



INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER  
**PANEL STUDY OF INCOME DYNAMICS**  
UNIVERSITY OF MICHIGAN



## Data Orientation

Quantitative Methods Working Group

Department of Sociology

University of Pennsylvania

February 9, 2023

Paula Fomby, [pfomby@sas.upenn.edu](mailto:pfomby@sas.upenn.edu)

# Orientation overview

- Overview of PSID design and content
- Youth-focused supplements and genomic data
- Tools for mapping relationships
- PSID Data Center
- Working with PSID data




# PSID Overview

- Study background
- Design and content
- Sample membership
- Data for demographic research



# PSID in the beginning

- Started in 1968
  - Extension to the 1967 Survey of Economic Opportunity
  - Founded by James Morgan at University of Michigan
- Originally planned as a 5-year study to understand the causes and consequences of poverty in the US
  - Measuring poverty *dynamics*
- 4,802 American families with a low-income oversample

|   |  |
|---|--|
| (INTERVIEW NUMBER)  | STUDY OF FAMILY ECONOMICS<br>Project 768   |
|  SURVEY RESEARCH CENTER<br>INSTITUTE FOR SOCIAL RESEARCH<br>THE UNIVERSITY OF MICHIGAN | (Do not  |
| 1. Interviewer's Label  | 2. P.S.U. _____<br>3. Your Interview # _____<br>4. Date _____<br>5. Length of Interv _____ |

A1. Are you the head of this household?  
 YES  NO - (ASK FOR HEAD)



# PSID's genealogical design

- Original householders followed until death/attrition, even if they move or enter separate households
  - Longitudinal/intragenerational
- Householders' adult children become PSID respondents themselves when they establish their own households
  - Multigenerational
- Periodic immigrant sample refreshment (1997/99, 2017/19)
  - Ongoing national representation



# Structure of core interview

- Respondent
  - One household informant reports for family unit
- Periodicity
  - Annual interviews 1968-1997
  - Biennial interviews 1997-present (March-December 31)
- Mode
  - By telephone since 1973 (CATI since 1993)
  - Web/CATI introduced in 2021
- Length
  - 75-80 minutes in recent waves



# Content of core interview

- Interview pertains to people living in a shared family household
  - Reference person = first householder (in different-sex couple-headed households, always male)
  - Spouse/cohabiting partner = second householder (in different-sex couple-headed households, always female)
  - Other family unit members (OFUMs)
    - Big OFUMs = 16+ years
    - Little OFUMs = 0-15 years



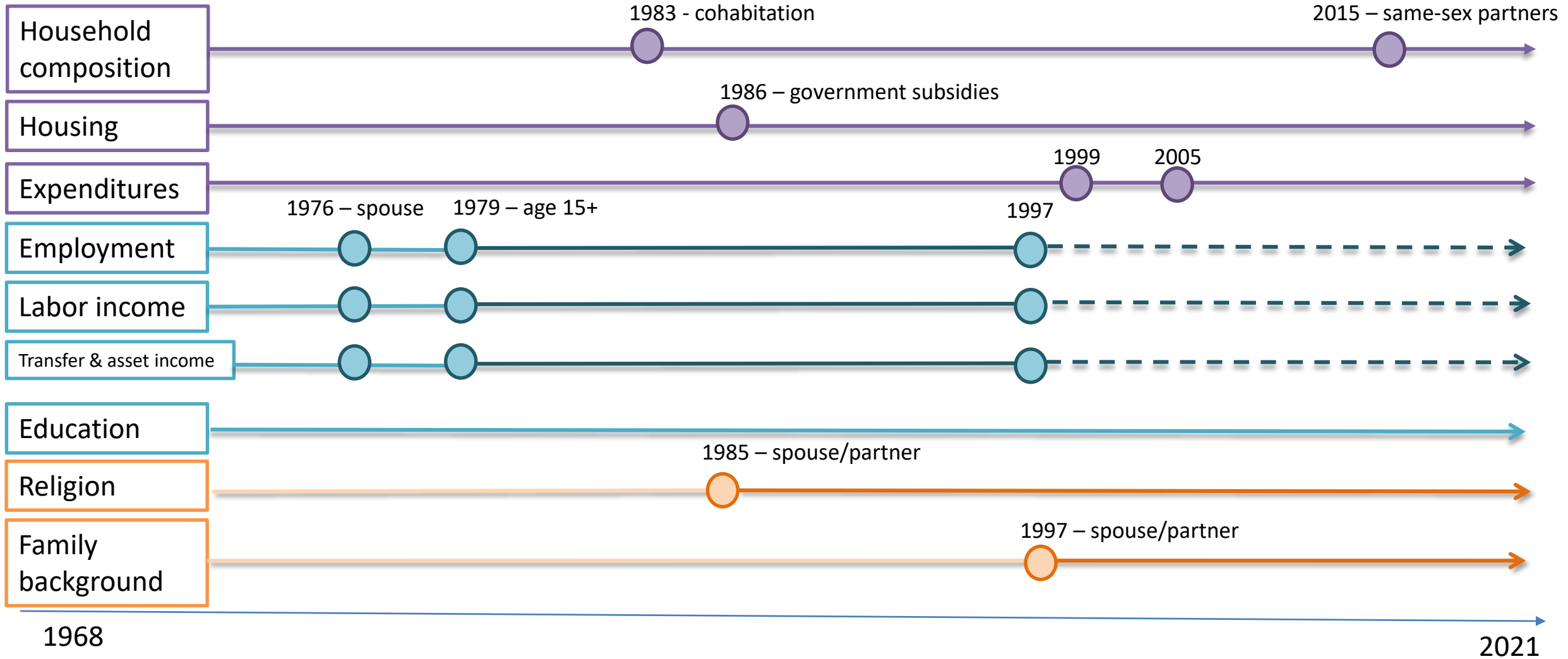
# Content of core interview

- Interview pertains to people living in a shared family household
  - Family-level characteristics
    - Fixed or aggregated information applies to everyone
    - Housing characteristics, total family income
  - Individual-level characteristics
    - Characteristics observed to vary between family members
    - Age, sex, years of education, labor income, marriage/birth history
    - Includes OFUMs (age-dependent content)
  - Householder characteristics
    - Measured at the individual level but observed only for householders (reference person and spouse/partner)
    - Health detail, 2-year employment history

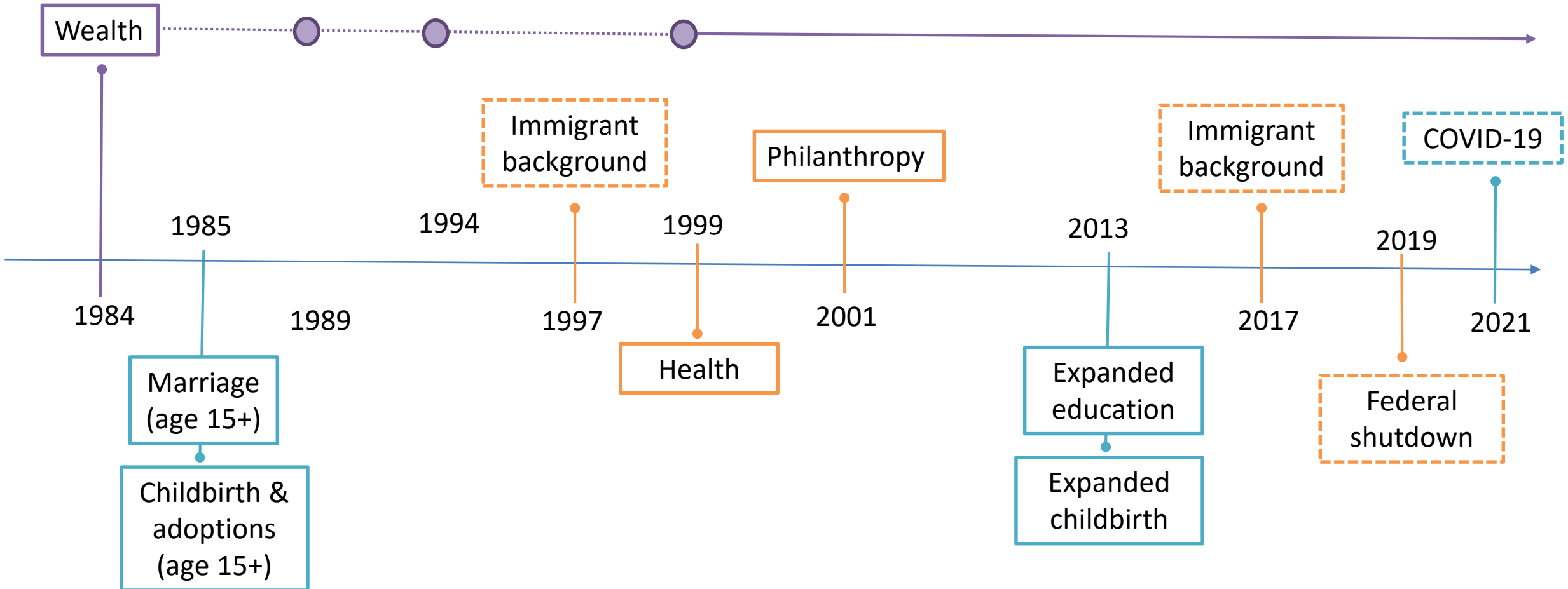




# Content of core interview



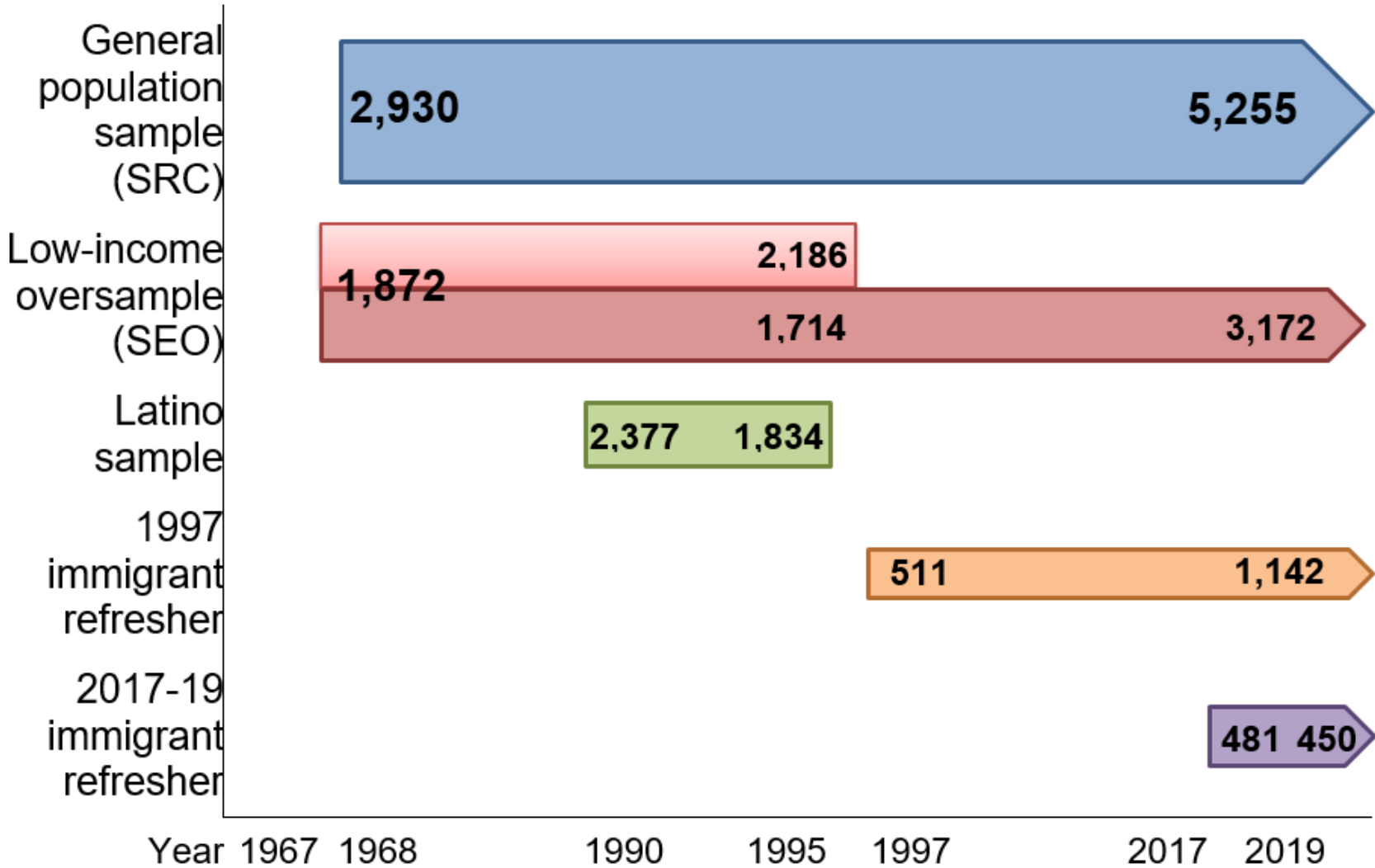
# Added core interview content



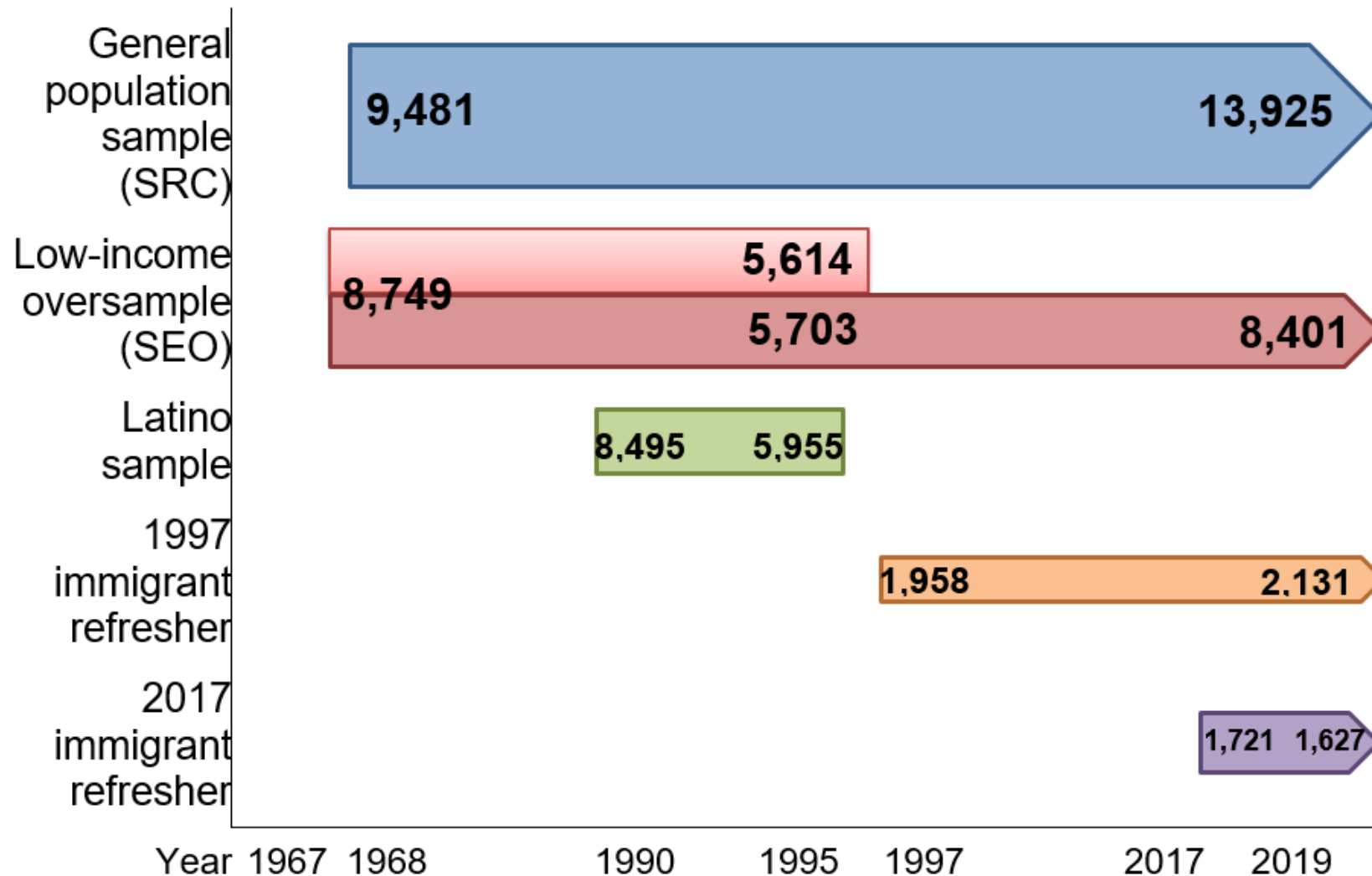
Solid box = ongoing content

Dotted box = special/one-time topic

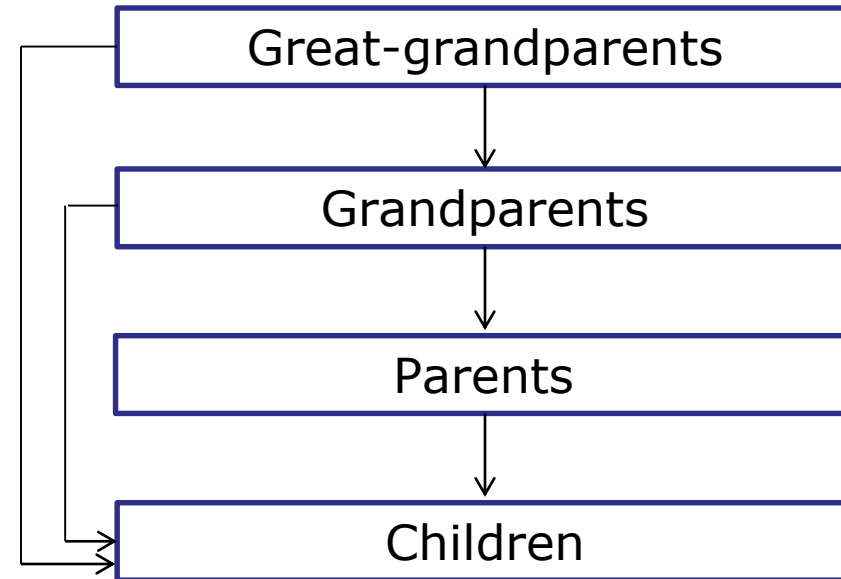
# Sample composition – family unit



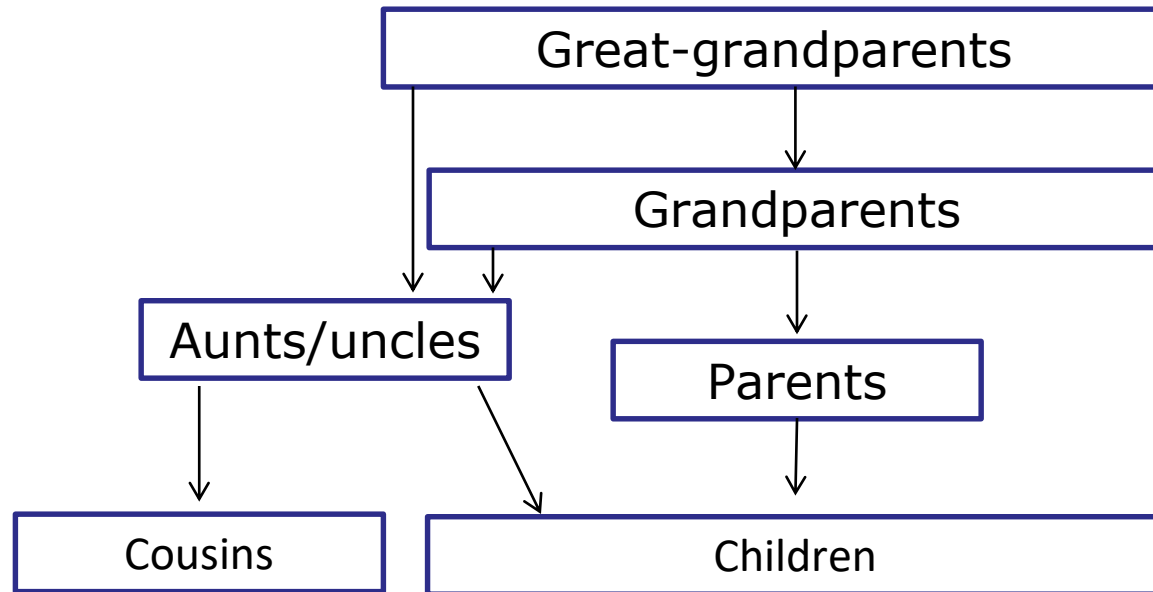
# Sample composition – individuals



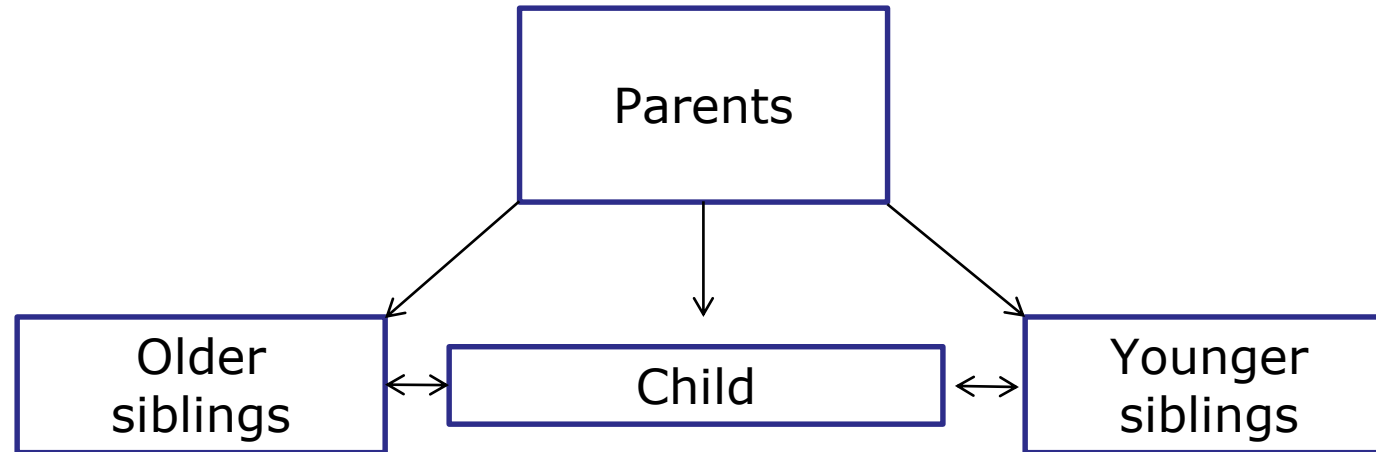
# Multigenerational Studies



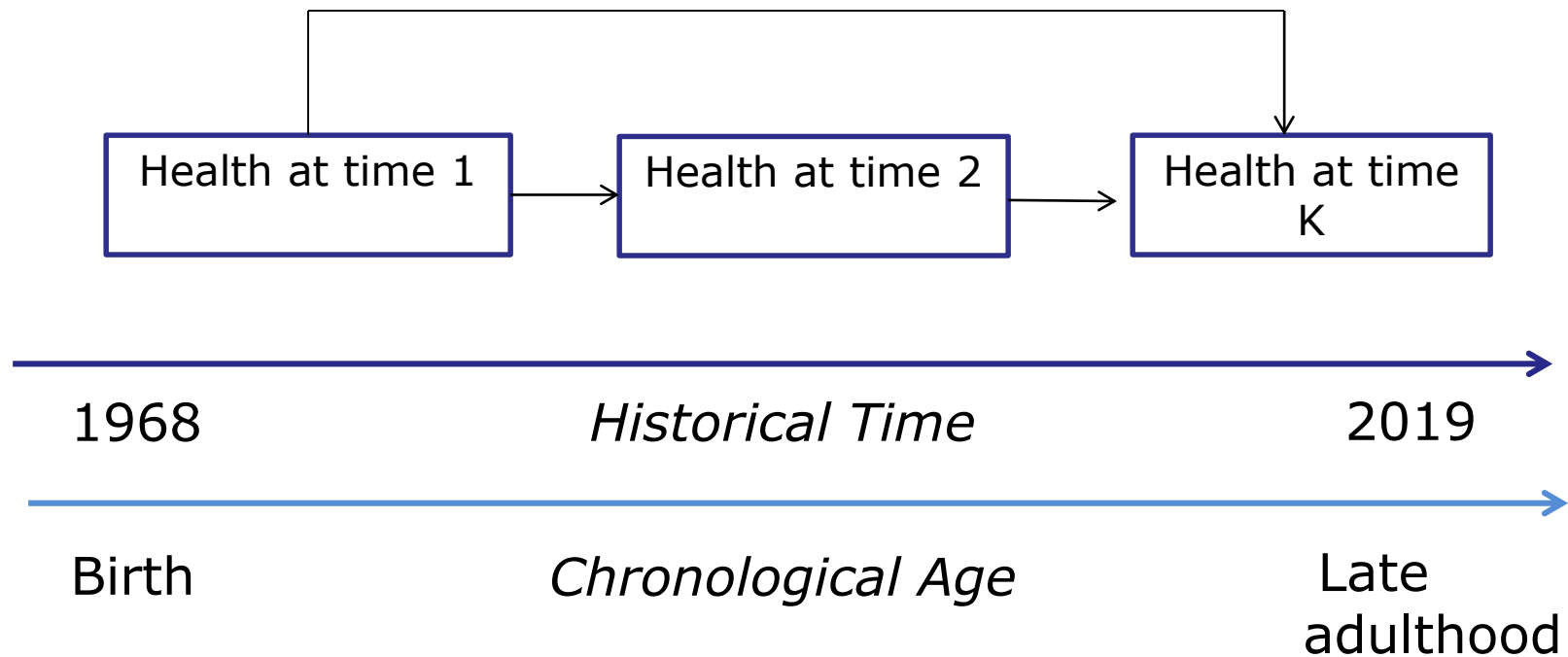
# Kinship Studies



# Within-Family Variation

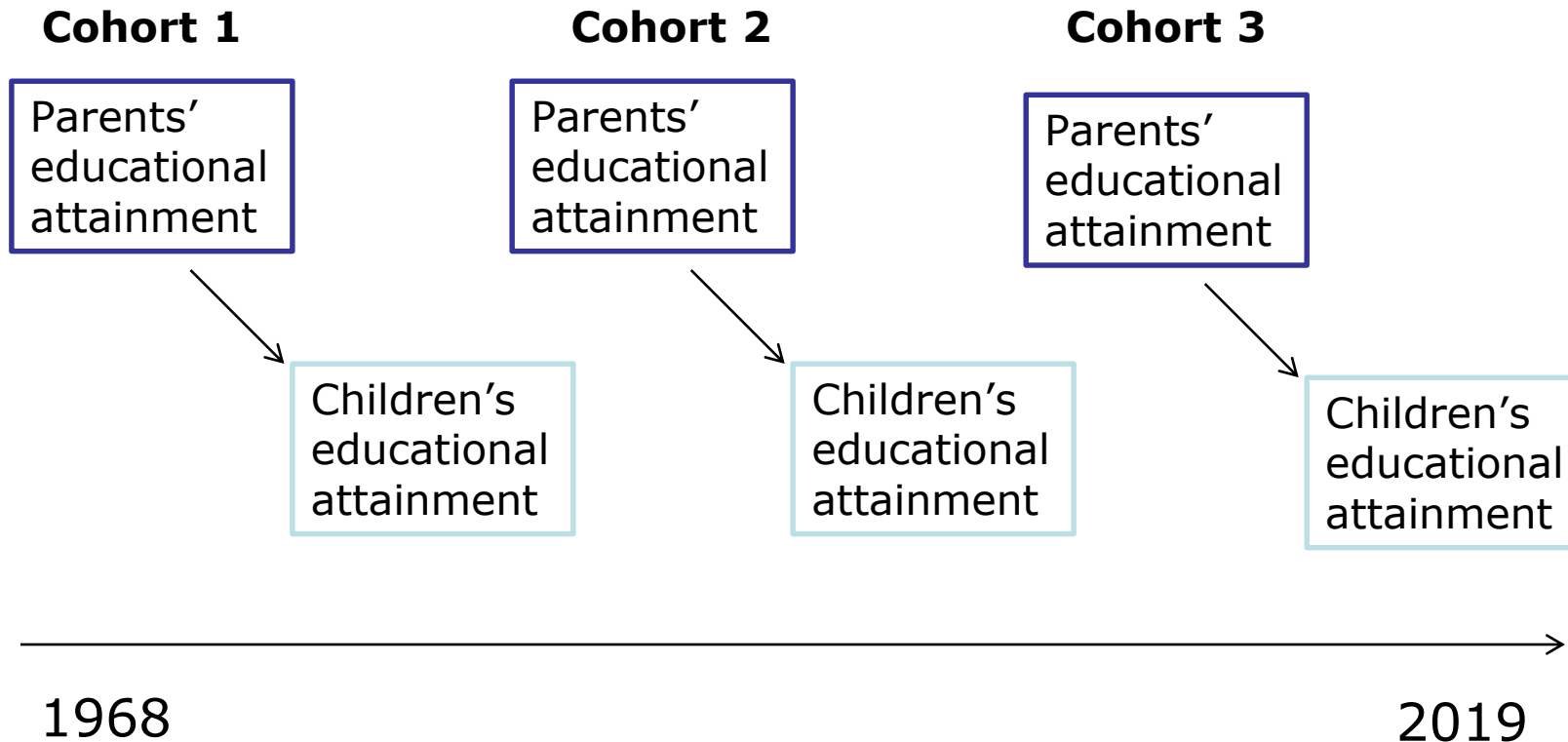


# Individuals' Life Course Trajectories





# Cohort Comparisons

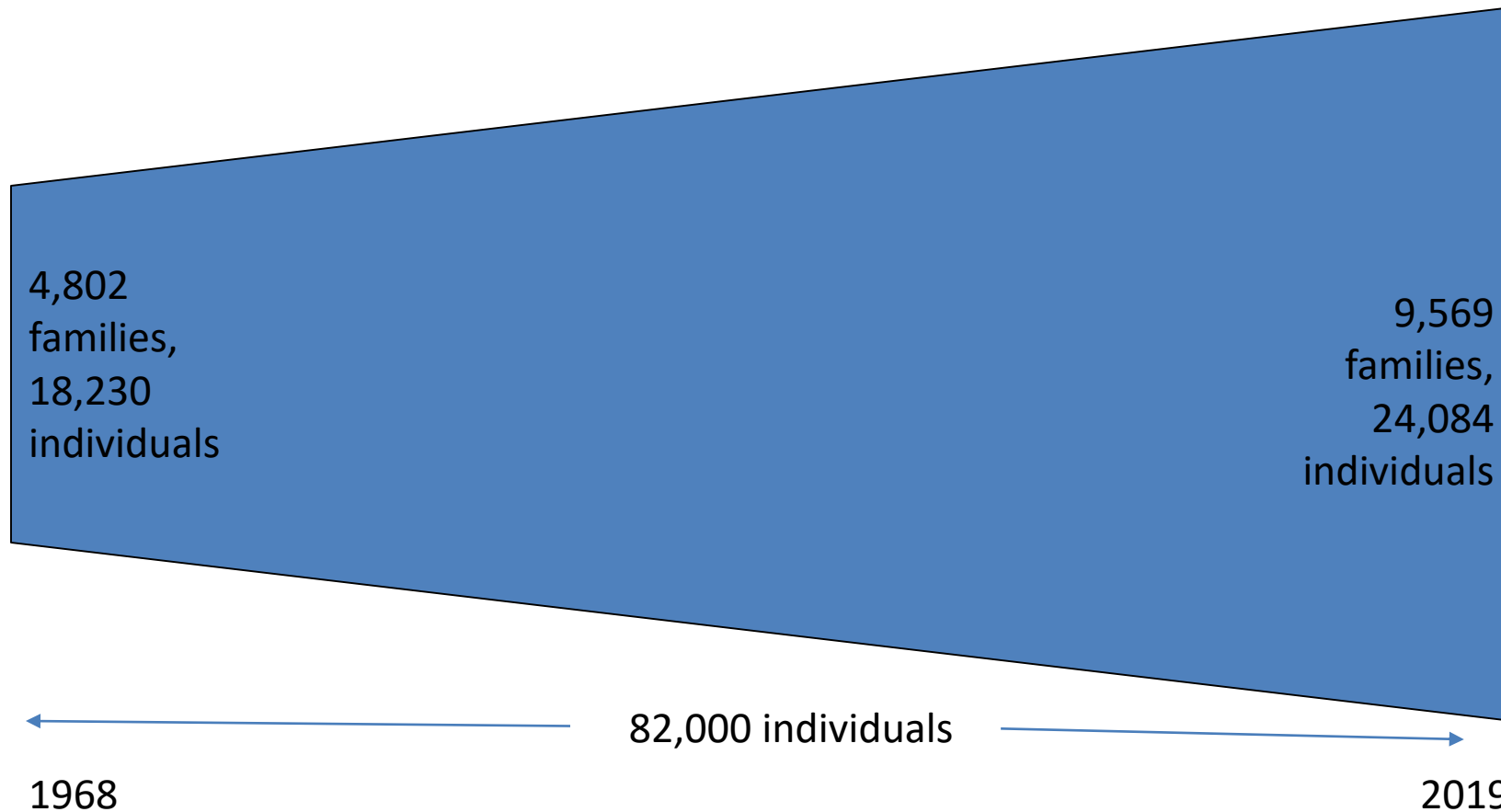


# PSID's Genealogical Design in Practice

- Who gets the “PSID gene”?
- How are gened (“Sample”) individuals followed over time?
- Which genealogical relationships can be observed?



# Who is in PSID?

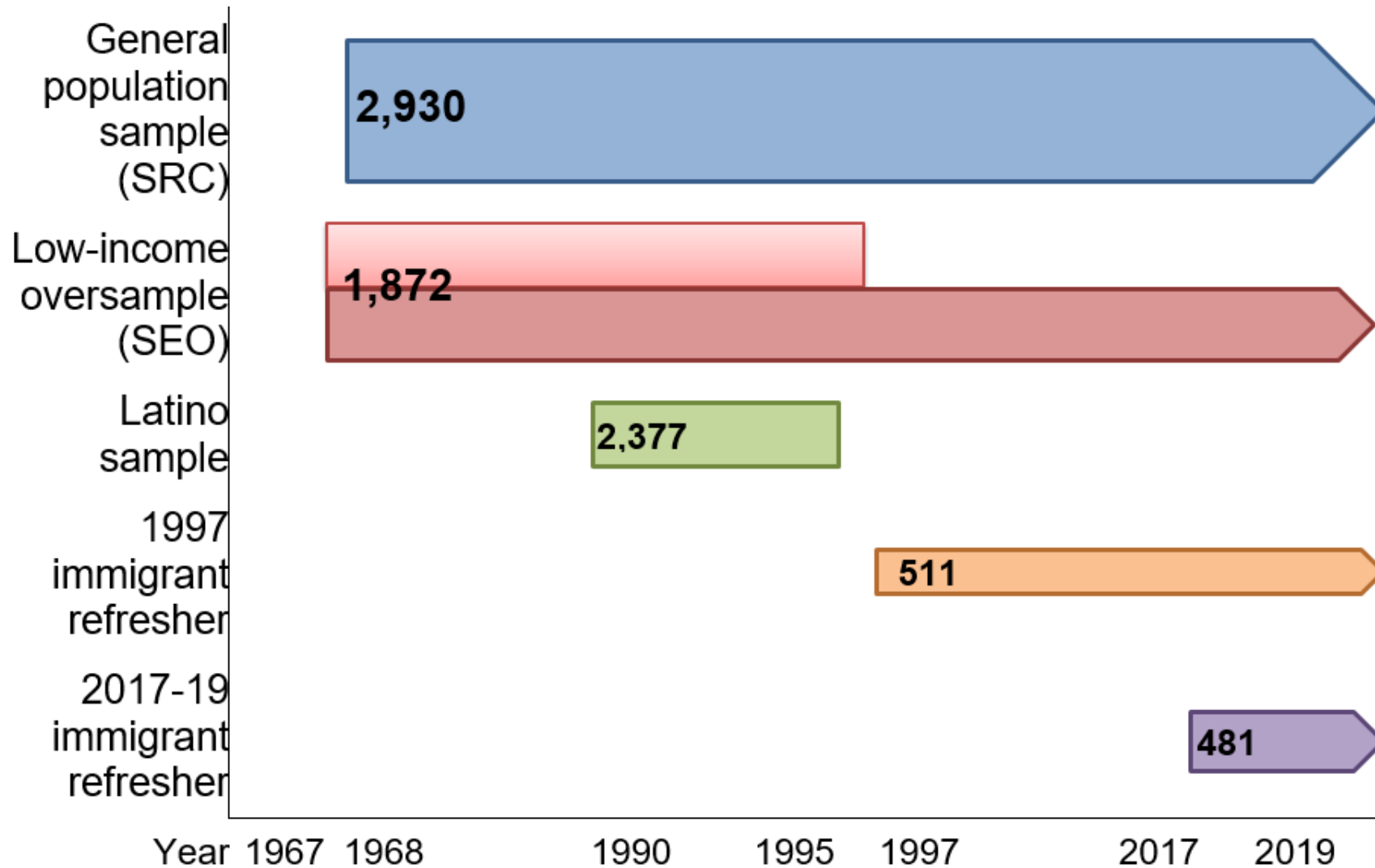


# Who is in PSID?

- Some terminology
  - Sample types (strata)
  - Clans
  - Family units
  - Family unit members
  - Sample members



# PSID Strata



# Who is in PSID?

- Clans
  - All individuals ever present in or descended from one of the originally sampled household in any stratum
    - $2830 + 1872 + 2377 + 511 + 481 = 8071$  separate clans
  - All individuals who share a common “1968 ID”
    - Variable name = ER30001

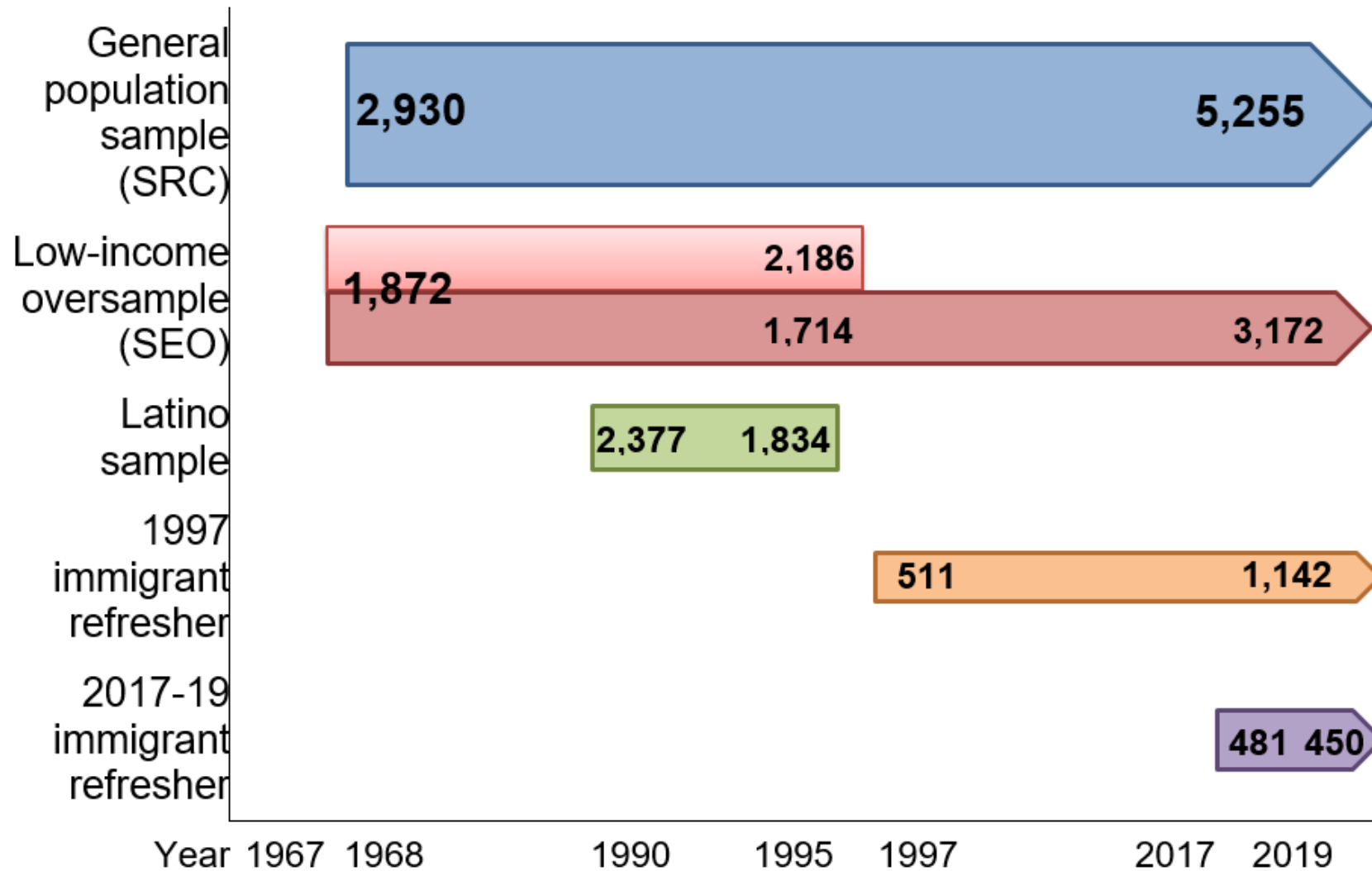


# Who is in PSID?

- Family units
  - Economically independent households whose members share household finances
    - Clan members who have moved into their own dwellings
  - Rostered members are usually related by birth/adoption, marriage or cohabitation
  - Biennial PSID interview pertains to members of family unit (not to clan members in other family units)
  - Composition of family unit can and often does change
    - “Clanship” is fixed
    - No time-invariant indicator of family unit membership



# PSID family units





# Who is in PSID in a given year?

- Four determinants
  - Demographic change (birth, death, union formation)
  - Attrition
  - Study design changes
  - “Sample” status and related following rules



# Who is a followable PSID sample member?

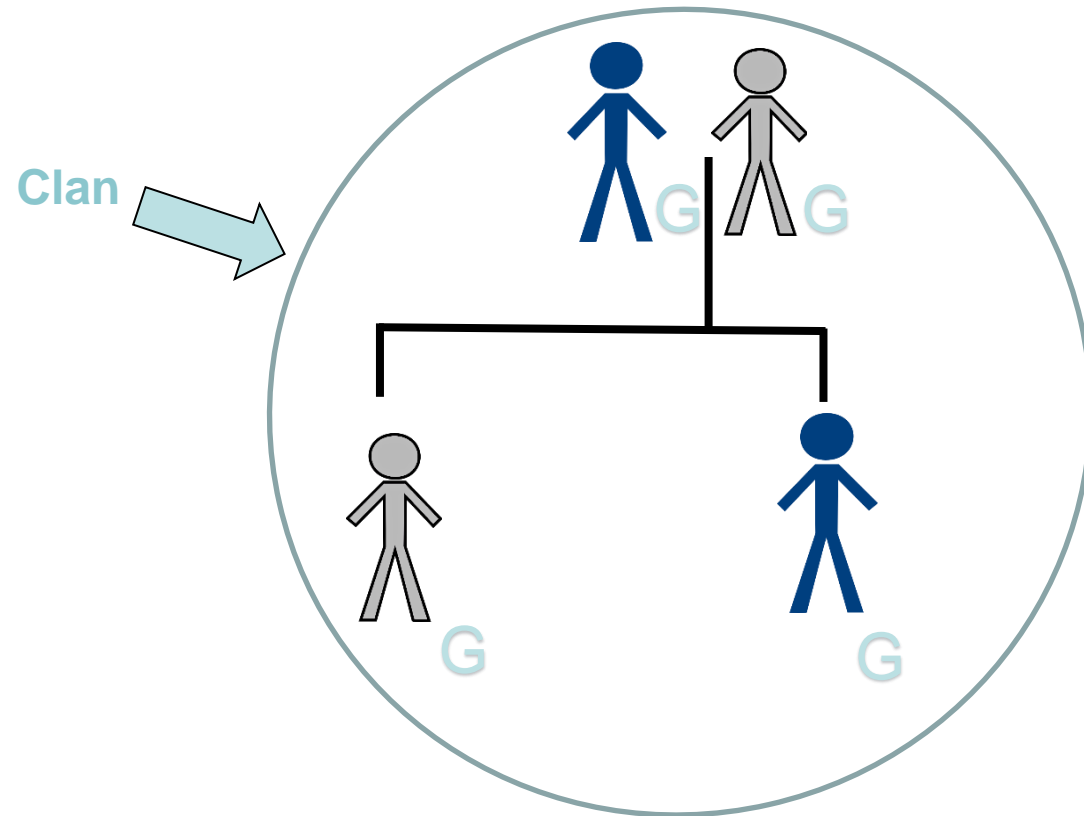
- Everyone who was living in an SRC or SEO family household in 1968 or in an immigrant refresher family in 1997/99 or 2017/19
- Anyone subsequently born to or adopted by an original family household member or their lineal descendant(s)
  - Includes children born outside of the household who move in later
  - Excludes children who are born and move out between waves
- Missed no more than two sequential waves (may re-enter by moving into an active family unit)
- Others deemed eligible by study directors (rare)



# PSID Design

1<sup>st</sup> Generation

2<sup>nd</sup> Generation



G = individual has  
the PSID “gene”  
(followable)

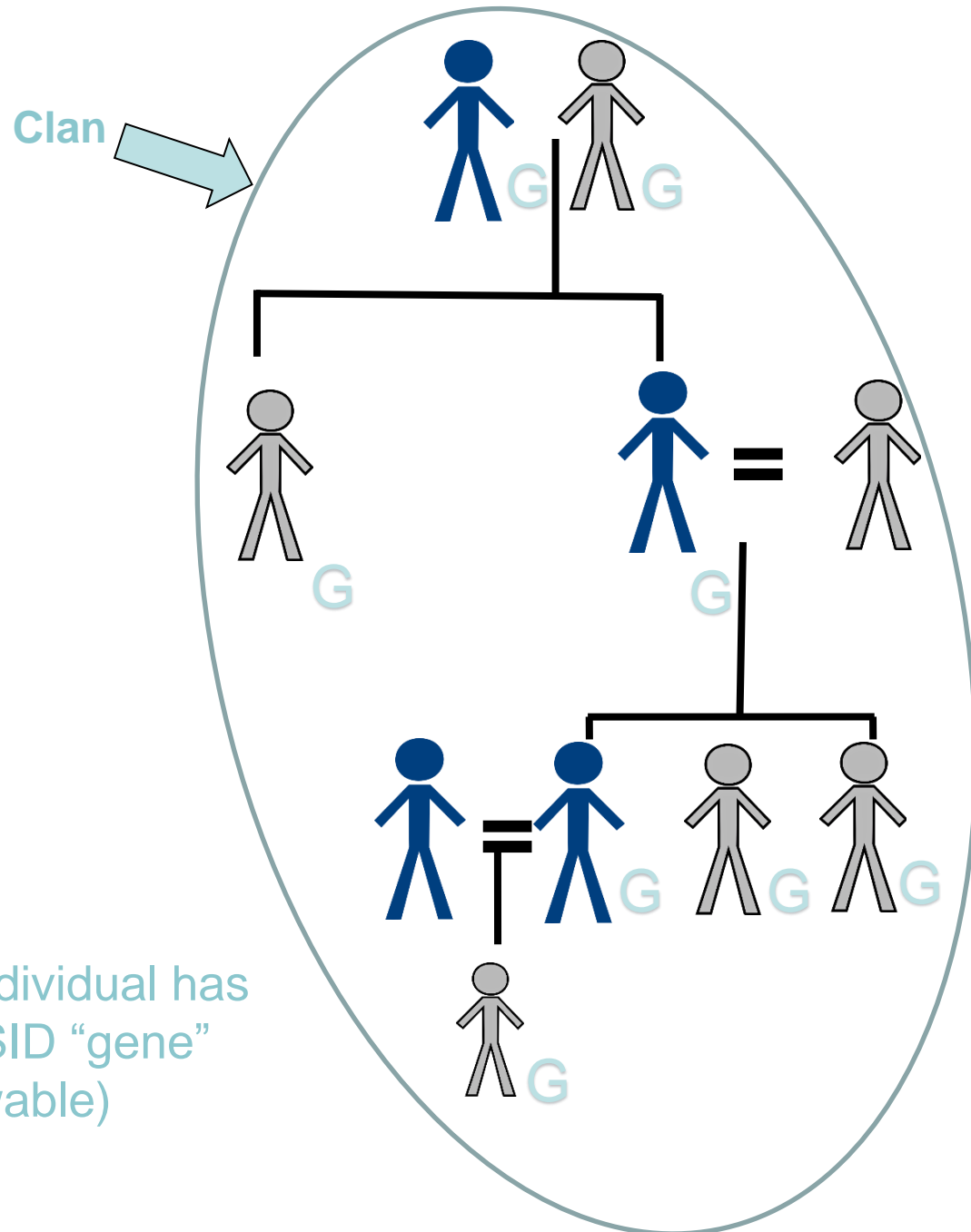
# PSID Design

1<sup>st</sup> Generation

2<sup>nd</sup> Generation

3<sup>rd</sup> Generation

4<sup>th</sup> Generation



G = individual has the PSID "gene" (followable)

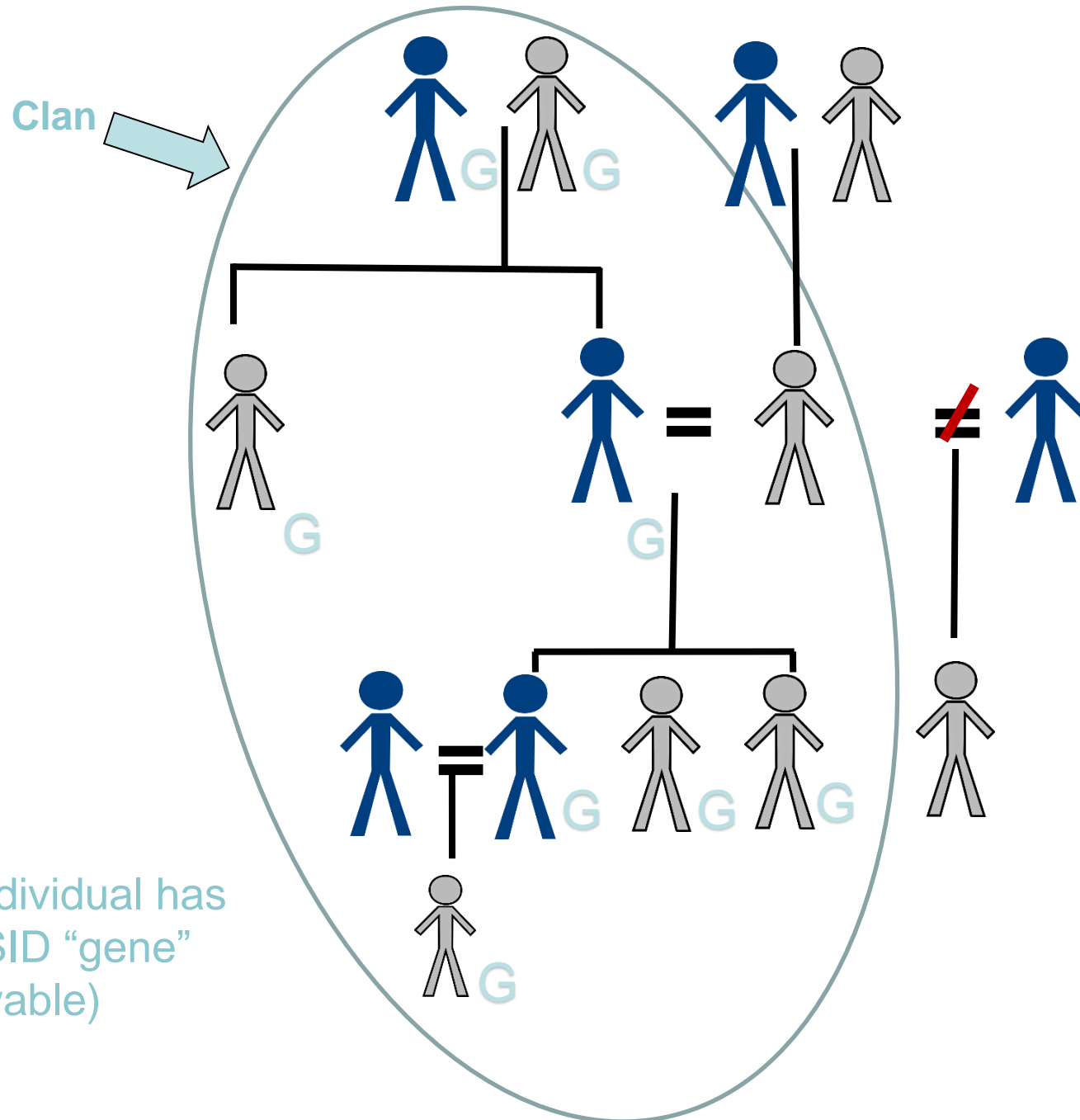
# PSID Design

1<sup>st</sup> Generation

2<sup>nd</sup> Generation

3<sup>rd</sup> Generation

4<sup>th</sup> Generation



G = individual has the PSID "gene" (followable)

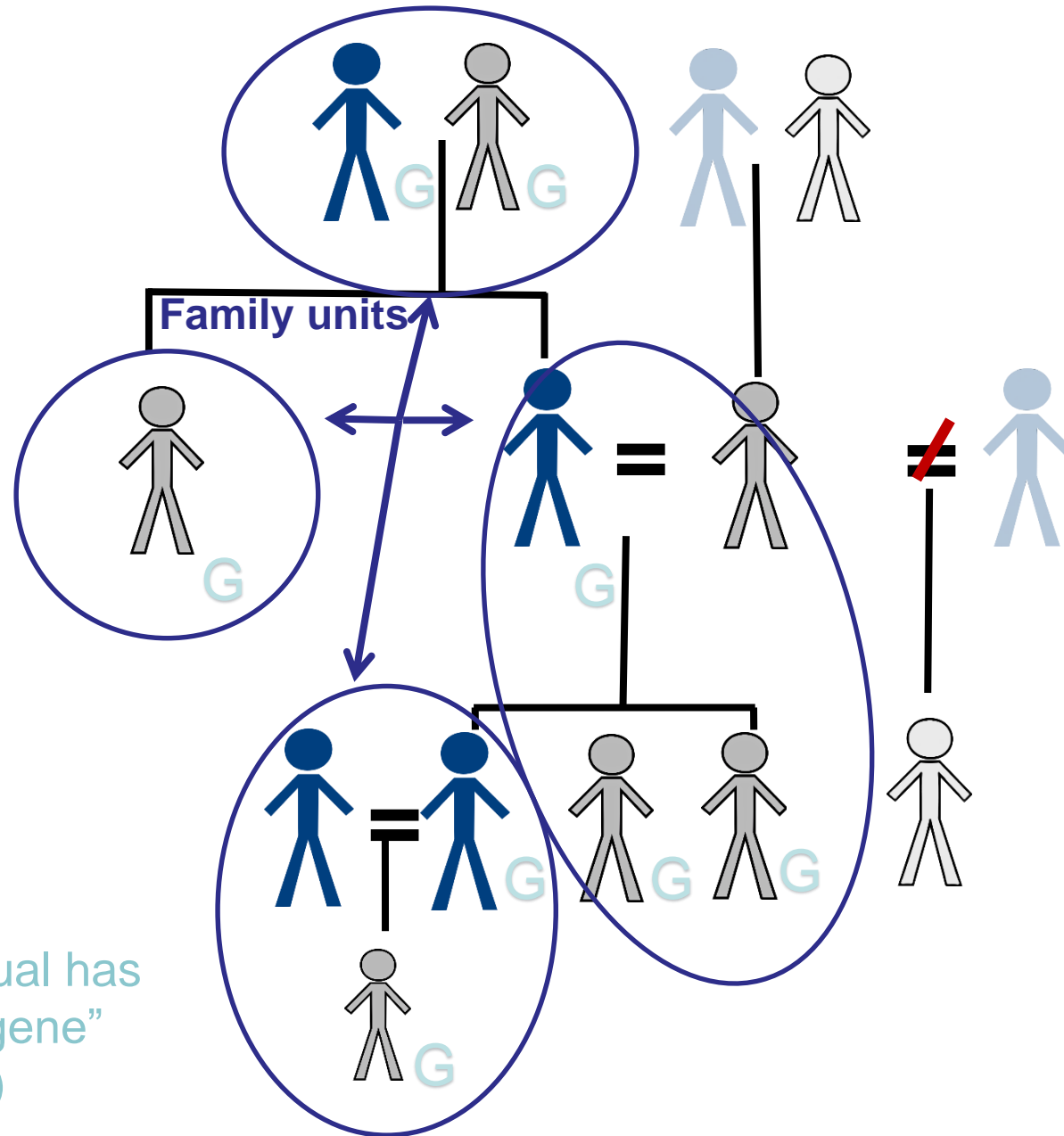
# PSID Design

1<sup>st</sup> Generation

2<sup>nd</sup> Generation

3<sup>rd</sup> Generation

4<sup>th</sup> Generation



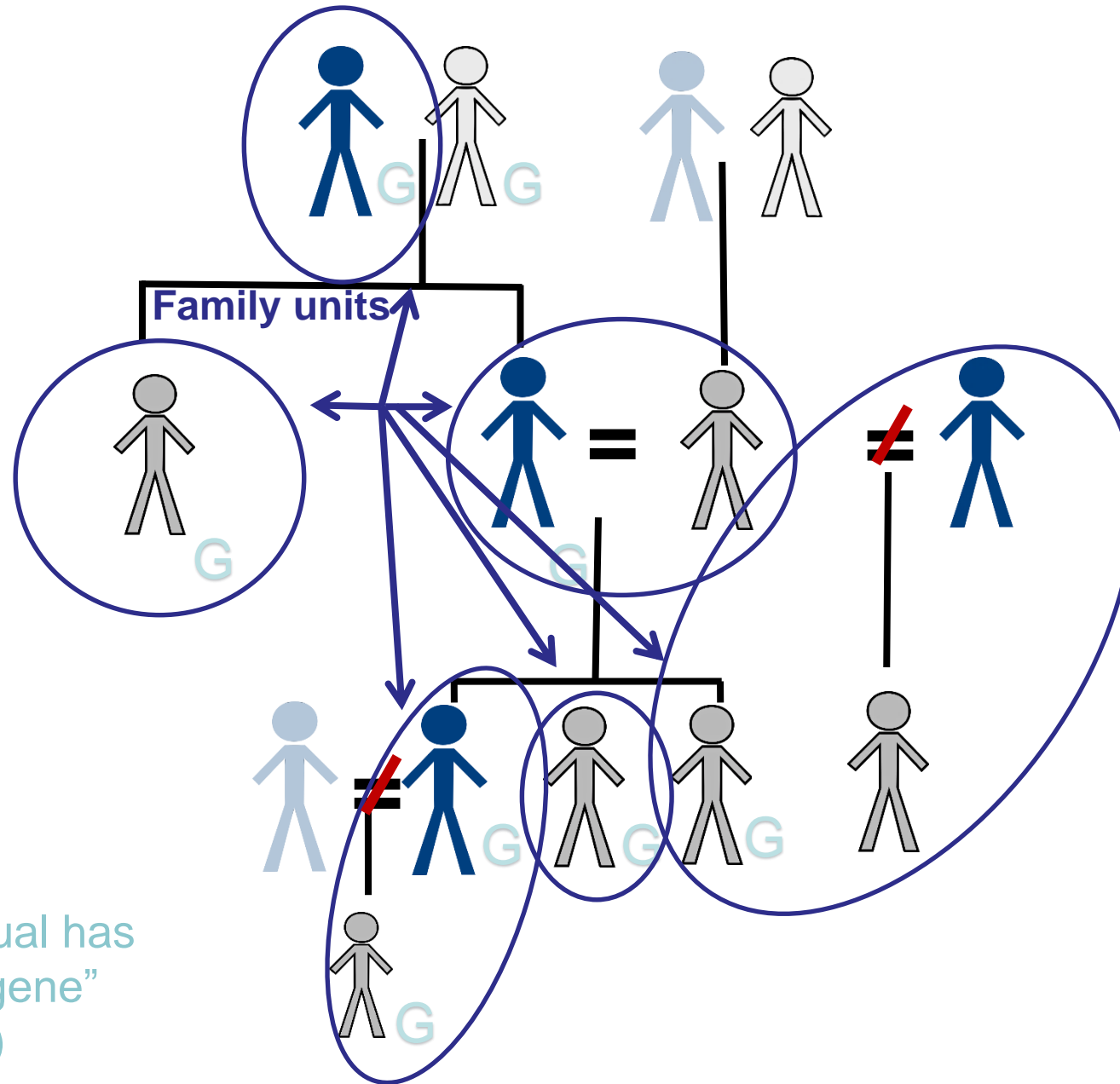
# PSID Design

1<sup>st</sup> Generation

2<sup>nd</sup> Generation

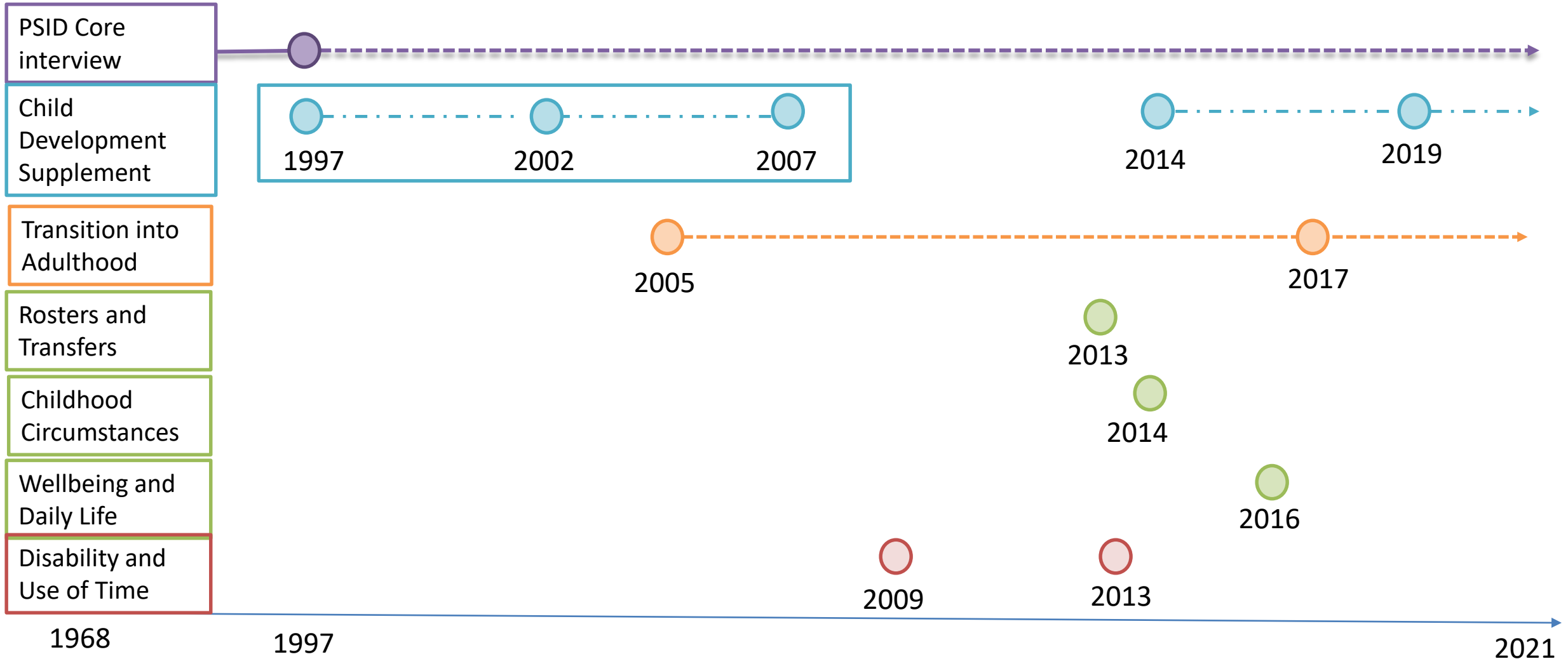
3<sup>rd</sup> Generation

4<sup>th</sup> Generation



G = individual has the PSID "gene" (followable)

# Supplemental Studies





# Tools to identify PSID participants

- In the upcoming lab session, we'll discuss
  - Time-invariant clan and person identifiers to follow people over time
  - Family unit and roster position identifiers to situate people in the family households they occupy at a given wave
  - Distinguishing followable sample members
  - Tools to connect people to parents, grandparents, and siblings across households and time

