LINKING FERTILITY EXPECTATIONS TO FERTILITY OUTCOMES BIOMETRICALLY:

ANALYSIS OF NLSY79 COMPLETED FERTILITY DATA

Joseph Lee Rodgers, Vanderbilt University
Mike Hunter, University of Oklahoma HSC
Hans-Peter Kohler, University of Pennsylvania

Motivation

- In the mid-1990's, it appeared that no one was looking for genetic variance in fertility and related variables
- In part, because of Fisher's Fundamental Theorem of Natural Selection (the FTNS)
  - Summary – Fitness traits must have no genetic variance, and therefore zero heritability, in the absence of perturbing forces
- But the world has always been filled with "perturbing forces," especially the modern world
  - Mutation
  - Frequency Dependent Selection
  - Contraception
  - Etc, etc
- So lots of us started looking

- Rodgers, Rowe, & Buster, *Social Biology*, 1999, found significant heritability in age at first intercourse measures in the NLSY, for both males and females
- Simple heritability study, with ACE modeling
- Dunne et al, *Demography*, and Miller et al, *Social Biology*, found the same kinds of patterns in AFI in papers in the same year

- Rodgers, Kohler, Kyvik, & Christensen, 2001, *Demography*, found a small but statistically significant link in Danish twin data between fertility motivation and fertility outcomes
  - This is not all that surprising, we expect motivation and outcomes to be related
  - Except for one feature: the link was through the **genetic variance** underlying fertility motivation and outcomes

BEHAVIOR GENETIC MODELING OF HUMAN FERTILITY: FINDINGS FROM A CONTEMPORARY DANISH TWIN STUDY*

JOSEPH LEE RODGERS, HANS-PETER KOHLER, KIRSTEN OHM KYVIK, AND KAARE CHRISTENSEN

Behavior genetic designs and analyses can be used to address issues of central importance to demography. We use this methodology to document genetic influence on human fertility. Our data come from Danish twin pairs born from 1953 to 1959, measured on age at first attempt to get pregnant (FirstTry) and number of children (NumCh). Behavior genetic models were fitted using structural equation modeling and DF analysis. A consistent medium-level additive genetic influence was found for NumCh, equal across genders; a stronger genetic influence was identified for FirstTry, greater for females than for males. A bivariate analysis indicated significant shared genetic variance between NumCh and FirstTry.

The research presented here combines a design and an analytic approach that are not typically found in demographic literature with an outcome variable that is of central interest to demographers. To study fertility, we present a set of behavior genetic analyses based on a twin design. In this paper we address two basic questions. First, are there genetic influences on fertility outcomes? Second, if so, what are the theoretical origins of those influences?

Behavior genetic designs are not often used in demography. During the past decade, however, increasing attention has been given to integrating biological explanations with the type of social models used in demography and sociology, particularly in relation to human reproduction (e.g., Adams et al. 1990; Rossi 1994; Udry 1995; Wood 1994). It is unfortunate that behavior genetic methods are not standard tools in demography, because they provide a powerful set of designs and analyses to address questions of interest to demographers. Further, the term *behavior genetics* is a misnomer, to some extent: Plomin and Rende (1991) commented, "The power of behavioral genet
well as nature-
sources of indivi

Even if a re
environmental in
for genetic confi
Scarr and Grajek
sue, suggesting tl
lations arise in bi
the research liter
374). Rodgers, R
approach that co
of influence, allo
tion of the other.

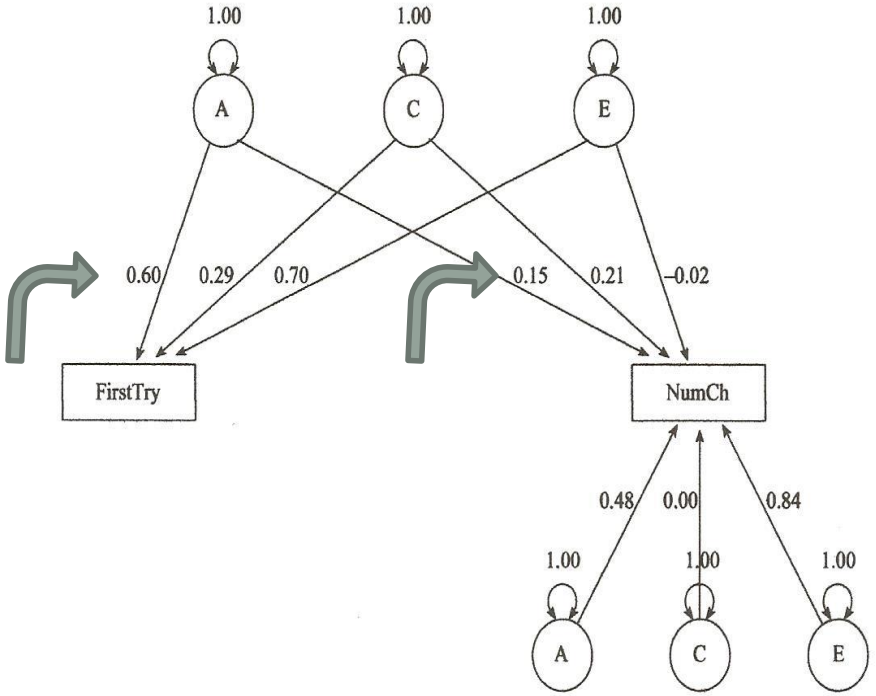
The primary
ior genetic analy
stantive interpret
concomitant goal
tool to introduce
havior genetic m
not attempt exha
tions that underli
descriptions and
ports those descr
ences on human
ment of human e
of natural selecti
haustive coverag
treatment along
which interested

Obviously th
that are not typi
work, however,
graphic research
example. Using a
twin data that
Christensen (199
efficients associa
from 1870 to 19
analysis showed
counting for fert
time, ranging fro
of around .40. (T
40% of the obser
with genetic varia
family environme
fashion. It is a ma

*Joseph Lee Rodgers, Department of Psychology, University of Oklahoma, Norman, OK 73019; E-mail: jrodgers@ou.edu. Hans-Peter Kohler, Max Planck Institute for Demographic Research; Kirsten Ohm Kyvik, Section for Epidemiology, Institute for Public Health, and The Danish Twin Registry, University of Southern Denmark; Kaare Christensen, Section for Epidemiology, Institute for Public Health, and the Danish Center for Demographic Research, University of Southern Denmark. The authors acknowledge contributions from reviewers and the editors that substantially improved this paper. The first author was supported by NIH Grant RO1-HD2-1973 during the course of this research. Research support and space were also graciously provided to him by the Danish Center for Demographic Research at Odense University and by the Max Planck Institute for Demographic Research in Rostock, Germany during a sabbatical leave in 1998. The second author acknowledges the support provided by the Max Planck Institute for Demographic Research. The third and fourth authors were supported by Sygekassernes Helsefond (Grant 11/209-93). Administrative support from Jim Vaupel and Hans Christian Johansen was particularly valuable in stimulating the collaboration that led to this paper.

Demography, Volume 38-Number 1, February 2001: 29–42

FIGURE 3. ACE BIVARIATE MODEL, SEPARATING VARIANCE IN NUMBER OF CHILDREN (NumCh) THAT OVERLAPS WITH AGE AT FIRST ATTEMPT TO GET PREGNANT (FIRSTTRY) AND THAT IS INDEPENDENT OF FIRST TRY; ESTIMATED PARAMETERS FOR FEMALE-FEMALE PAIRS



- Rodgers, Bard, & Miller, 2007, *Behavior Genetics*, using NLSY fertility data up to 2002, defined female fertility outcomes at different age intervals, and decomposed genetic and environmental variance in correlations across those ages.
  - Two different genetic factors were identified, one for fertility up to age 20, another for fertility after age 20
  - One shared environmental factor was identified, for early fertility

Behav Genet (2007) 37:345–361
DOI 10.1007/s10519-006-9137-9

ORIGINAL PAPER

Multivariate Cholesky Models of Human Female Fertility Patterns in the NLSY

Joseph Lee Rodgers · David E. Bard · Warren B. Miller

Received: 29 April 2005 / Accepted: 30 November 2006 / Published online: 5 January 2007

Abstract Substantial evidence now exists that variables measuring or correlated with human fertility outcomes have a heritable component. In this study, we define a series of age-sequenced fertility variables, and fit multivariate models to account for underlying shared genetic and environmental sources of variance. We make predictions based on a theory developed by Udry [(1996) Biosocial models of low-fertility societies. In: Casterline, JB, Lee RD, Foote KA (eds) Fertility in the United States: new patterns, new theories. The Population Council, New York] suggesting that biological/genetic motivations can be more easily realized and measured in settings in which fertility choices are available. Udry's theory, along with principles from molecular genetics and certain tenets of life history theory, allow us to make specific predictions about biometrical patterns across age. Consistent with predictions, our results suggest that there are different sources of genetic influence on fertility variance at early compared to later ages, but that there is only one source of shared environmental influence that occurs at early ages. These patterns are suggestive of the types of gene–gene and gene–environment interactions for which we must account to better understand individual differences in fertility outcomes.

Keywords Fertility · Fisher's theorem · FTNS · Heritability · Shared Environment · Multivariate models · Phenotypic plasticity · Hox genes · Life history theory

Edited by John Hewitt and Wendy Slutske

J. L. Rodgers (✉)
Department of Psychology, University of Oklahoma, Norman, OK 73019, USA
e-mail: jrodgers@ou.edu

D. E. Bard
Department of Pediatrics, Health Sciences Center, University of Oklahoma, Norman, OK, USA

W. B. Miller
Transnational Family Research Institute, Aptos, CA, USA

Fisher's (1930) Fundamental Theorem of Natural Selection (the FTNS) has been mis-interpreted by many researchers for many years. The traditional interpretation is that fitness traits and behaviors strongly affected by natural selection will "lose" their genetic variance in the long run. The implication is that fertility and fertility precursors should have little or no genetic variance, and thus zero heritability. This inference is correct if natural selection is the only process at work. But it is not. The FTNS has been interpreted by many to mean that, by definition, for fertility (and other traits related to fitness), $h^2 = 0$. But it does not.

Hughes and Burleson (2000) suggested a number of different processes that re-introduce genetic variance into fitness traits, even while natural selection is washing it out. These include mutation (the most important), frequency-dependent selection, heterozygote advantage (overdominance), sexual antagonism, and environmental perturbations. Fisher (1930) clearly did not intend or participate in this misinterpretation of the FTNS. In the same source in which he presented the FTNS, he also discussed the role of contraception (an environmental perturbation) in re-introducing genetic variance into human fertility (see also Houle 1992, for additional and more modern discussion of "perturbing forces" from

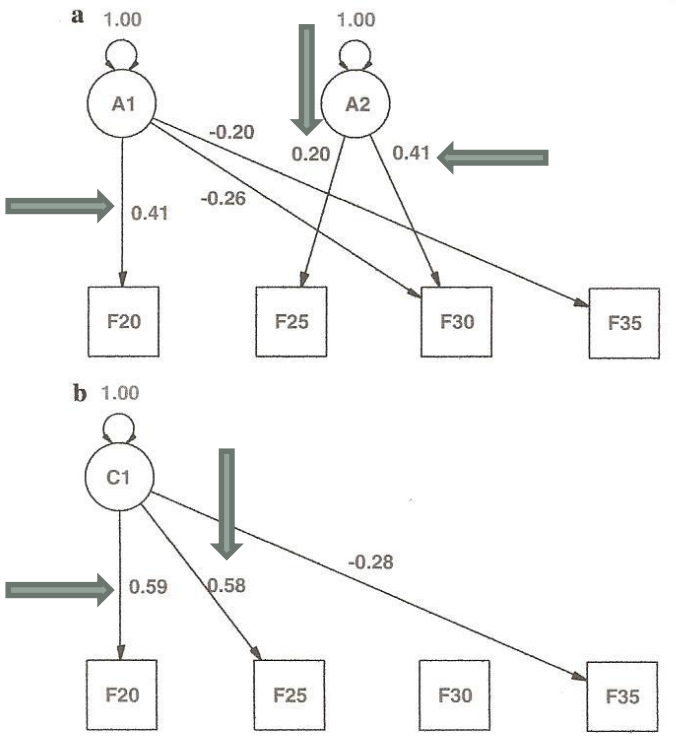


Fig. 3 (**a**) Best-fitting Cholesky multivariate model, additive genetic (A) part of the ACE model; the coefficients are standardized coefficients (corresponding to the unstandardized *a*'s in Table 6); to get h^2 and c^2 values, these coefficients would have to be squared; all individual coefficients are significant ($P < 0.05$) or marginally significant ($P < 0.10$). (**b**) Best-fitting Cholesky multivariate model, shared environmental (C) part of the ACE model; the coefficients are standardized coefficients (corresponding to the unstandardized *c*'s in Table 6); to get c^2 values, these coefficients would have to be squared; all individual coefficients are significant ($P < .05$)

- Fertility among the NLSY79 cohort is now complete (in 2012, NLSY79 respondents were 47-55 years old)
- The NLSY79 contained longitudinal measures of “Fertility Expectations” as well as completed fertility
- We have recently completed updating the NLSY79 kinship links, providing reliable linking information for 95% of all kinship pairs in the NLSY79
- These set the stage for a comprehensive biometrical analysis of fertility in the NLSY
- Advantages over previous studies:
  - Completed fertility
  - Male-male and female-female pairs
  - Family biometrical design, based on a probability sample (with half siblings, full siblings, and twin pairs, at approximately representable levels from US 1979 households)
  - Many external variables available

NLSY79 Kinship Links

- In 1979, NLSY79 had thousands of siblings in the survey
- But they weren't separately identified as adoptive sibs, half-sibs, full-sibs, or twins
- In the mid-1990's, we developed an algorithm to distinguish these categories, using indirect information in the NLSY data
  - Birthdates (for twins)
  - Shared household with biological father and mother (for siblings)
  - Distance away from biological father (for siblings)
  - Other indicators
- Two dozen behavior genetic studies have used these links
- Updated in 2013, using direct responses collected in 2006
- http://liveoak.github.io/NlsyLinks/

Measurement

- Fertility expectations: "How many (more) biological children do you expect to have?" (added to previous children)
  - We took measures at age 21 (or as close as possible)
  - We took measures at age 26 (or as close as possible)
- Fertility Outcomes: "How many total (biological) children have you had?"
  - We took completed fertility through 2012, ages 47-55
- In this study, we used half-siblings, full-siblings, and twins
  - There are thousands of full sib pairs, hundreds of half sib pairs, and dozens of MZ and DZ twin pairs
- http://liveoak.github.io/NlsyLinks/
  - Files, source code, and vignette examples, NLSY-79 & NLSY-Children

Summary Statistics

(Note: Approximately Representative of 1979 US cohort, up to attrition and non-response)

	N	Mean	StdDev
MM			
Exp21	819/1045	**2.46/2.44**	1.21/1.32
Outcomes	1140/1140	**1.84/1.76**	1.54/1.50
FF			
Exp21	1446/1820	**2.37/2.41**	1.21/1.23
Outcomes	1939/1939	**1.99/1.92**	1.55/1.54

Note: First score is Sibling 1, second is Sibling 2

Empirical Starting Point, Biometrical Modeling – Some ACE Models

- ACE models decompose kinship correlations into variance components associated with
  - Genetic variance (segregating genetic variance, additive)
  - Shared environmental variance (creating sibling similarities)
  - Nonshared environmental variance (creating sibling differences; usually confounded with measurement error)
- These kinship correlations are the starting point for both conceptualizing a biometrical study, and for getting started with the modeling

NLSY79 Kinship Correlations

	MMEx21	MMOut	FFEx21	FFOut
HalfSibling	.05	.03	-.08	-.16
FullSiblings	.11	.08	.12	.19
MZ-twins	.48	.28	.95	.89

Note: MM = male-male pairs; FF = female-female pairs;
Ex = Fertility Expectations; Out = Fertility Outcomes;

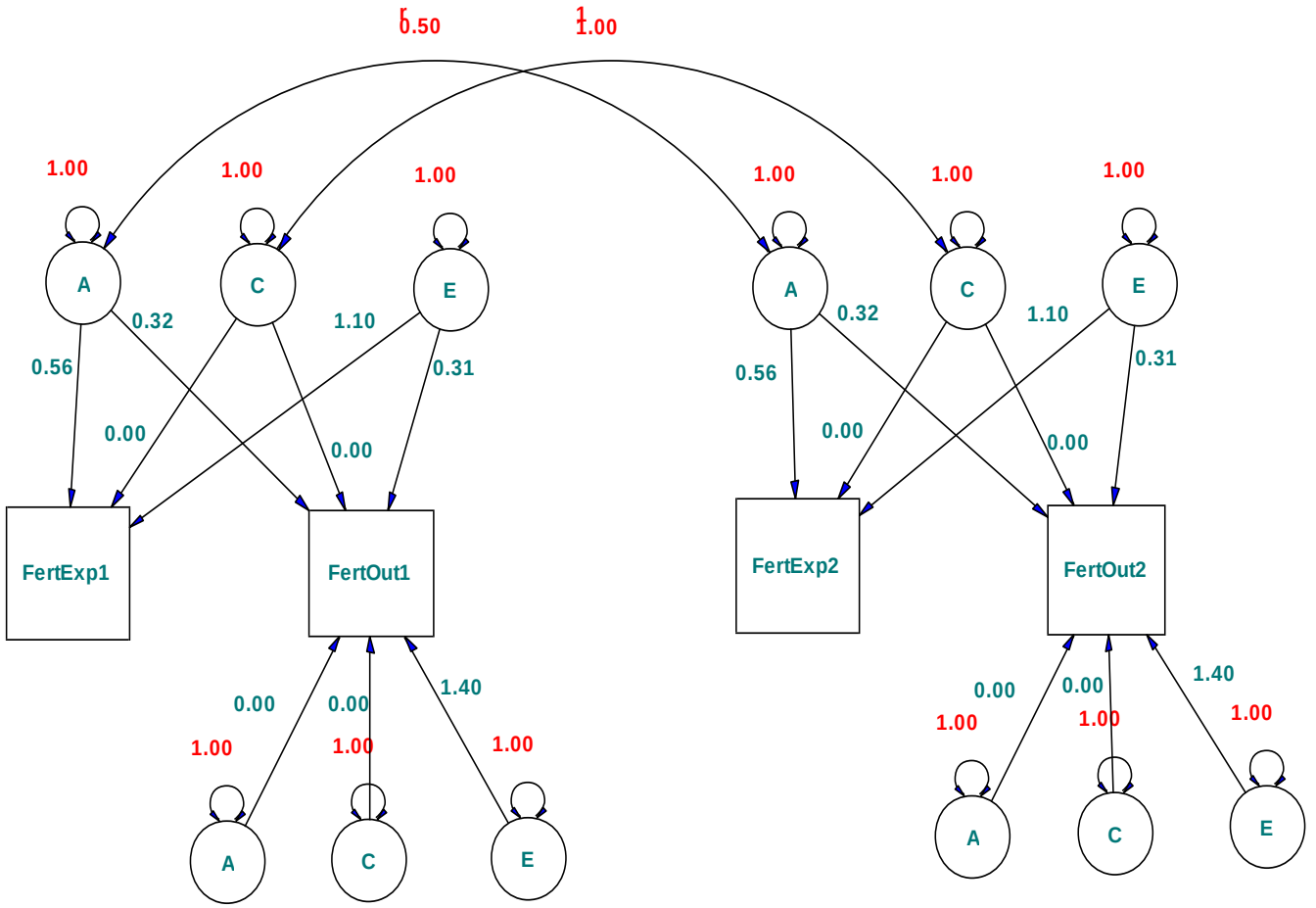
ACE, Univariate Results

	AGE 21 FertExp	FertOut
Male-Male		
A	.18	.15
C	*.03*	*.00*
E	*.80*	*.85*
Female-Female		
A	*.05*	*.38*
C	*.09*	*.00*
E	*.86*	*.62*

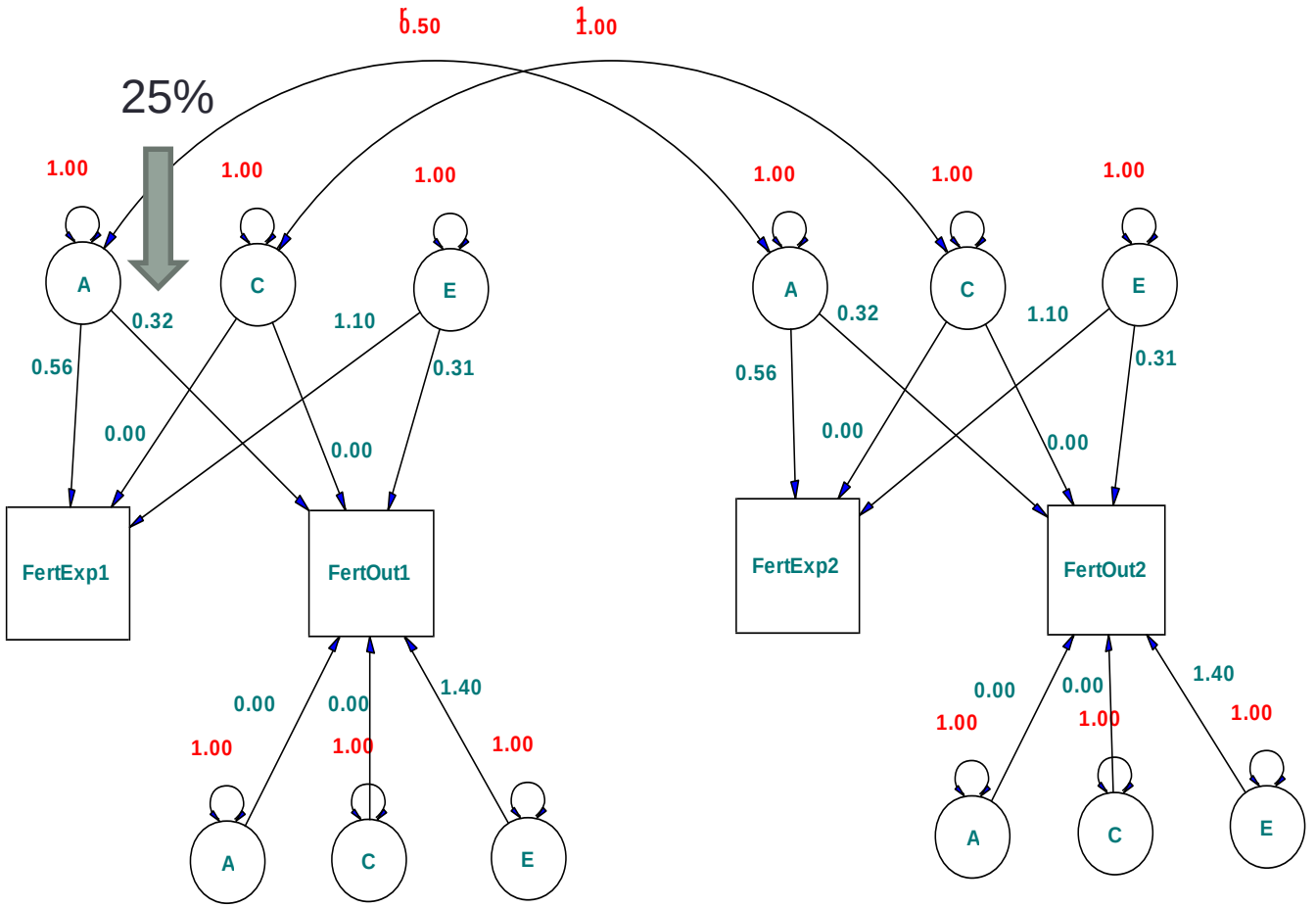
Multivariate Cholesky Models

- In multivariate Cholesky models, there's an ordering
  - User specifies the first variable, the second, etc
  - Time can be used to order
  - In our case, it's conceptual: Fertility Expectations are presumed to precede Fertility Outcomes
- Same biometrical decomposition as in univariate ACE models, but we decompose cross-variable correlations instead of within-variable correlations

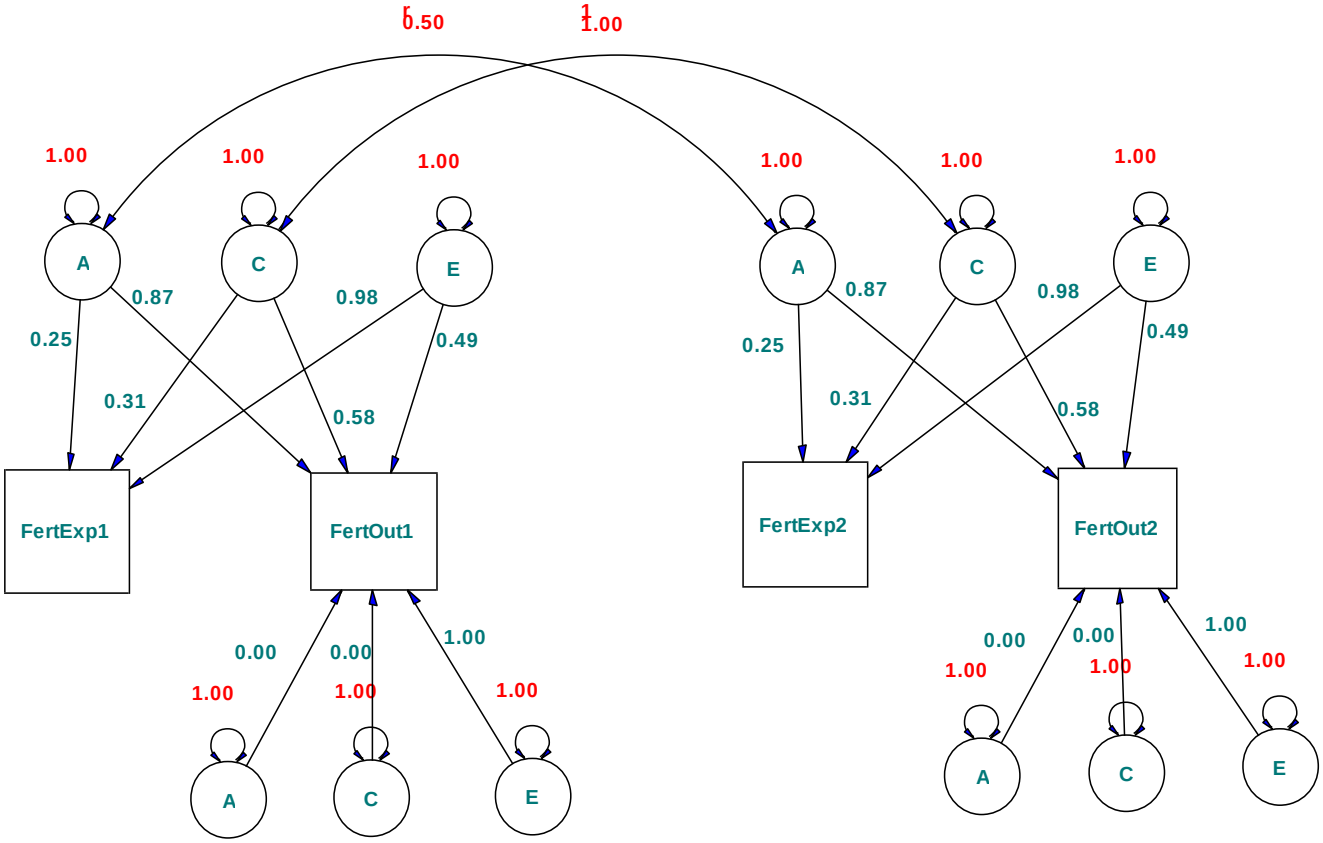
Male-Male


0.50
1.00
1.00
1.00
1.00
1.00
1.00
1.00
A
C
E
A
C
E
0.32
1.10
0.32
1.10
0.56
0.31
0.56
0.31
0.00
0.00
0.00
0.00
FertExp1
FertOut1
FertExp2
FertOut2
0.00
0.00
1.40
0.00
0.00
1.40
1.00
1.00
1.00
1.00
1.00
1.00
A
C
E
A
C
E

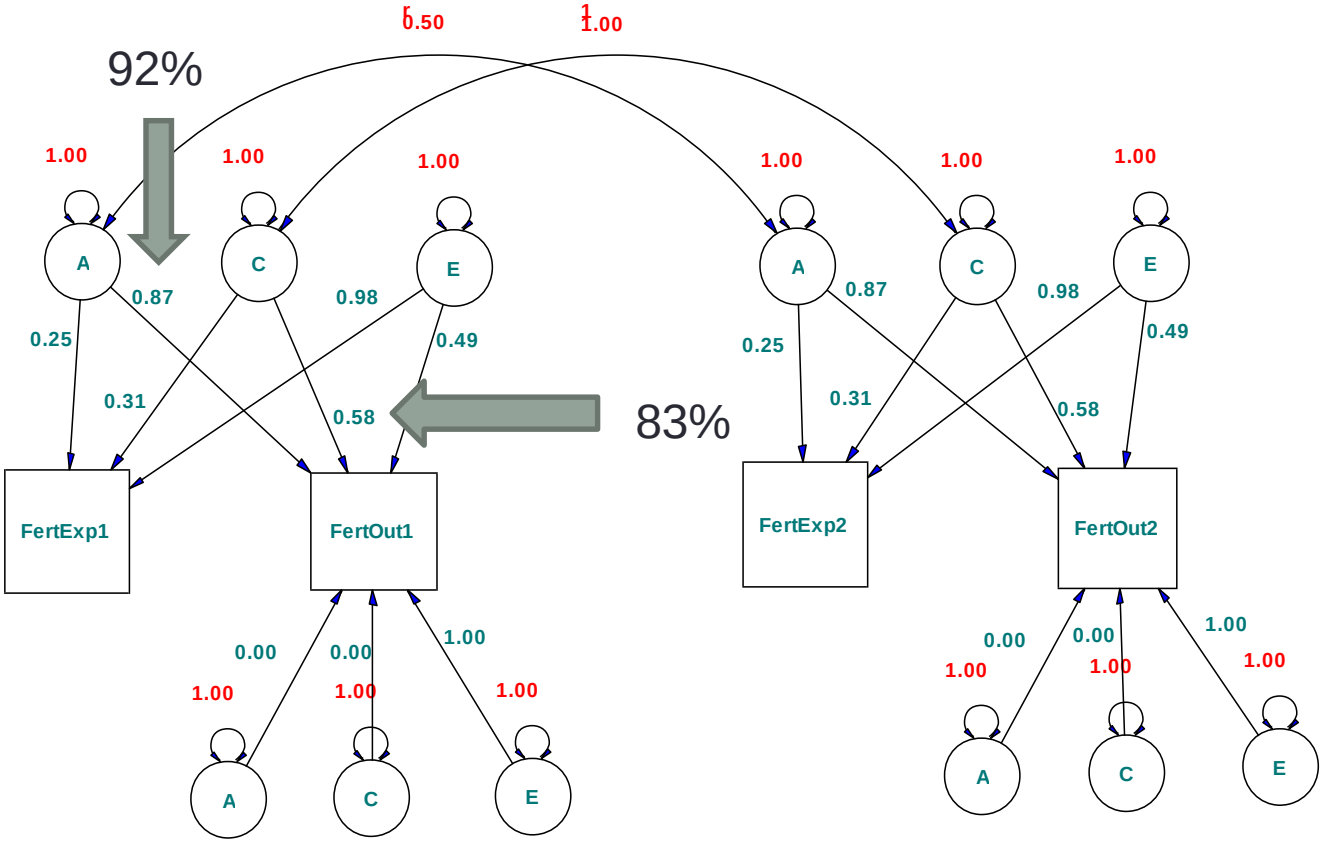
Male-Male


0.50
1.00
25%
1.00
1.00
1.00
1.00
1.00
1.00
A
C
E
A
C
E
0.32
1.10
0.32
1.10
0.56
0.31
0.56
0.31
0.00
0.00
0.00
0.00
FertExp1
FertOut1
FertExp2
FertOut2
0.00
0.00
1.40
0.00
0.00
1.40
1.00
1.00
1.00
1.00
1.00
1.00
A
C
E
A
C
E

Female-Female



Female-Female



Summary

- For both males and females, within these (simplified) models, Fertility Expectations overlap substantially in their genetic variance with Fertility Outcomes.
- Fertility Expectations have some of their own variance
  - 25% of fertility expectations genetic variance overlaps with fertiliy outcomes for Males
  - 92% of Fertility Expectation genetic variance overlaps with Fertility outcomes for females;83% of shared environmental variance in fertility expectation overlaps with fertility outcomes for females
- In these NLSY Data, variance in the fertility precursors are much more highly shared with fertility outcomes than in the Danish twin data

- Udry's theory: Genetic variance can only manifest in settings with strong fertility options
  - In natural fertility settings, with high family size and low contraceptive use, there is little chance for genetic variance in fertility expectations to show through
  - In the modern US – with efficacious contraception, low infant mortality, etc. – women and men have lots of fertility choices; genetic variance in fertility preferences can be achieved

Future Work

- Look at expectations as they develop across other ages
  - Previous work (using the NLSY) justifies that there are different biometrical patterns in fertility for late adolescence compared to young adults
  - We did run cholesky models for Age 26 expectations – there were different patterns, though not substantially different
- Look at reverse causation – exchange expectations and outcomes in the Cholesky ordering
- Refine the above models to more precisely estimate effects
- http://liveoak.github.io/NlsyLinks/