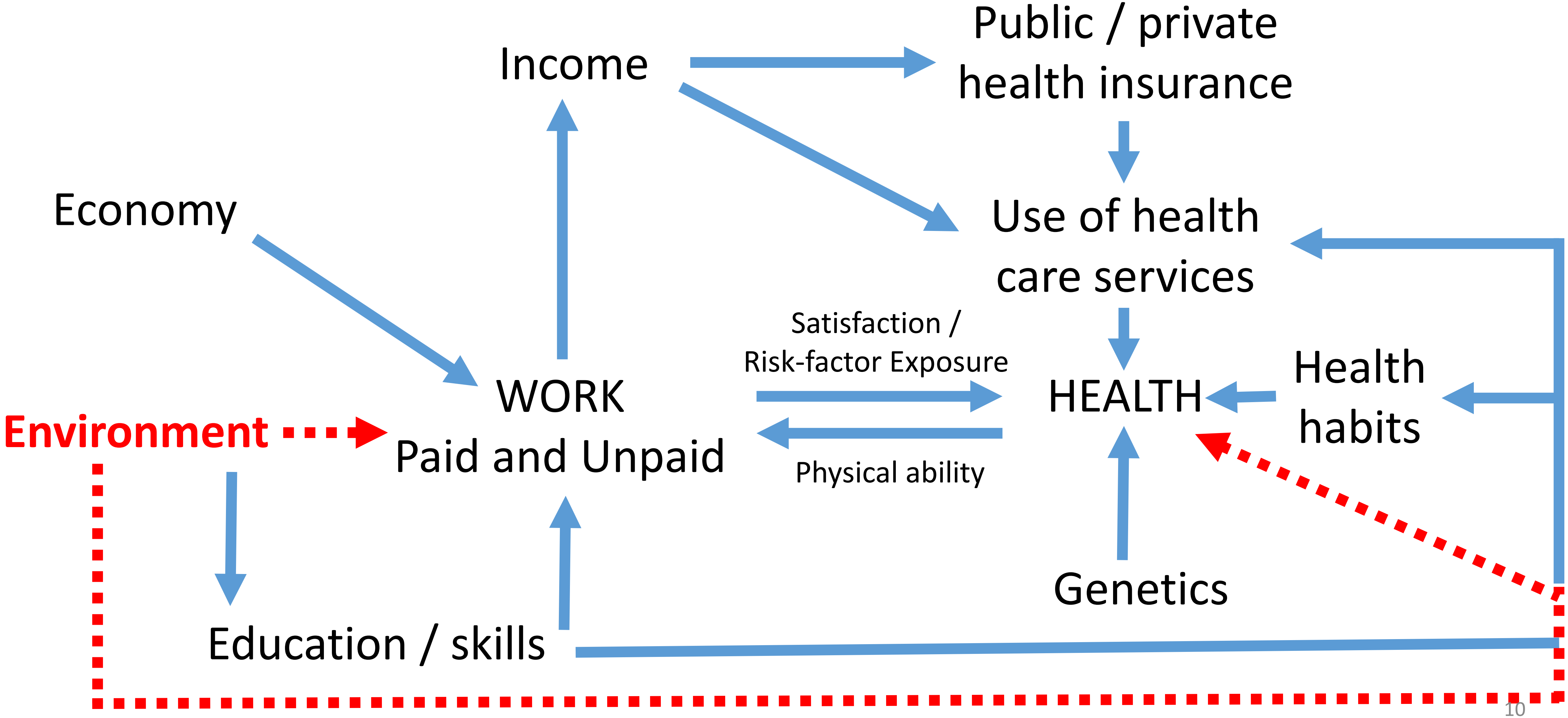


Mediator (endogenous) Variables (in red)

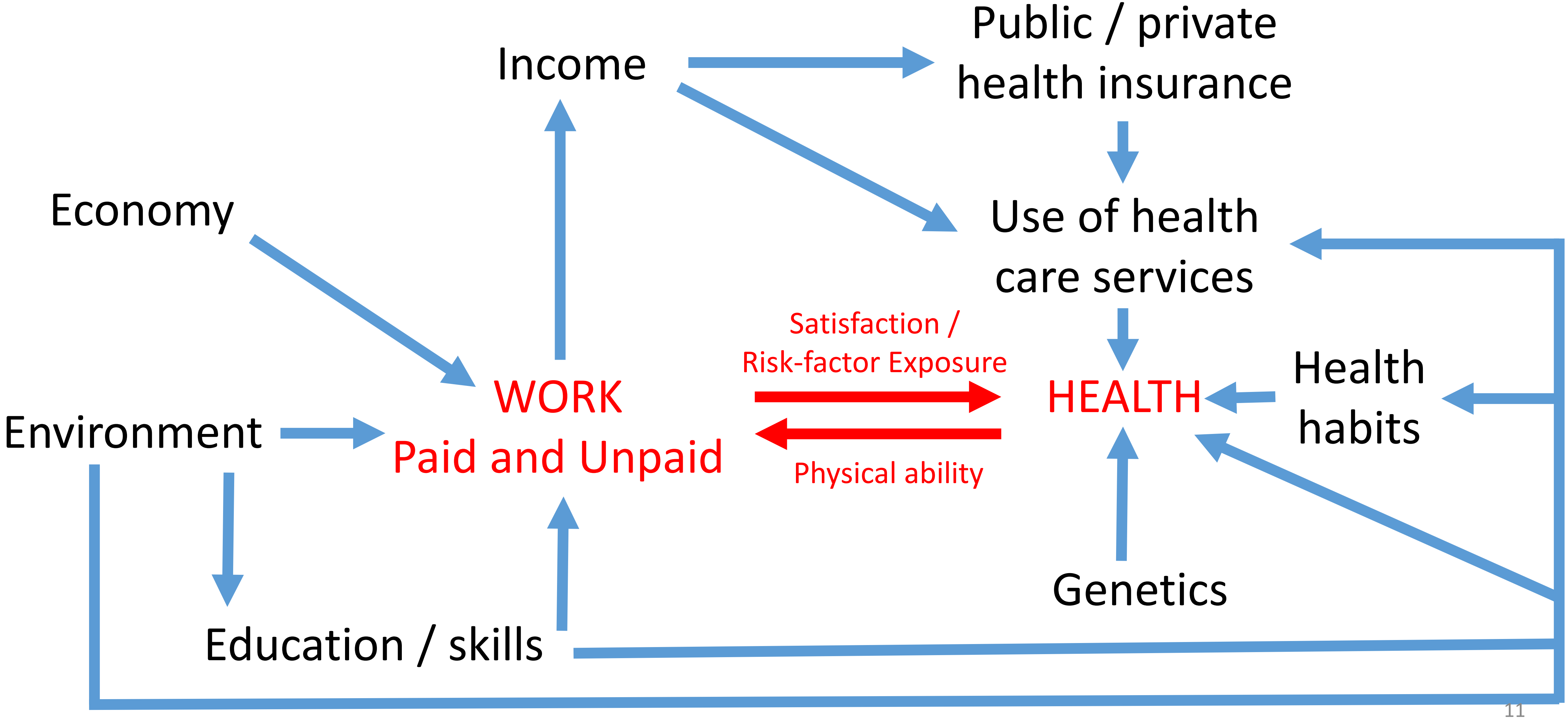


How omitted variable bias (OVB) might arise

A dimension of environment is unobserved

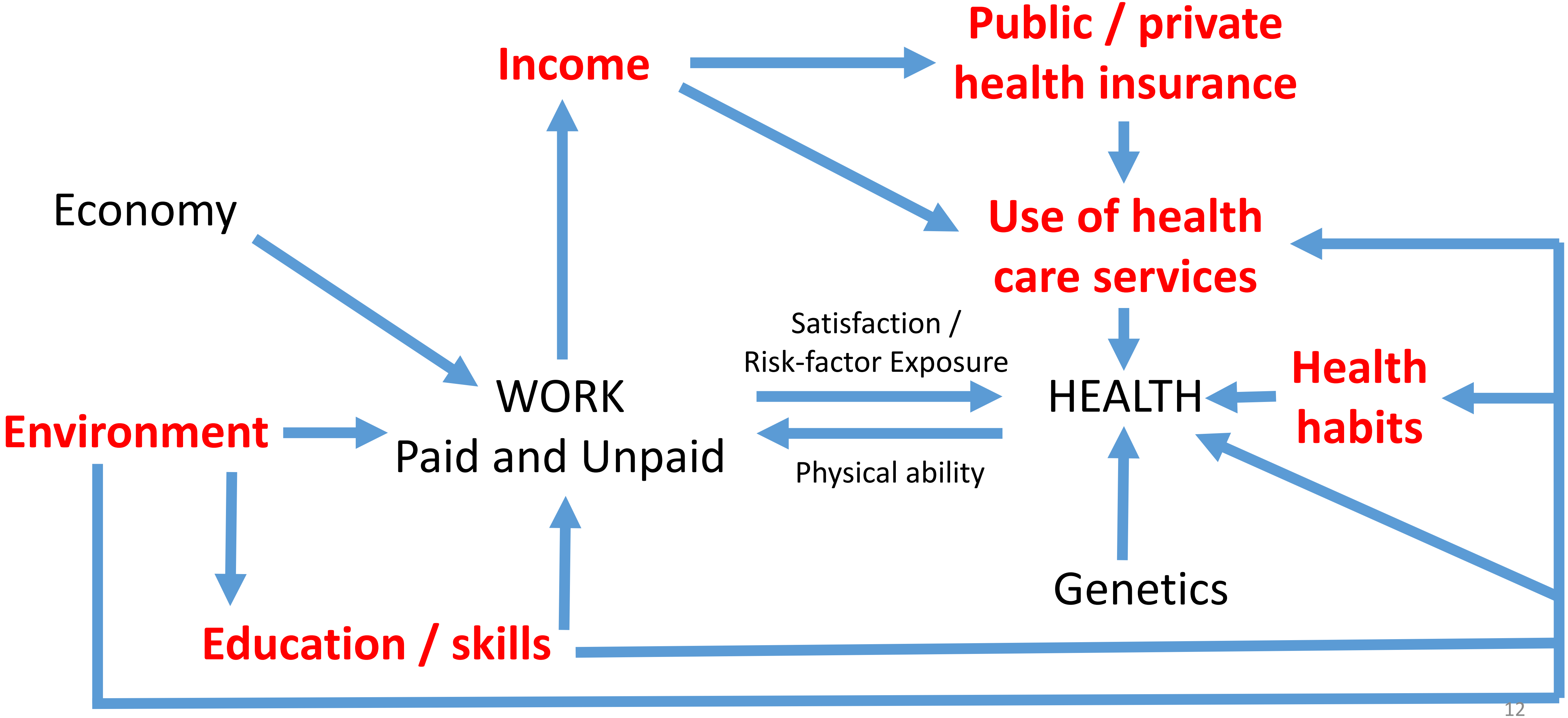


How Reverse Causality Might Arise



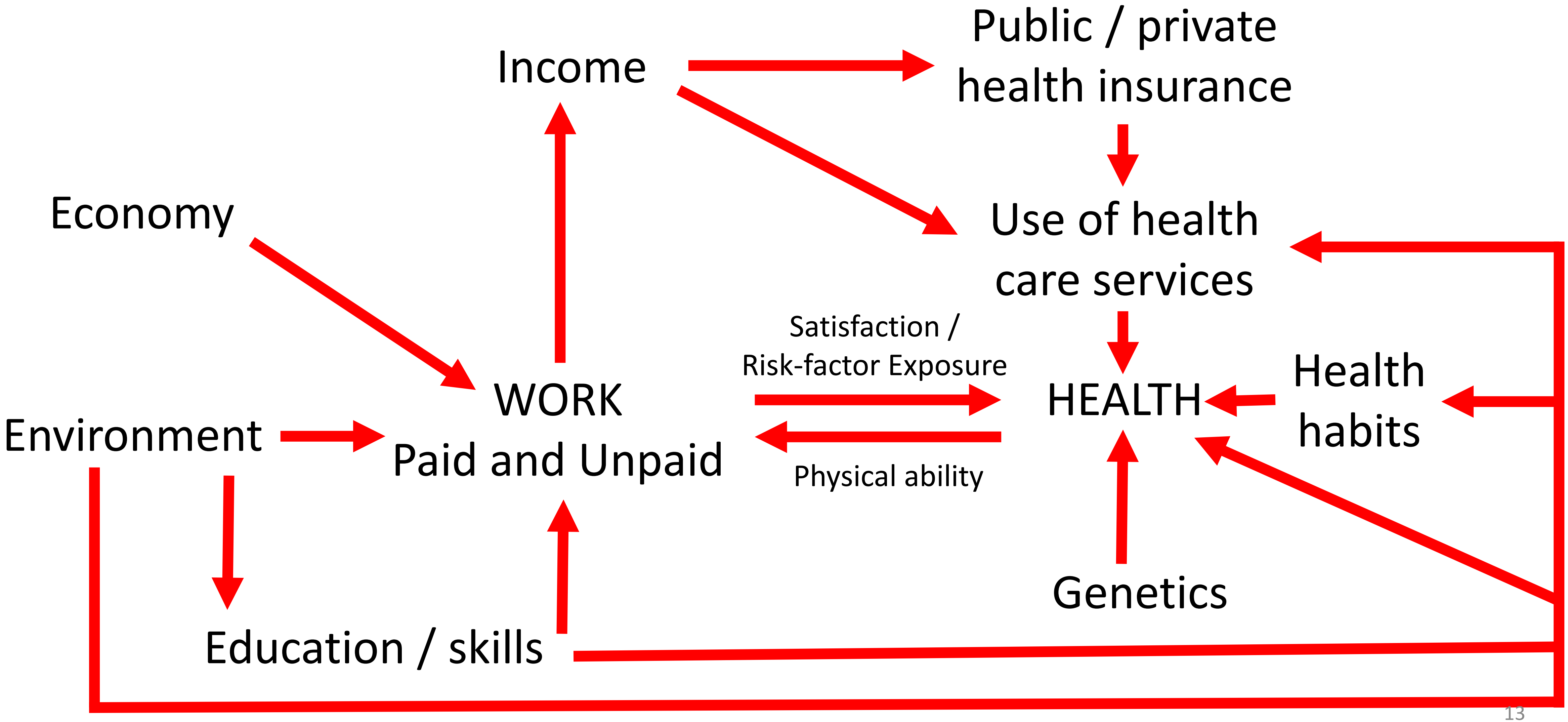
Variables subject to discrimination (in red)

Discrimination on the basis of age, race/ethnicity, and gender could alter the values of these variables.



Relationships subject to discrimination (in red)

Discrimination on the basis of age, Race/ethnicity, and gender could modify these relationships among variables.



The Blinder-Oaxaca decomposition

Laying aside the issues of OVB and reverse causality for a moment, suppose we observe that a health outcome of interest is different for men and women.

Suppose further that we have a regression equation relating the outcome of interest to some explanatory variables (**X**).

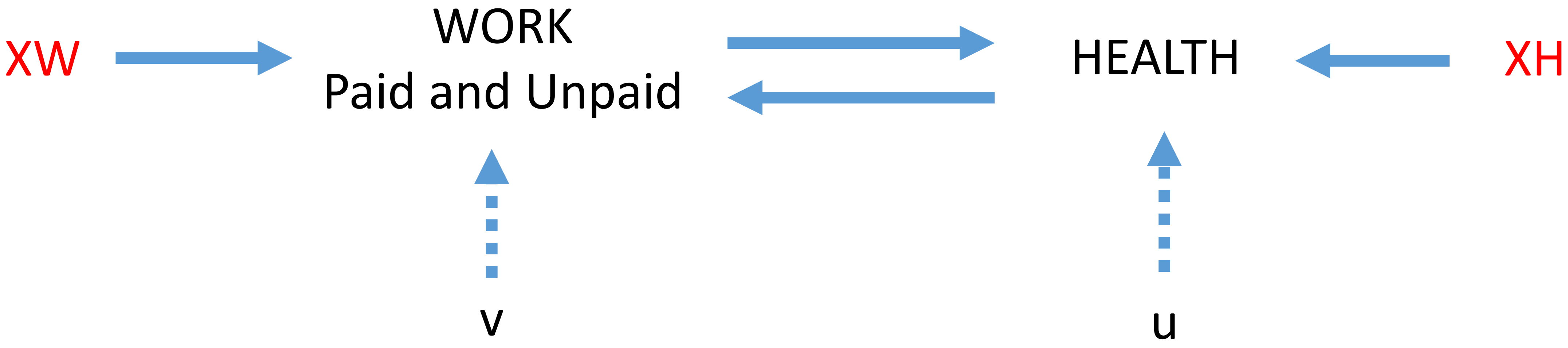
$$\text{Outcome} = \beta_0 + \boldsymbol{\beta}_{\mathbf{x}}\, \mathbf{X} + u$$

Question: Is the difference in outcomes for men and women due to differences in values of **X**, or is it due to differences in the βs?

The Blinder-Oaxaca approach allows us to decompose the difference in outcomes into those two components.

Correcting for OVB and/or Reverse Causality

The essential idea is to have a variable (instrument) that changes the value of ("shocks") the (endogenous) explanatory variable, but has no direct effect on the outcome variable.



XW "shocks" WORK, but has no direct effect on HEALTH.
XH "shocks" HEALTH, but has no direct effect on WORK.

In theory, XW and XH could be *randomization,* but…

Problems with RCTs

- They're expensive and time consuming.
- Randomization often is technically difficult (e.g., randomizing people to "HEALTH")
- Randomization can be unethical (denying patients a plausibly effective treatment)
- Pearl (*Causality,* 2000) has constructed a dataset with heterogeneous treatment effects in which an RCT can't distinguish between a treatment that saves your life or kills you.

Problems with RCTs (cont.)

- The most important question is: "To whom do you want to draw inference?"

- If you plan to administer a successful intervention randomly to the population or mandate it for everyone, then the average treatment effect (ATE) from a well-conducted RCT may give you valuable information.

- If you plan to allow individuals to opt into the treatment, then a research design based on observational data may provide a more accurate estimate of the effect of the treatment on the people who voluntarily select into the treatment group (average effect of the treatment on the treated, or ATT).

Estimation Strategies

<u>Methods that control for **both** OVB and reverse causality</u>:

- Randomized controlled trials (RCTs)*
- Instrumental variables (including simple ratios, natural experiments, two- or three-stage least squares)*

<u>Additional methods that control **only** for OVB</u>:

- Sample selection models*
- Regression discontinuity*
- Difference-in-differences (as long as omitted variables are time-invariant)

<u>Methods that **don't** control for **either** OVB or reverse causality</u>:

- Matching on observables (propensity scores, entropy balancing, synthetic controls, inverse probability weighting)*

Vector autoregression (VAR) is helpful for exploring reverse causality in panel data in the absence of omitted variable bias.

* = Needs only cross-sectional data

Instruments

Instruments can be:

Found:

- Naturally occurring events
  - Earthquakes, floods, famines, rainfall
  - Legislated HEALTH policy changes (Medicaid expansion, anti-discrimination legislation, wage subsidies)
- Unrelated events
  - The Vietnam draft lottery randomized young men to military service.
  - Choice of residence made without regard to distance to providers offering different treatment (controversial)
  - Distance to, or cost of, a job training program.

Made:

- Randomization by the analyst. (Imperfect randomization can become an instrument!)
- Encouragement design (letters of encouragement to participate in the treatment group that are sent to a random sample of potential participants).
- Other encouragement, e.g., discount coupons for a job training program.
- Crossing a threshold (regression discontinuity)

Choice of IV method can depend on the type of instrument (binary vs continuous).

A Comparison of Estimation Strategies

WORK → HEALTH

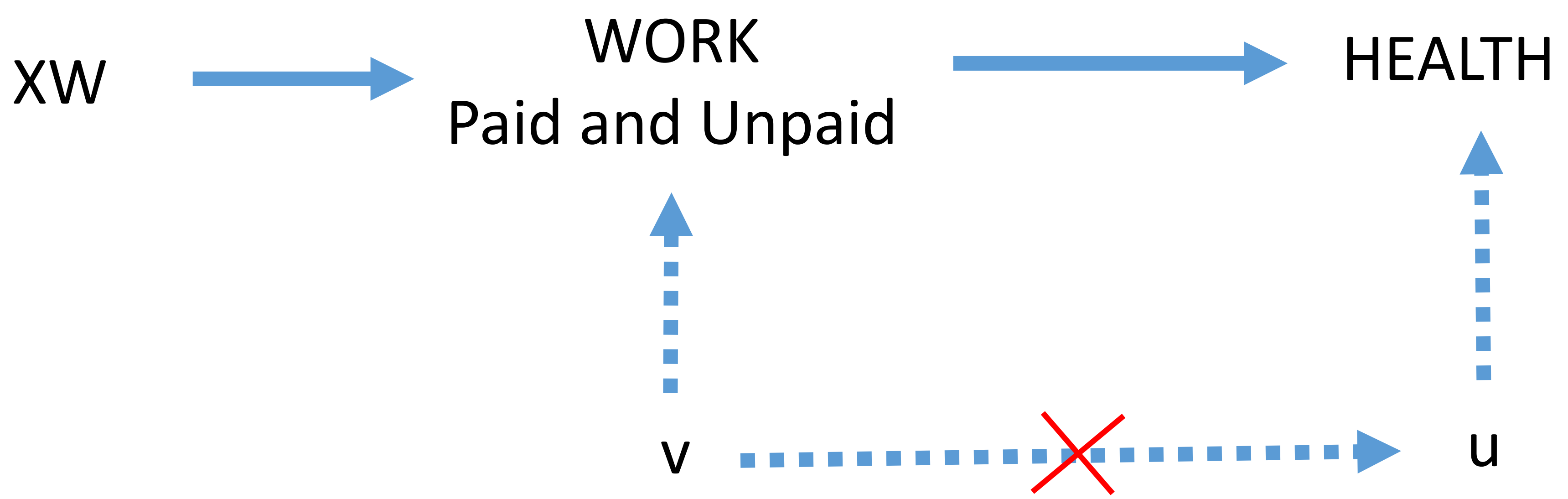
1. Simple IV ratios

2. Sample selection models

3. Two and three stage least squares

4. Regression discontinuity

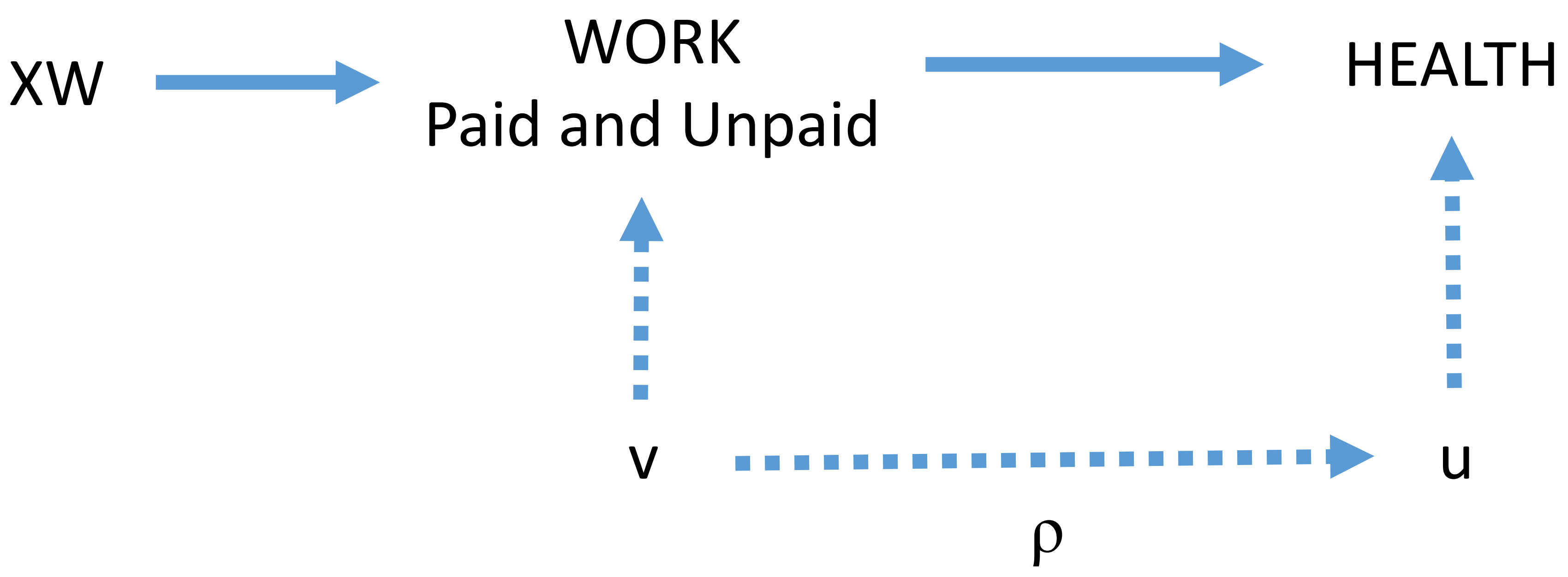
1. A Simple IV Ratio



Effect of WORK on HEALTH when XW is binary (and uncorrelated with u):

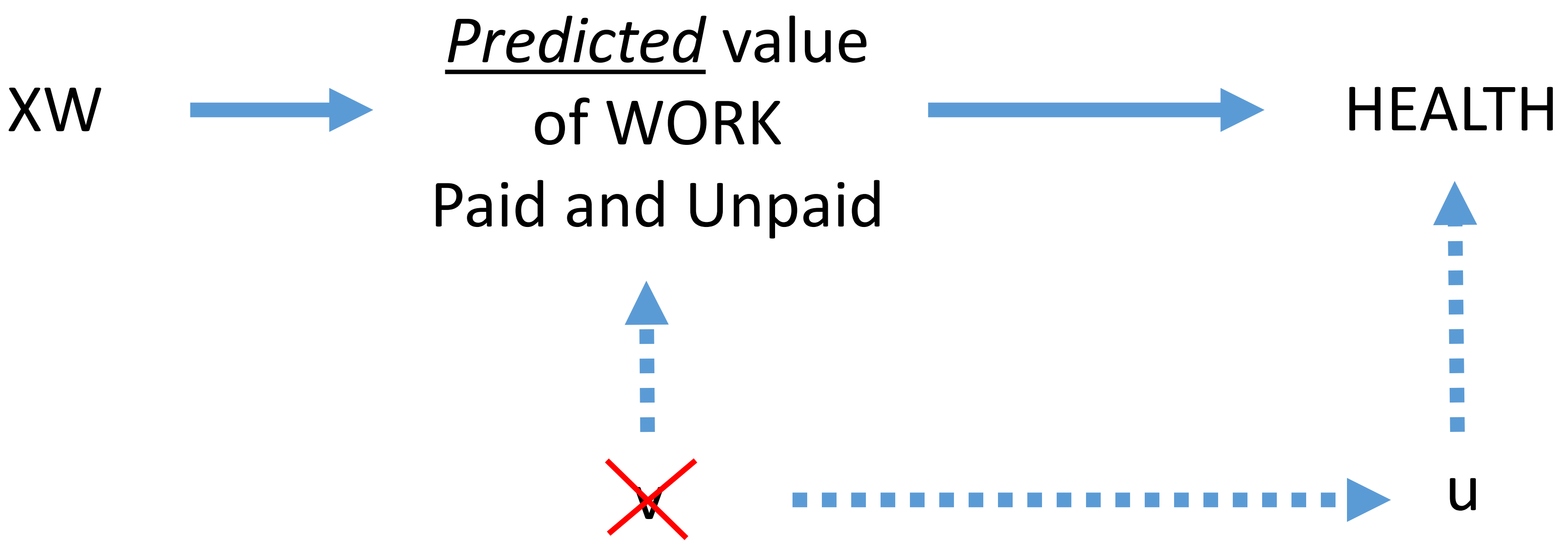
$$\frac{(HEALTH \mid XW=1) - (HEALTH \mid XW=0)}{(WORK \mid XW=1) - (WORK \mid XW=0)}$$

2. Sample Selection Models



The correlation of v and u (ρ) is estimated along with the other parameters in the model.

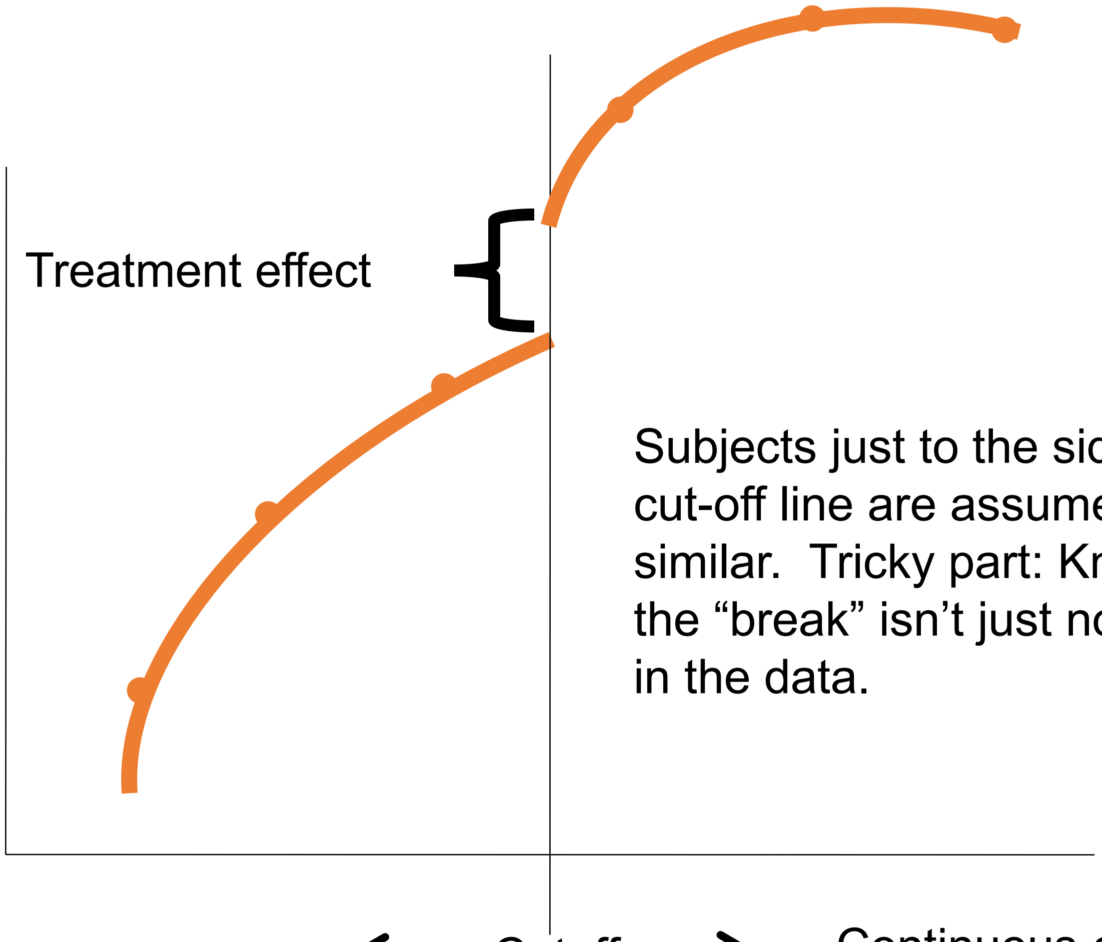
3. Two- and Three-stage Least Squares



The *predicted* value of WORK is purged of the effects of v.

4. Regression discontinuity

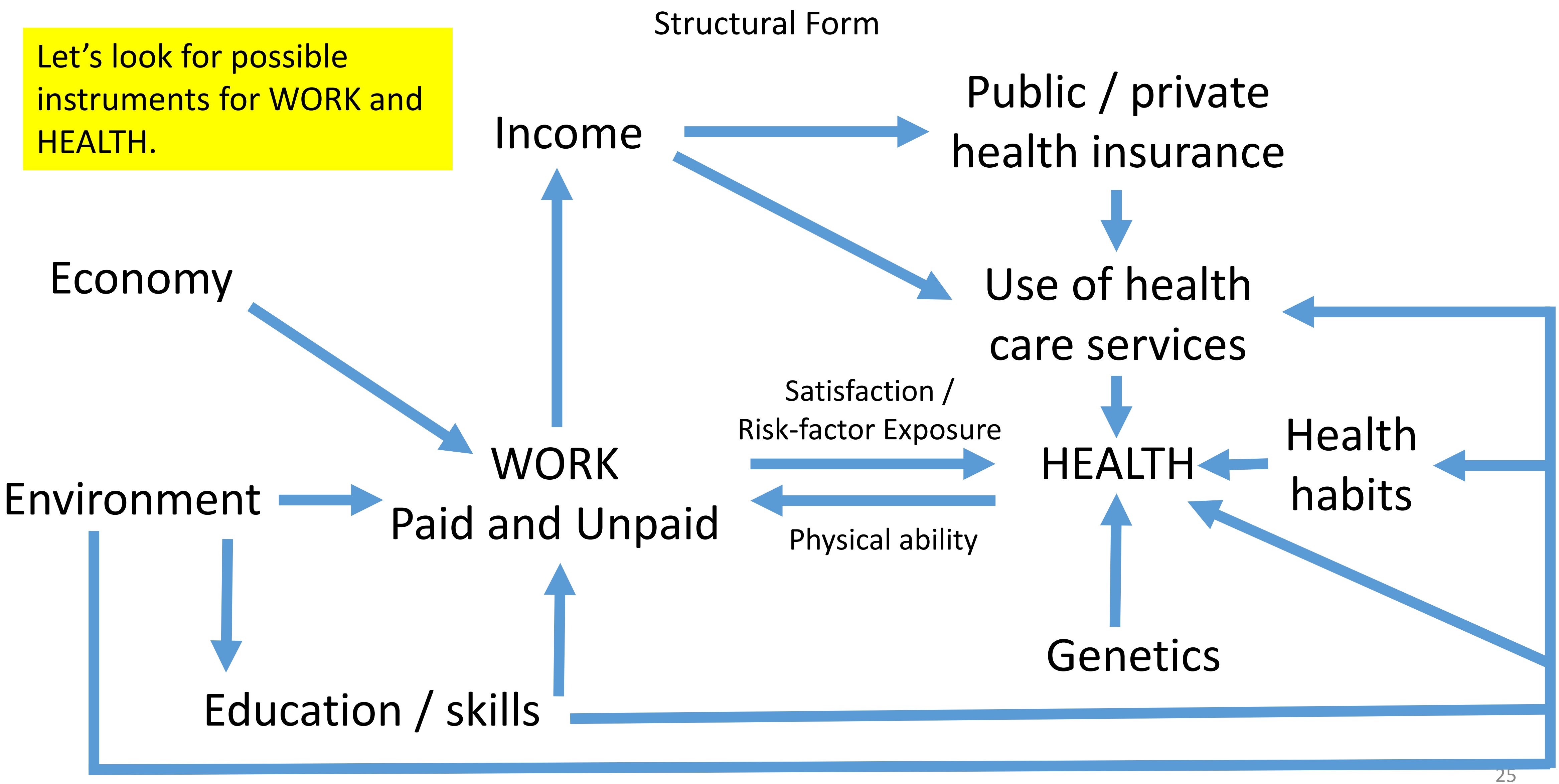
Evaluation of an intervention



Subjects just to the side of the cut-off line are assumed to be similar. Tricky part: Knowing that the “break” isn’t just non-linearity in the data.

Returning to Our Model

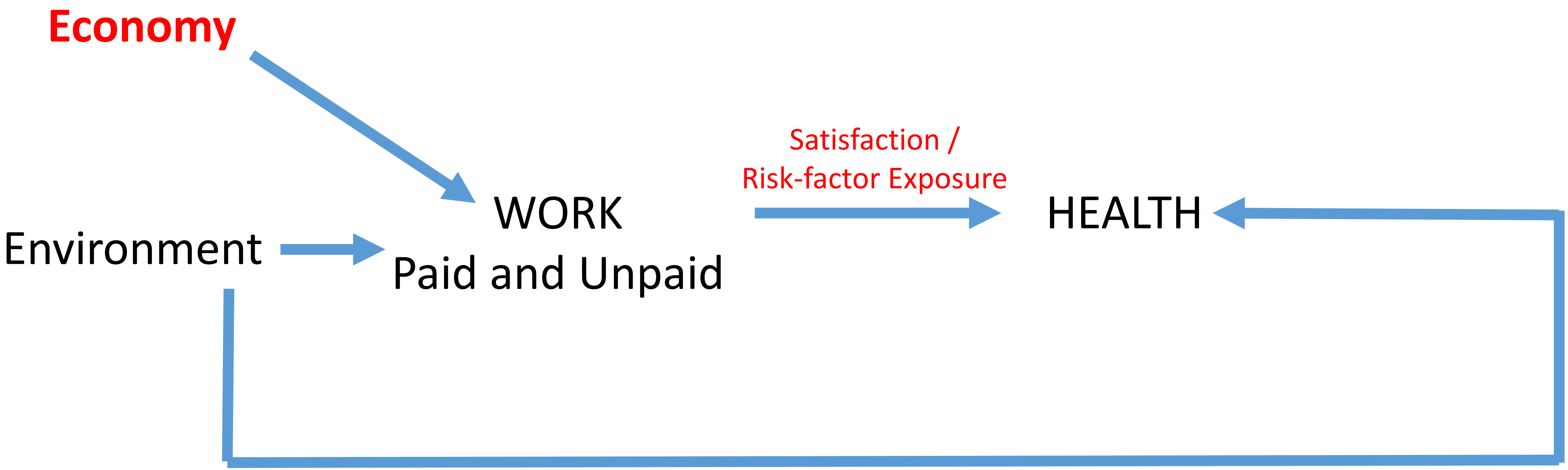
Let's look for possible instruments for WORK and HEALTH.



WORK → HEALTH

WORK → HEALTH is subject to reverse causality (and possibly omitted variable bias).

Economy is a good instrument for WORK because Economy has no direct effect on health.
Environment is not a good instrument. Environment has a direct effect on HEALTH.



HEALTH → WORK

HEALTH → WORK is subject to reverse causality and possibly omitted variable bias.

Genetics is a good instrument for HEALTH but may be difficult to separate from race/ethnicity and gender.

WORK
Paid and Unpaid

Physical ability

HEALTH

Genetics

Some helpful references

Bertrand, Marianne, Duflo Esther, and Sendhil Mullainathan. "How Much Should We Trust Differences-in Differences Estimates?" *Quarterly Journal of Economics,*119:1 (February 2004) 249-275.
Pearl, Judea. 2000. Causality: Models, Reasoning and Inference. Cambridge, U.K.: Cambridge University Press.

Bhattacharya, J. and W. Vogt. "Do Instrumental Variables Belong in Propensity Scores?" NBER Technical Working Paper 343. http://www.nber.org/papers/t0343

Brooks JM and RL Ohsfeldt. "Squeezing the Balloon: Propensity Scores and Unmeasured Covariate Balance," *Health Services Research* 48:4 (August 2013) 1487-1507

Cook, T.D. "Waiting for Life to Arrive": A history of the regression-discontinuity design in Psychology, Statistics and Economics. Journal of Econometrics 142:2 (2008) 636-654.

Daw, Jamie R. and Laura A. Hatfield. "Matching in Difference in Differences: between a Rock and a Hard Place," *Health Services Research* 53:6 (December 2018) 4111-4117

Harris, Katherine and D.K. Remler. "Who is the Marginal Patient?" Health Services Research 33:5 Part 1, (December 1998) 1337-1360.

Some helpful references (cont.)

Holland, Paul. “Statistics and Causal Inference.” Journal of the American Statistical Association 81:396 (December 1986) 945-960.

Imbens, GW and JD Angrist. “Identification and Estimation of Local Average Treatment Effects,” Econometrica. 62:2 (1994) 467-475.

Imbens, Guido W. and Donald B. Rubin. “Estimating Outcome Distributions for Compliers in Instrumental Variables Models,” Review of Economic Studies 64 (1997) 555-574.

King G and R Nielsen. “Why propensity scores should not be used for matching,” *Political Analysis* (2019) DOI: 10.1017/pan.2019.11.

Kline P and Walters CR. “On Heckits, LATE, and numerical equivalence,” Econometrica 87:2 (March 2019) 677-696.

Pearl, J. *Causality*. Cambridge University Press, 2000.

Ryan, Andrew M., Burgess, James F., Dimick, Justin B. “Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences,” *Health Services Research* 50:4 (August 2015) 1211-1235.

Supplementary information

Terminology

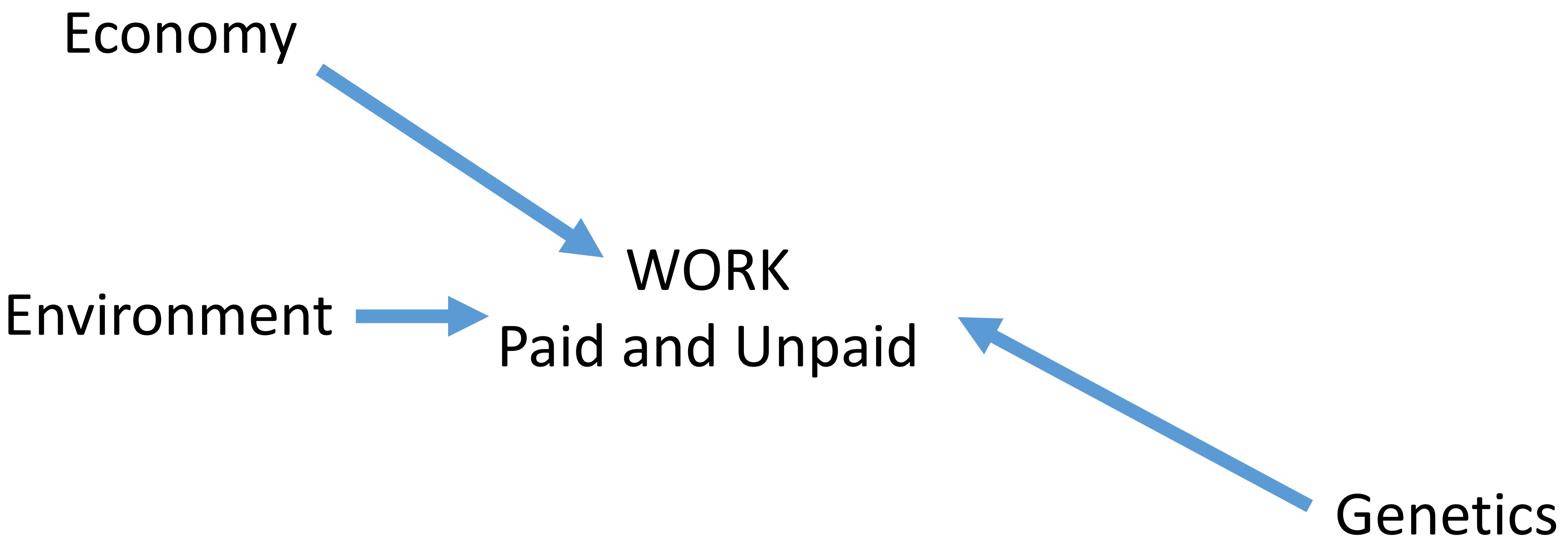
- A <u>structural form</u> equation includes mediators. **Work** is <u>exogenous</u>, or "predetermined." **Risk Factor Exposure** and **Health** are <u>endogenous</u>.

Work $\longrightarrow$ Risk Factor Exposure $\longrightarrow$ Health

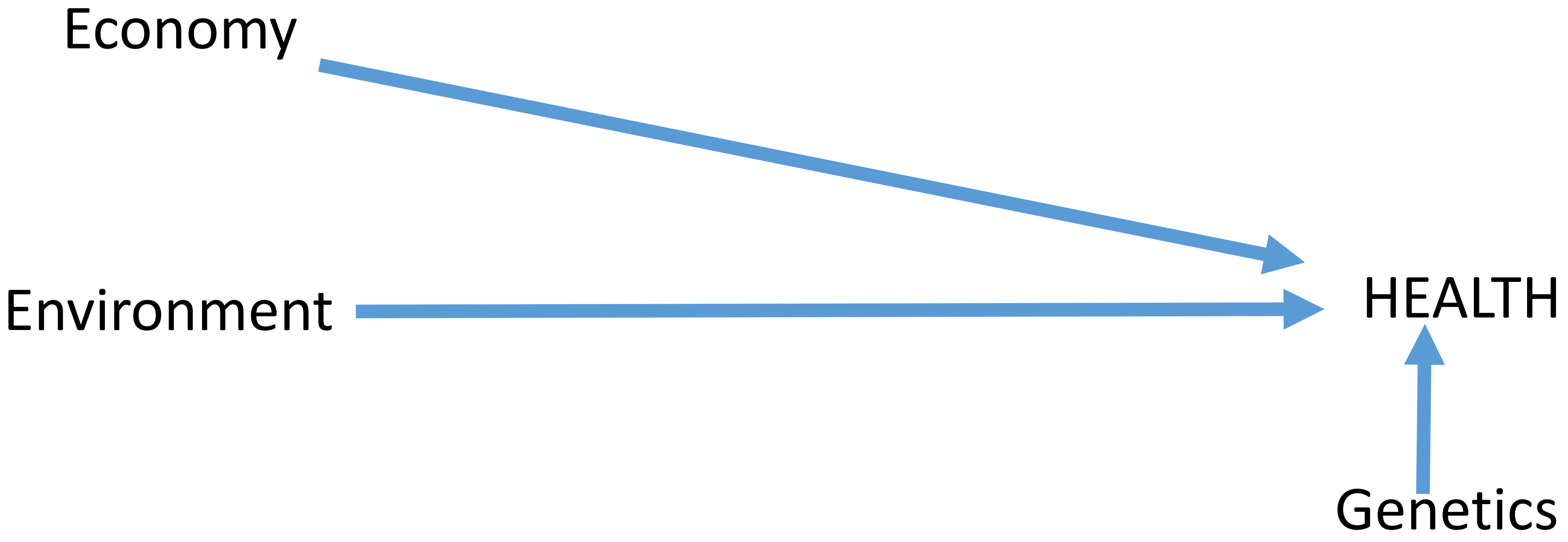
- A <u>reduced form</u> equation (all the explanatory variables are exogenous).

Work $\longrightarrow$ Health

Reduce form WORK equation



Reduced form HEALTH equation



A methodological controversy about disparity variables

- Holland and Rubin's insistence that causal effects can be investigated only for <u>manipulatible</u> variables seems designed to advance the argument that experimental research designs (randomized trials) are superior to observational data. Or that searching for "effects of causes" is superior to searching for "causes of effects." You find the same controversy in the econometrics literature.

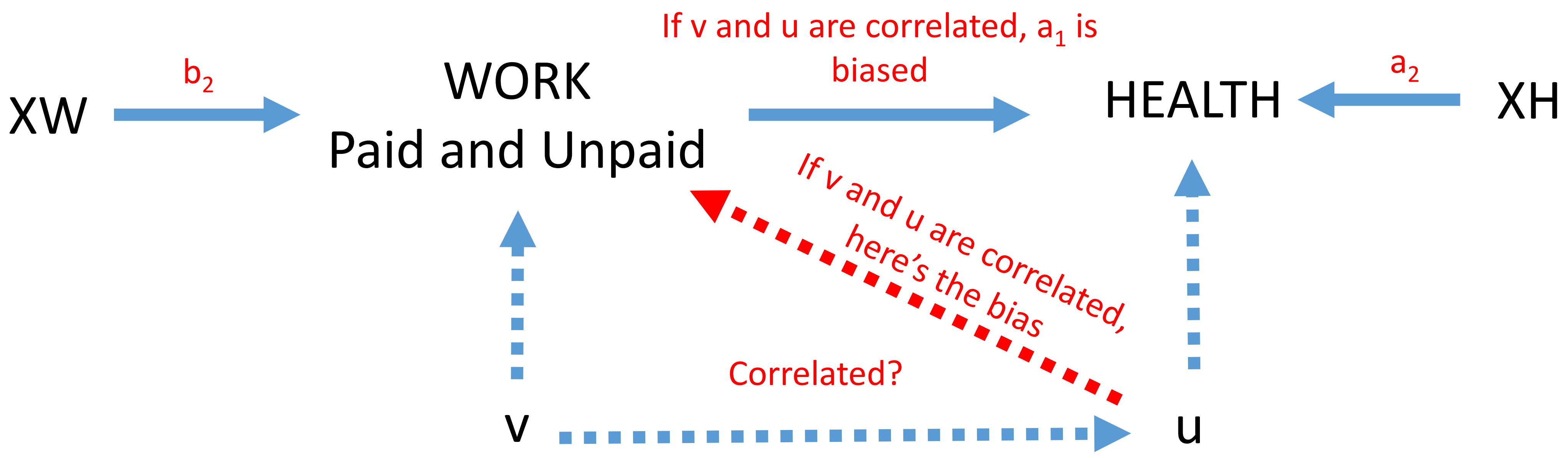
- But few people would deny that (non-manipulatable) earthquakes or pandemics have causal effects on outcomes of interest that can be empirically investigated. Or that race carries a genetic causal component. Or that age or gender carry physiological causal components.

- And in dire situations, we often search desperately for causes of effects (Legionnaires disease).

Why OVB produces biased estimates

Every variable with an arrow pointing to it has its own equation.

Bias arises when an explanatory variable is correlated with the error term.



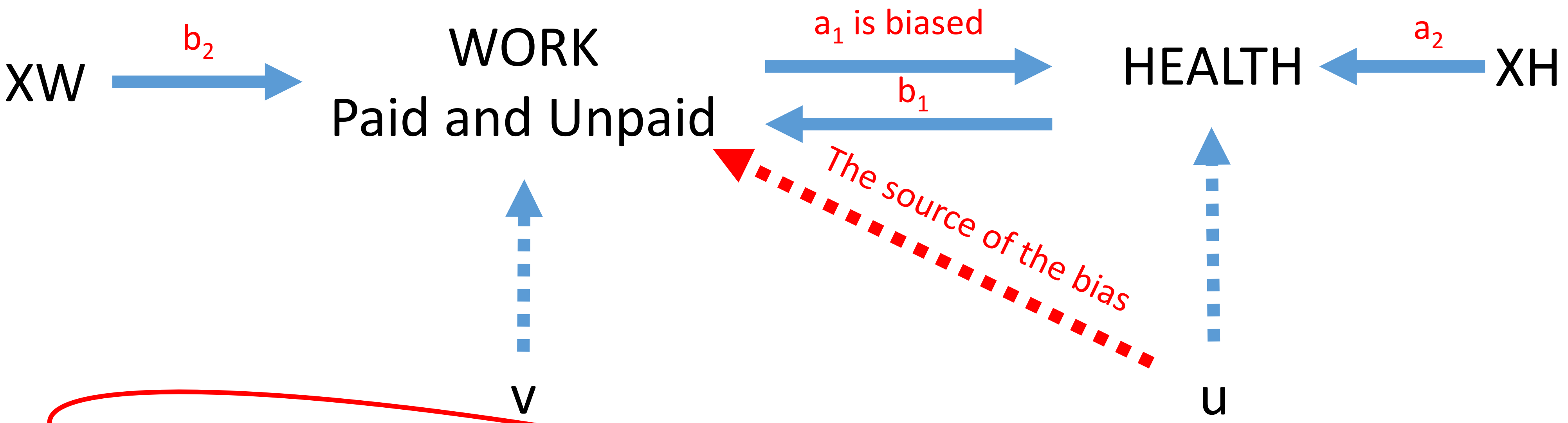
$$\text{HEALTH} = a_0 + a_1\ \text{WORK} + a_2\ \text{XH} + u$$

$$\text{WORK} = b_0 + b_1\ \text{XW} + v$$

The blue arrow shows that u is correlated with WORK.

Why reverse causality produces biased estimates

The "HEALTH" equation



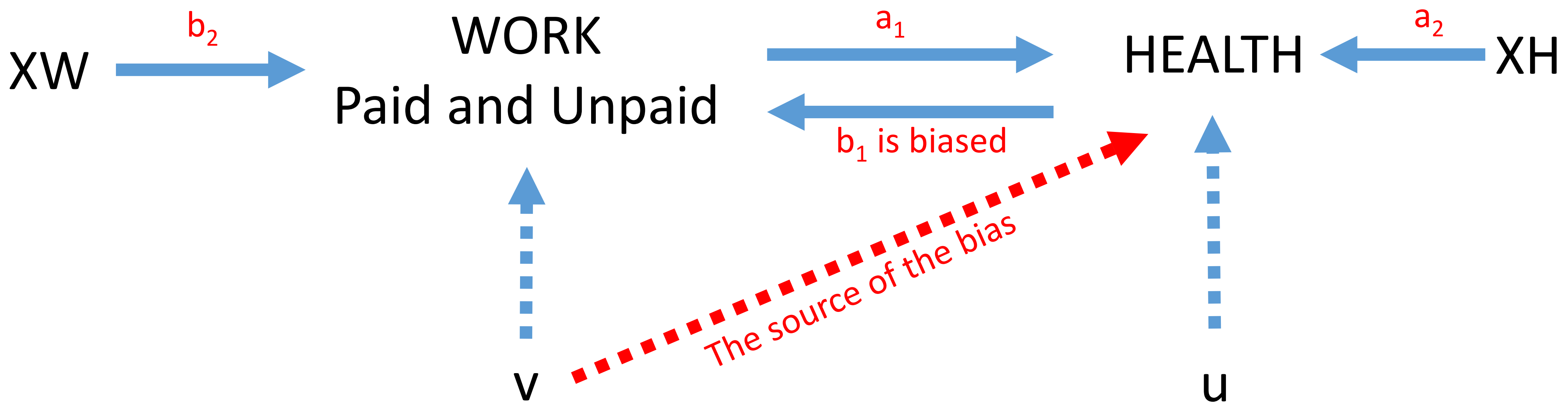
$$HEALTH = a_0 + a_1\ WORK + a_2\ XH + u$$

$$WORK = b_0 + b_1\ HEALTH + b_2\ XW + v$$

The blue arrow shows that u is correlated with WORK.

Why reverse causality produces biased estimates

The “WORK” equation



$$HEALTH = a_0 + a_1\ WORK + a_2\ XH + u$$

$$WORK = b_0 + b_1\ HEALTH + b_2\ XW + v$$

The blue arrow shows that v is correlated with HEALTH.

Instrumental Variables

There typically are 4 kinds of people in the data:

1. People who always will choose the treatment regardless of the instrument's value.
2. People who never will choose the treatment regardless of the instrument's value.
3. People for whom the value of the instrument determines whether they will choose the treatment.
4. People who choose not to get the treatment when the instrument makes it more attractive ("Defiers" usually are ruled out in IV analyses).

IV estimates the treatment effect for group 3, above, since that's the group for whom (we hope) the instrument is working like randomization. That's the *local* average treatment effect or LATE (Harris and Remler, 1998).

A good instrument is hard to find

<u>The usual problems</u>:

- The instrument for WORK (XW) doesn't really affect WORK, or at least not for some people.
- XW is correlated with u.

<u>The less usual problems:</u>

- XW doesn't affect WORK for the "right" (policy relevant) subjects, possibly because …
- The effect of WORK on HEALTH is different for different subjects – heterogeneous treatment effects – and the subjects for which XW → WORK are not the same subjects for which WORK → HEALTH.

<u>Cautions</u>:

- Poor instruments can produce more biased estimates than uncorrected estimates

But there also are problems with methods that *don't* control for OVB or reverse causality

<u>In addition to the possibility of biased estimates due to OVB</u>:

1. If you match on all *observed* variables in observational data, then people must be choosing the treatment group for *unobserved* reasons, and if those reasons are correlated with the error term in the equation of interest, matching *increases* the bias.

2. If you accidently include a good instrument in the matching variables, then you will have reduced the "good" variation in the endogenous explanatory variable (the part that is uncorrelated with the error term in the equation of interest) and emphasized the "bad" variation (the part that *is* correlated with the error term in the equation of interest).

But there also are problems with methods that *don't* control for OVB or reverse causality (cont.)

3. If the outcome variable has different means in the treatment and control groups and you match on pre-treatment values of the outcome variable, pre-treatment trends in the outcome variable, or variables related to the pre-treatment trend in the outcome variable, you will create biased difference-in-differences estimates.

The Blinder-Oaxaca decomposition

This is an approach to studying disparities.

Suppose we want to know if men are paid more than women.

Write the wage equation as:

$Y_i = X_i\beta + u_i$

where: Y = wages

X = explanatory variables like type of job, skill level, etc.

The Blinder-Oaxaca decomposition

$Y_i = X_i\beta + u_i$

Typically, we say that there are two reasons why women's wages might be lower than men's wages.

1. Women might have different X values than men. This is not necessarily the result of discrimination, but you have to be careful. We should ask, "Why are the X values different for men and women?"

2. The coefficients on the X variables might be different for women than for men. This is sometimes referred to as "returns to skill" and differences often are viewed as discrimination.

The Blinder-Oaxaca decomposition

Following Jann's notation and letting women=A and men=B, we can write that difference in mean wages as R, where:

$$R = E(Y_A) - E(Y_B) = E(X_A)\beta_A - E(X_B)\beta_B$$

Which can be rewritten as:

$$R = \{E(X_A) - E(X_B)\}\beta_B + E(X_B)(\beta_A - \beta_B) + \{E(X_A) - E(X_B)\}\,(\beta_A - \beta_B)$$

The Blinder-Oaxaca decomposition

$$R = \{E(X_A) - E(X_B)\}\beta_B + E(X_B)(\beta_A - \beta_B) + \{E(X_A) - E(X_B)\}\,(\beta_A - \beta_B)$$

Let's look at each part of this expression:

$$E\ (\text{“}endowment\ effect\text{”}) = \{E(X_A) - E(X_B)\}\beta_B$$

This part represents the difference in wages dues to different Xs, holding the men's βs constant.

$$C\ (\text{“}coefficient\ effect\text{”}) = E(X_B)(\beta_A - \beta_B)$$

This part represents the difference in wages due to different βs, holding the men's Xs constant.

The Blinder-Oaxaca decomposition

$$I\ (\text{“}interaction\ effect\text{”}) = \{E(X_A) - E(X_B)\}\ (\beta_A - \beta_B)$$

This part represents an interaction of different X values and different βs.

So the difference in means can be expressed as:

$$R = C + E + I$$

And:

$$\hat{R} = \bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\hat{\beta}_B + \bar{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right) + (\bar{X}_A - \bar{X}_B)\left(\hat{\beta}_A - \hat{\beta}_B\right)$$

The Blinder-Oaxaca decomposition

$$\hat{R} = \bar{Y}_A - \bar{Y}_B = (\bar{X}_A - \bar{X}_B)\hat{\beta}_B + \bar{X}_B(\hat{\beta}_A - \hat{\beta}_B) + (\bar{X}_A - \bar{X}_B)(\hat{\beta}_A - \hat{\beta}_B)$$

So we run the men's wage equation and the women's wage equations separately to obtain estimates of $\hat{\beta}_A$ and $\hat{\beta}_B$. Then we apply the men's coefficients to the women and the men's Xs to the women's βs. That gives us the components we need for the decomposition.

The decomposition is canned in Stata as *oaxaca*.

Ben Jann (2008) "The Blinder–Oaxaca decomposition for linear regression models," *The Stata Journal* 8:4; 453–479.

The Blinder-Oaxaca decomposition

The decomposition is somewhat more complex for non-linear models like logit and probit because you can't just plug in the means of the Xs. The decomposition has to be computed separately for each person and the results averaged.

Stata code: *nldecompose*

Mathias Sinning, Markus Hahn, and Thomas K. Bauer. (2008) "The Blinder–Oaxaca decomposition for nonlinear regression models," *The Stata Journal.* 8:4; 480–492.