The Dynamics of Health Behaviors, Pregnancies, and Birth Outcomes

Jacob Klimek

University of North Carolina at Chapel Hill

October 13, 2021

Question

What role do the histories of smoking, marijuana use, and alcohol consumption play in determining birth outcomes?

- Does timing of behavior cessation prior to pregnancy matter?
  - The answer informs women's fertility behavior
  - It informs current public health guidelines that suggest that life-cycle health is important
  - It informs researchers studying women's fertility decisions
- Might the dynamic effects of behaviors differ across the life cycle?

Background

What do we know about low birth weight?

- Low birth weight births are expensive (Goldenberg and Culhane, 2007)
- Birth outcomes impact later life outcomes (Behrman and Rosenzweig, 2004; Black et al., 2007, Currie and Moretti, 2007)
- Behavior modification during pregnancy is associated with better birth outcomes (Harris et al., 2015)

Motivation

Why is continued study of behaviors and birth outcomes important?

- Poor birth outcomes are becoming more common nationally (Goldenberg et al., 2008; NCHS, 2019)
- Differential trends in marijuana use and smoking (Gnofam et al., 2020; NCHS, 2018)
- Prenatal care uptake is increasing
- Proportion of pregnancies deemed as "high-risk" is increasing (BCBS, 2020)
- Lack of studies regarding the causal impacts of health behaviors *prior* to *pregnancy* on birth outcomes (Witt et al., 2012; Strutz et al. 2012)

Contribution

- Estimate a model that captures the dynamic effects of health behaviors during fecundity on pregnancy and birth outcomes
- Separate the direct and indirect effects that health behaviors have on birth outcomes
- Measure the different impacts of health behaviors at different ages
  - Do smoking and marijuana use before pregnancy have different impacts on birth outcomes?
  - Are there any interactions between behavior cessation?

Behaviors and Outcomes

What health behaviors are modeled:

- I observe smoking, marijuana use, and binge drinking
- I do not observe nutrition, exercise, preventive and curative medical care, or mental health
    - But I do observe body mass index (BMI) and self-reported health status (SRHS)

Behaviors and Outcomes

What health behaviors are modeled:

- I observe smoking, marijuana use, and binge drinking
- I do not observe nutrition, exercise, preventive and curative medical care, or mental health
    - But I do observe body mass index (BMI) and self-reported health status (SRHS)

What birth outcomes are modeled:

- Main interest: probability of a low birth weight (LBW) birth
- Other outcomes of interest:
    - Probability of pregnancy
    - Probability of a live birth conditional on being pregnant
    - Gestational age at birth conditional on live birth

Data Features

National Longitudinal Survey of Youth 1997

- Panel data following a nationally representative group aged 14-18 in 1997
- Surveyed annually for 15 years (then biennially for 3 more waves)
- Information on health behaviors, health measures, and pregnancy in each survey wave
- Geographic extract allows me to link market level information from a variety of sources

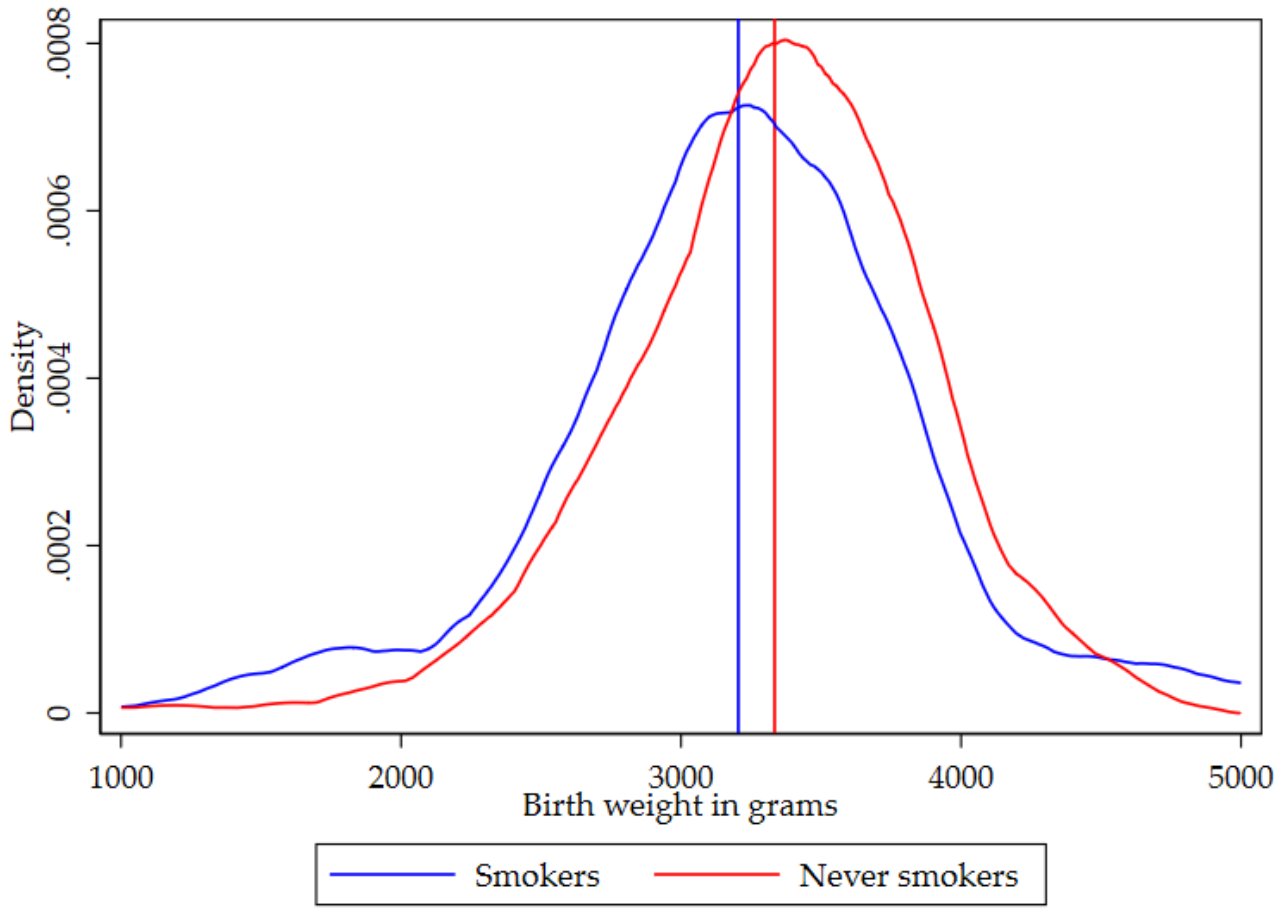
Sample Description

The estimation sample includes all years of data for women who:

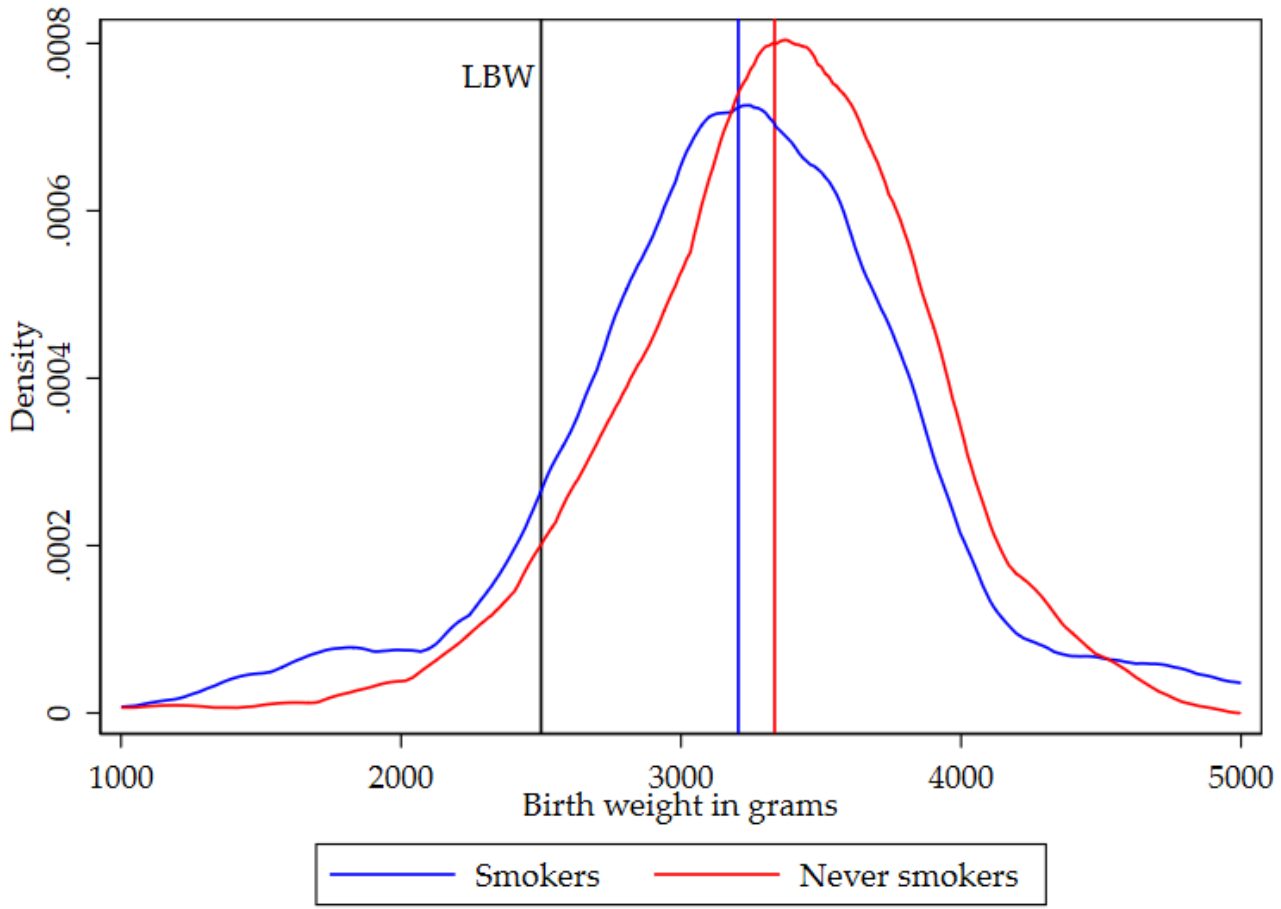
- Provide at least two consecutive years of data
- Have not been pregnant prior to the first observed year of data
- And until the first year of non-response to important dependent variables
- Only the first 15 survey waves are used

The estimation sample contains 3,675 women (84% of full sample) who contribute 40,797 consecutive person-years.

Distribution of birth weight (NLSY97)



Distribution of birth weight (NLSY97)



Empirical observations from NLSY97

Table 1: LBW probability for different ages and histories of smoking

	Never smoker	Past smoker	Recent smoker	Current smoker
Age 25 and up	6.42%	6.25%	9.59%	11.79%
Under age 25	9.07%	6.62%	10.26%	12.09%

Empirical observations from NLSY97

Table 1: LBW probability for different ages and histories of smoking

	Never smoker	Past smoker	Recent smoker	Current smoker
Age 25 and up	6.42%	6.25%	9.59%	11.79%
Under age 25	9.07%	6.62%	10.26%	12.09%

- Recent smokers have higher likelihood of LBW than past smokers
- Association between LBW and smoking history may change with age

Empirical observations from NLSY97

Table 2: Correlation of health behaviors with LBW probability

	Indicator only	No covariates	Covariates
Smoked entering t	0.037***	0.013	0.011
	(0.014)	(0.019)	(0.018)
Ever smoked		0.025	0.037**
		(0.021)	(0.019)
Smoking cessation (years)		-0.006*	-0.007**
		(0.003)	(0.003)

Empirical observations from NLSY97

Table 2: Correlation of health behaviors with LBW probability

	Indicator only	No covariates	Covariates
Smoked entering t	0.037***	0.013	0.011
	(0.014)	(0.019)	(0.018)
Ever smoked		0.025	0.037**
		(0.021)	(0.019)
Smoking cessation (years)		-0.006*	-0.007**
		(0.003)	(0.003)

Compared to someone who smokes while pregnant:

- A woman who never smokes has a 4.8 pp lower probability of having a LBW birth
- A woman who quit smoking three years before she got pregnant has a 1.6pp lower probability of having a LBW birth

Empirical observations from NLSY97

Table 2: Correlation of health behaviors with LBW probability

	Indicator only	No covariates	Covariates
Smoked entering t	0.037***	0.013	0.011
	(0.014)	(0.019)	(0.018)
Ever smoked		0.025	0.037**
		(0.021)	(0.019)
Smoking cessation (years)		-0.006*	-0.007**
		(0.003)	(0.003)
Used marijuana entering t	0.041**	0.038	0.040*
	(0.018)	(0.024)	(0.024)
Ever used marijuana		0.0099	0.0076
		(0.020)	(0.019)
Marijuana cessation (years)		0.002	0.003
		(0.003)	(0.003)
Observable characteristics	N	N	Y

Preview of Model

I jointly estimate a set of dynamic equations.

- Equations approximate demand and production functions
- Allows for unobserved correlation within and across time periods
- Estimate using full information maximum likelihood (FIML)
- Disentangle direct and indirect effects histories of health behaviors have on probability of LBW

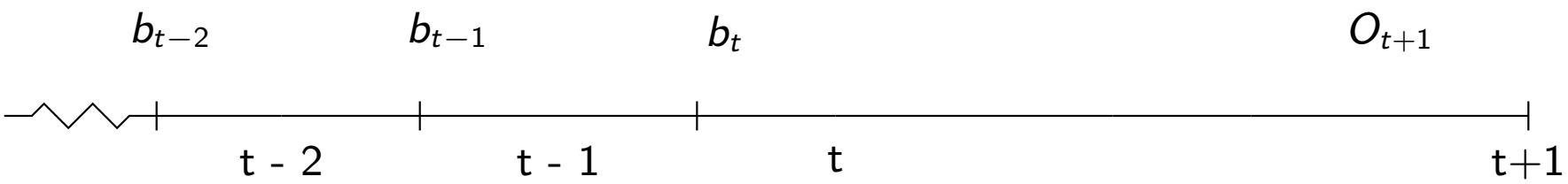
Motivating Model

Let the probability of LBW entering $t+1$ (O_{t+1}) be defined by behaviors during and prior to pregnancy:

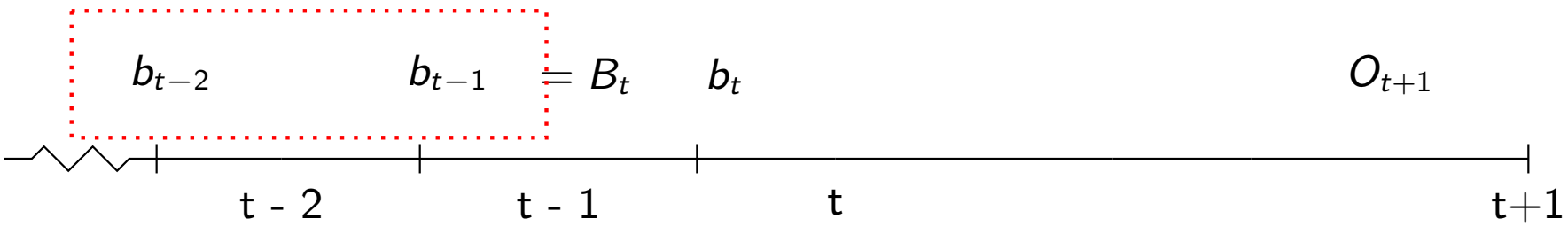
$$Pr(O_{t+1} = o | P_t, L_{t+1}) = f^O(b_t, B_t, X_t, u_{it}^O) \quad \forall t, \quad o = 0, 1 \quad (1)$$

- Smoking, marijuana use, and binge drinking are captured by $b_t = (b_t^1, b_t^2, b_t^3)$
- B_t captures the history of behaviors up to period t for all behaviors
- X_t is a placeholder for other observed characteristics
- P_t is pregnancy, L_{t+1} is live birth
- u_{it}^O is unobserved to the econometrician

Timing

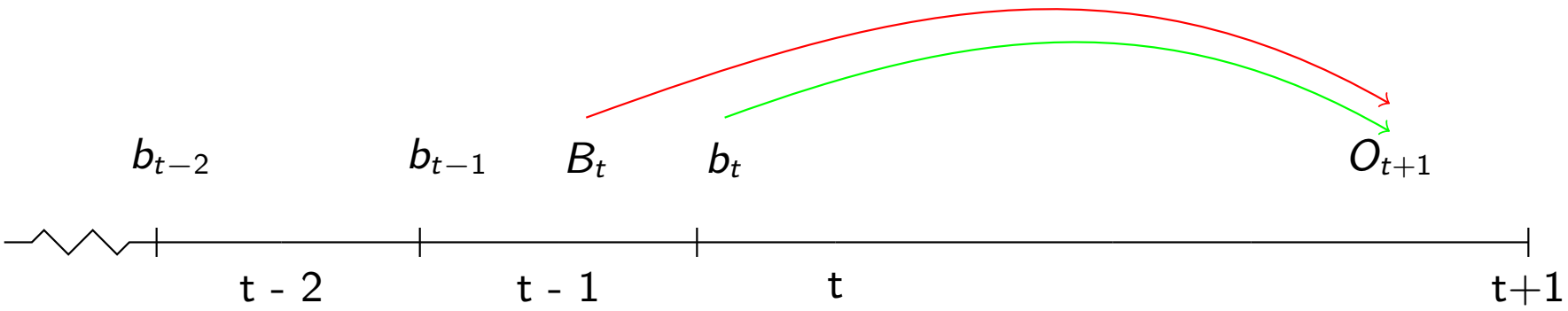


Timing



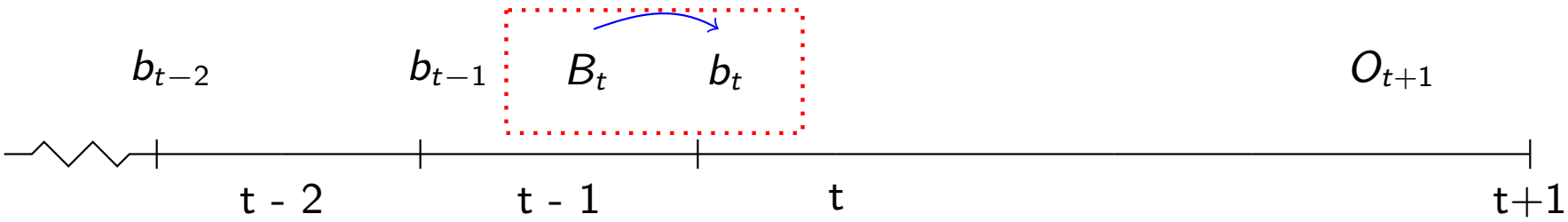
B_t describes the history of all behaviors $(b_1, ..., b_{t-1})$ in a parsimonious way, and is observed entering period t.

Timing



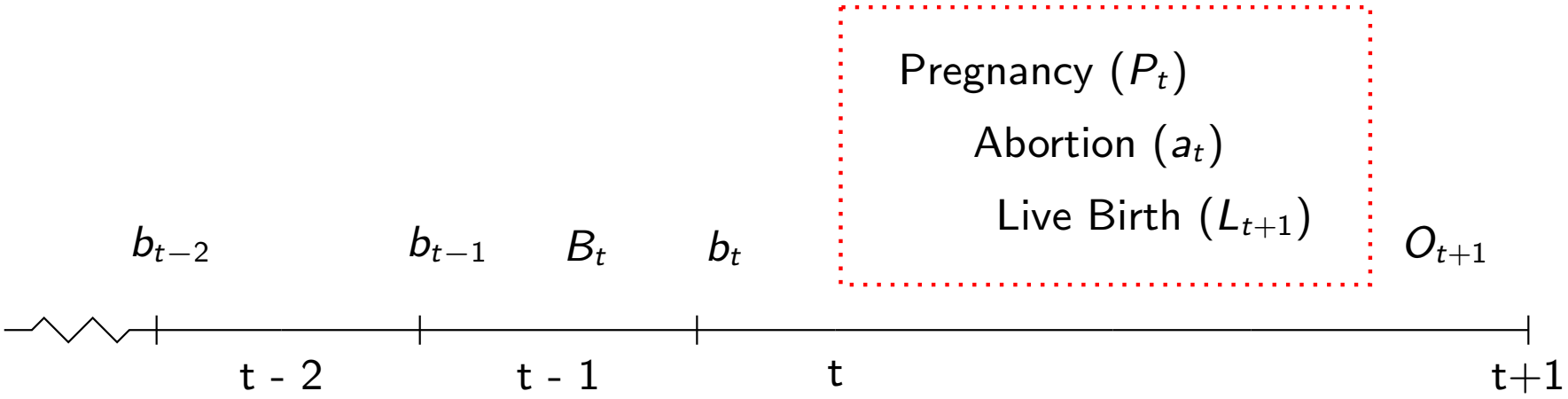
I want to distinguish the effects of b_t and B_t.

Timing



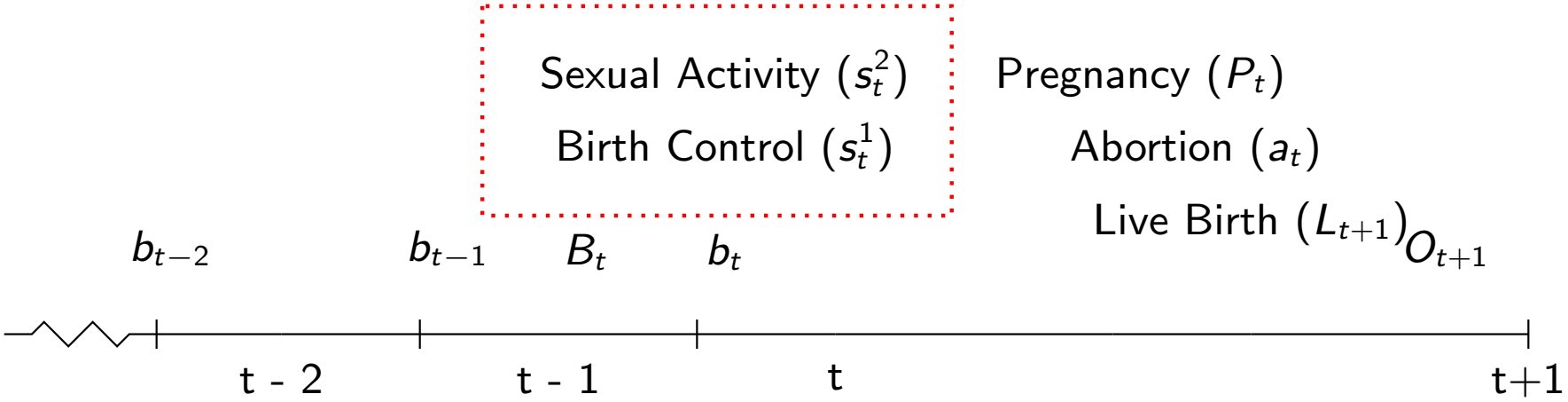
Yet, B_t influences b_t.

Timing



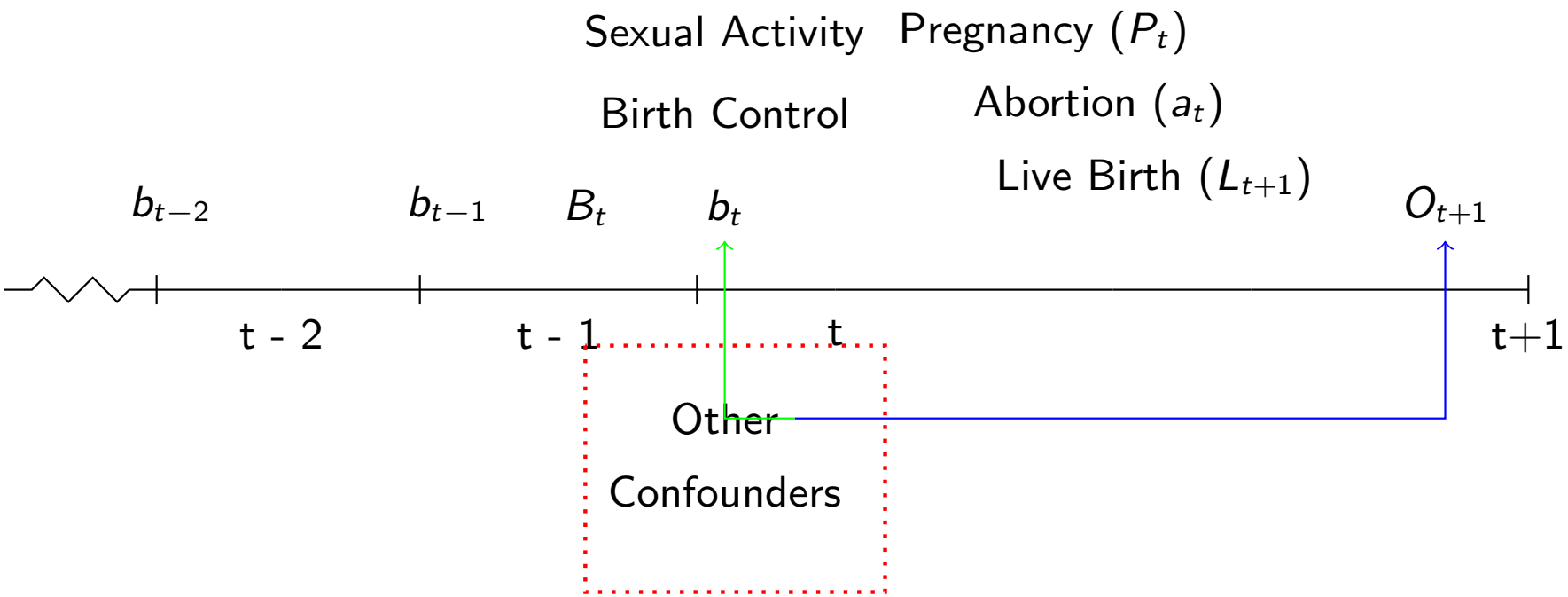
It's also the case that behaviors like smoking influence selection into pregnancy/birth.

Timing

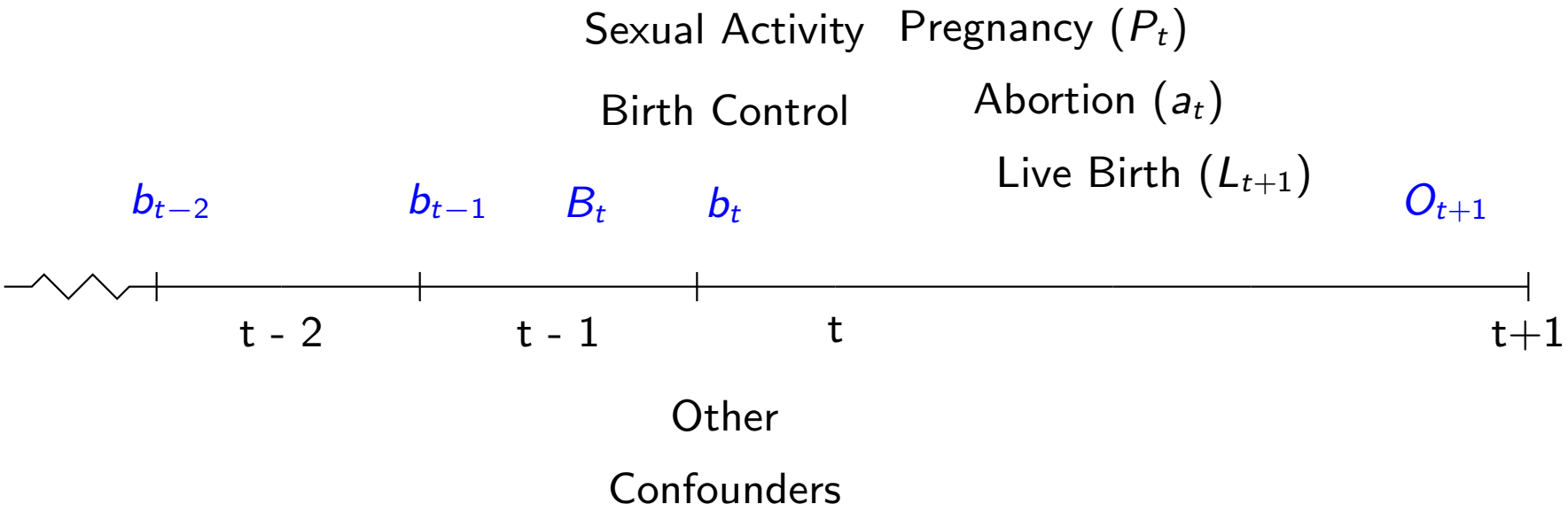


All health (b_t) and pregnancy-related (s_t) behaviors are determined simultaneously.

Timing


Sexual Activity
Pregnancy (P_t)
Birth Control
Abortion (a_t)
Live Birth (L_{t+1})
b_{t-2}
b_{t-1}
B_t
b_t
O_{t+1}
t - 2
t - 1
t
t+1
Other
Confounders

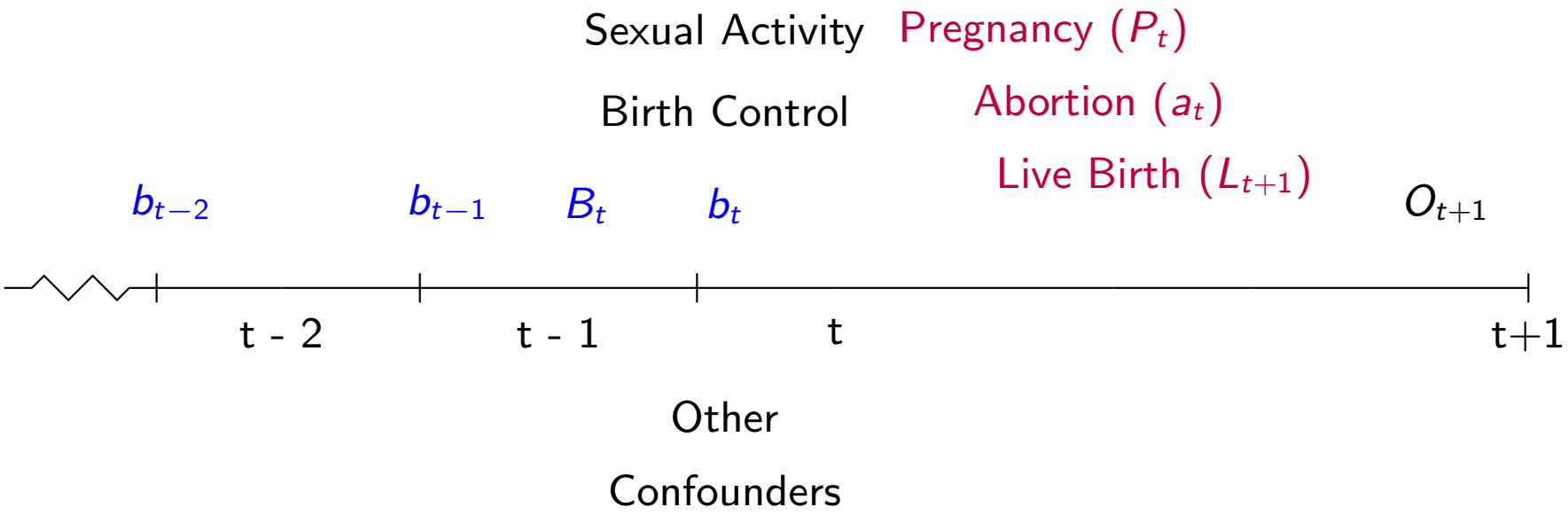
Timing



Recap:

1. Endogeneity

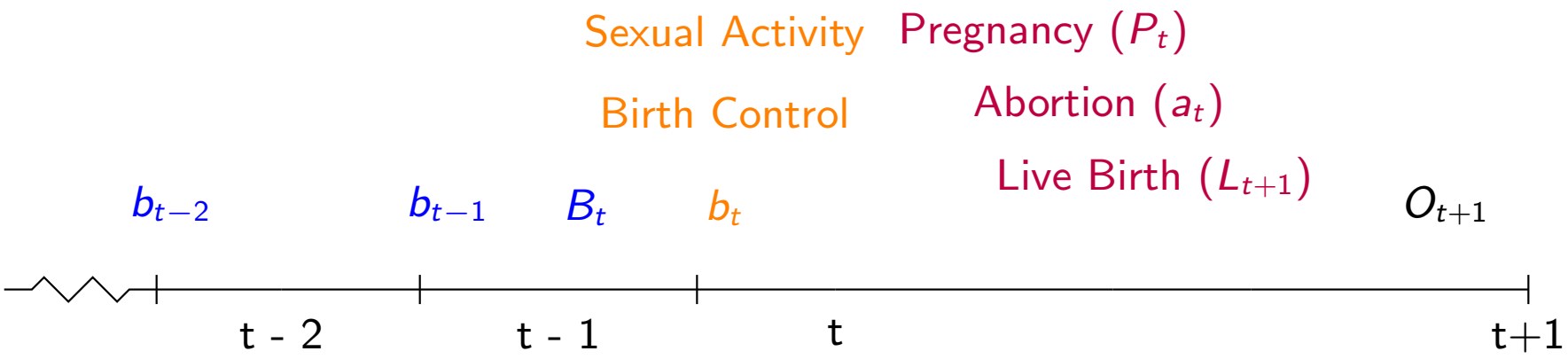
Timing



Recap:

1. Endogeneity
2. Selection

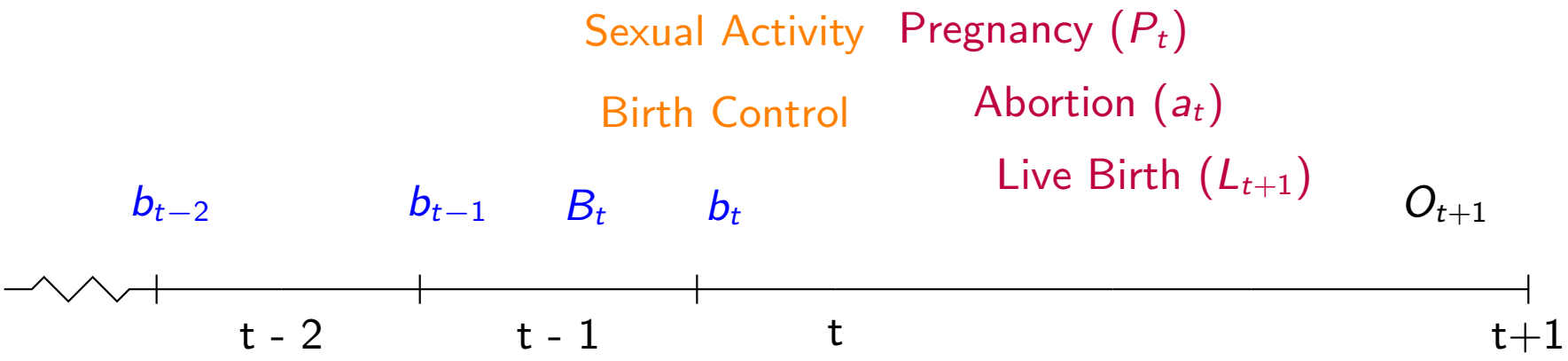
Timing



Recap:

1. Endogeneity
2. Selection
3. Simultaneity

Timing



Recap:

1. Endogeneity
2. Selection
3. Simultaneity
4. Confounding

Estimation Equations

What's included in the model?

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking
- Sexual activity and birth control use

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking
- Sexual activity and birth control use
- Other confounders: relationship status, school enrollment, and employment

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking
- Sexual activity and birth control use
- Other confounders: relationship status, school enrollment, and employment
- Pregnancy/Birth: Pregnancy, abortion, live birth, gestational age at birth, and LBW status

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking
- Sexual activity and birth control use
- Other confounders: relationship status, school enrollment, and employment
- Pregnancy/Birth: Pregnancy, abortion, live birth, gestational age at birth, and LBW status
- Health measures: BMI (continuous), and self-reported health status

Estimation Equations

What's included in the model?

- Health behaviors: smoking, marijuana use, binge drinking
- Sexual activity and birth control use
- Other confounders: relationship status, school enrollment, and employment
- Pregnancy/Birth: Pregnancy, abortion, live birth, gestational age at birth, and LBW status
- Health measures: BMI (continuous), and self-reported health status
- Attrition
- Initial conditions for health behaviors and health measures

Unobserved Heterogeneity

I decompose the unobserved component of equation j, (u_{it}^j), into three additively separable parts.

- A permanent component, μ^j
- A time-varying component, ν_t^j
- An i.i.d. component, (ϵ_{it}^j)

$$u_{it}^j = \mu_j + \nu_t^j + \epsilon_{it}^j$$

where ϵ_{it}^j is distributed normally for continuous outcomes and Extreme Value for dichotomous/polychotomous outcomes.

Estimating the distribution of UH (Mroz, 1999)

I estimate a discrete distribution of each component (μ, ν_t), allowing for a non-linear relationship across equations.

- I discretize the support of each component and estimate mass points and weights
- Relaxes typical assumption of normality
- Method sometimes referred to as discrete factor method (DFM) or discrete factor random effects (DFRE)

Estimation and Identification

I estimate the parameters of a multiple equation dynamic empirical model via full information maximum likelihood.

Identification:

- Exclusion restrictions - motivated by theory- in the behavior equations (health behaviors, abortion) and selection (pregnancy, live birth)

Estimation and Identification

Exclusions for behaviors:

- Shift demand for behaviors
- Unrelated to probability of LBW conditional on those behaviors
- Includes time-varying supply-side market conditions and prices (Gilleskie et al., 2017; Darden et al., 2021)

Estimation and Identification

Exclusions for behaviors:

- Shift demand for behaviors
- Unrelated to probability of LBW conditional on those behaviors
- Includes time-varying supply-side market conditions and prices (Gilleskie et al., 2017; Darden et al., 2021)

Exclusions for pregnancy:

- Time-varying male-female ratios at county level

Exclusions for abortion:

- Time-varying access to abortion clinics and abortion rate per capita at state level

Estimation and Identification

I estimate the parameters of a multiple equation dynamic empirical model via full information maximum likelihood.

Identification:

- Exclusion restrictions - motivated by theory- in the behavior equations (health behaviors, abortion) and selection (pregnancy, live birth)
- Non-linear functional form of demand (behaviors) and production (outcome) equations (Guilkey and Lance, 2014; Lewbel, 2019; Mroz, 1999)
- Variation in histories of exogenous variables that explain the lagged outcomes in the dynamic equations (Bhargava and Sargan, 1983; Arellano and Bond, 1991)

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^{O}(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^{O}, u_t^{O})$$

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

- B_t describes the history of behaviors entering t Details

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

- B_t describes the history of behaviors entering t Details
- H_t are health measures entering period t

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

- B_t describes the history of behaviors entering t Details
- H_t are health measures entering period t
- N_t characterizes the history of pregnancy and birth outcomes Details

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

- B_t describes the history of behaviors entering t Details
- H_t are health measures entering period t
- N_t characterizes the history of pregnancy and birth outcomes Details
- b_t, s_t and c_t are period t endogenous behaviors

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

- B_t describes the history of behaviors entering t Details
- H_t are health measures entering period t
- N_t characterizes the history of pregnancy and birth outcomes Details
- b_t, s_t and c_t are period t endogenous behaviors
- X_t is exogenous demographic information

Low Birth Weight Specification

$$Pr(O_{t+1} = o | L_{t+1} = l) = f^O(B_t, H_t, N_t, b_t, s_t, c_t, X_t, Z_t^O, u_t^O)$$

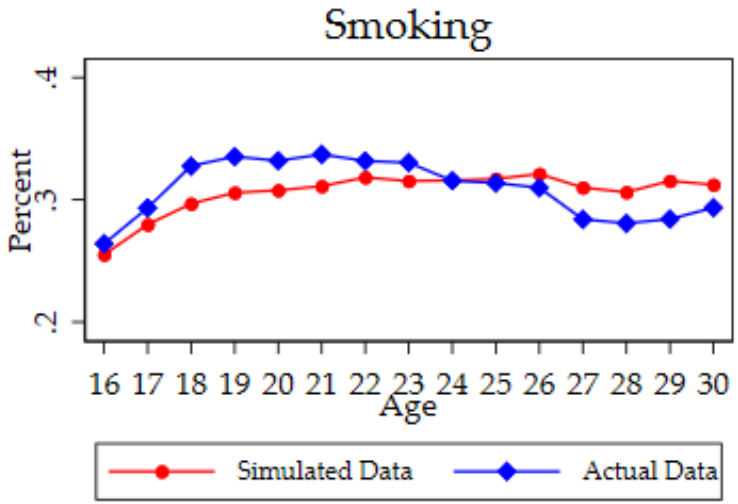
- B_t describes the history of behaviors entering t Details
- H_t are health measures entering period t
- N_t characterizes the history of pregnancy and birth outcomes Details
- b_t, s_t and c_t are period t endogenous behaviors
- X_t is exogenous demographic information
- Z_t^O represents exogenous location-specific period t supply-side information

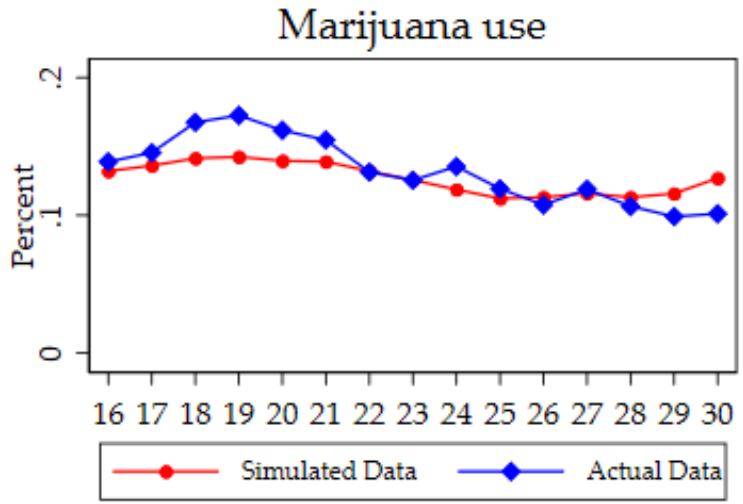
Model Fit

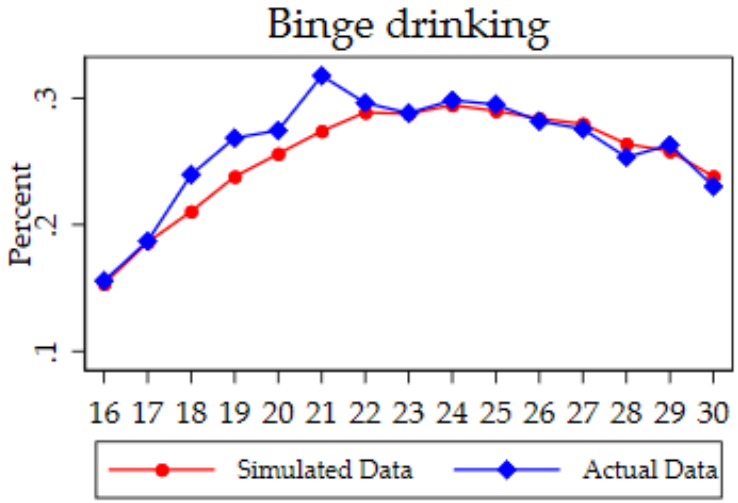
I use the estimated parameters of the model to simulate behavior for all individuals from age 14 to age 32.

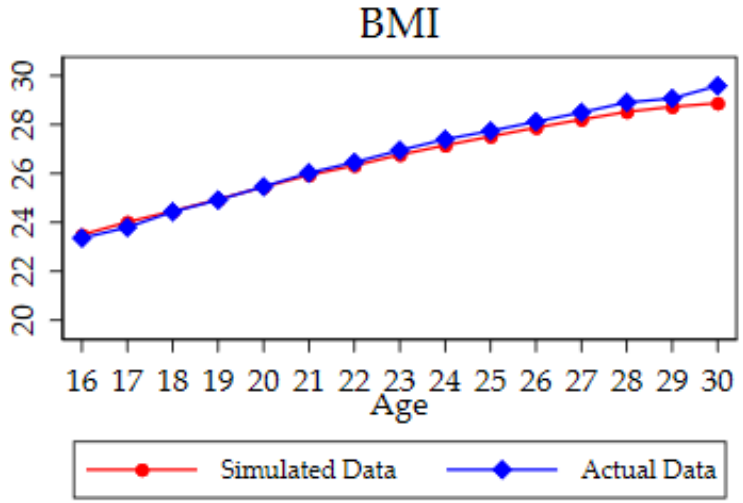
- Each individual-year is replicated 500 times
- All endogenous variables update over time based on parameter estimates and random error draws

Model fit of health behaviors and BMI

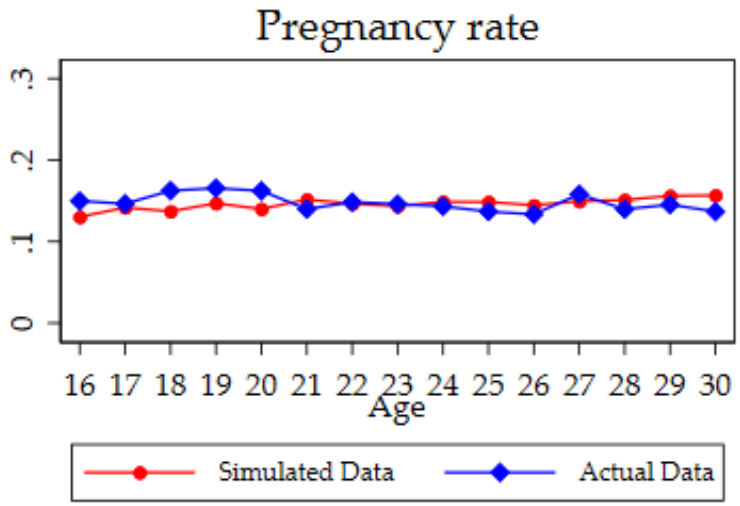


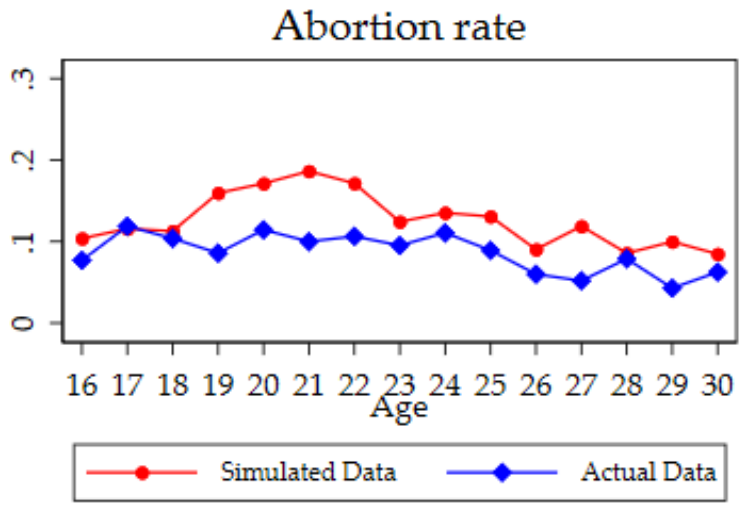


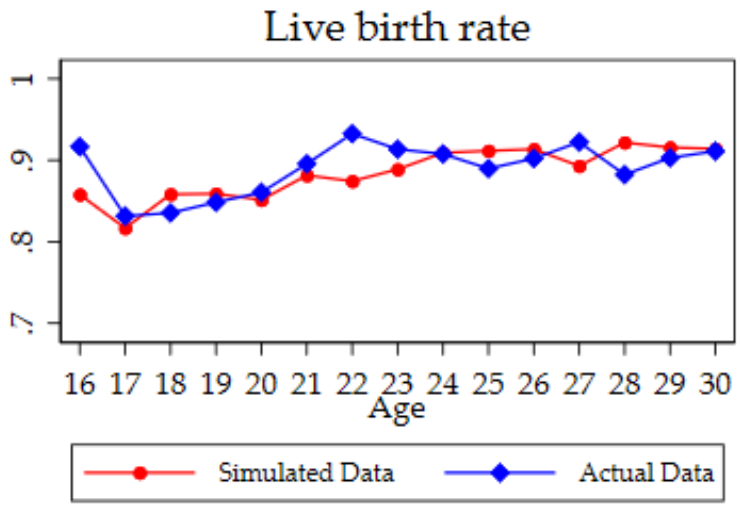


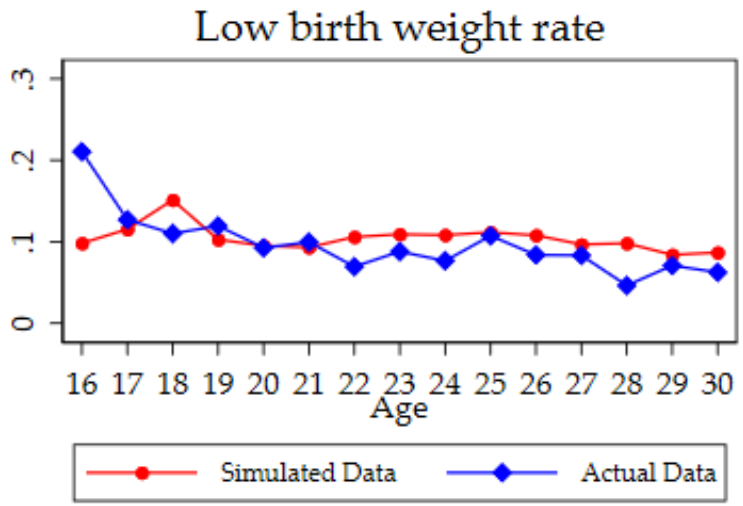


Model fit of pregnancy, abortion, live birth, and LBW









Simulation

I conduct several simulations that alter smoking (or marijuana use) behaviors over time.

- Each individual-year is replicated 500 times
- Smoking (or marijuana use) is imposed
- All other endogenous variables update over time based on parameter estimates, random error draws, and updated values of the endogenous history of variables based on simulated values

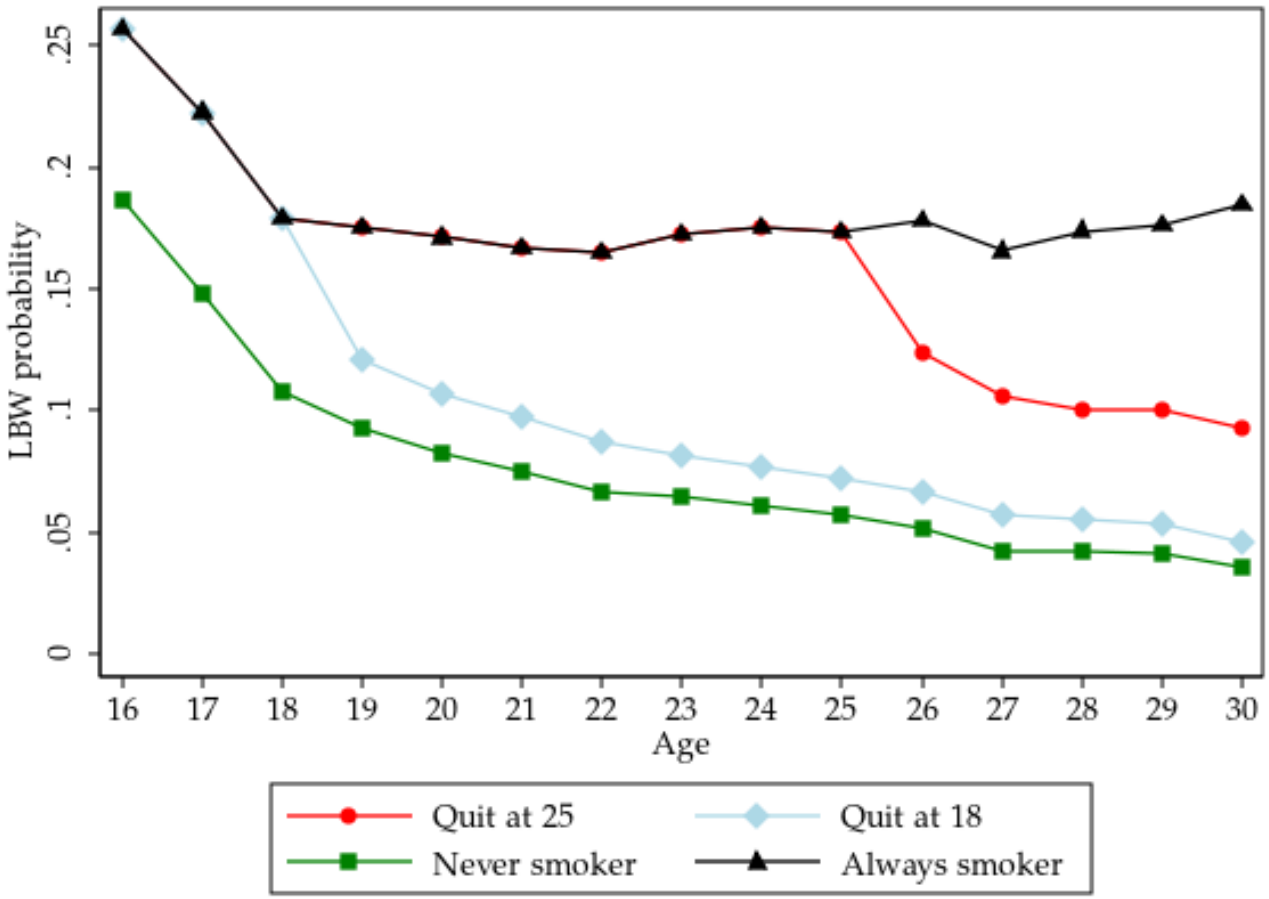
Simulation 1

Fix smoking patterns to reflect four different histories of smoking:

- Always smoke
- Quit smoking at age 25
- Quit smoking at age 18
- Never smoke

and evaluate LBW probability from age 16-30.

Impact of different smoking patterns on LBW prob



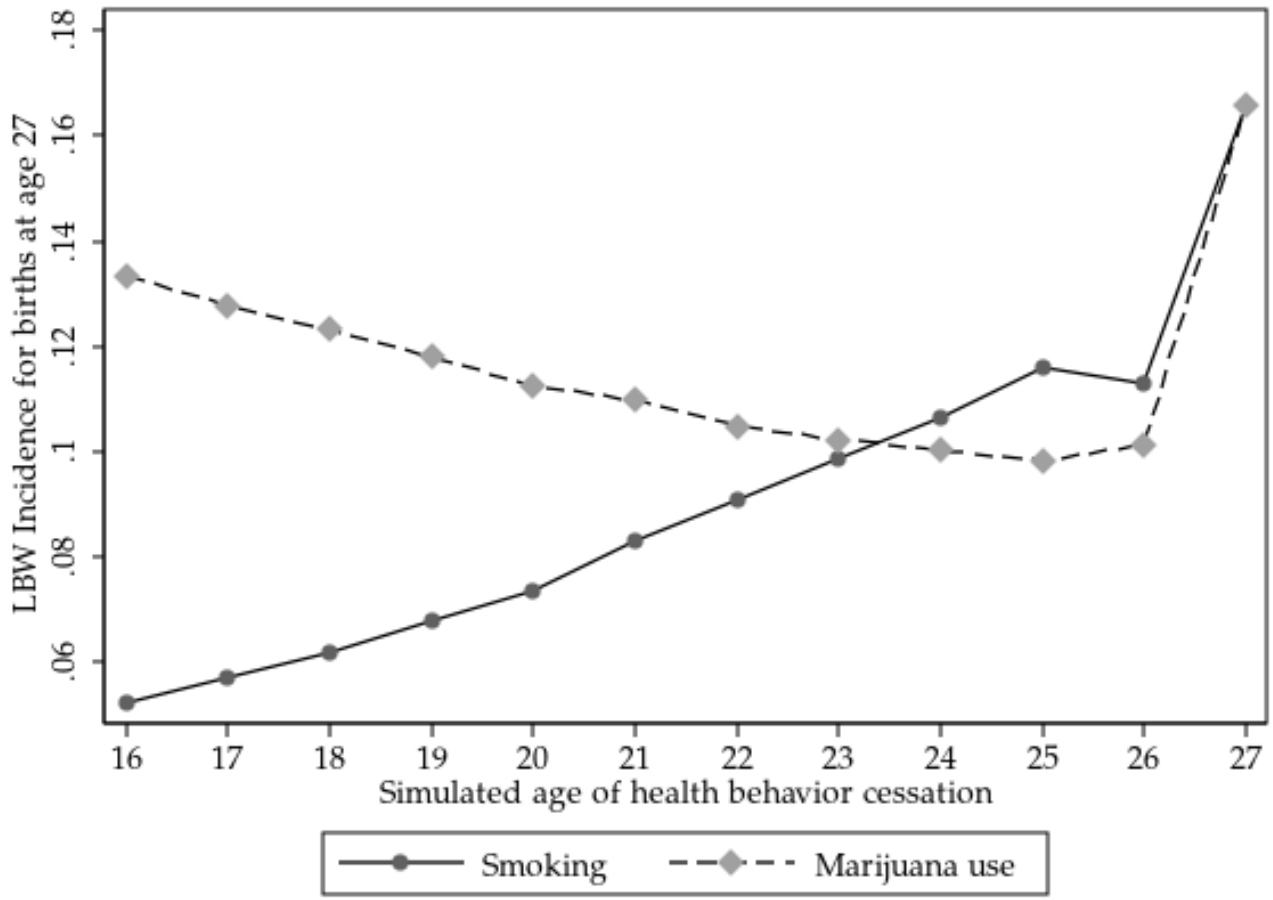
Note: Conditional on pregnancy, no abortion, and live birth

Simulation 2

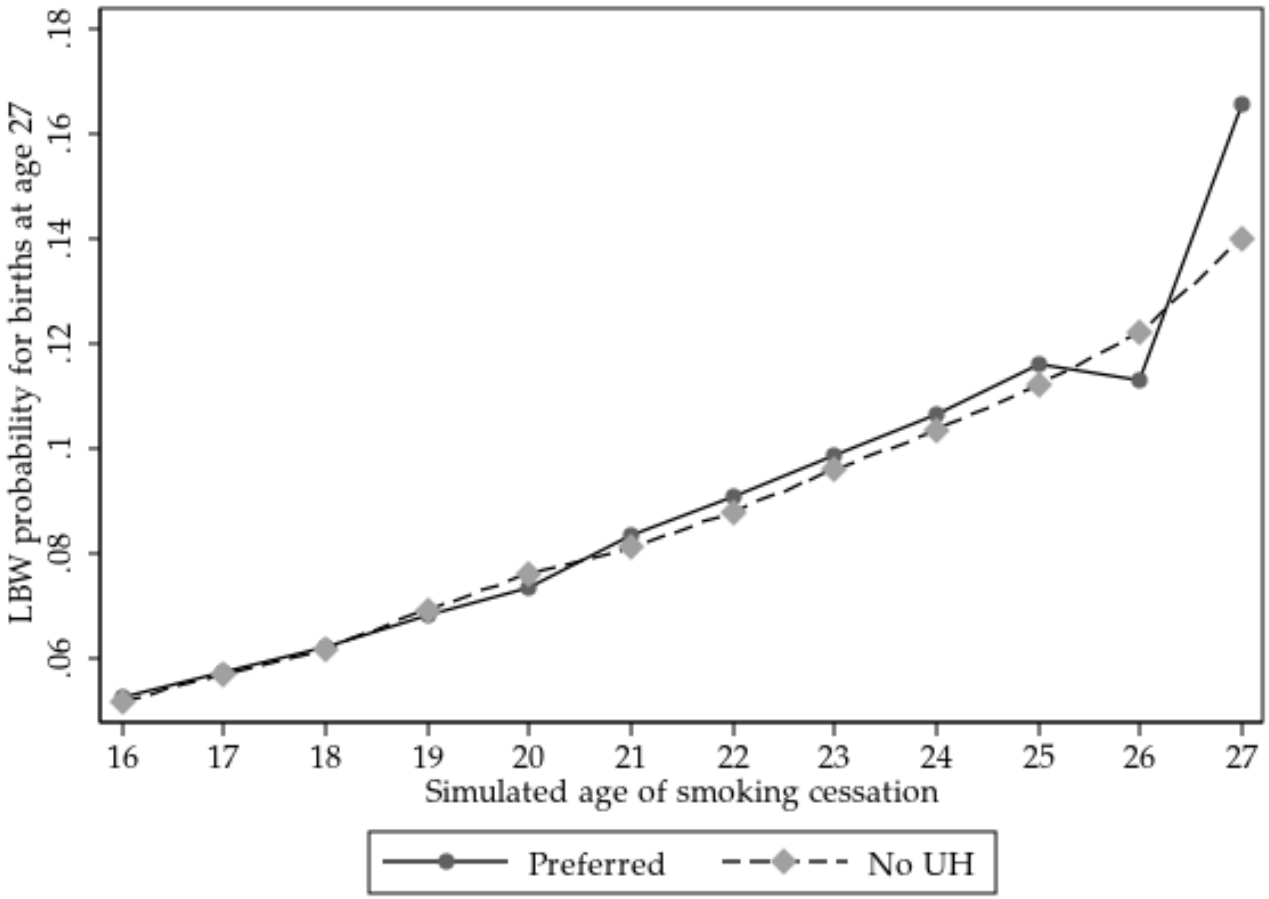
Simulate women to quit smoking (or marijuana use) permanently at each age starting with age 16.

- I evaluate births for women who are 27 years old
- I simulate women to quit at age 16 and summarize LBW probability for women who give birth at 27
- Then I simulate women to quit at age 17 and so on...

Age-specific cessation effects on LBW: smoking vs. marijuana use



What difference does modeling UH make?



Conclusion

- A history of smoking impacts the probability of LBW
  - There is a direct effect of past smoking behavior *conditional* on smoking behavior while pregnant
  - There is an indirect effect of past smoking through its effect on other behaviors/outcomes and selection into an observed birth
- There are substantial dynamic differences in how smoking and marijuana use affect the probability of LBW
  - The impact of smoking cessation increases gradually over time
  - The impact of marijuana use is concentrated to more recent use

Questions

Details on B_t

The history of health behaviors entering period t is summarized by four variables for each behavior:

- Indicator of participation in the behavior in $t-1$
- Indicator of ever having particpated in the behavior up to t
- Number of years of cessation conditional on ever having done the behavior
- Age interacted with lagged behavior to allow for different effects across age

Back

Details on N_t

The history of pregnancy, abortion, and birth outcomes entering period t is characterized by four variables:

- Number of pregnancies up to t
- Number of live births up to t
- Indicator of a LBW birth up to t
- Indicator of an abortion up to t

Back