PROPENSITY SCORES AS A MEASURE OF SELECTION

SPENCER L. JAMES
ASSISTANT PROFESSOR
SCHOOL OF FAMILY LIFE

FAMILY STUDIES CENTER
OCTOBER 11, 2013

TWO PERSPECTIVES ON COHABITATION AND MARITAL QUALITY

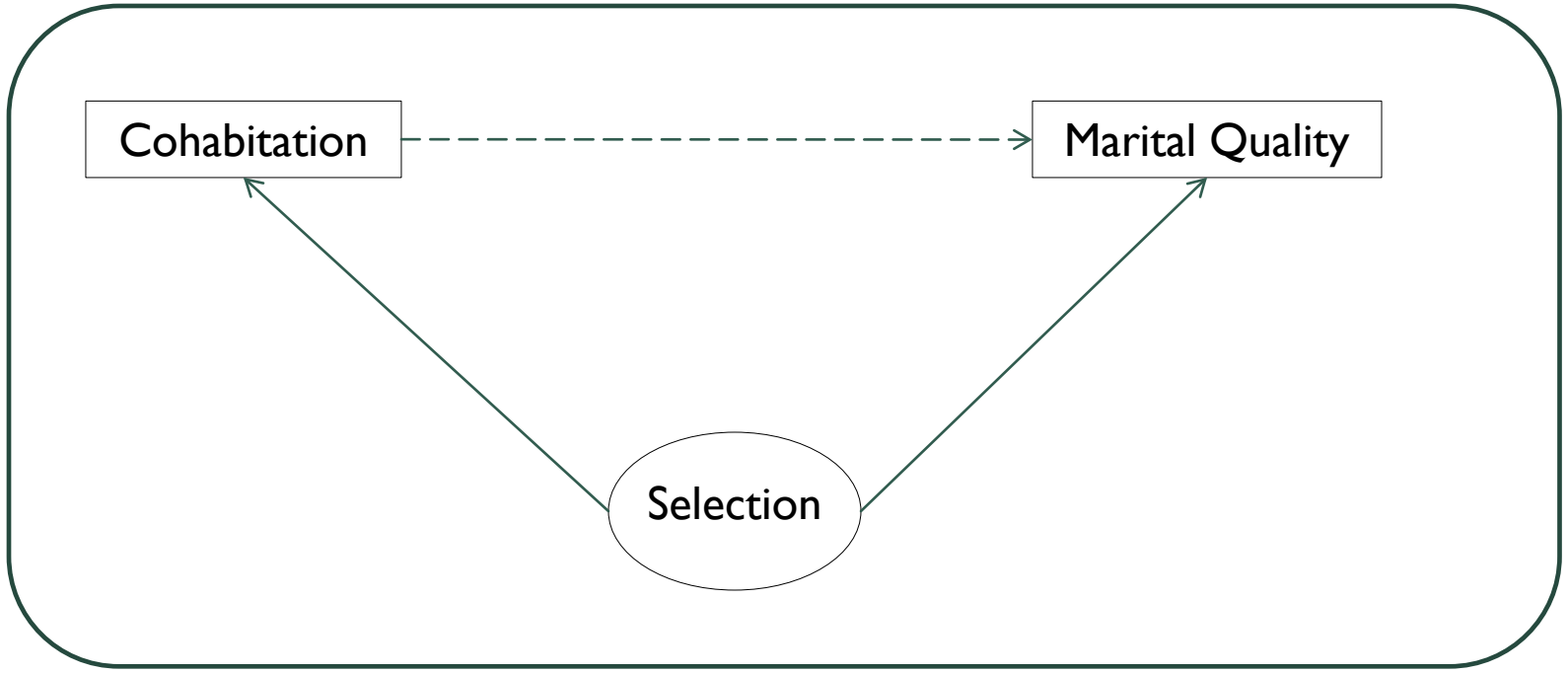
- Social Selection
  - Differences between cohabitors and non-cohabitors are responsible for the observed negative marital outcomes (Glezer, 1997; Stanley, Whitton & Markman, 2004; Bumpass, Sweet, & Cherlin, 1991; Brown & Booth 1996)
- Experience
  - Cohabitation itself may influence the risk of marital outcomes beyond one's characteristics at the beginning of the union (Axinn and Thornton, 1992; Thornton et al., 1992; Nock, 1995; Kamp-Dush, Cohan, & Amato, 2003; James and Beattie 2012; Jose et al., 2010)

MODEL SPECIFICATION

- Zero-Order (Bivariate)

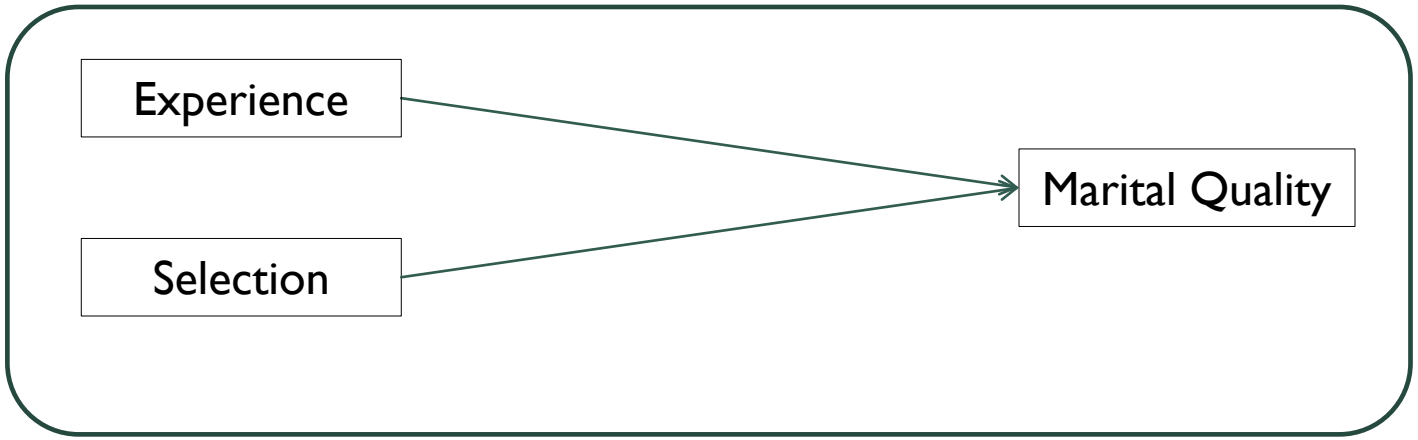
MODEL SPECIFICATION

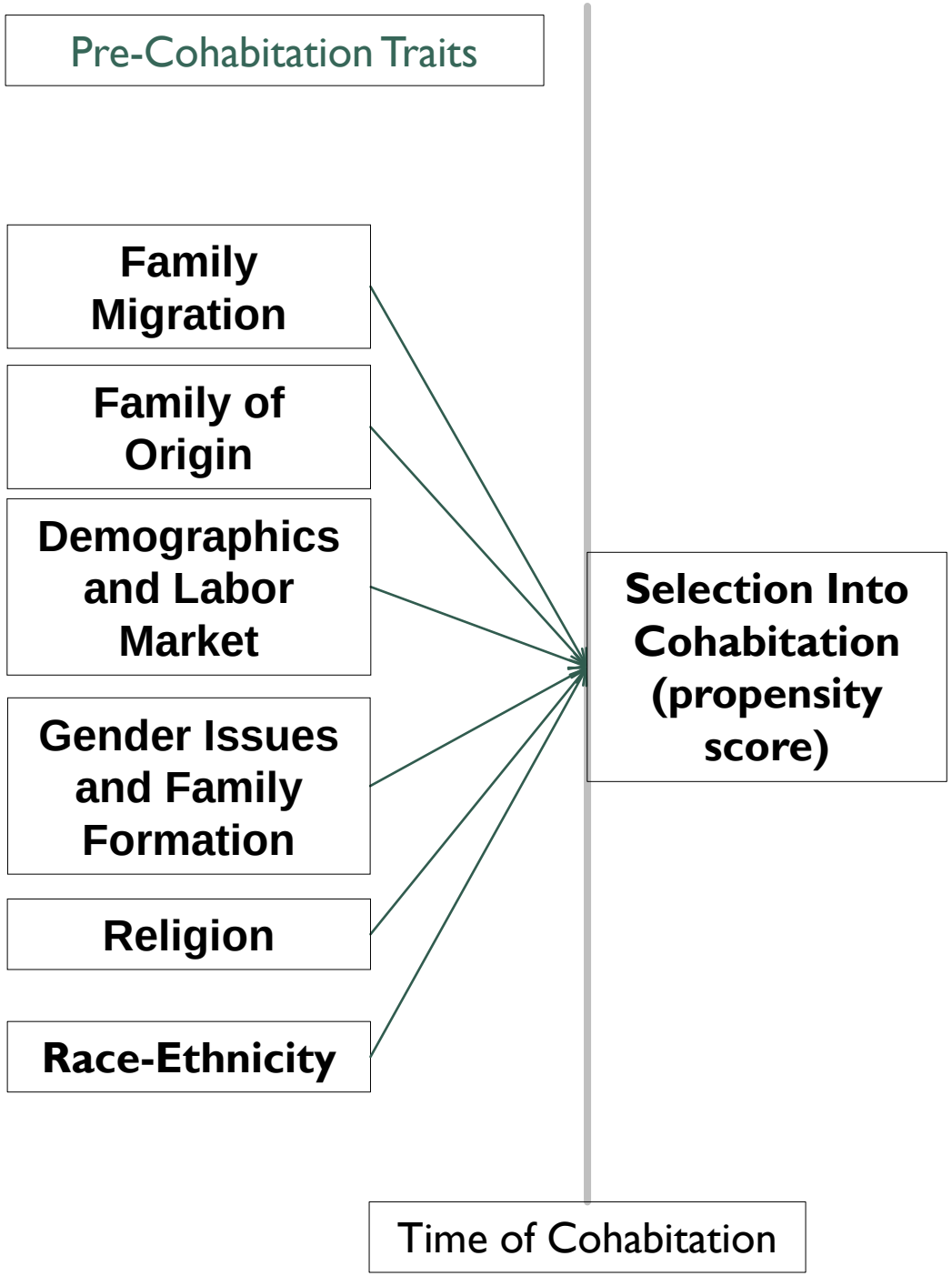
- Spurious

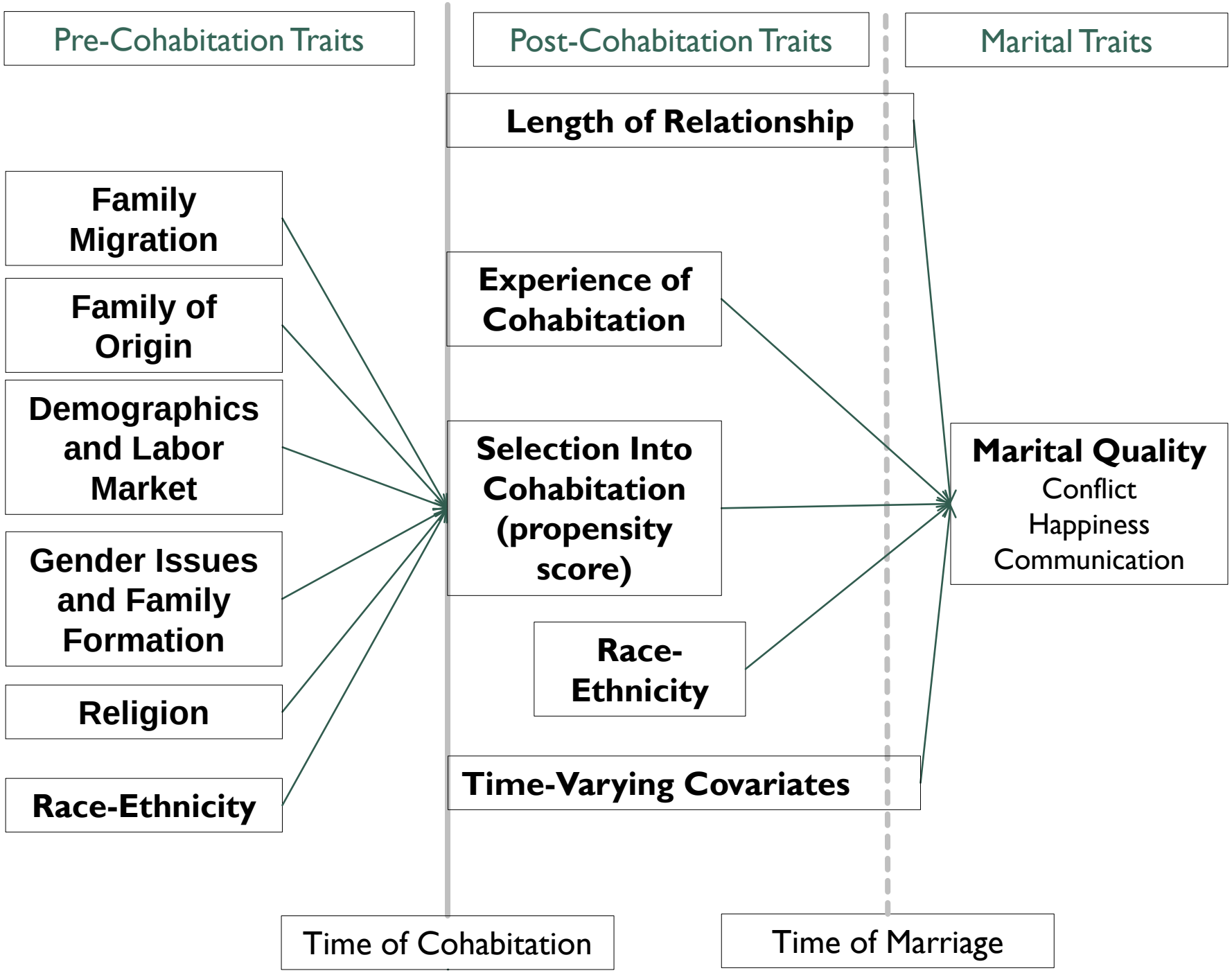


MODEL SPECIFICATION

- Must Include Both Selection and Experience




Pre-Cohabitation Traits
Family Migration
Family of Origin
Demographics and Labor Market
Gender Issues and Family Formation
Religion
Race-Ethnicity
Selection Into Cohabitation (propensity score)
Time of Cohabitation


Pre-Cohabitation Traits
Post-Cohabitation Traits
Marital Traits
Length of Relationship
Family Migration
Family of Origin
Demographics and Labor Market
Gender Issues and Family Formation
Religion
Race-Ethnicity
Experience of Cohabitation
Selection Into Cohabitation (propensity score)
Race-Ethnicity
Time-Varying Covariates
Marital Quality
Conflict
Happiness
Communication
Time of Cohabitation
Time of Marriage

VARIABLES

Independent Variables

Selection

- Propensity score varying between 0 and 1

Experience

- Coded 1 if respondent cohabited with spouse, 0 if they did not

Variables in the Propensity Score Equation

Family migration features (3 variables)
Attributes of the family of origin (7 variables)
Demographic and labor market characteristics (9 variables)
Attitudes toward gender issues and family formation (6 variables)
Religious orientation (11 variables)

Variables in the Random Effects Model

R's Experience of Cohabitation
R's Selection into Cohabitation
R's Relationship Length
R's Income
R's Education
R's # of Children
R's # of Hours Worked
R is Black
R is Hispanic

METHODS

- Estimating the Propensity Score

$$\Pr(Premarital\ Cohabitation_i = 1|X_i) = \frac{\exp(x_i\beta_{i)}}{1 + \exp(x_i\beta_{i)}}$$

- Estimating the Random Effects Model

$$MarQual = \alpha + b_1(Experience_i) + b_2(Selection) + b_3(Relationship\ Length_{it}) + b_4(Income_{it}) + b_5(Education_{it}) + b_6(\#\ of\ Children_{it}) + b_7(\#\ of\ Hours\ Worked_{it}) + b_8(Black_i) + b_9(Hispanic_i) + \varepsilon_{it}$$

$$Where\ \varepsilon_{it} = \zeta_i + e_{it}$$

Table 3. Random Effects Model Predicting Women's Marital Quality (Conflict, Happiness, and Communication), with Selection Into and Experience of Cohabitation as Predictors, NLSY79 1992-2008, (n=2898)

	Marital Conflict	Marital Happiness	Marital Communication
	(1)	(2)	(3)
Independent Variables			
Experience of Cohabitation	0.106***	-0.037**	-0.057*
	0.028	0.013	0.027
Selection into Cohabitation	-0.082	-0.119**	0.119
	0.088	0.043	0.080
Length of Relationship	-0.015***	-0.009***	-0.009***
	0.002	0.001	0.002
Random Effects			
ζ_i-Variance of the Intercept	0.29***	0.07***	0.23***
e_{it}-Residual Variance	0.69***	0.20***	0.76***
N	2898	2898	2898
Person-Years	21,245	21,245	21,245

Note: Model controls for respondent's income, education, race-ethnicity, the number of hours respondent worked during year, and the number of children living with the respondent. Missing values imputed.
*** p <.001, ** p <.01, * p <.05.

VARIABLES INCLUDED IN THE PROPENSITY SCORE EQUATION

Family Migration

R is Foreign Born

R's Father is Foreign-Born

R's Mother is Foreign-Born

Attributes of the Family of Origin

2 Parent, Biological Family

Father Employed

Mother Employed

Father's Education

Mother's Education

R's Number of Siblings

R Lived in City @ Age 14

Respondents' Demographic and Labor Characteristics

R's Education

R's Income < $15K

R's Income >15K & <40K

R's Income > $40K (Reference)

R is Black

R is Hispanic

R is non-Black, non-Hispanic (reference)

R's Age at First Interview

R's Hours Worked-Past Year

Children in the Household

Attitudes toward Gender Issues and Family Formation

Gender Attitudes

R's Desired Parity

R Expects to Marry (ref: not marry):

- before Age 20
- between 20 and 24
- between 25 and 29
- after Age 30

Religious Orientation

R was Raised Protestant

R was Raised Baptist

R was Raised Episcopalian

R was Raised Lutheran

R was Raised Methodist

R was Raised Presbyterian

R was Raised Catholic

R was Raised Jewish

R was Raised No Religion

R was Raised Other Religion (reference)

R Attends Church Rarely

R Attends Church Monthly

R Attends Church at least Weekly (reference)

POTENTIAL DISCUSSION QUESTIONS

- How do we address the possible tension between statistical parsimony and methodological rigor?
- How do other approaches to modeling selection (Heckman's method, instrumental variables, etc.) compare?
- In what instances would you choose one method over the other?
- Is it possible to *approximate* an instrumental variable approach (i.e., obtain similar parameters) using a well-estimated propensity score[1]?

I suppose this would be an 'instrumental-variables-by-committee' approach…