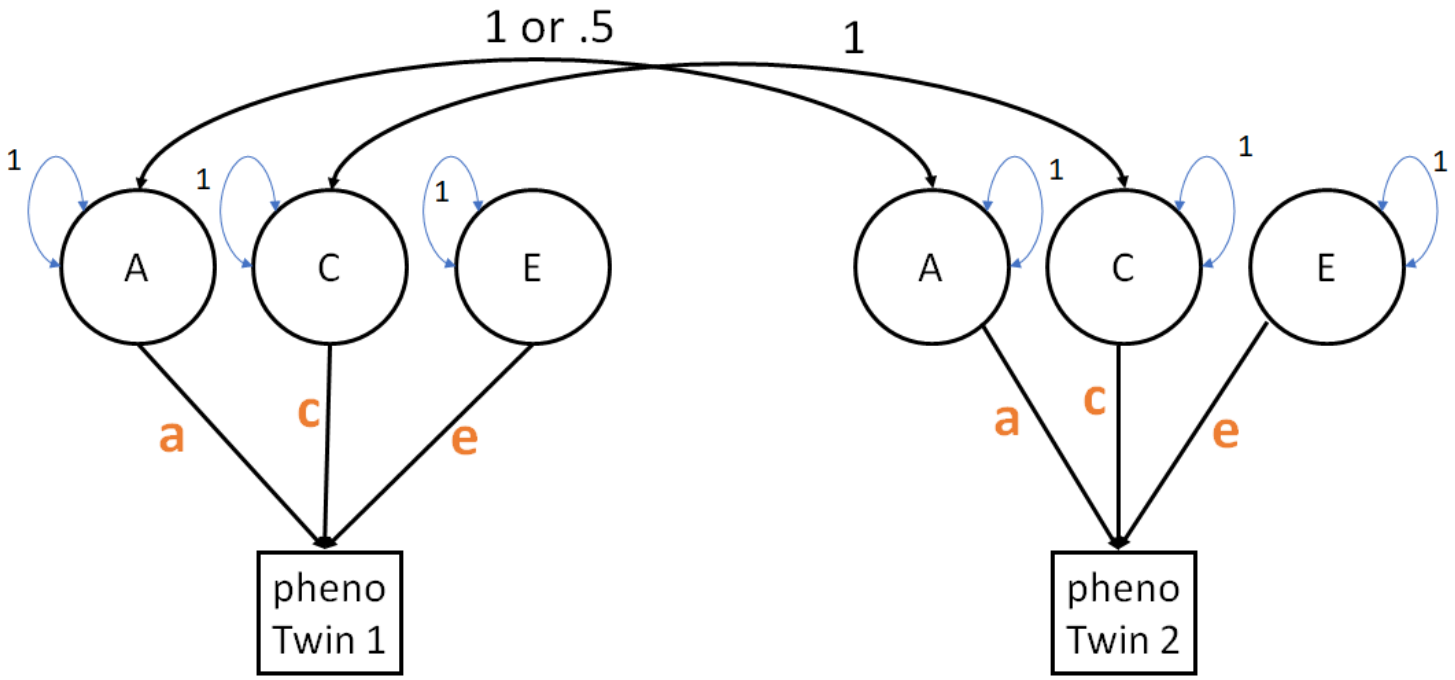
Introducing polygenic risk scores into the twin design

Estimating r_{AC}
and
MR twin model

Conor V. Dolan

Boulder 2020



What is var(A) again ? (in a simple model assuming linkage equilibrium)

(in a simple model assuming linkage equilibrium)

$pheno_i = \mathbf{a_0} + \mathbf{a_1}*QTL_{A1i} + \mathbf{a_2}*QTL_{A2i} + \ldots + \mathbf{a_N}*QTL_{ANi} + e_i$

$var(A) = \boldsymbol{s^2_{Ph\_QTL(A)}} = \mathbf{a}_1^2*s^2_{QTLA1} + \mathbf{a}_2^2*s^2_{QTLA2} + \ldots + \mathbf{a}_N^2*s^2_{QTLAN}$

Suppose you have measured QTL_1 to QTL_{10} and
you have reliable GWAS estimates of $\mathbf{a_1}$ to $\mathbf{a_{10}}$: $\mathbf{\hat{a}_1}$ to $\mathbf{\hat{a}_{10}}$

$PRS_i = \mathbf{\hat{a}_1}*QTL_{A1i} + \mathbf{\hat{a}_2}*QTL_{A2i} + \ldots + \mathbf{\hat{a}_{10}}*QTL_{ANi}$

1 to 10

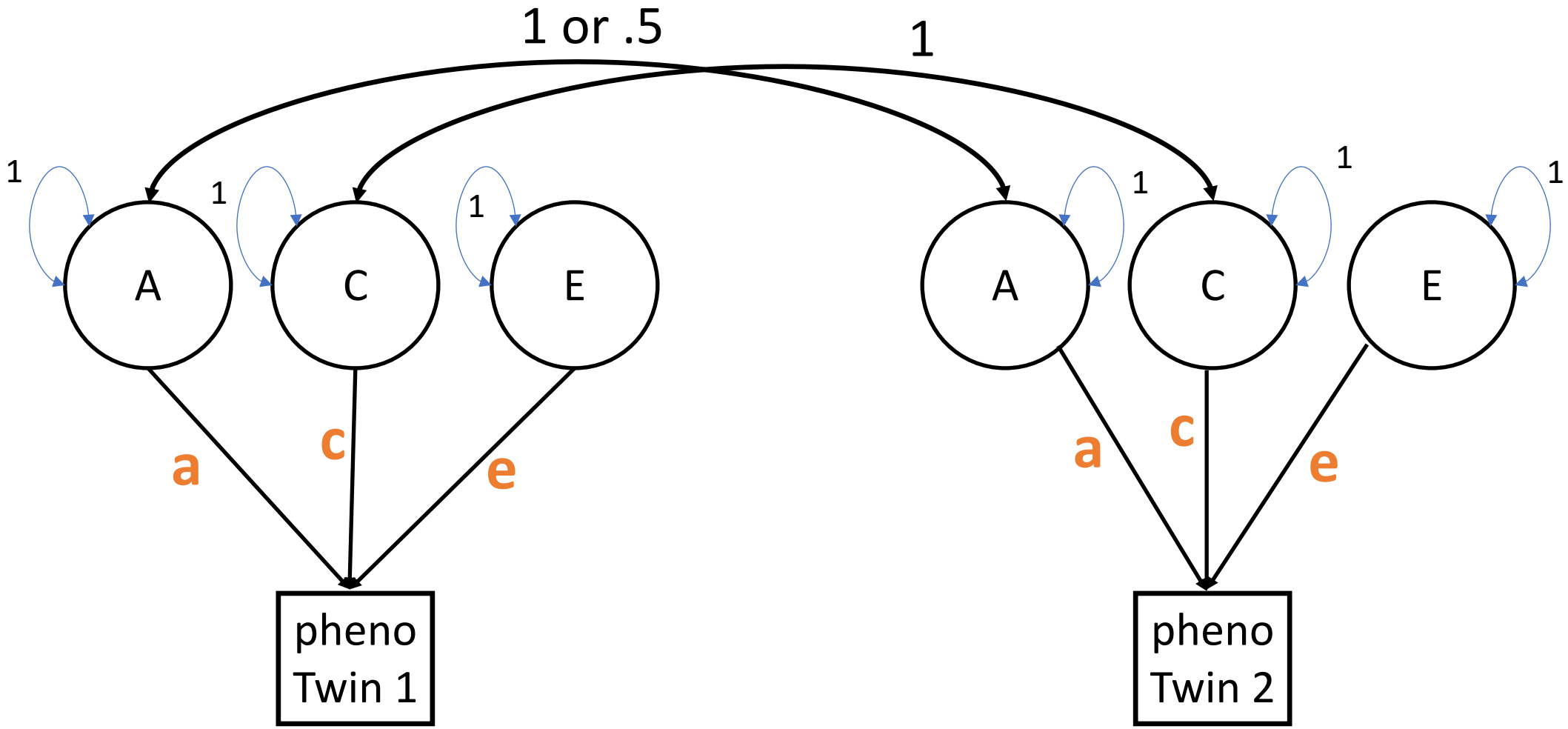
$var(A^*) = \mathbf{a}_{11}^2*s^2_{QTLA11} + \mathbf{a}_{12}^2*s^2_{QTLA12} + \ldots + \mathbf{a}_N^2*s^2_{QTLAN}$

11 to N

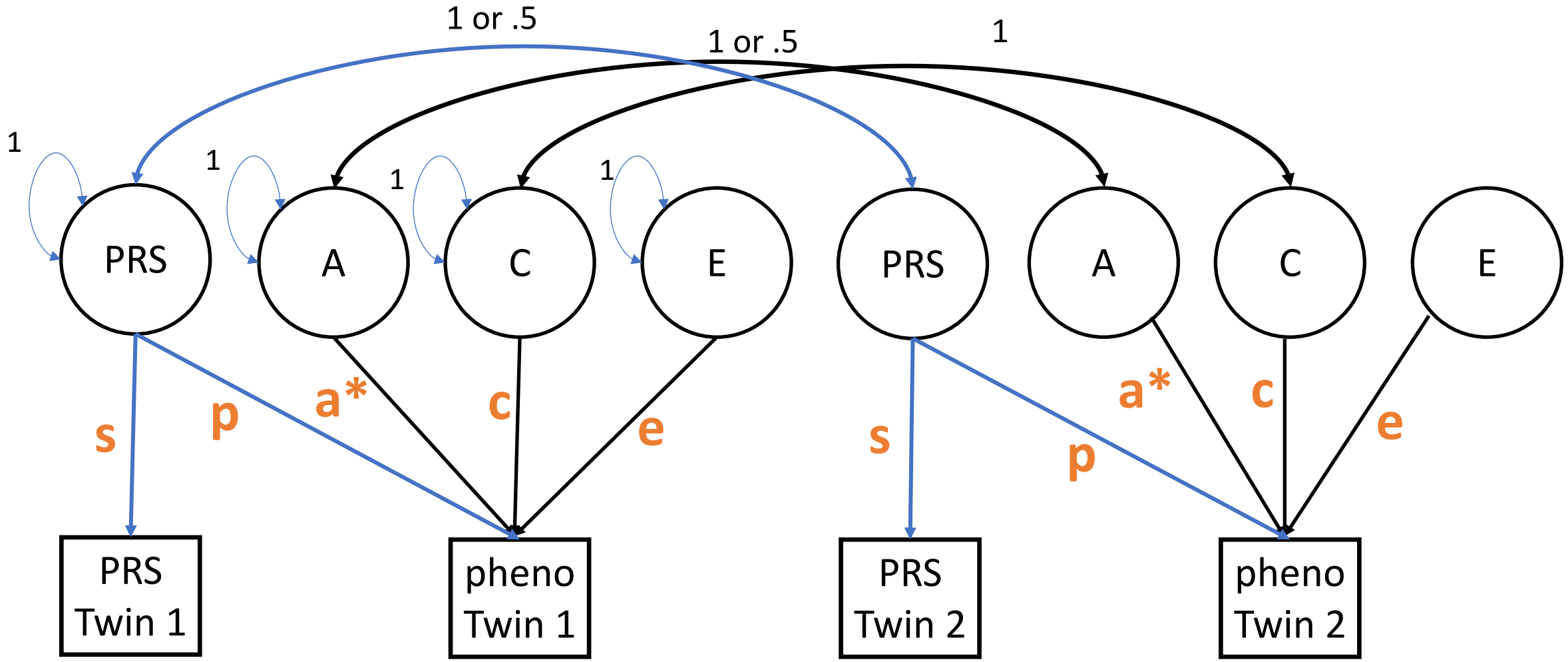
$var(A) = var(P) + var(A^*)$

The risk score PGS renders part of the **latent** A **observable**.

Include polygenic risk scores in the ACE twin model to estimate r_{AC}

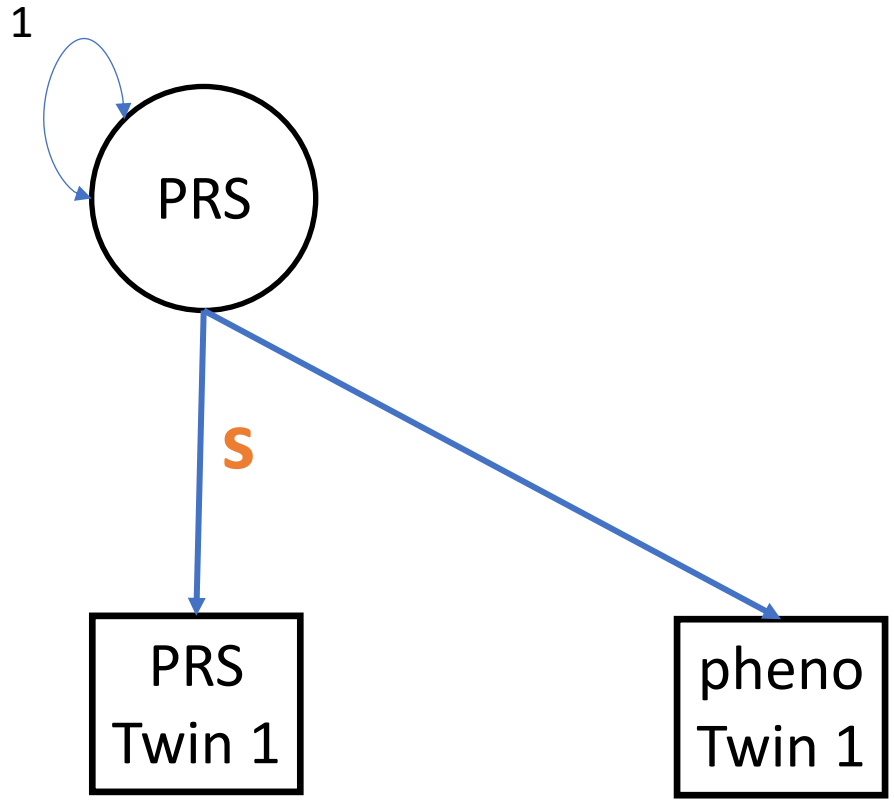


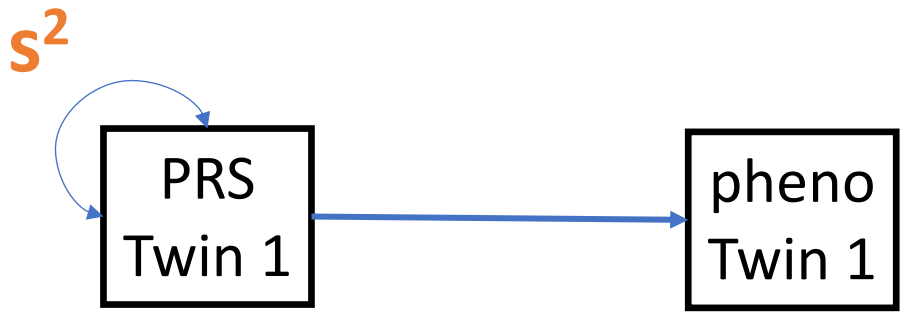
$$var(pheno) = a^2 + c^2 + e^2$$



$$var(pheno) = a^{*2} + p^2 + c^2 + e^2$$

Identical models





the standard deviation of PRS is s

the variance of PRS is s^2

Nmz = Ndz = 1000 exact data simulation

Standard model ACE

	LB	est	UB
ACE.A[1,1]	0.490	**0.600**	0.714
ACE.C[1,1]	0.198	**0.300**	0.399
ACE.E[1,1]	0.471	**0.500**	0.532

var(pheno) = .60 + .30 + .50

Standard model ACE with P

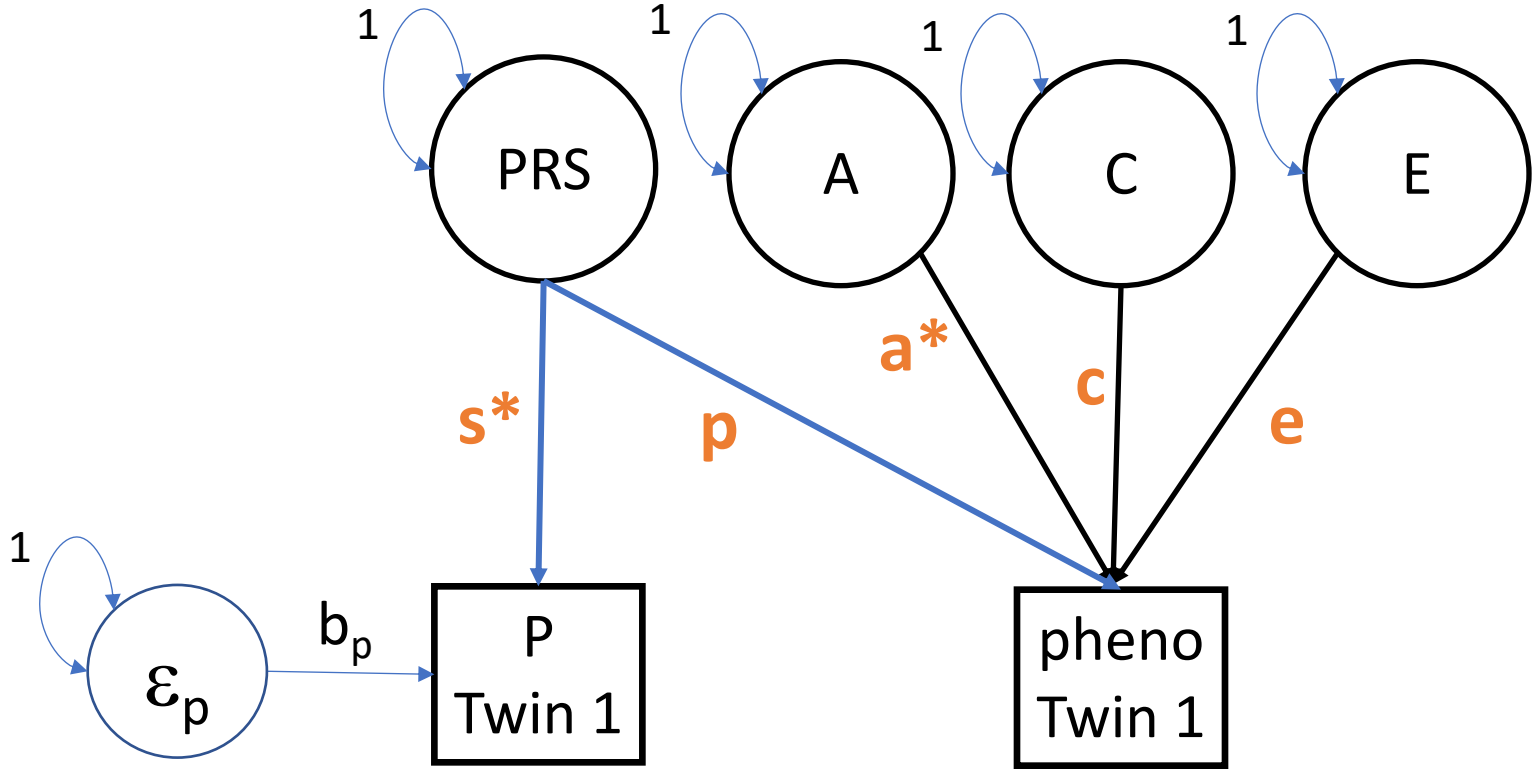
	LB	est	UB
A*+P	0.492	**0.600**	0.711
C	0.202	**0.300**	0.396
E	0.471	**0.500**	0.532
P	0.047	**0.060**	0.075
A*	0.433	**0.540**	0.650

var(pheno) = .54 + .06 + .30 + .50

.60 in standard model is decomposed into .060 (P) and .540 (A)

R^2 = .06/60 genetic level 10%

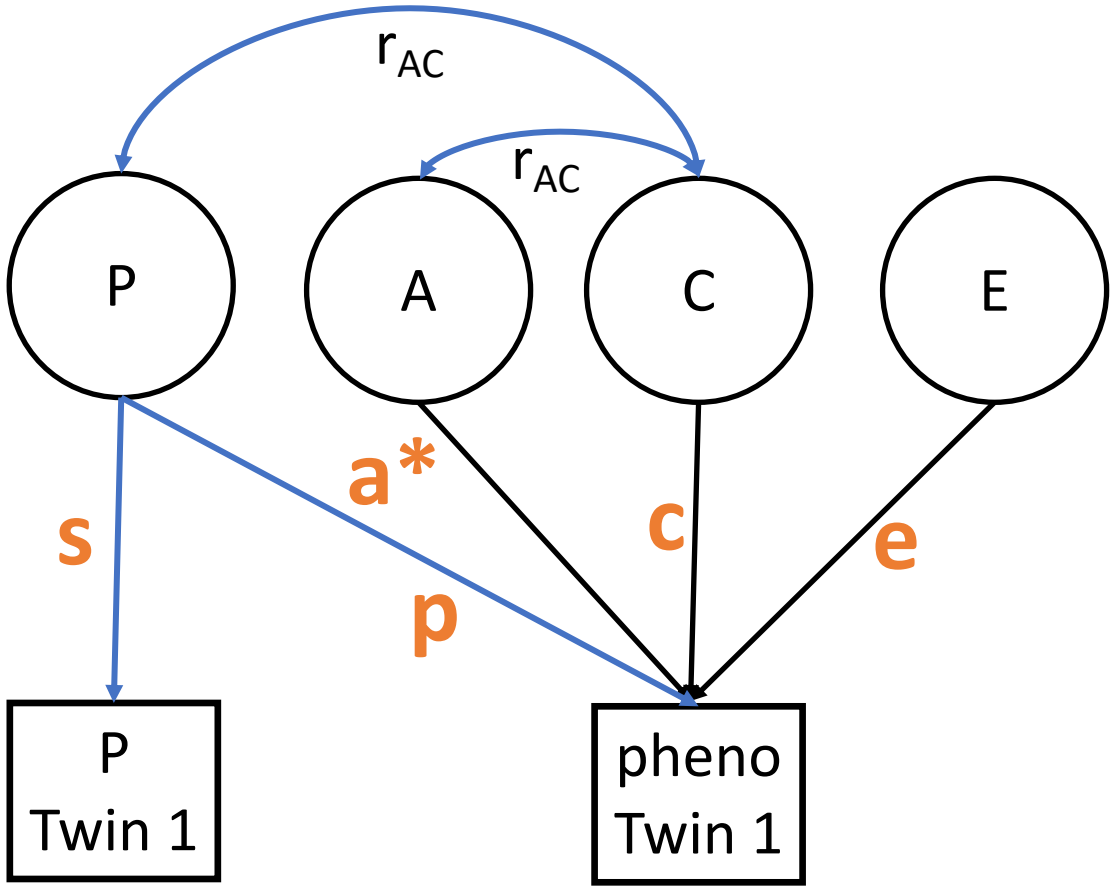
R^2 = .06/1.4 R^2 phenotypic level 4.28%



Phenotype usually includes measurement error which is absorbed by E.
Here we assume that the PRS is error free. If it is not and we know the reliability (r_{pp}) , we can include that info. The parameter b is fixed to equal:

$$b_p = \sqrt{\{var(P) * (1-r_{pp})\}}$$

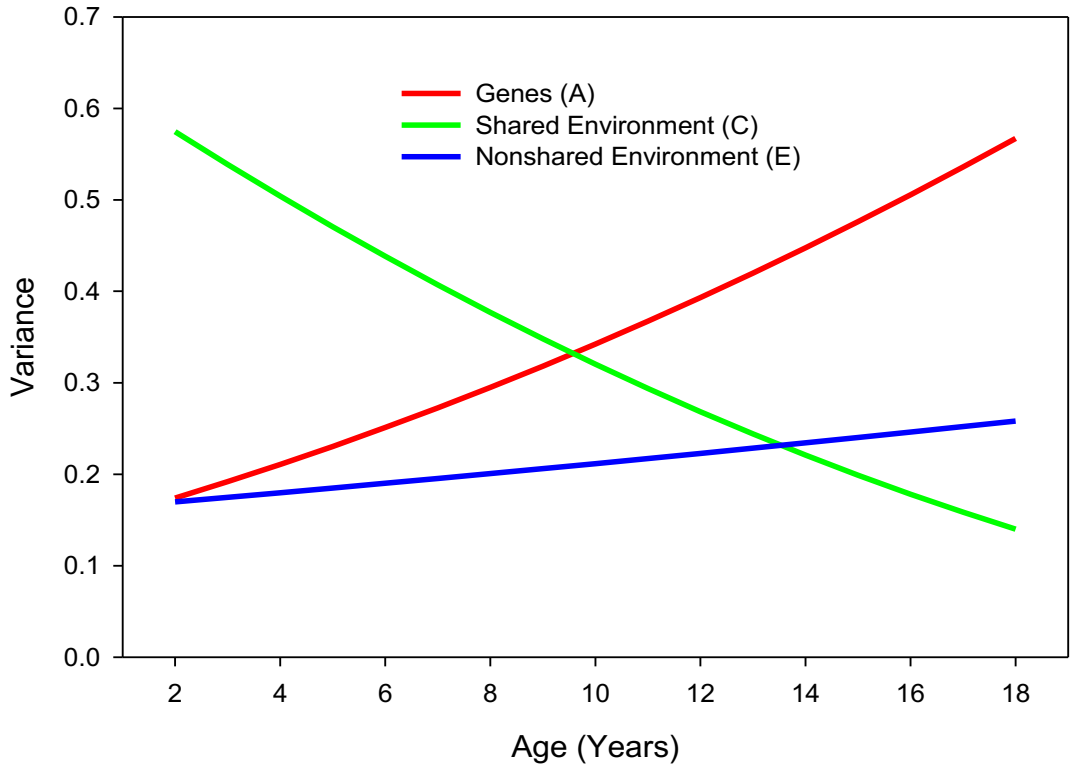
Can we estimate r_{AC}, which not identified in the standard ACE model?



What motivation?

Motivation: Fading C... Is it fading parental influence?

Is the decrease in C due to decrease in C proper **+ decrease in r_{AC}**



Tucker-Drob & Bates (2015)

$c^2 + 2*a*c*r_{AC}$

What is decreasing?

c^2 or r_{AC} in $2*a*c*r_{AC}$

probably both

If $r_{AC} > 0$ the C variance in the ACE model is overestimated

If A is correlated with (rather than interacting with) an environmental variable, say C, with correlation r_{AC} then the expected trait variance is $Var(T)=a^2 + c^2 + 2ac \times r_{AC} + e^2$ and the expected twin covariances are

$$Cov(T_1, T_2) = a^2 Cov(A_1, A_2) + c^2 Cov(C_1, C_2) + e^2 Cov(E_1, E_2) + acCov(A_1, C_2) + acCov(A_2, C_1)$$

$$= a^2 + c^2 + 2ac \times r_{AC} \text{ for MZ twins}$$

$$= a^2/2 + c^2 + 2ac \times r_{AC} \text{ for DZ twins}$$

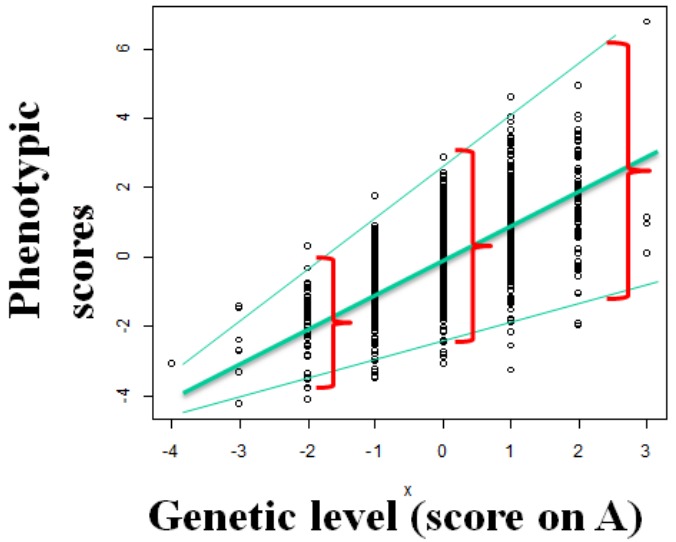
as $Cov(A_1, C_2) = Cov(A_2, C_1) = r_{AC}$.

Homework: Check

Purcell (2002) Twin Research Volume 5 Number 6 pp. 554-571

"INTERPLAY"	unmodeled biases
r(AC)	**C**
r(AE)	A
AxC	A
AxE	E

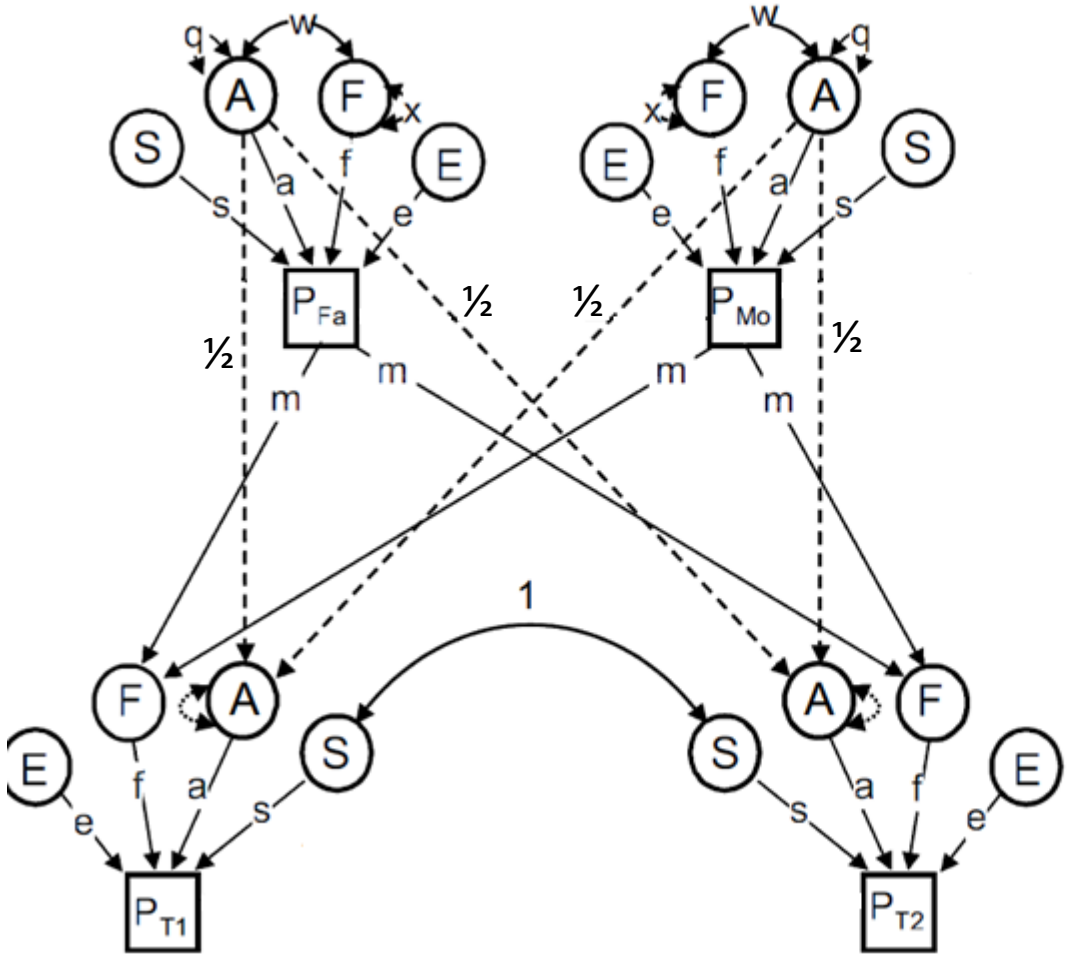
AxE??? what's that? (Slide 14 mod)

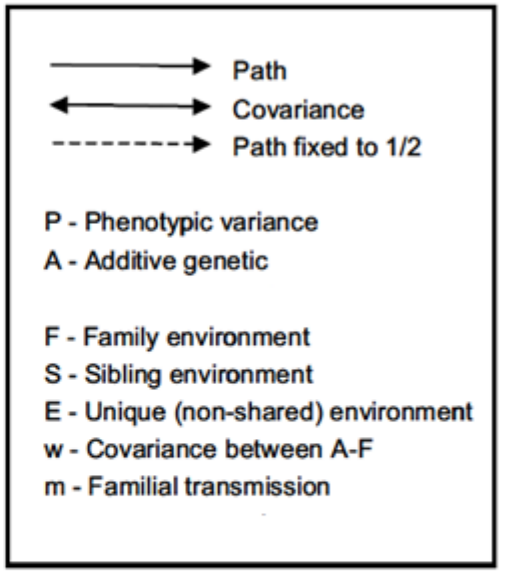


"don't reify variance components"

you can have a long discussion about
the magnitude of variance of A in IQ
(but a very short discussion about the proposition var(A) = 0)

Motivation: Is $r_{AC}>0$ of interst? Parent influence the environmental of children (recent interest via '*transmitted*' – '*non transmitted alleles*" design), **gives rise to r_{AC} in the offspring (w).**



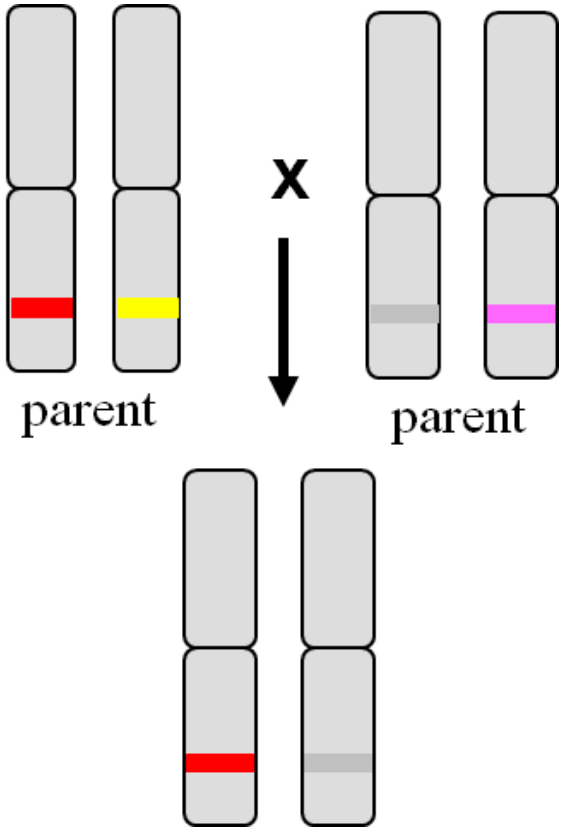


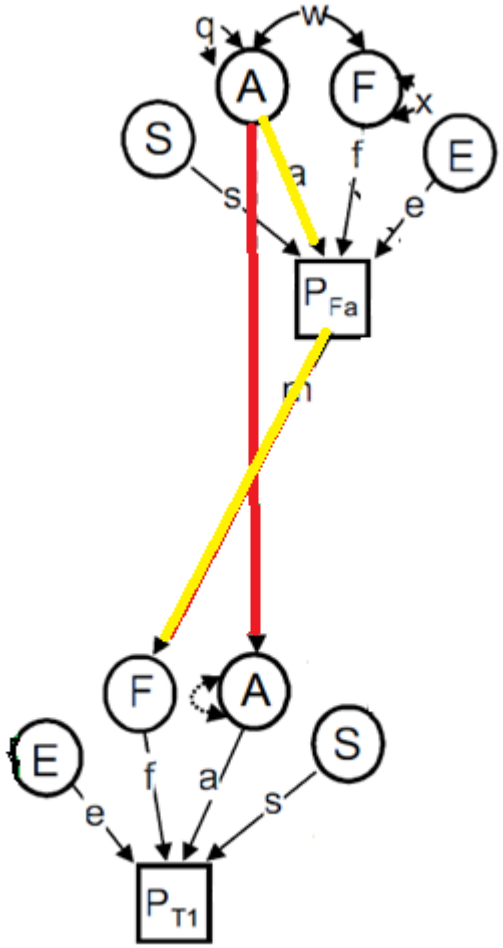
Random mating version

Keller et al (2009) **Twin Research and Human Genetics** Volume 12 Number 1 pp. 8–18.

whence w?

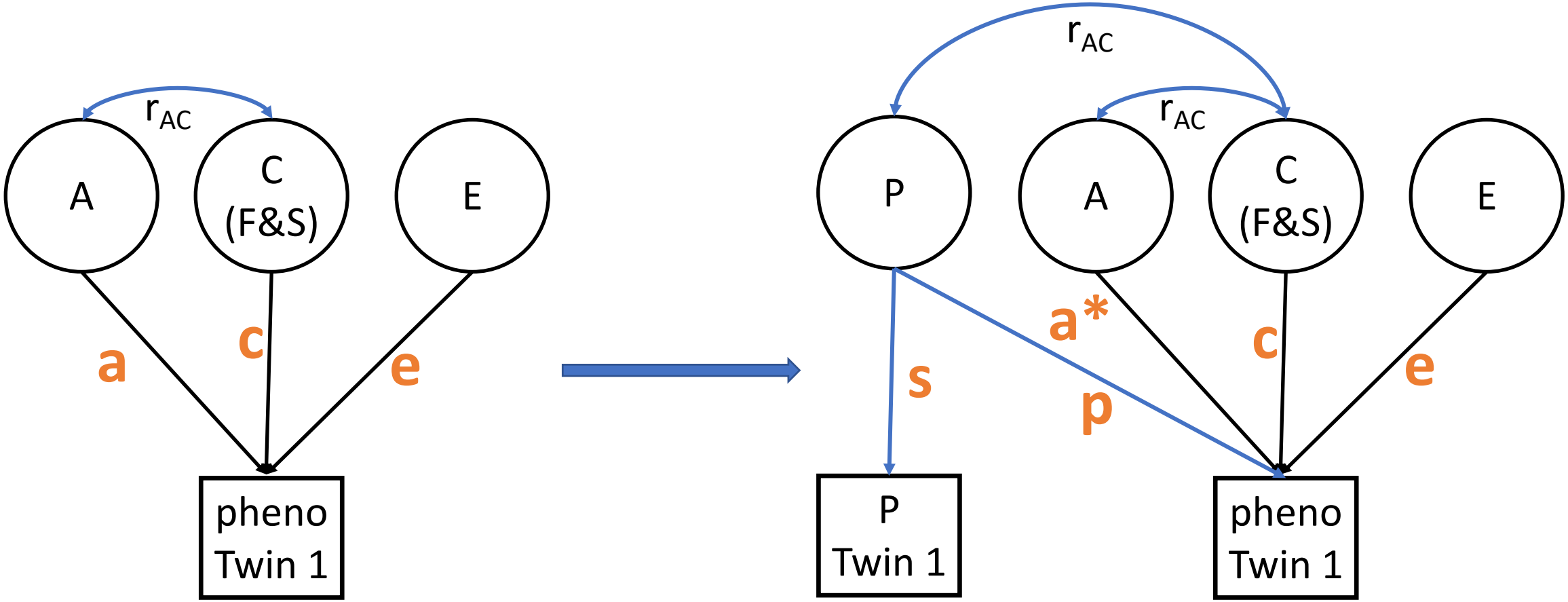
recent interest via '*transmitted*' – '*non transmitted alleles*" design





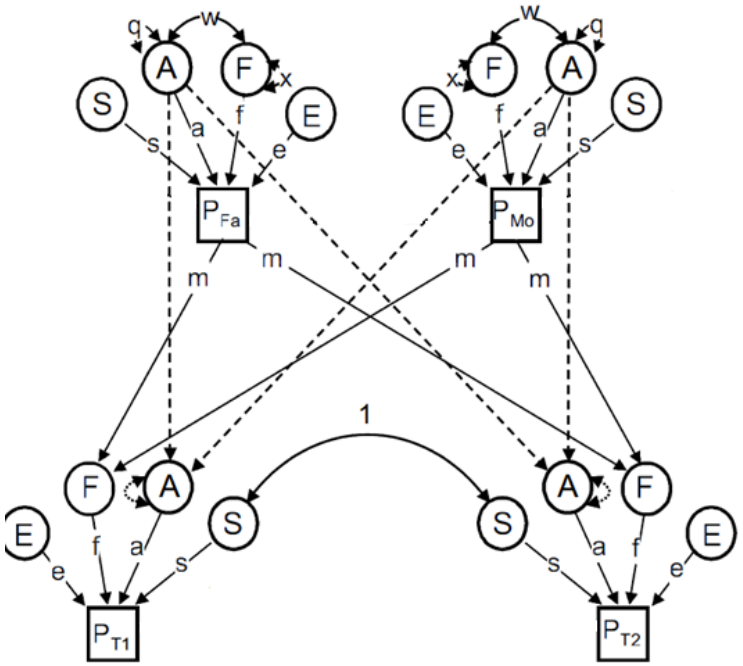
What are the transmitted alleles, what are the untransmitted alleles?

Upshot in the offspring (C = F&S in slide 15)

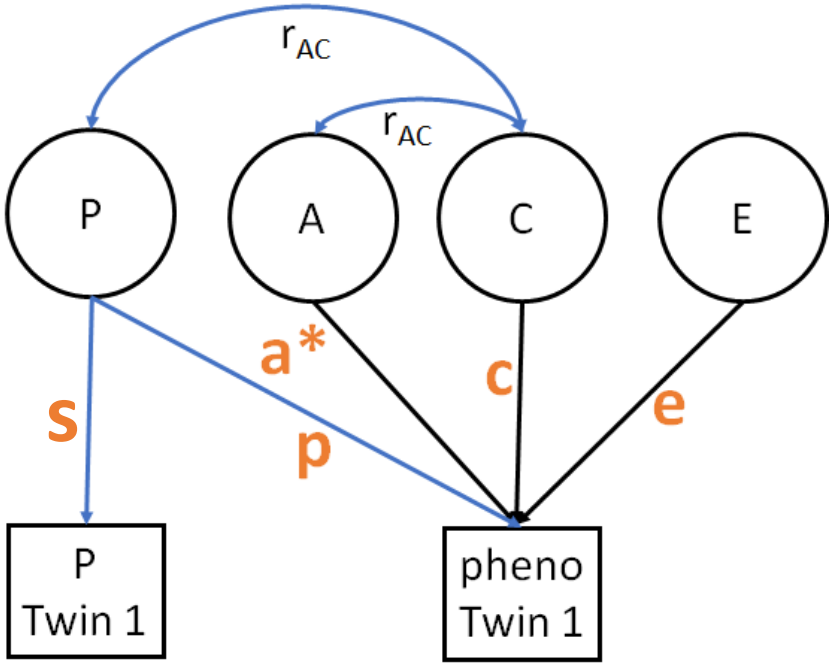


r_{AC} not identified

r_{AC} identified ?



Simulation model



Fitted model

here r_{AC} = .105 (model right) and R^2 of the PGS is .03 (that is .03 effect in the full model)

DZ covariance matrix (r=.542)

	T1	T2
T1	1.639	0.888
T2	0.888	1.639

MZ covariance matrix (r=.694)

	T1	T2
T1	1.639	1.139
T2	1.139	1.639

Lots of C!

2*.542 - .694 = .39

“39% is due to C”

	LB	est	UB
ACE.A[1,1]	0.349	0.500	0.658
ACE.C[1,1]	0.489	0.638	0.784
ACE.E[1,1]	0.459	0.500	0.546

Nmz = Ndz = 1000

	LB	est	UB
rac	-	fixed 0	-
E	0.459	0.500	0.546
P	0.057	0.079	0.106
A*	0.306	0.453	0.607
C	0.452	0.596	0.737

	LB	est	UB
rac	**-.003**	**0.105**	**0.242**
E	0.459	0.500	0.546
P	0.023	0.050	0.088
A*	0.302	0.450	0.604
C	0.319	0.505	0.678
A	0.349	0.500	0.658

Fix r_{AC} = 0. Likelihood ratio T(1) = 3.61057, p = 0.0574
Power to detect r_{AC} here is **.4762** given parameters and Nmz=Ndz=1000

Bad model? Maybe....requires a lot more analyses! (increase r_{AC}, increase %variance due to PGS)

DZ cov matrix

	T1	p1	T2	p2
T1	1.638	0.298	0.888	0.186
p1	0.298	1.000	0.186	0.500
T2	0.888	0.186	1.638	0.298
p2	0.186	0.500	0.298	1.000

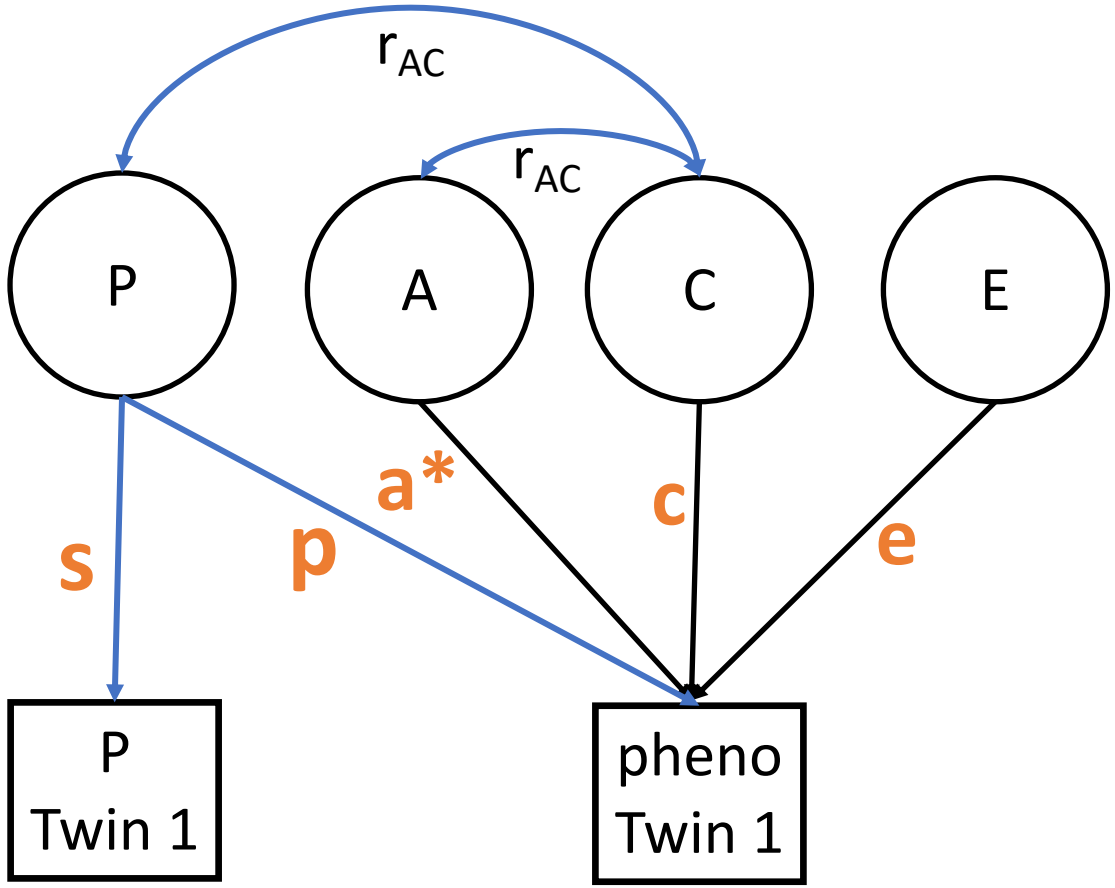
	LB	est	UB
rac	-.003	0.105	0.242
E	0.459	0.500	0.546
P	0.023	0.050	0.088
A*	0.302	0.450	0.604
C	0.319	0.505	0.678
A	0.349	0.500	0.658

The R^2 in regression of phenotype (T) on p (PGS) is .298^2 / 1.638 = .054 (5.4% explained)

The R^2 in the true model is .05 / 1.638 = .030 (3% of the variance)

3% vs. 5.4%

A question.....



Whence the discrepancy (5.4% vs. 3.0%)?

Extending Causality Tests with Genetic Instruments: An Integration of Mendelian Randomization with the Classical Twin Design

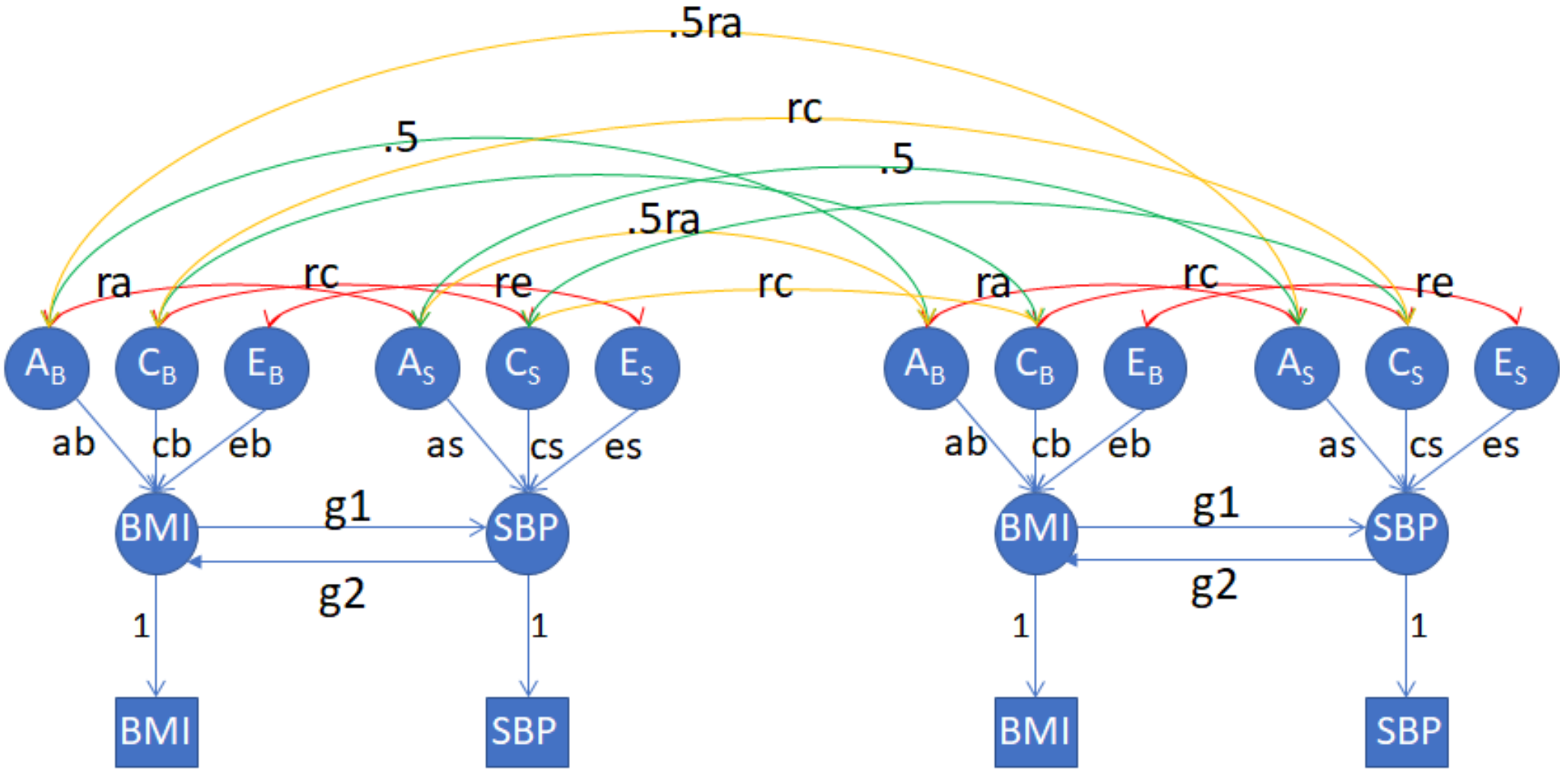
Camelia C Minca & Conor Dolan,

+

Eco de Geus, Dorret Boomsma, Michael C. Neale

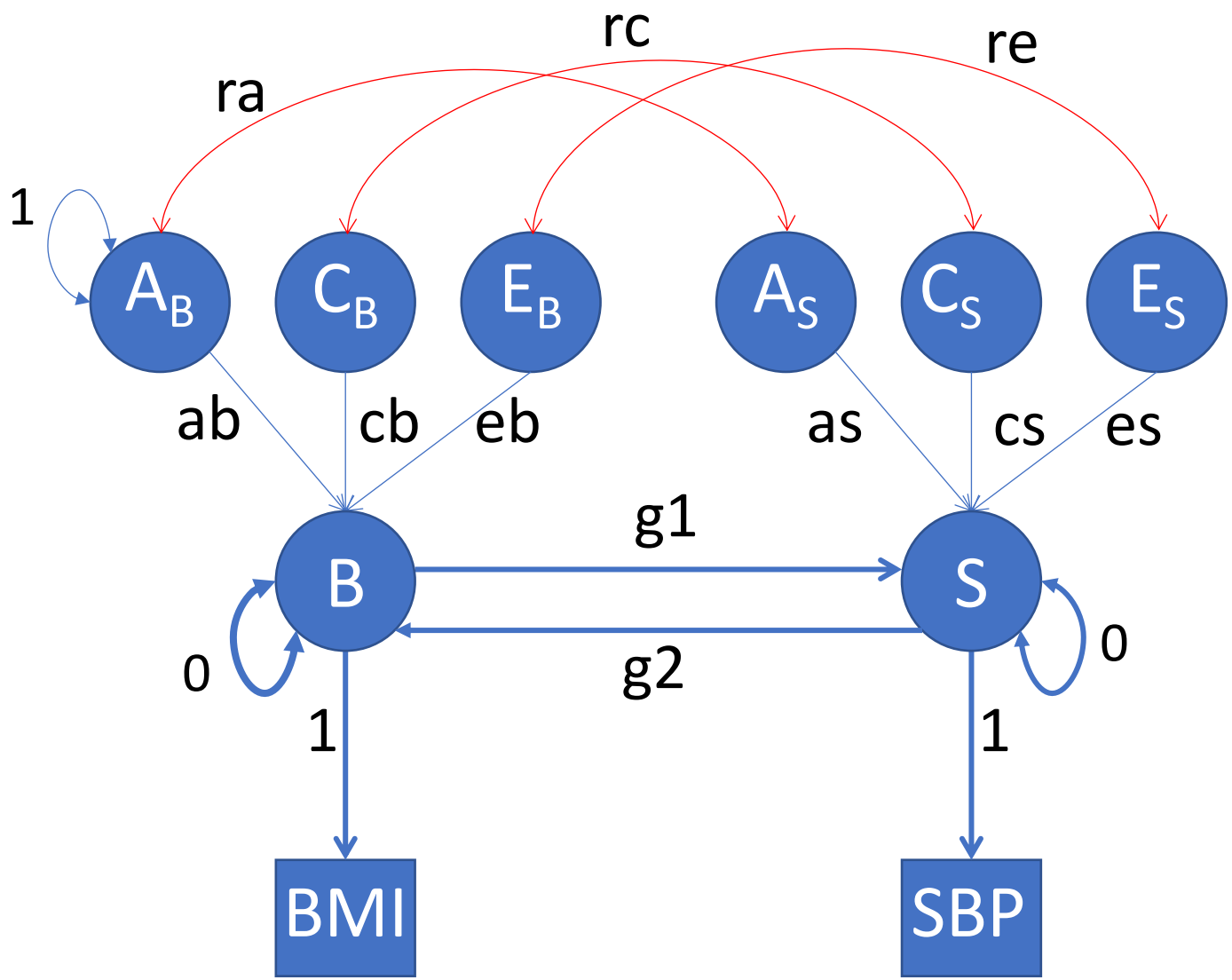
Thanks: David Evans

Modeling causality: Towards the Direction of Causation (DoC) model



not identified

Neale & Cardon, 1992 ; Heath et al., 1993; Duffy & Martin, 1994; Verhulst & Estabrook, 2012.



note “scaling”

Special cases (identified models)

Model 1: g1=g2=0; re≠0; rc≠0; ra≠0 (general bivariate model)

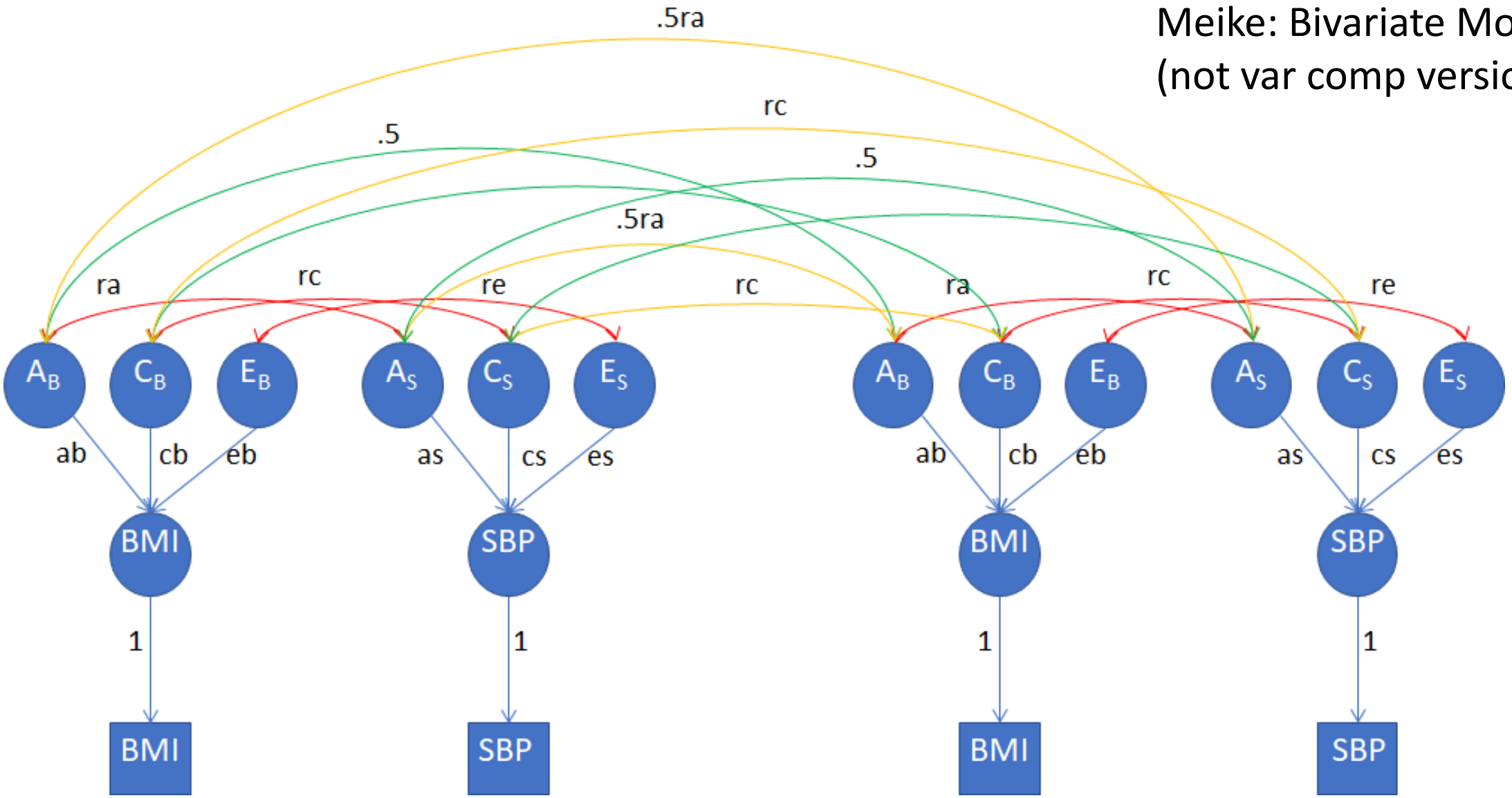
Model 2: BMI →SBP, g1 ≠ 0 & re=rc=ra=0 (unidirectional causation)

Model 3: SBP → BMI, g2 ≠ 0 & re=rc=ra=0 (unidirectional causation)

Model 4: BMI→SBP & SBP →BMI; re=rc=ra=0; g1≠0; g2≠0 (reciprocal causation)

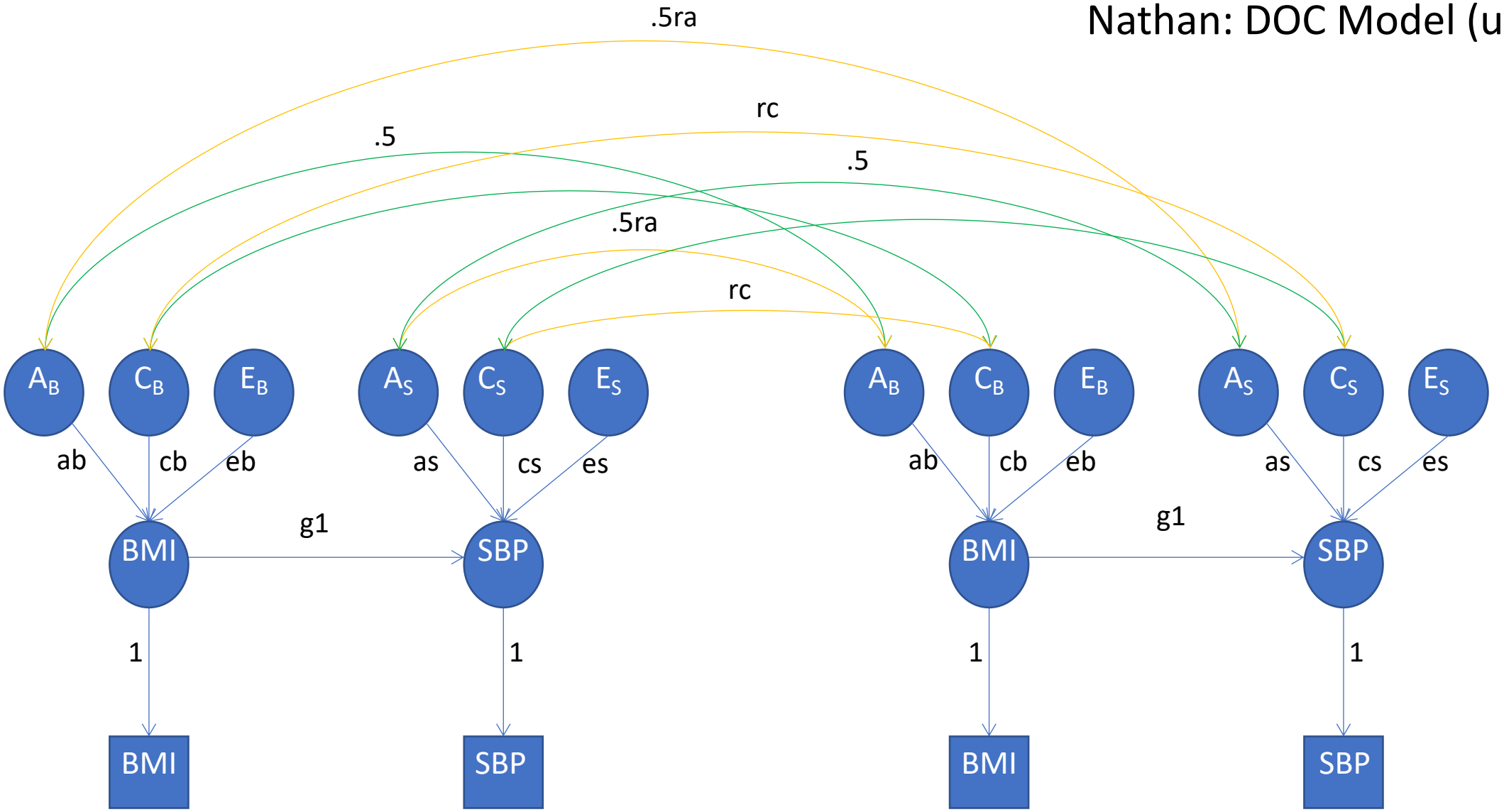
Model 5: no association between BMI and SBP

Meike: Bivariate Model
(not var comp version)



Model 1: g1=g2=0; re≠0; rc≠0; ra≠0 (general bivariate model)

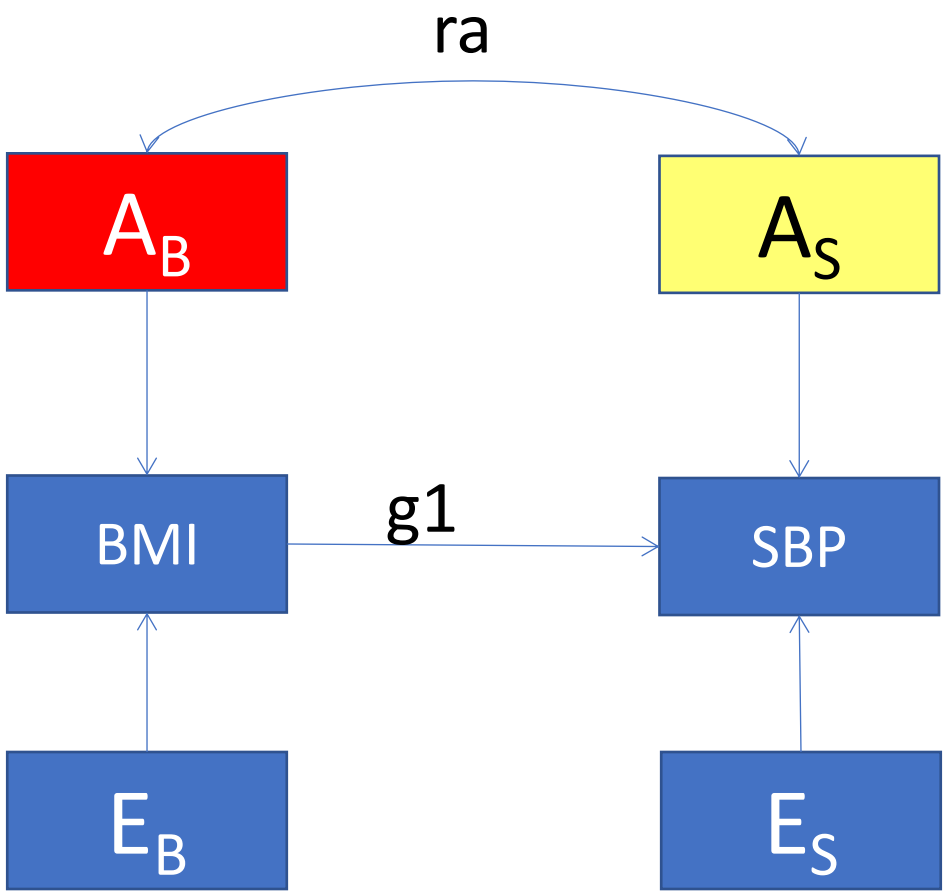
Nathan: DOC Model (unidir)



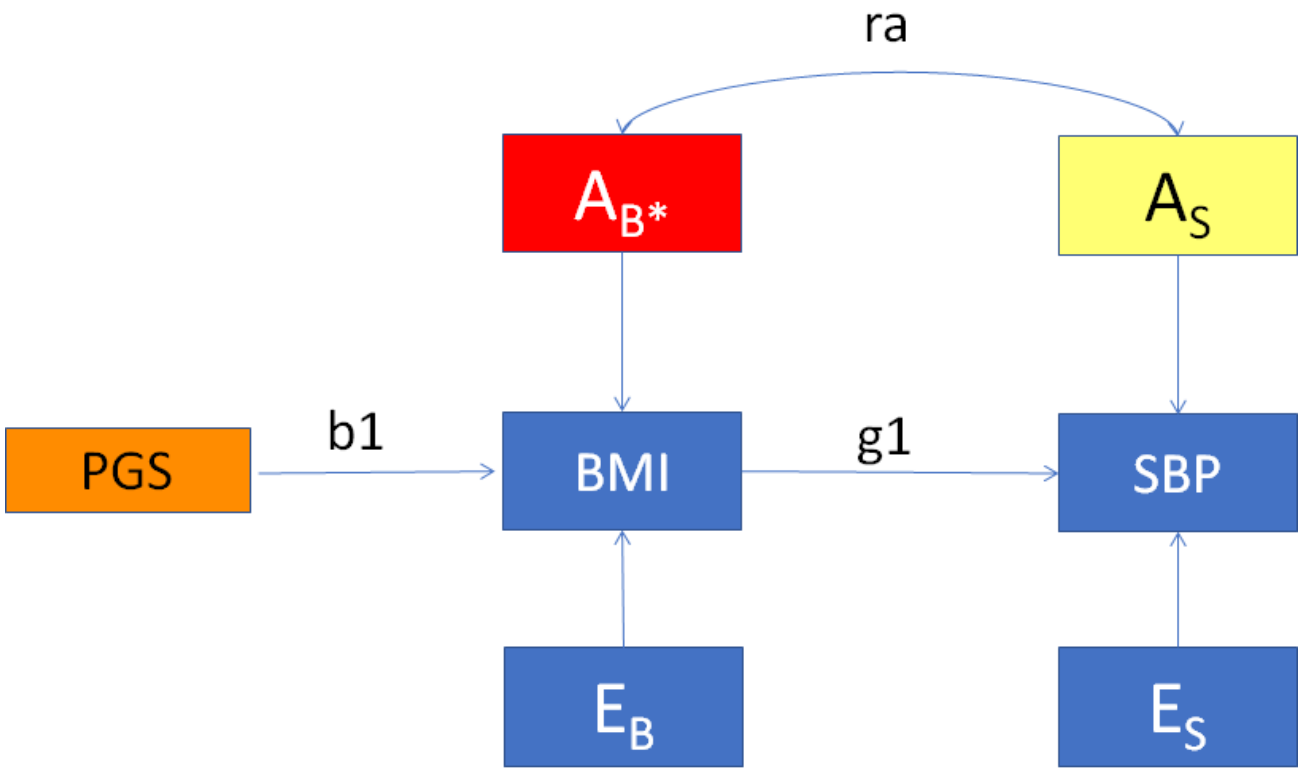
Model 2: BMI →SBP, g1 ≠ 0 & re=rc=ra=0 (overidentified unidirectional causation)

MR

Dave: MR

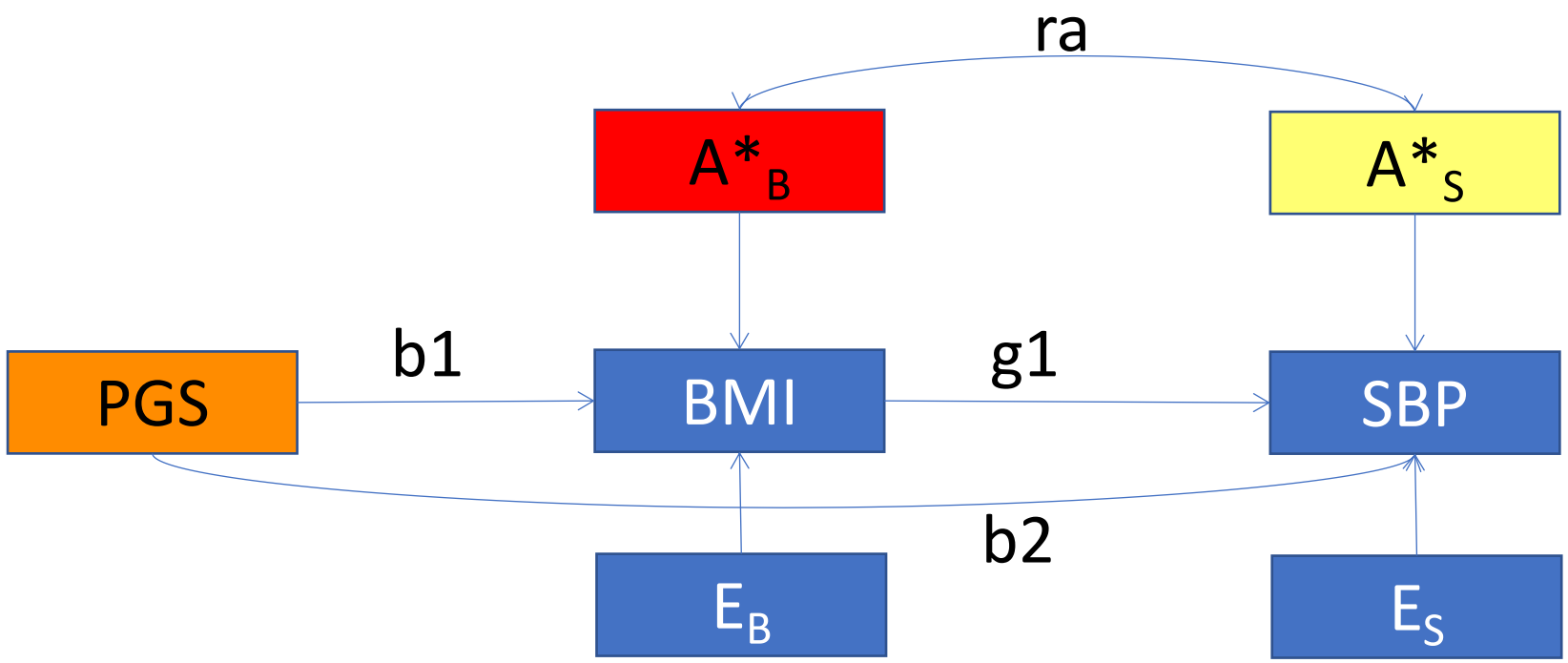


Modeling causality: Mendelian Randomization - model 1 with PGS & no pleiotropy

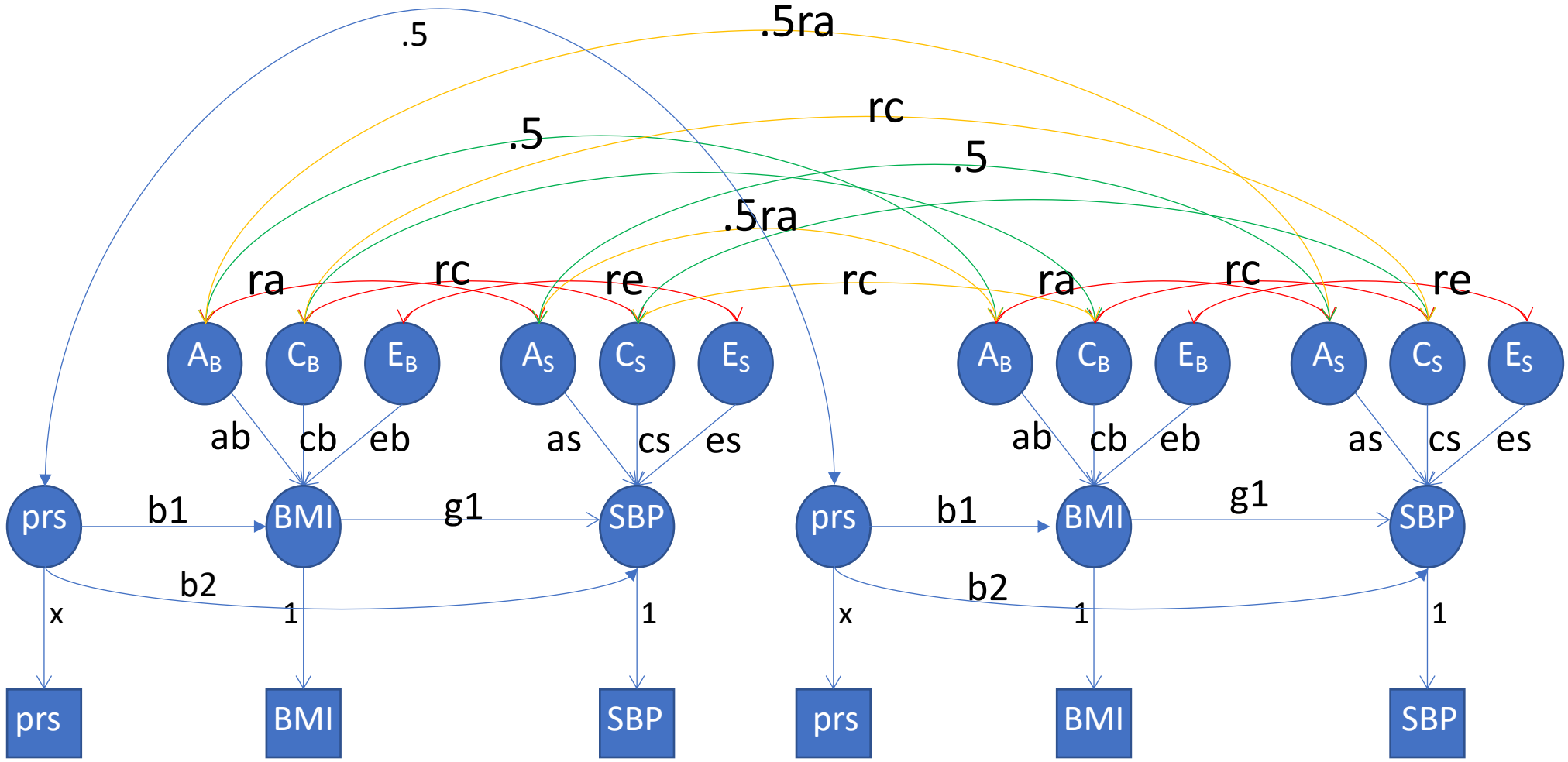


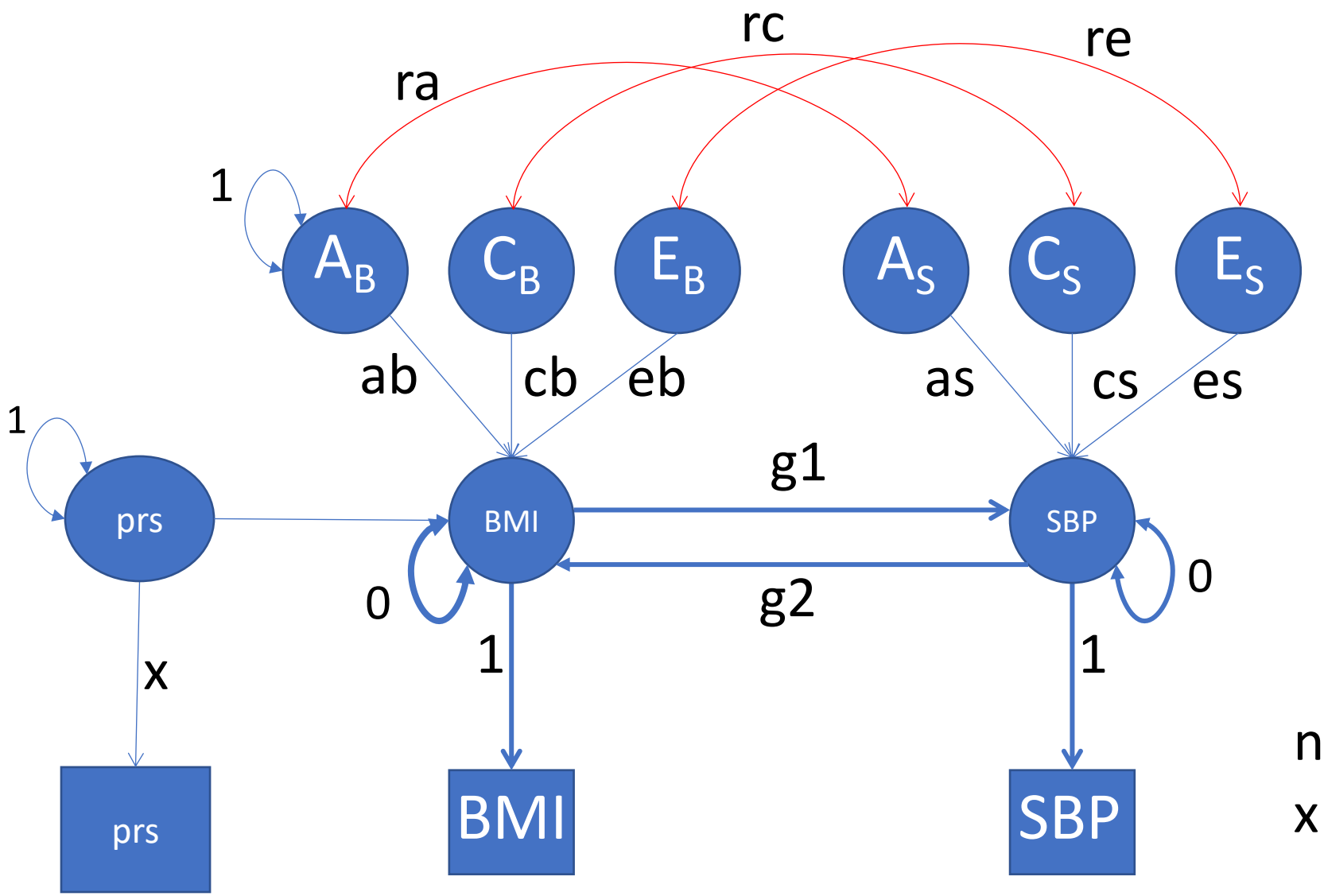
Presence of the "instrumental variable PGS" renders g1 identified – the larger the b1 (R^2 + p-value) the more power to reject g1=0

Modeling causality: Mendelian Randomization – model 2 with PGS & pleiotropy – violates IV assumption



Relaxing MR's assumptions: Mendelian Randomization meets the Classical Twin Design (MR-Twin model)... DZ model





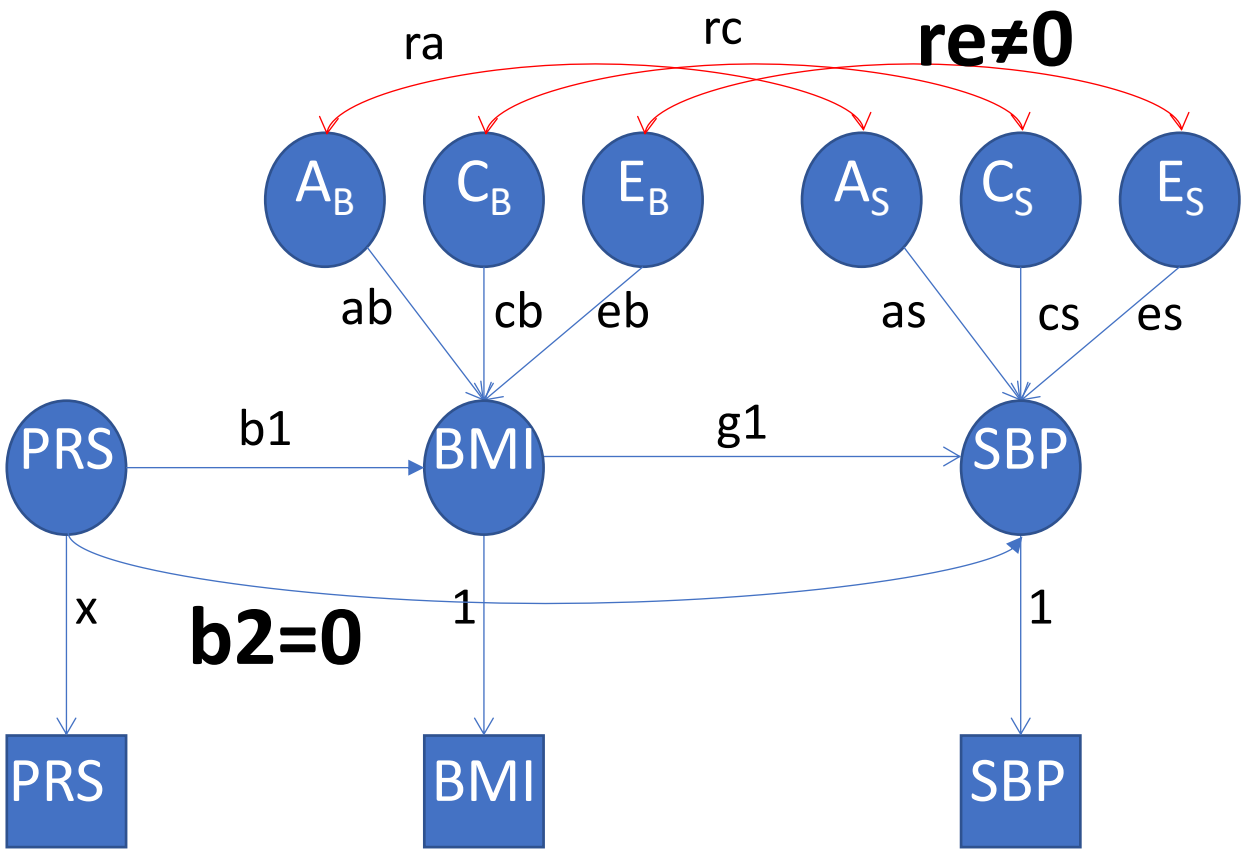
note “scaling”
x is the stdev of FTO

ab cb eb as cs es all freely estimated

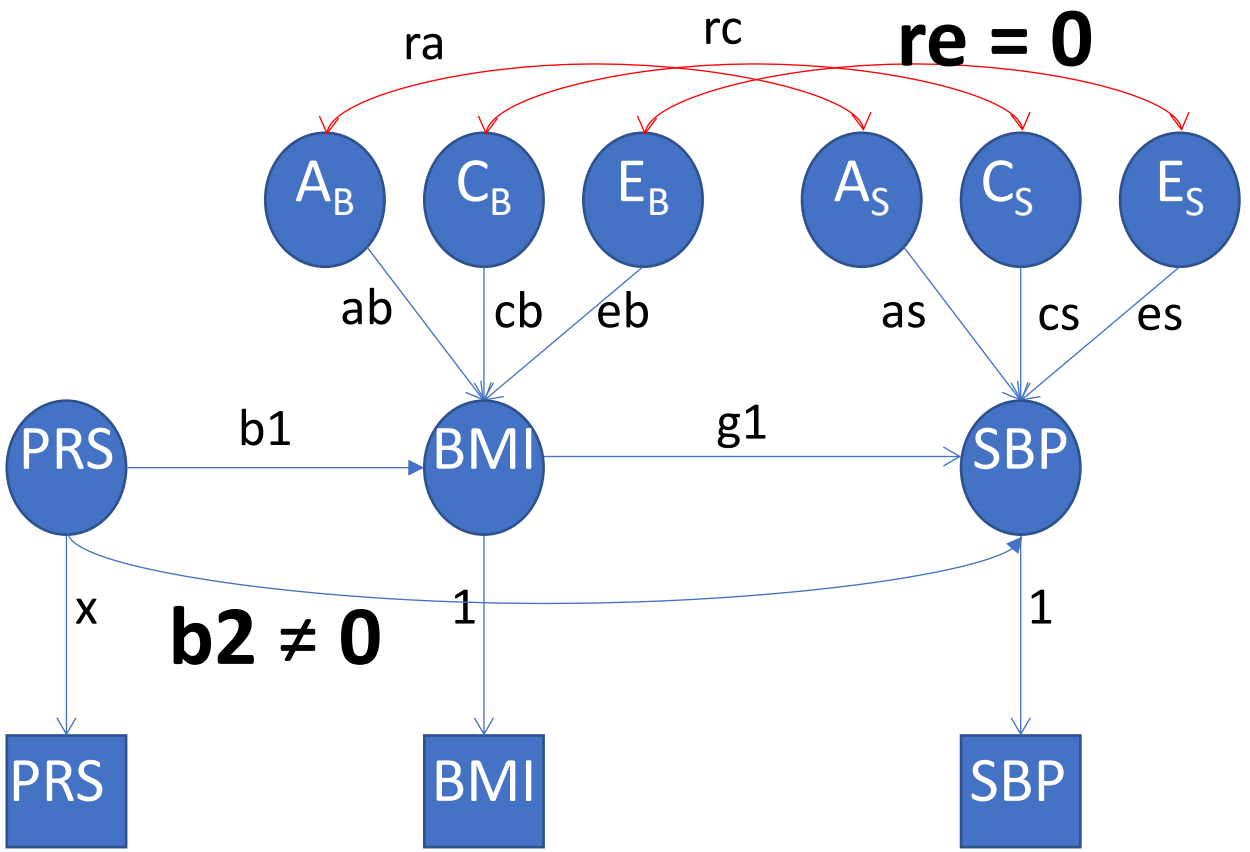
	ra	rc	re	b_1	b_2	g_1	ID?
1	fr	fr	fr	fr	fr	fr	No
2	fr	fr	0	fr	fr	fr	Yes
3	fr	fr	fr	fr	0	fr	Yes

Numerical check: mxCheckIdentification {OpenMx}
Symbolic check: MAPLE, Mathematica, Maxima (free)

Model 1: PGS & no pleiotropy



Model 2: PGS & pleiotropy



Behavior Genetics (2018) 48:337–349
https://doi.org/10.1007/s10519-018-9904-4

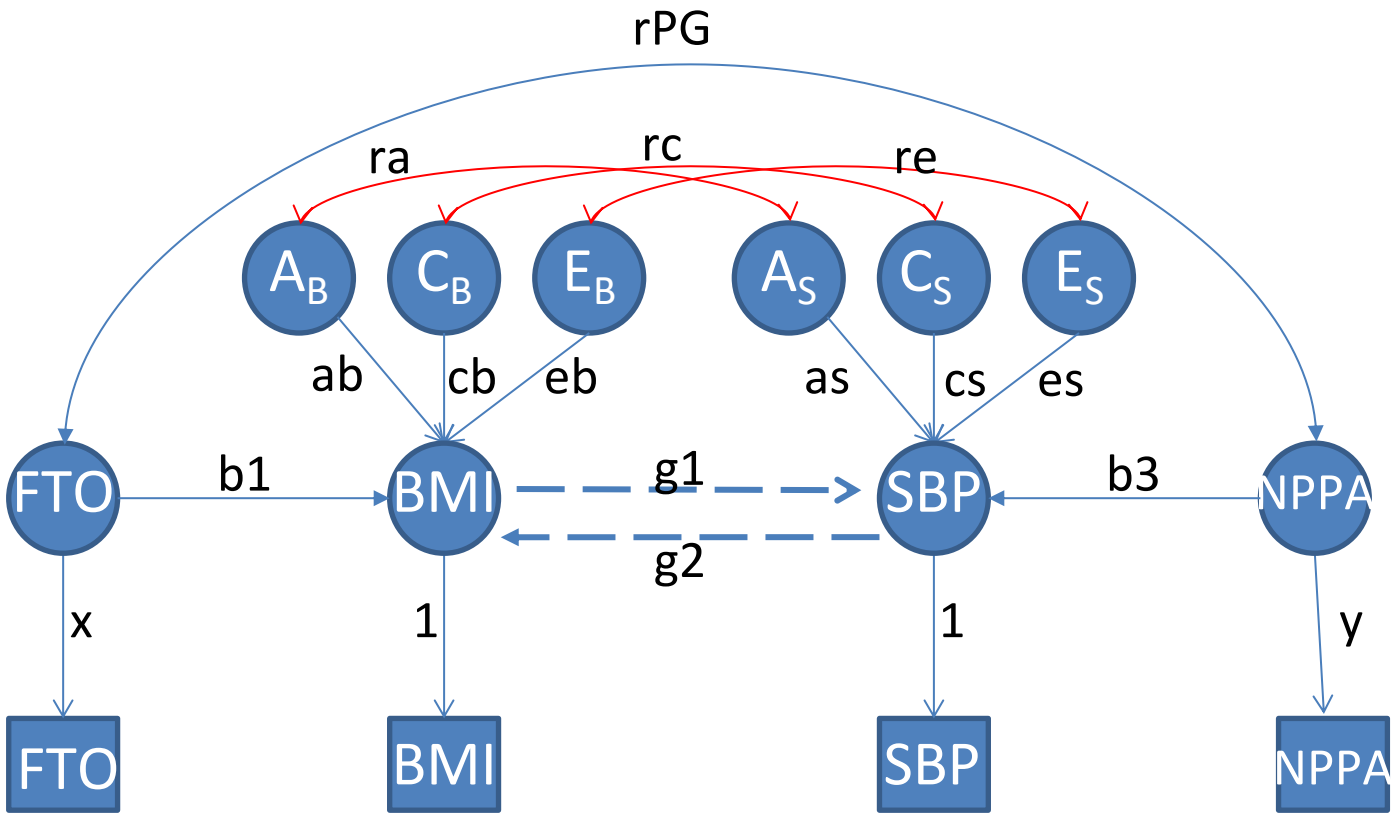
ORIGINAL RESEARCH

Extending Causality Tests with Genetic Instruments: An Integration of Mendelian Randomization with the Classical Twin Design

Camelia C. Minică[1] · Conor V. Dolan[1] · Dorret I. Boomsma[1] · Eco de Geus[1] · Michael C. Neale[1,2]

Adding the second IV for SBP (polygenic score NPPA)

Dave: MR



As depicted the model is identified with re≠0

To do: study resolution & power & pleiotropy

good paper:

Biodemography Soc Biol. 2011;57(1):88-141.

Social science methods for twins data: integrating causality, endowments, and heritability.

Kohler HP[1], Behrman JR, Schnittker J.

Camelia C. Minica

previously, a charming young lady

presently, a charming Nexus 6 Replicant advanced science model