"The Platform" (ELLP): Opportunities and Challenges – A Researcher's Perspective

Ross Finnie (rfinnie@uottawa.ca)
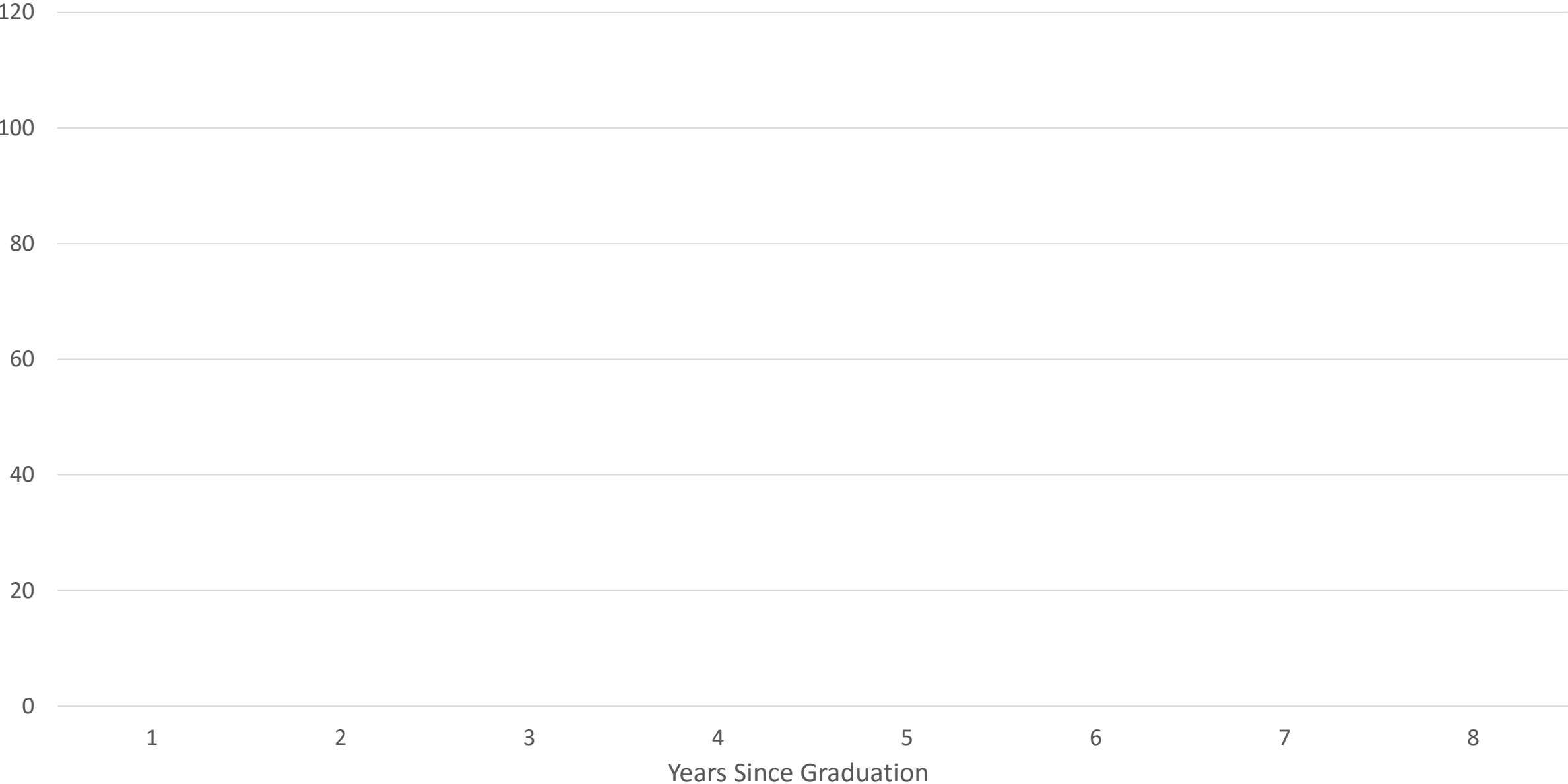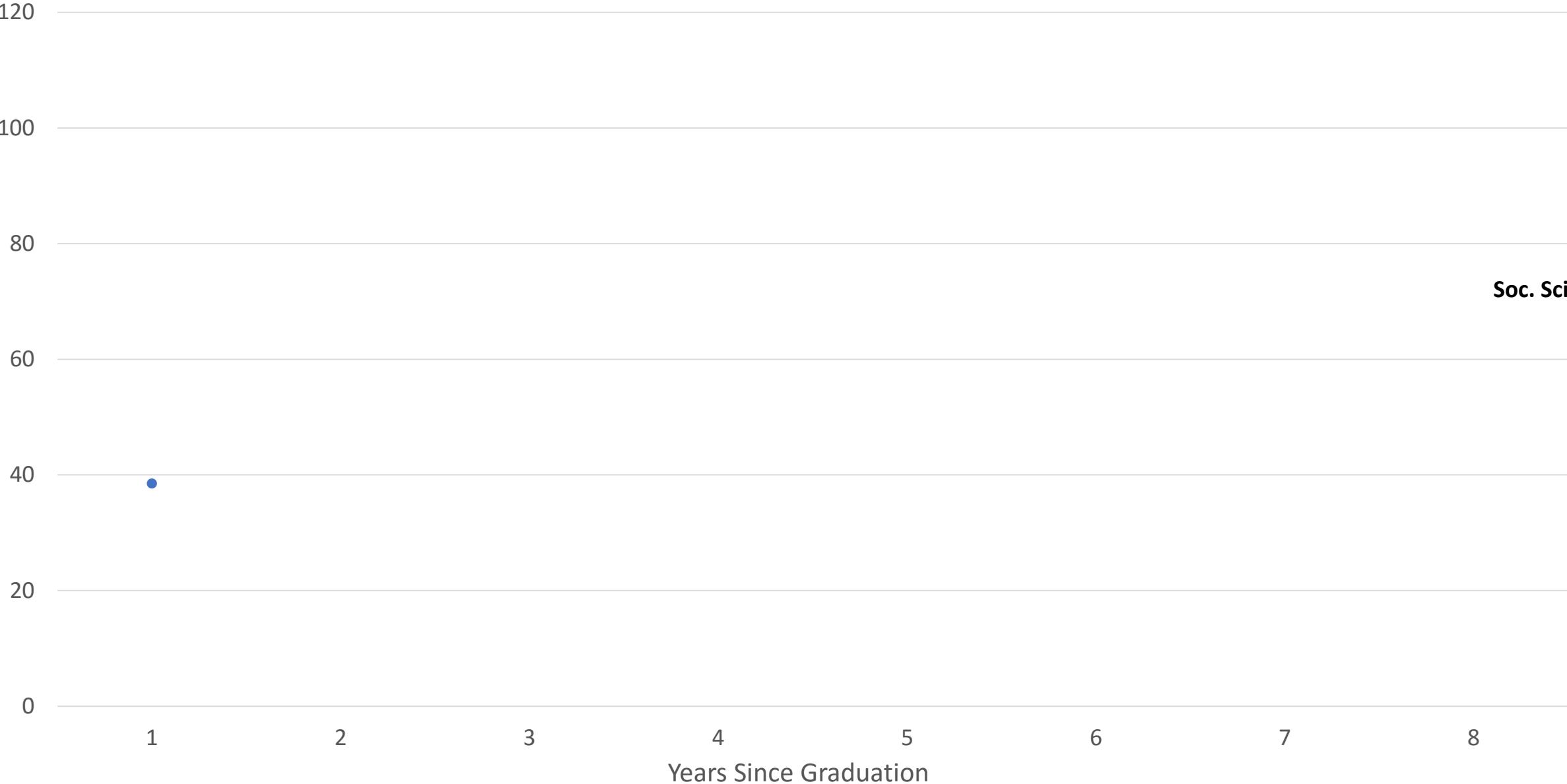
Context I: EPRI-ESDC Tax Linkage Project

"Barista or Better?"

Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)

Soc. Sci.

Years Since Graduation

Mean Earnings of Graduates by Field of Study ($2014)

Soc. Sci.

Years Since Graduation

Mean Earnings of Graduates by Field of Study ($2014)

Soc. Sci.

Years Since Graduation

Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)
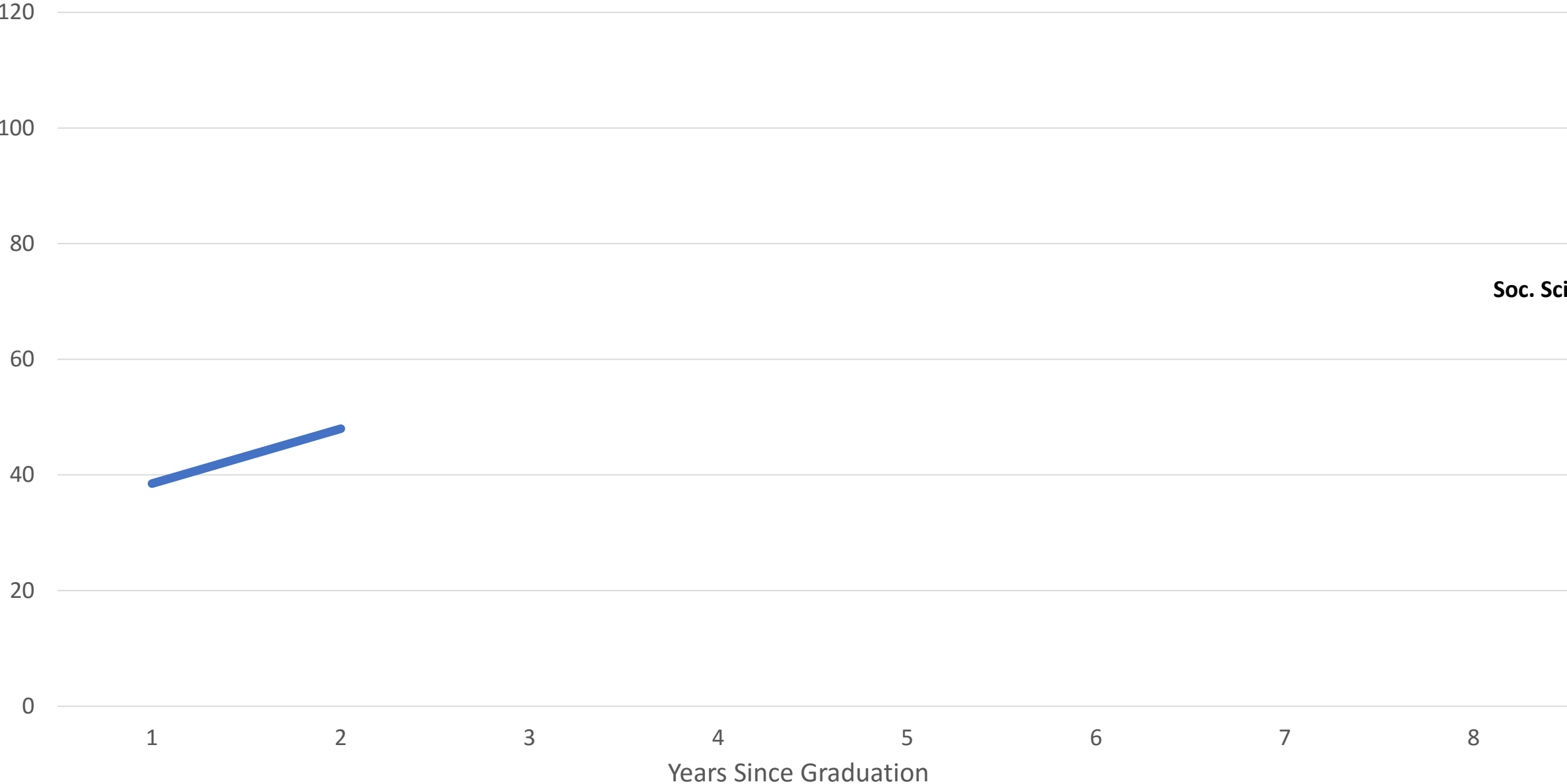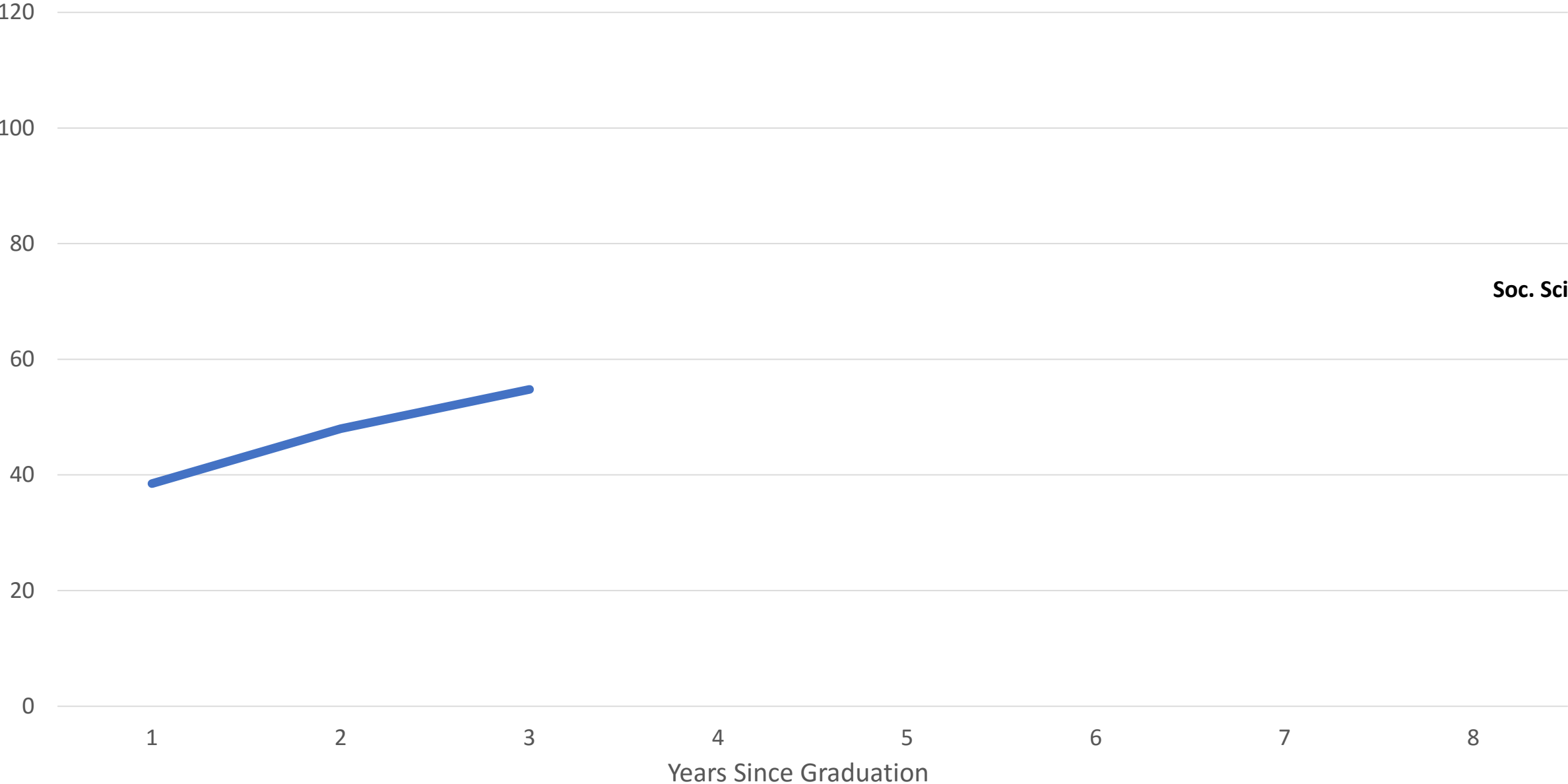
Mean Earnings of Graduates by Field of Study ($2014)
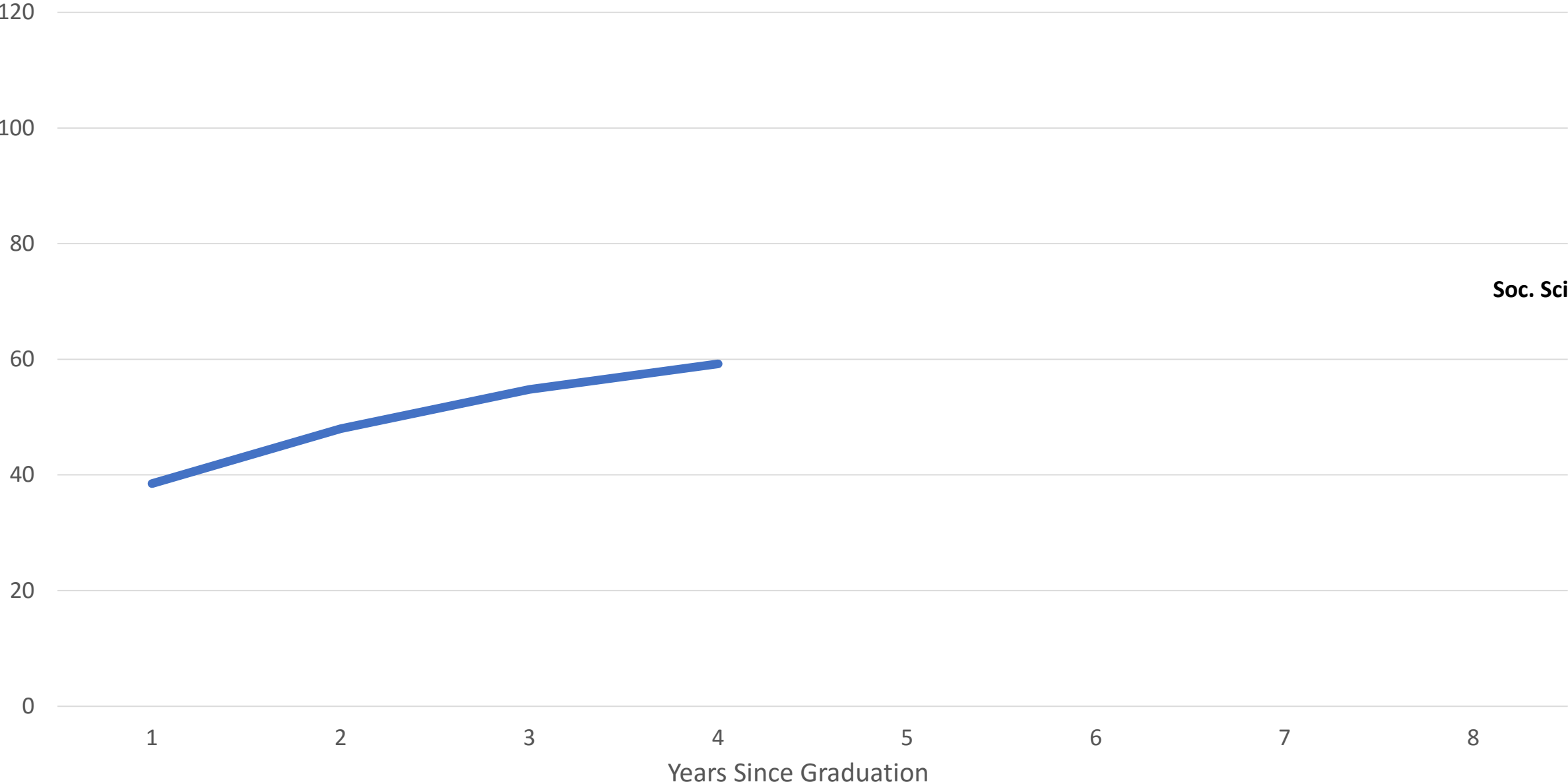
Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)
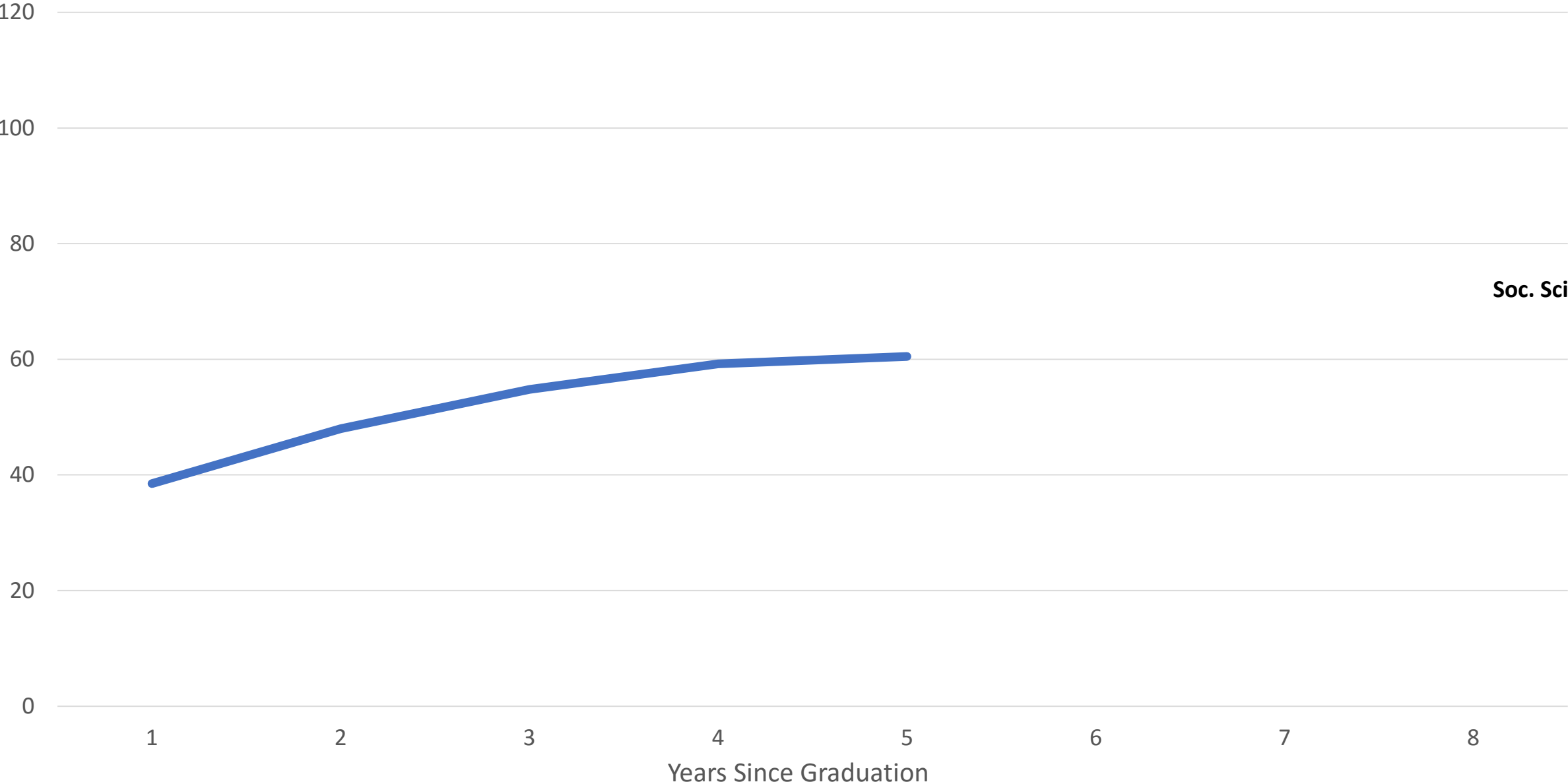
Mean Earnings of Graduates by Field of Study ($2014)

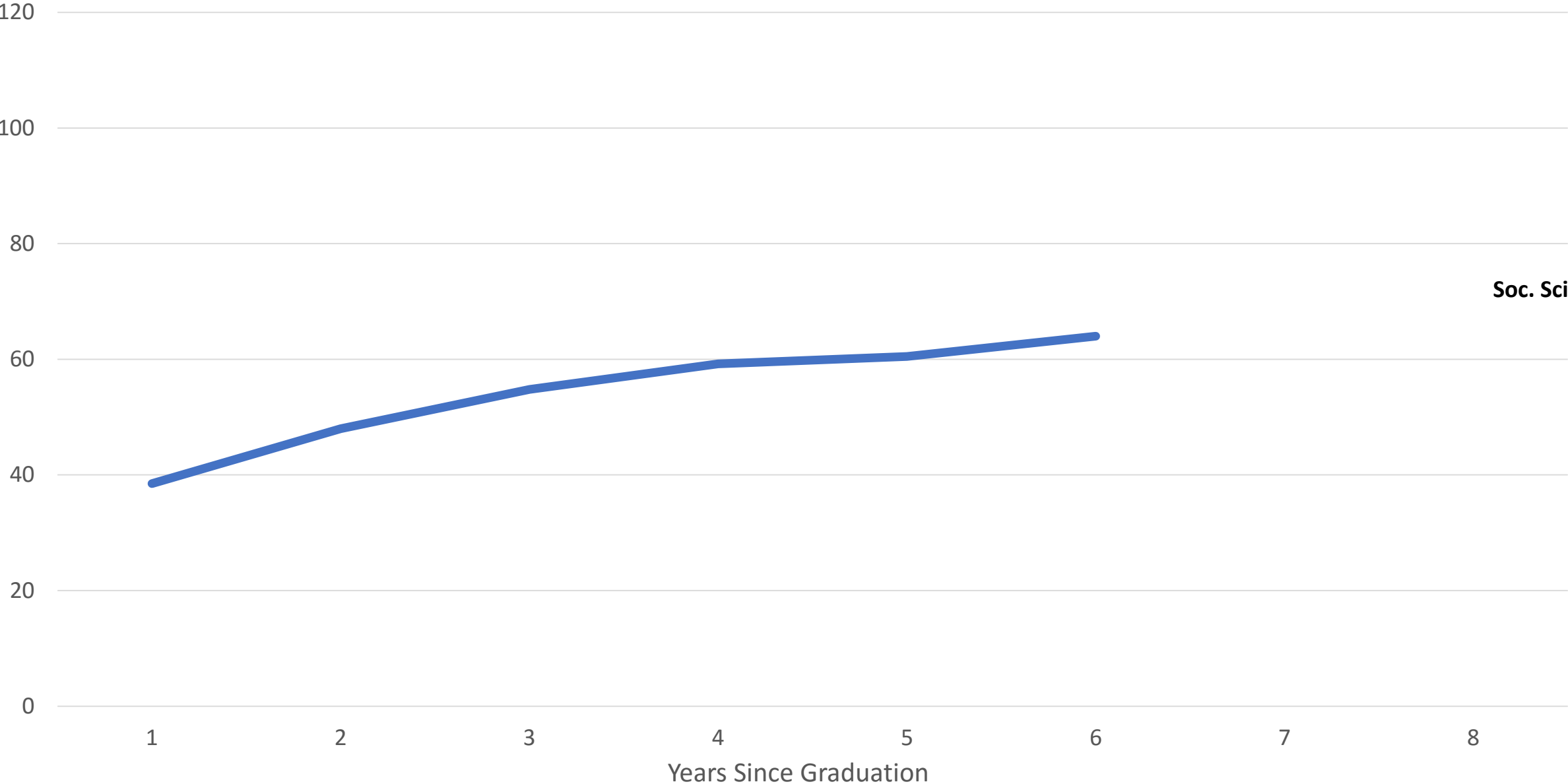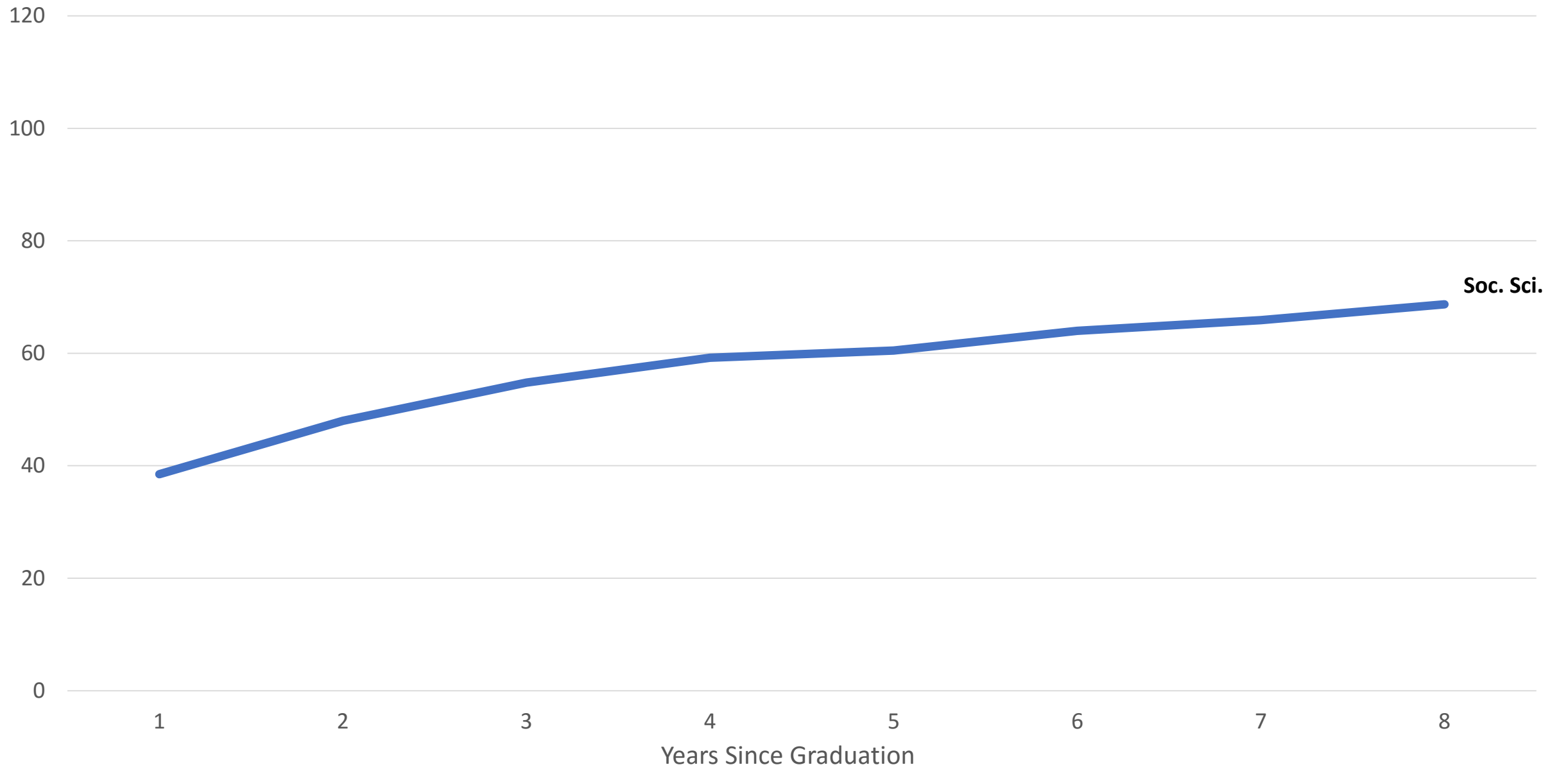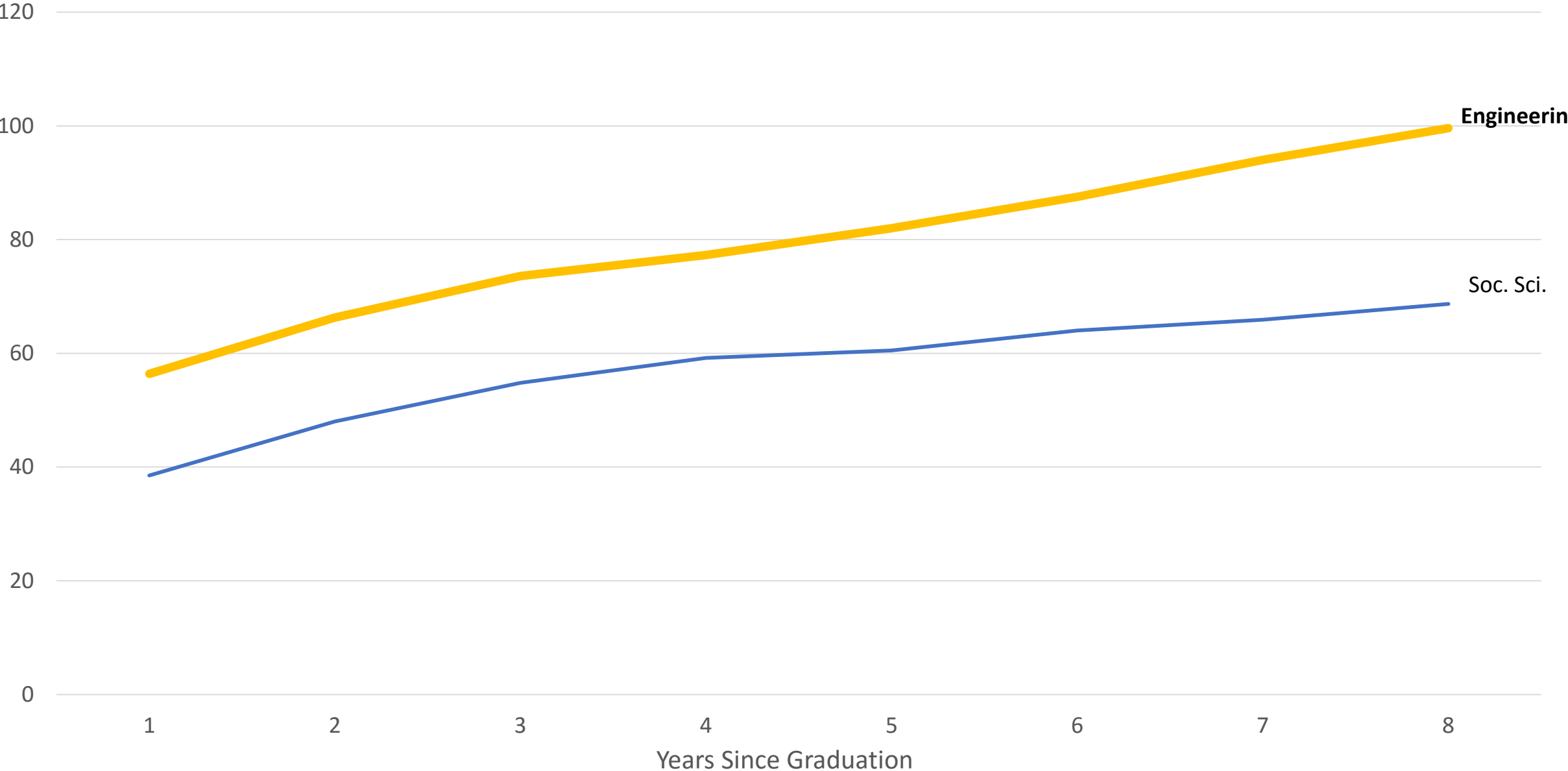# Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)

Mean Earnings of Graduates by Field of Study ($2014)
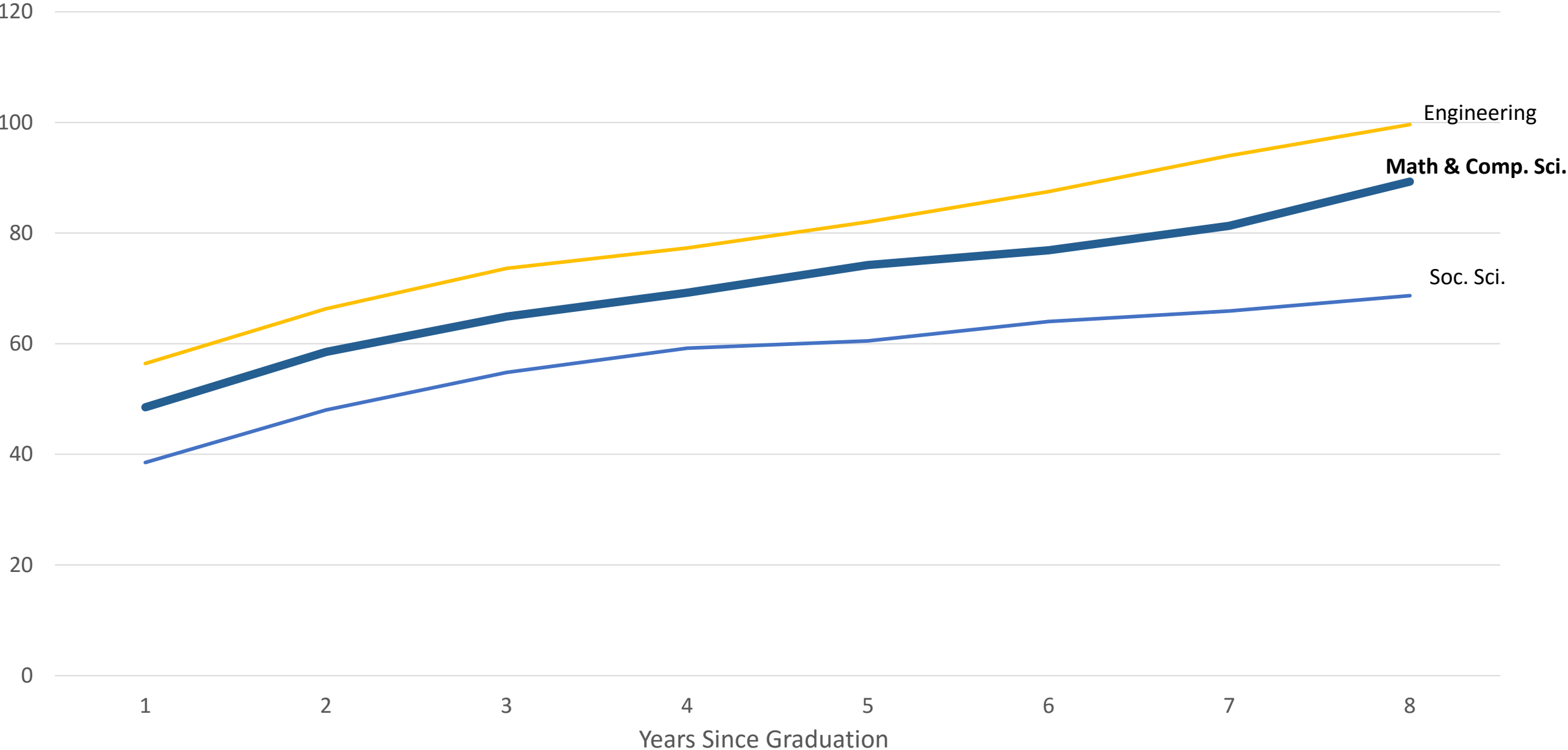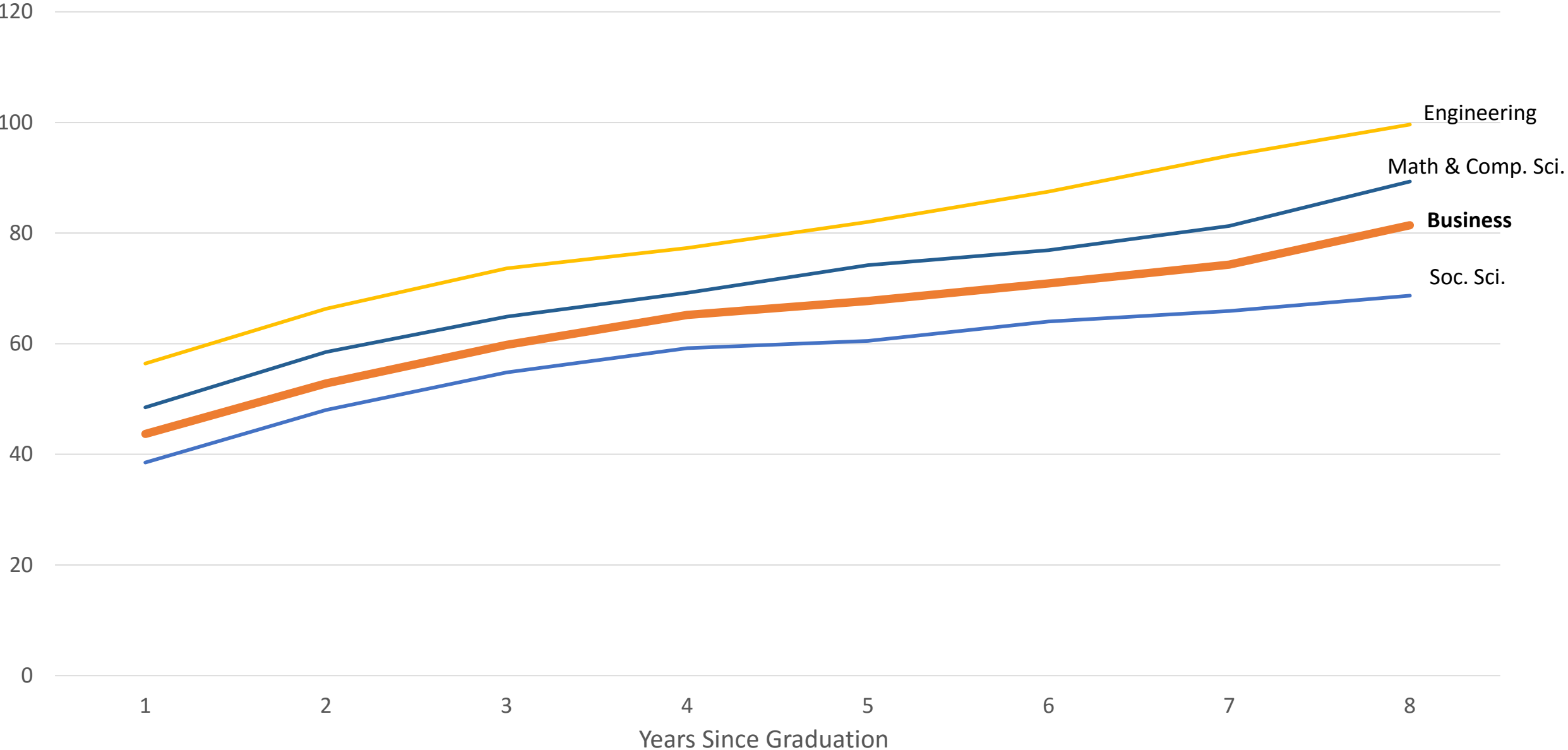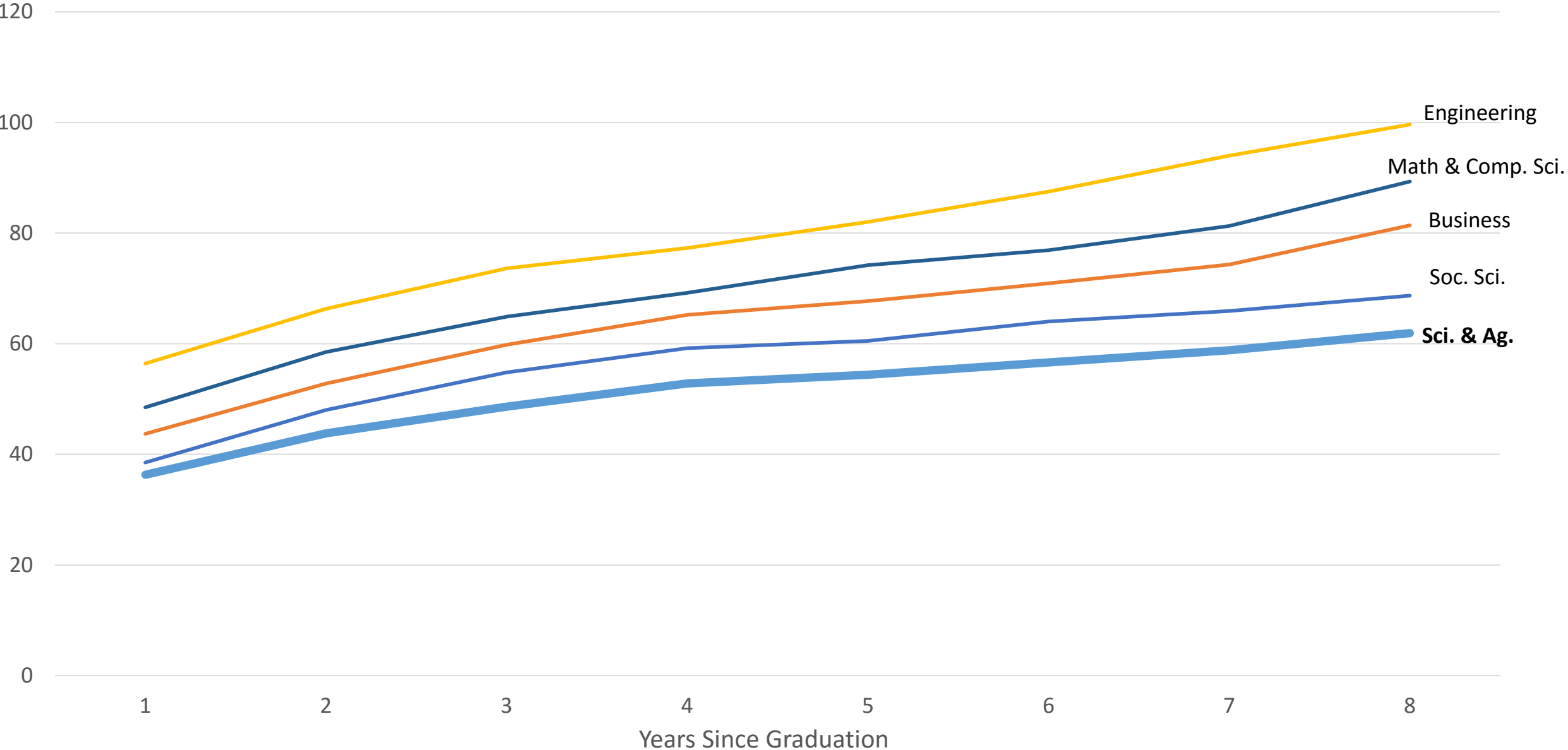
Mean Earnings of Graduates by Field of Study ($2014)

Engineering
Math & Comp. Sci.
Business
Soc. Sci.
Health
Sci. & Ag.
Humanities
Vis. & Perf. Arts

Years Since Graduation

Selected Cohorts – All Areas of Study

QICSS_june 13 (Montreal)
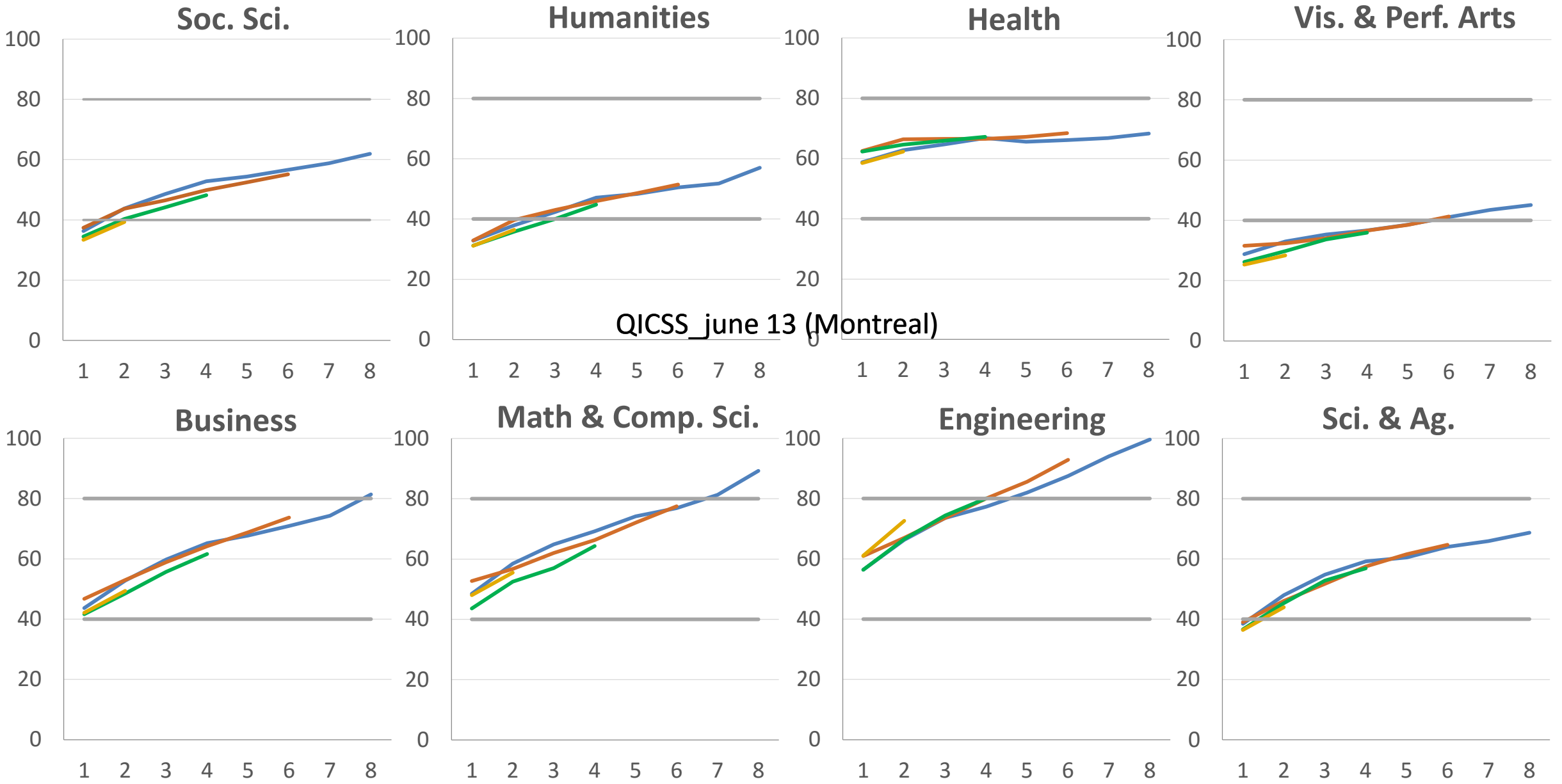
How Barista May Have Helped the ELLP

- Showed the potential/value of linked PSE-tax data: "Why wasn't this done before?"
- Trial in terms of how it would be received: privacy, etc.
- Demonstrated technical feasibility
- Developed one general starting point for ELLP research
- Got people interested in doing more

Context II: A Coordinated Research Approach in the Context of Complex Data and Overlapping Research Questions -

Context II: A Coordinated Research Approach in the Context of Complex Data and Overlapping Research Questions -

My Own Story (I will indulge!) –
But it Explains Some Things…

Wisconsin: The IRP - The Context

- IRP established at Wisconsin as a *Research Centre*

- Complex US longitudinal datasets emerging

- Overlapping research questions

- Government (and other) money being invested

- Wanted answers to real world research questions

Wisconsin: The IRP - What It Did

- Created sophisticated in-house research capacity & more

- Researchers then came to do their projects
  - Reduced barriers to entry
  - Research more efficient: building on expertise developed in past/on-going work

- It worked! - Led to more money coming in = more research projects & more answers to policy questions

The MESA Project

- Substantial contract with CMSF (Finnie-Sweetman-Mueller + Alex Usher)
- General focus on PSE access and persistence
- "The IRP/MESA/EPRI Approach":
    - Established research expertise (the YITS)
    - Rounds of projects – put out calls
    - Researchers plugged into our team
    - Efficient/coordinated research program
    - Plus dissemination – changed views?

Context III: The Genesis of the ELLP – A Personal Perspective (Encore!)

PSE-Tax Data and Linkages

- Long line of work with tax data (LAD): labour market outcomes
- Then work with PSE (PSIS) administrative data (CAMET)
- "Hey, why don't we link these?!"
- Got approval, but never moved
- Pushed the idea at every opportunity
- Sylvie Michaud: uOttawa pilot project
- Then ESDC-EPRI project: 14 institutions (as above)

Now - The Education Longitudinal Linkage Platform (ELLP or "ELMLP"?)

(Which is What You Came For)

The ELLP

- ESDC-StatsCan initiative

- "A relational data environment linking multiple datasets"
    * Postsecondary Information System (PSIS)
    * Registered Apprenticeship Information System (RAIS)
    * Tax information from the T1 Family File (T1FF)

- Truly a "platform": current form + future enhancements

1. First Track Students Through PSE Using PSE Administrative Data

K-12   **PSE**   **Post Graduation**
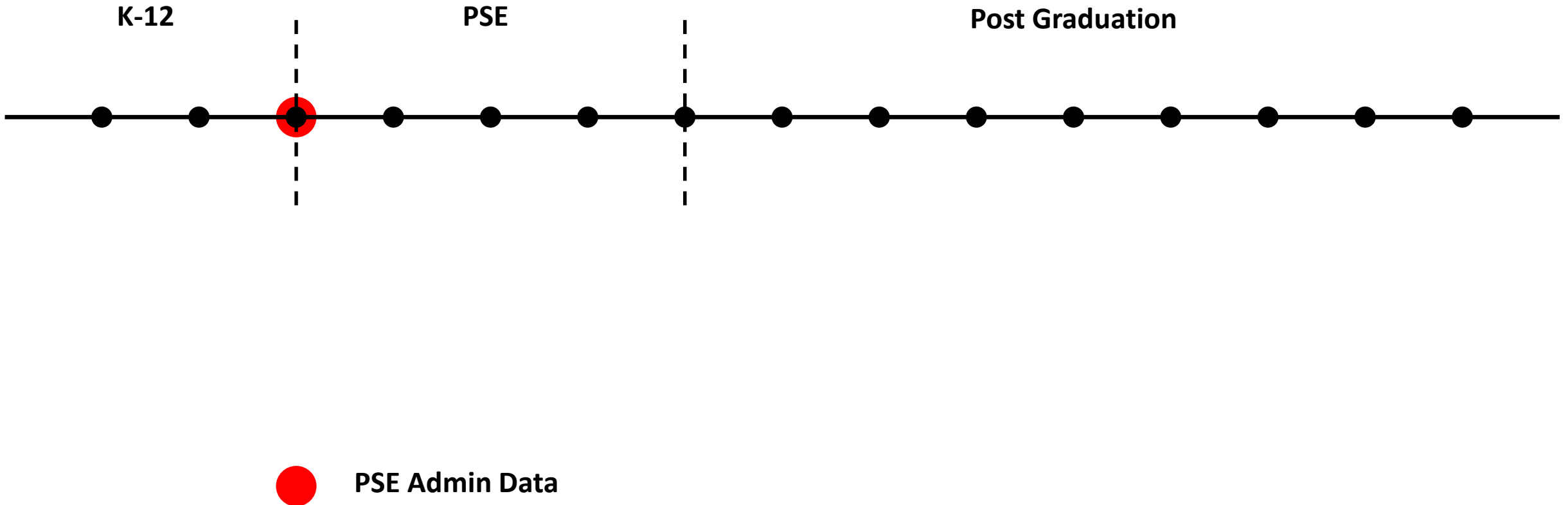
🔴 **PSE Admin Data**
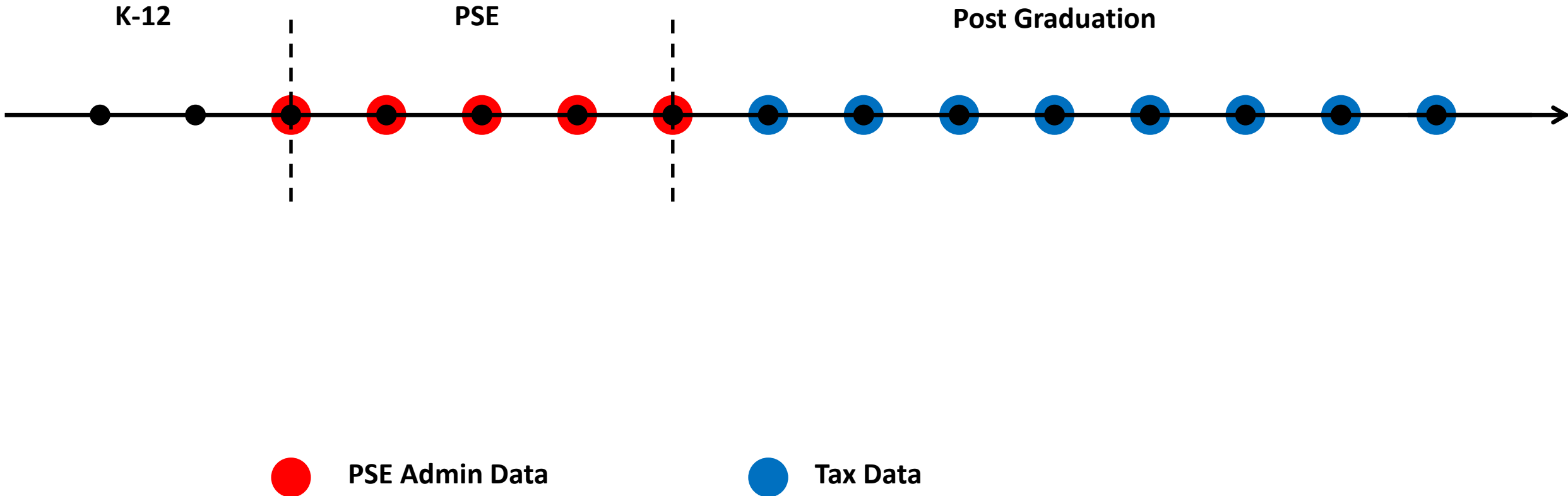
1. First Track Students Through PSE Using PSE Administrative Data

2. Then Follow Them Post-Graduation Using Tax Data

K-12   **PSE**   **Post Graduation**

PSE Admin Data   *Tax Data*
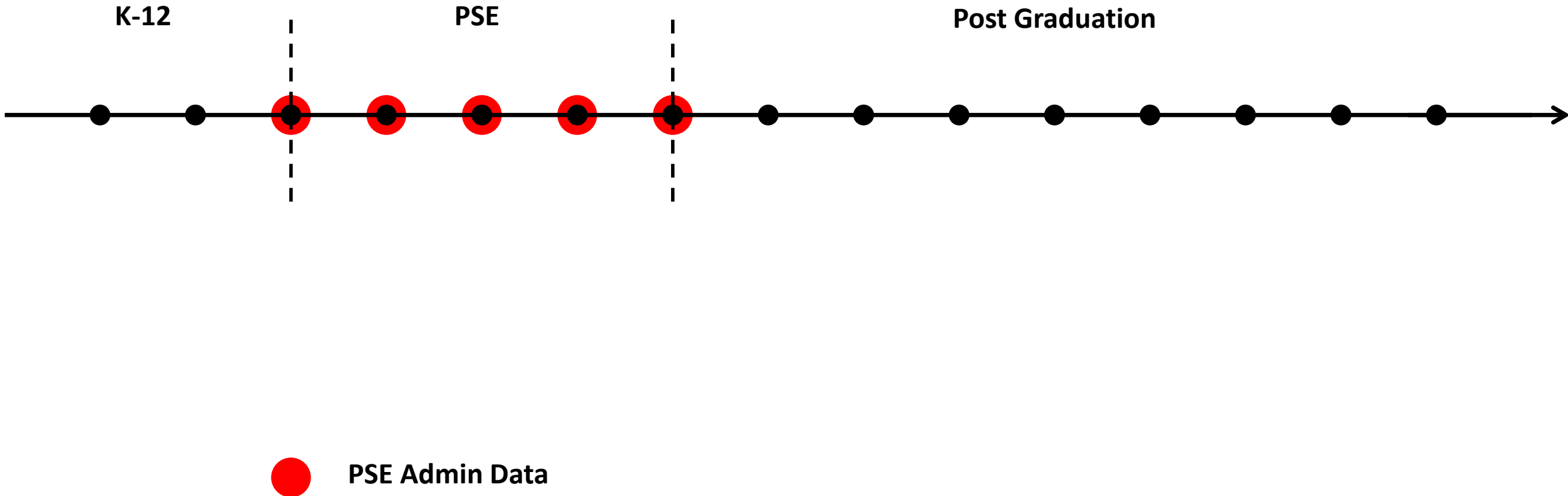
2. Then Follow Them Post-Graduation Using Tax Data
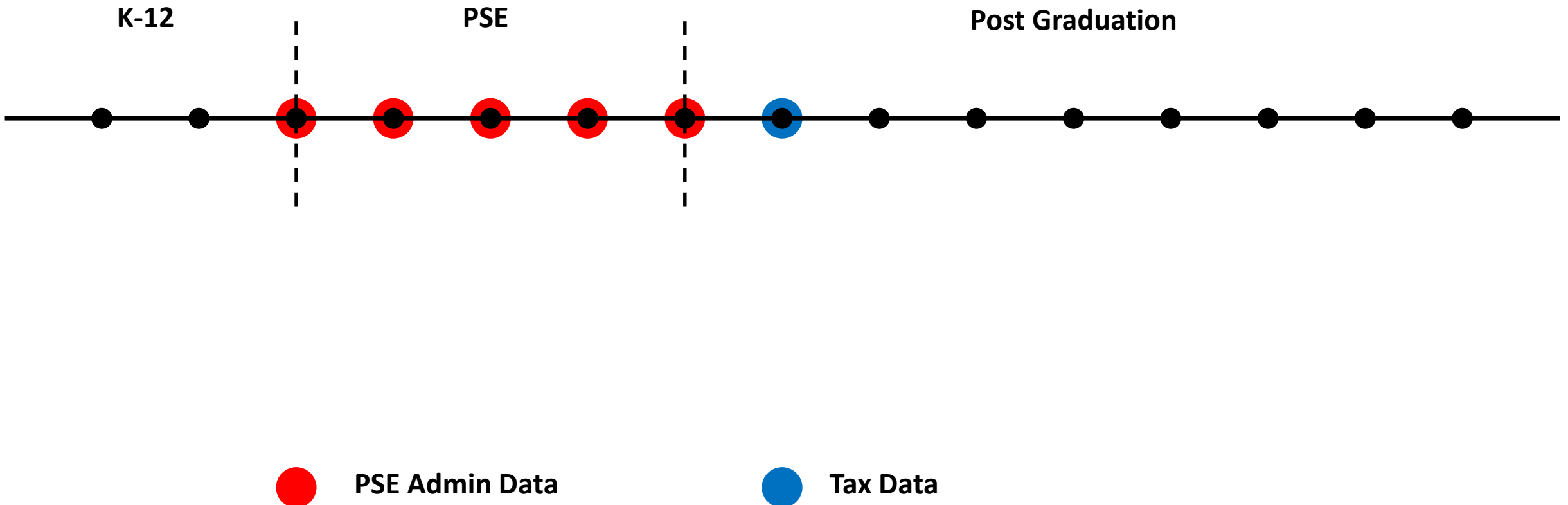
K-12 **PSE** **Post Graduation**

🔴 **PSE Admin Data**   🔵 **Tax Data**

3. Connect PSE Experiences to Post-Graduation Outcomes

K-12   PSE   Post Graduation

PSE Admin Data   Tax Data
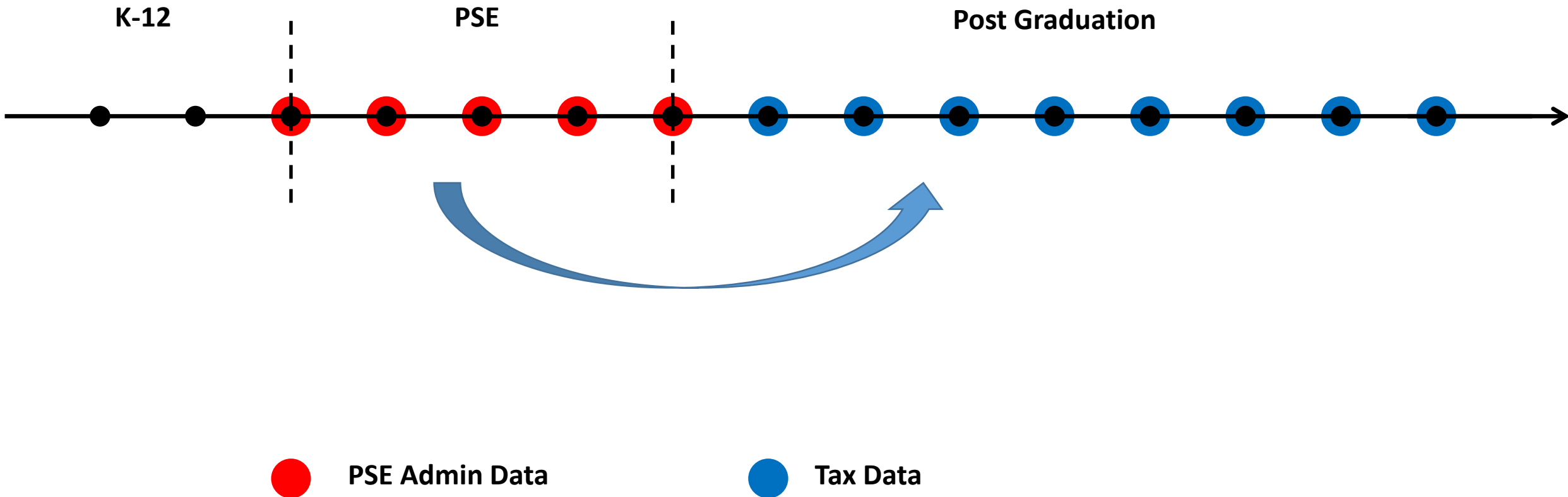
4. But Tax-Based Information Also for the PSE Years
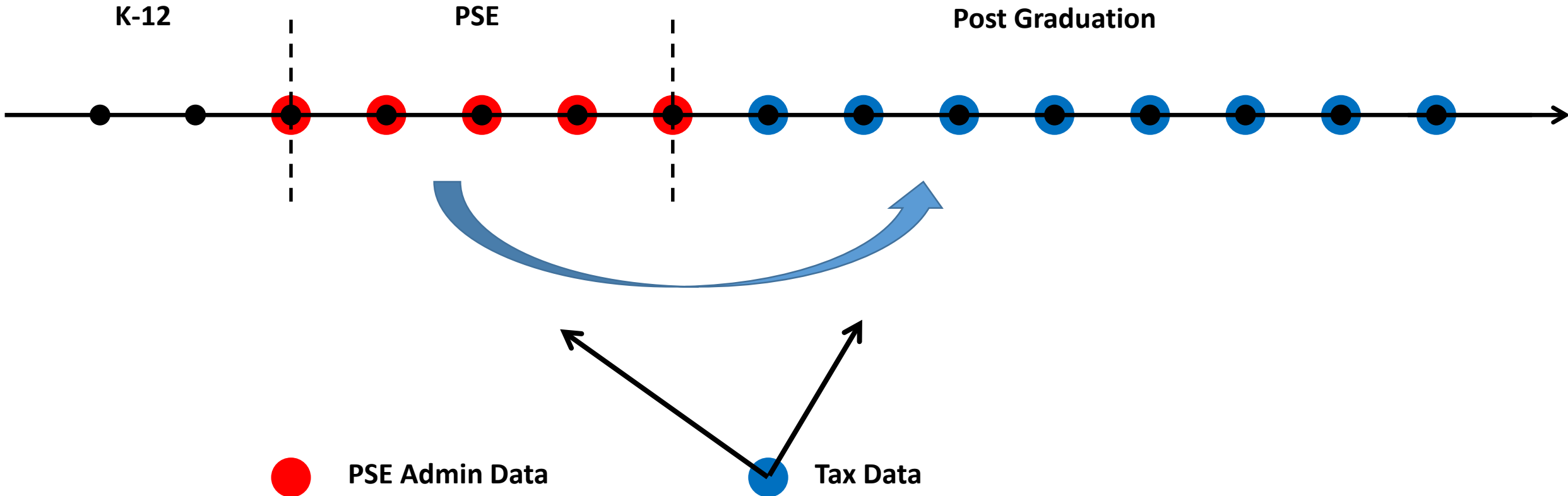
That's It - Like Never Before (National Level)
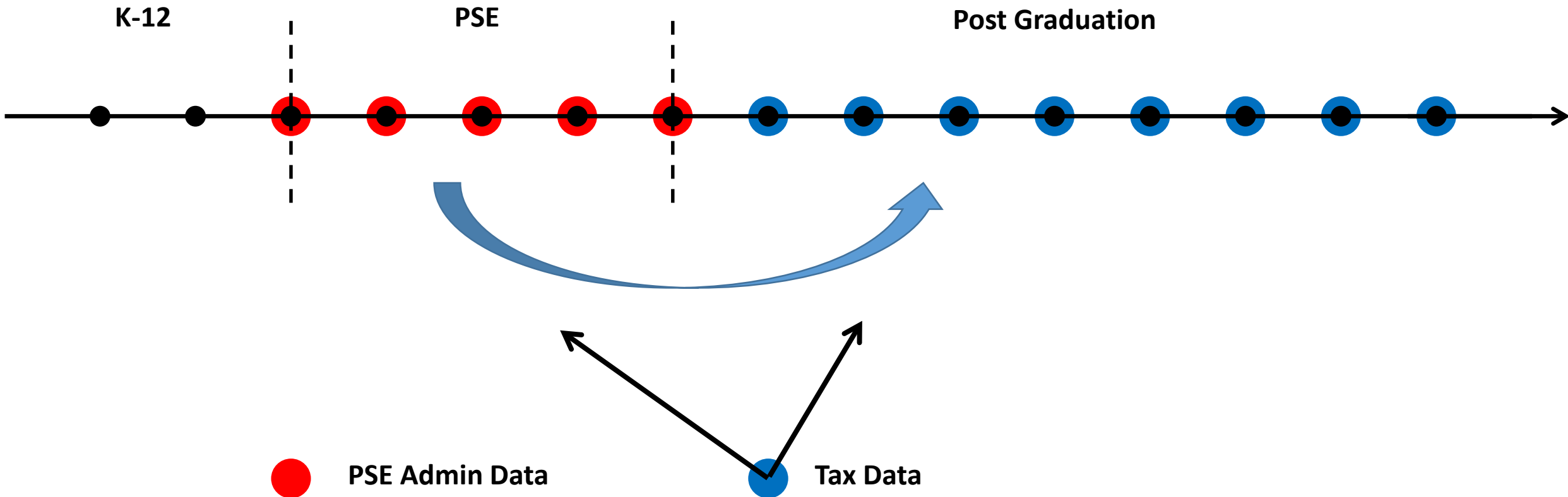
PSIS

- Centralized PSE administrative data

- Student & program information

- Unit of observation: program-person-year

- Annual, province-by-province basis since 2009

PSIS – Program Information Included

- Institution/type
- Credential
- Program CIP code & name
- (Original) start & end dates, graduation flag
- Major(s)
- Credits: cumulative to date, transferred, to graduate
- Co-op (a bit of an exception)

PSIS – Program Information *Not* Included

- Courses taken
- Minor(s)
- GPA: PSE, high school
- Distance education flag
- Class size
- Etc. – i.e. Other detailed information

PSIS – *Student* Information Included

- Birthdate
- Sex
- Mother tongue
- Immigration status/country
- Province & FSA (permanent & current)
- Indigenous/vis. min. (as reported)

- No other information included (disability, children, marital status…)

RAIS

- Similar to PSIS, also from 2009

- Adapted to trades & apprenticeships

- Fewer student characteristics

T1FF: Mostly, Think of Your Tax Form – and Others'

- Individual-level information (only)
- Earnings, other employment income, # T4 slips
- Self-employment income
- EI, SA & CTB
- PSE tax credits
- CPP/QPP & RPP contributions, union dues
- NAICS: main, second, #
- Province & tax prov.
- Sex & age

Extended Data Possibilities – in Principle ("Platform Concept")

- Add more data from underlying sources
  - PSIS & RAIS
  - T1FF
- Add data from institutions
  - currently available
  - constructed (e.g. co-op criteria)
  - developed (e.g. skills measures)
- New data linkages (see below)
- Socio-economic background ☺ (neighbourhood, family)

A Starter Research Agenda: A Few Suggestions/Illustrative Ideas

1. PSE Experiences & Outcomes

1. Start now
   - Persistence, mobility, completion, pathways
   - Specific schooling experiences (e.g., co-op)
   - All at system level – or the researcher's choice
2. When more data added
   - Other schooling experiences
   - Performance, Engagement & satisfaction
   - Skill level & development (skills!)
   - What else?

2. Post-Schooling Experiences & Outcomes

- Earnings
- Income from other sources
- Other outcomes (e.g. savings, family outcomes, loans)
- Place of residence & mobility
- Relate to
  - Student/individual characteristics
  - schooling experiences (credential, co-op, more…)
  - current situation (where living, local labour market…)
  - Socio-economic/other background

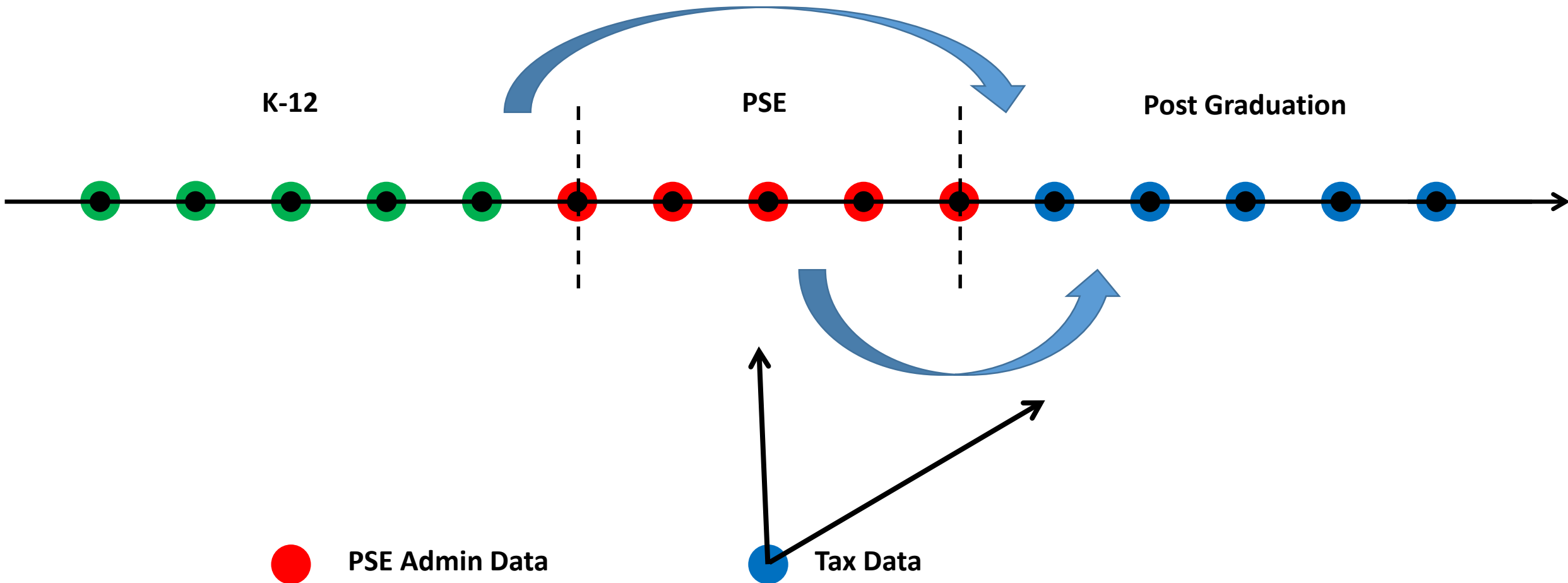# A Big Picture View
Including K-12 Data
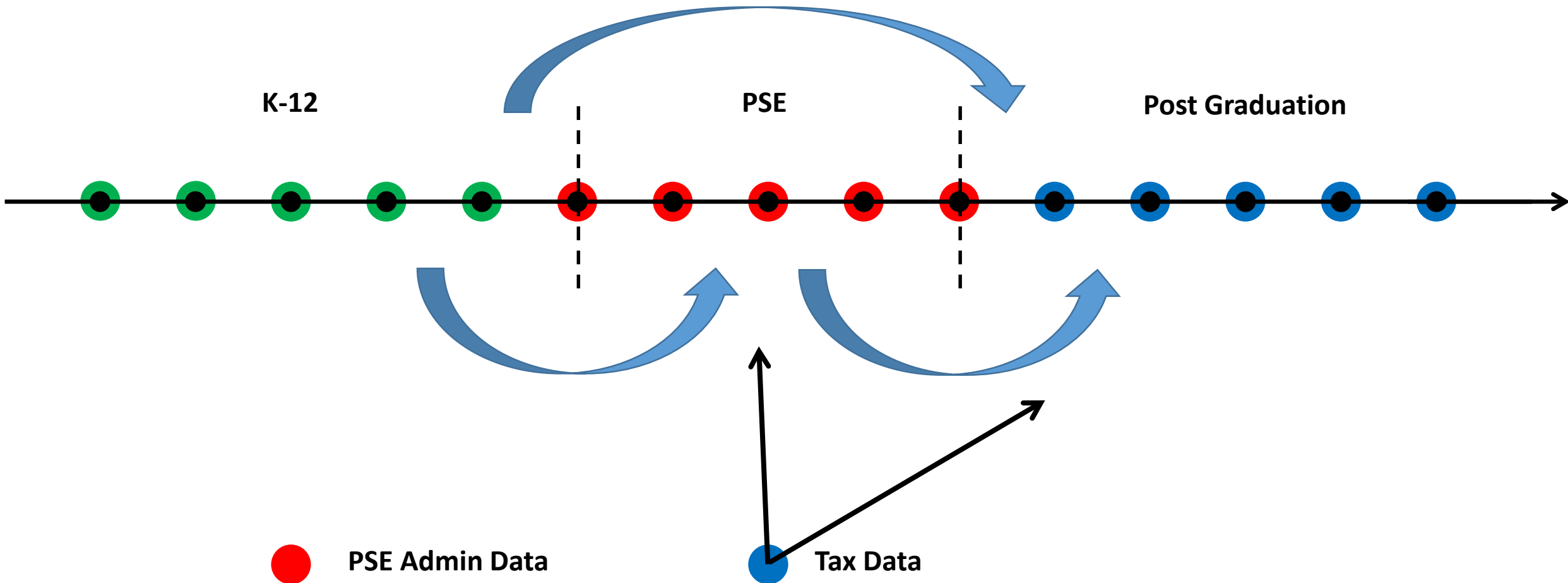
The Big Picture Now
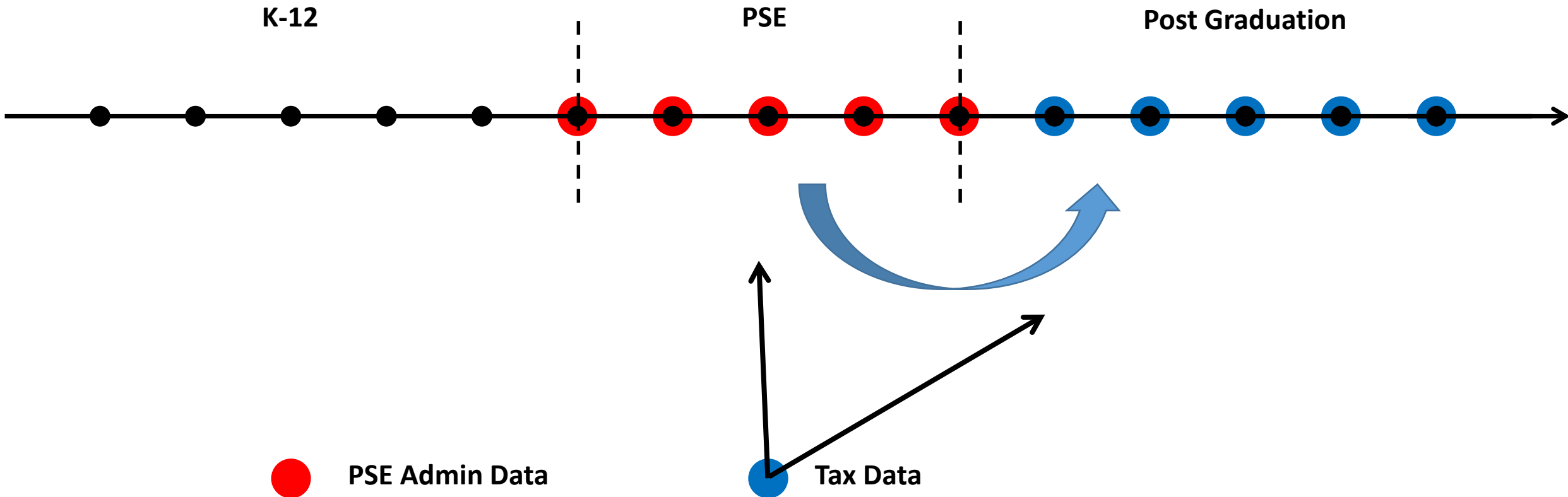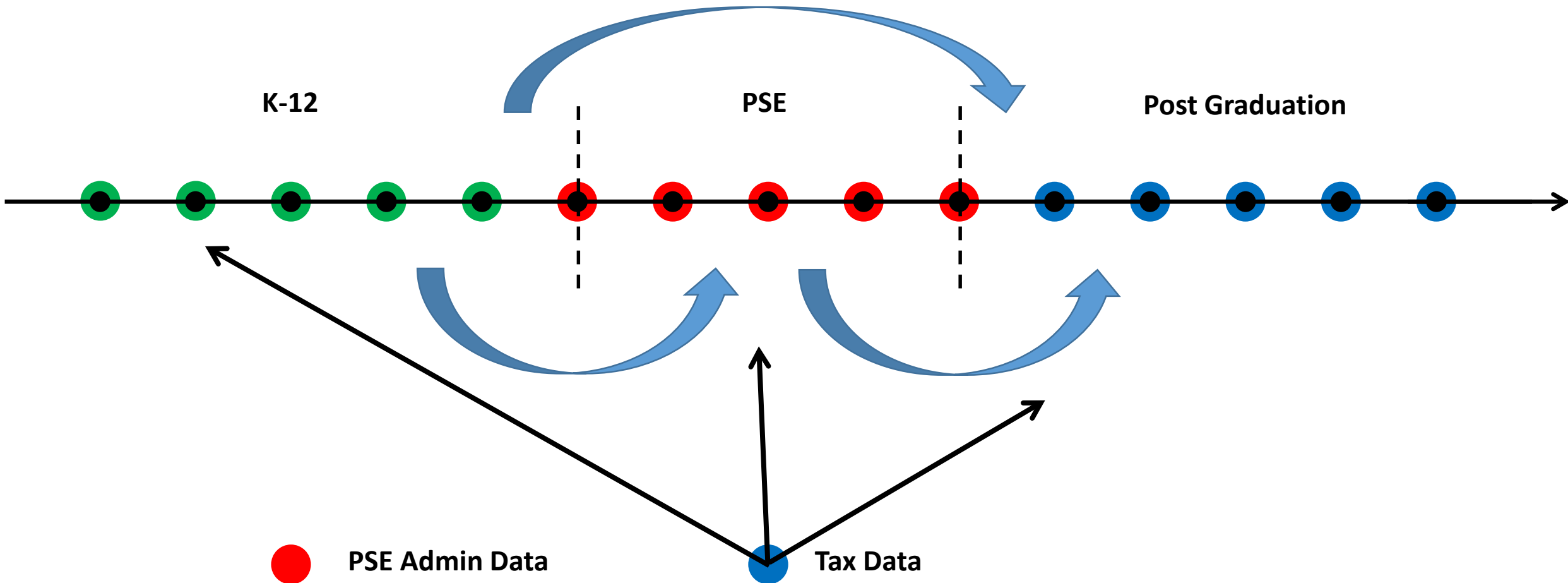
The Big Picture With K-12
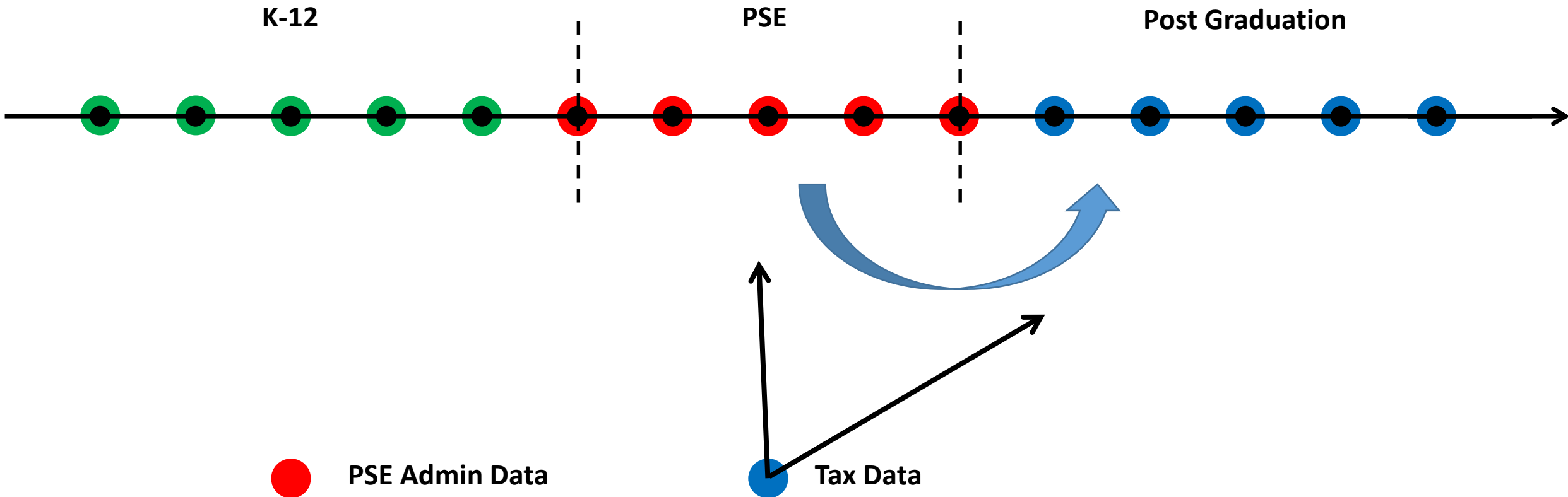
The Big Picture With K-12

The Big Picture With K-12

The Big Picture With K-12

Summary of the ELLP's Potential

- ELLP: new, unique, powerful
- Truly a "platform" to be built on
- Starting with PSE-tax data (like New Zealand)
- But many ways to extend
- All about better understanding and ultimately improving student experiences and outcomes
- And that's a good thing!

Oh, By the Way...The ELLP is Extremely Complex

Structure of the ELLP

- Not a dataset, but a relational data environment
  - PSIS & RAIS
  - T1FF
  - Anything else that is added

- PSE data: separate record each program each year
- T1FF: separate record each year

- PSE data complex + tax data complex = doubly complex

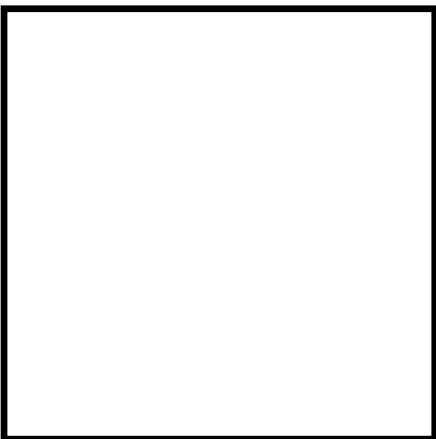# Tracking Students through PSE – PSIS & RAIS

Tracking – Complexity

- 1 record per program each year
- Multiple programs per year possible
- Need to link programs across years
    - individual's master ID
    - credential
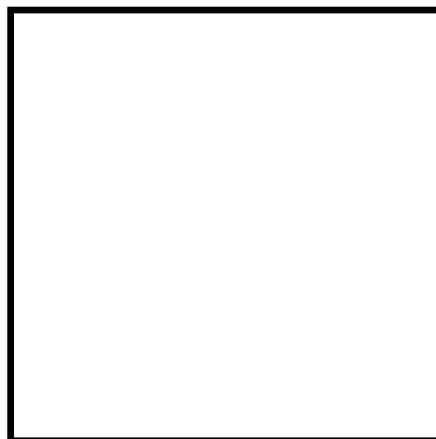   - program
    - start date

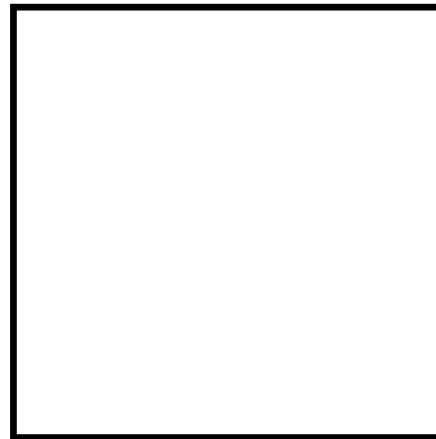# e.g., Pathways to Graduation (Should be Easy?...)
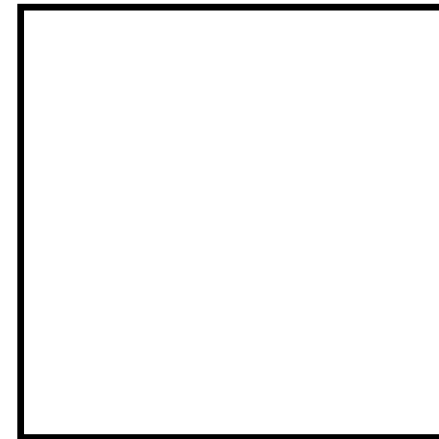
Pathways to Graduation

2009

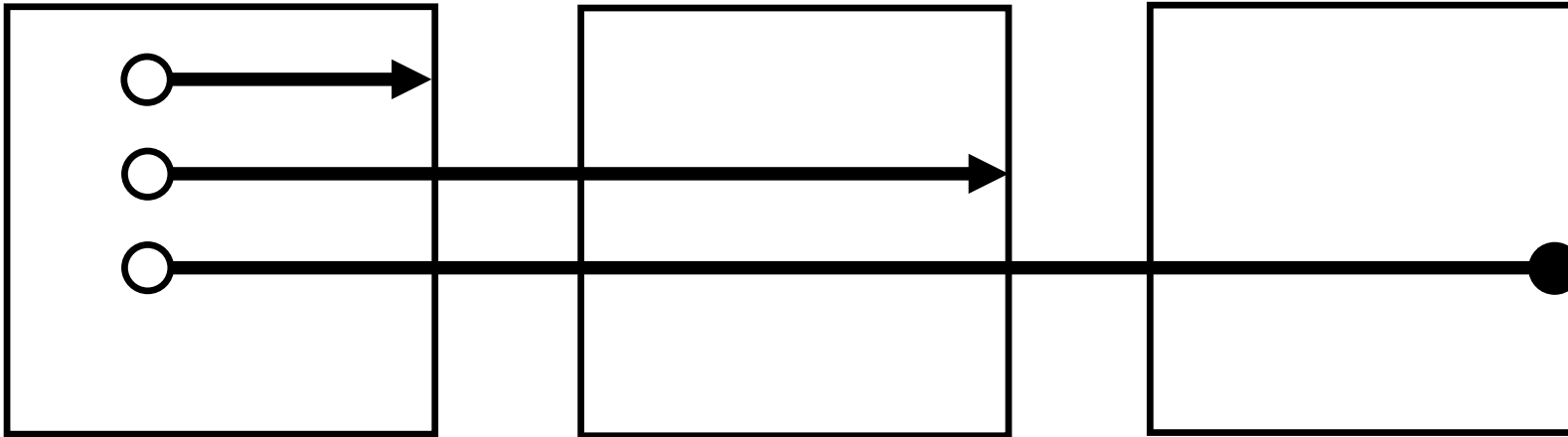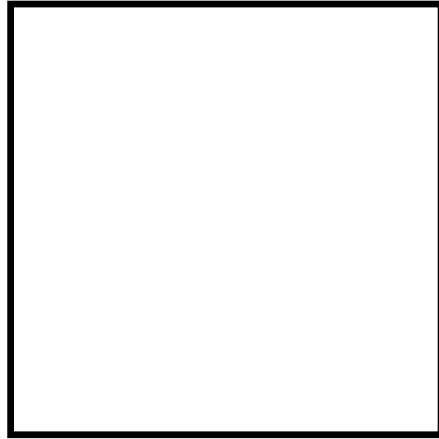**2010**
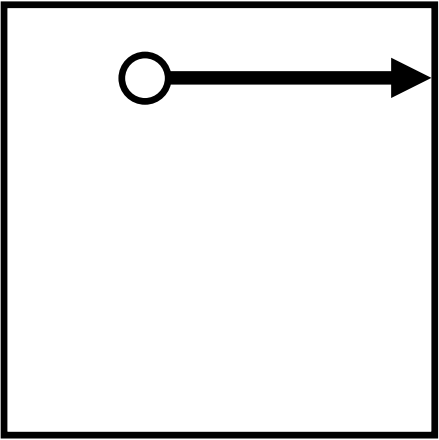
2011

• • •

2015
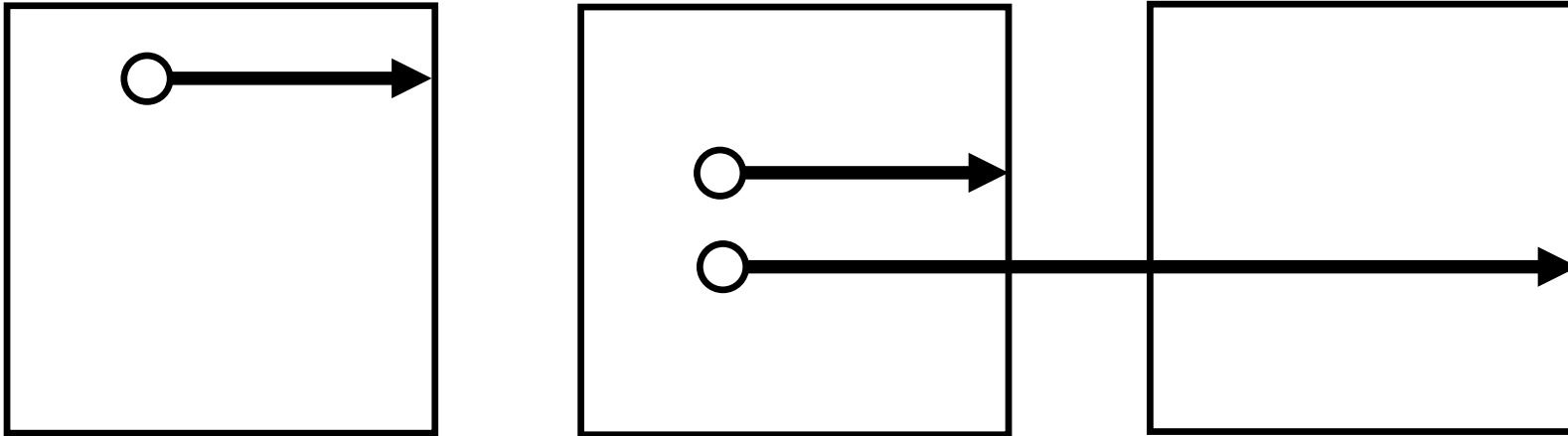
Linear Pathway to Graduation

Matched by:
- Master ID
- credential
- program
- start date

Withdrawal from PSE

Switching Programs

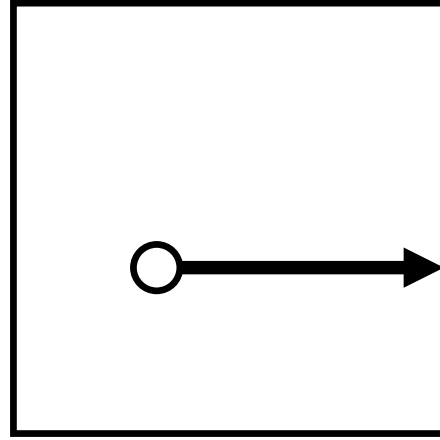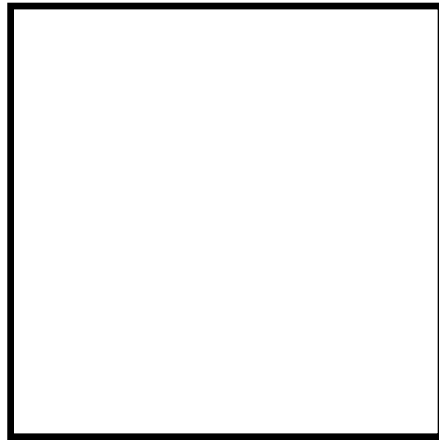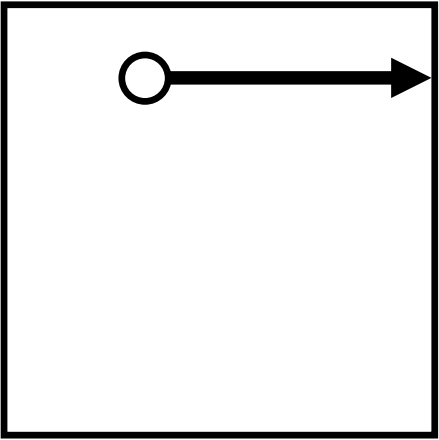- Same institution
- Different institution/type
- Same or different credential
- Combinations of the above

"Stopping Out" from PSE – Same Program

"Stopping Out" from PSE – Starting New Program

Graduation & New Program
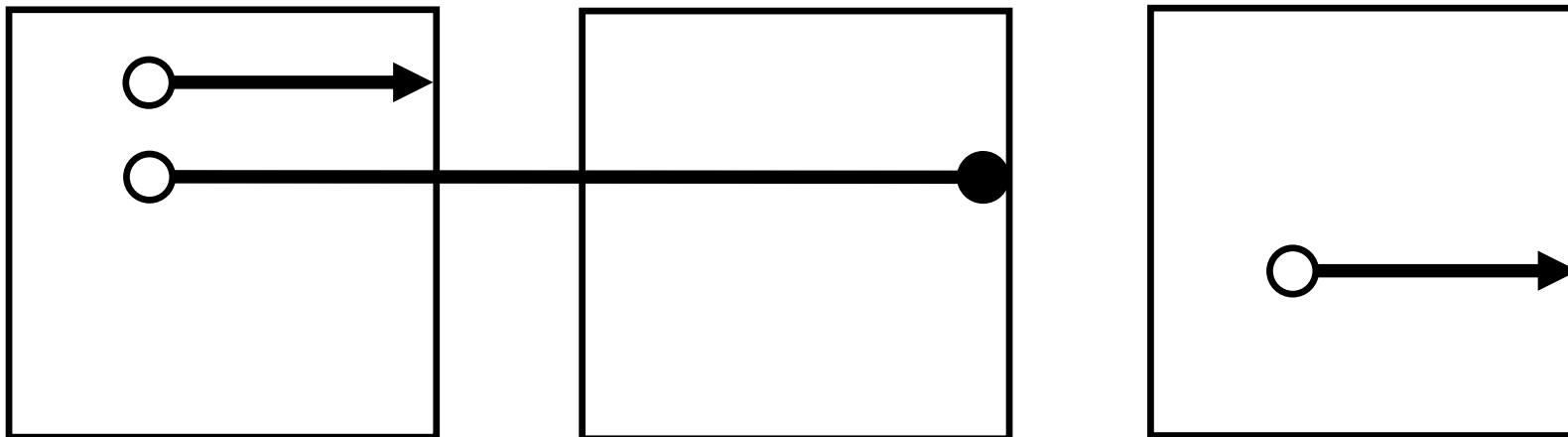
Multiple Programs in a Given Year: What's What?
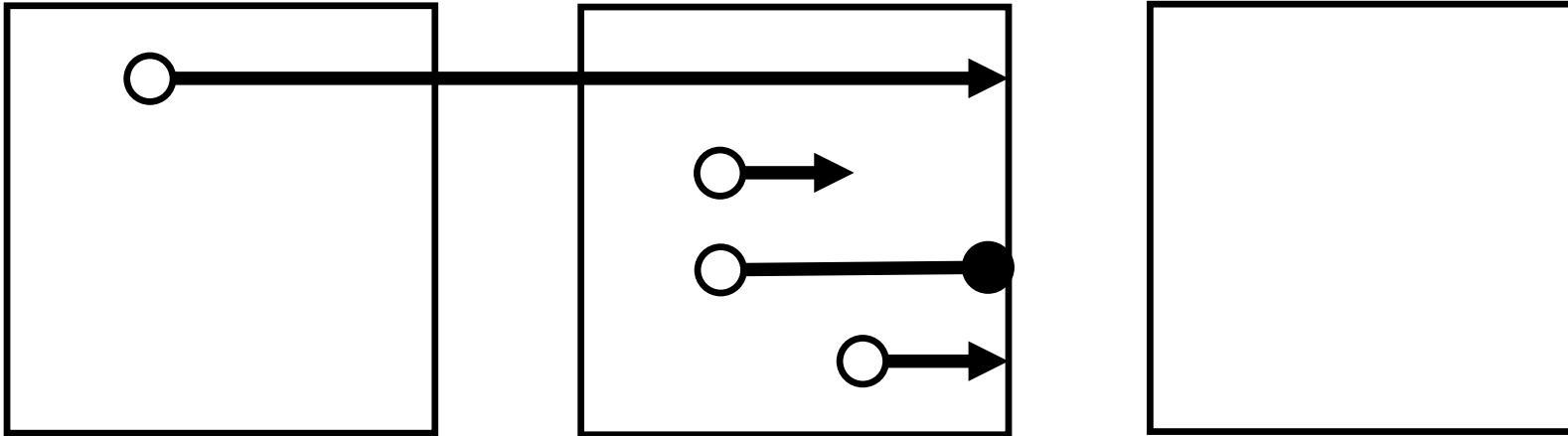
Tax Data Are Also Complex

- Need to understand definitions/meaning of each variable
- Data not necessarily consistent over time (death code...)
- Sometimes missing data – including whole years
- Variables for analysis may need to be constructed
- Sample selection issues (e.g. earnings)
- Ultimately data are limited (e.g. wage rate, labour supply): how to use to best advantage?

T1FF Data: Annual Basis
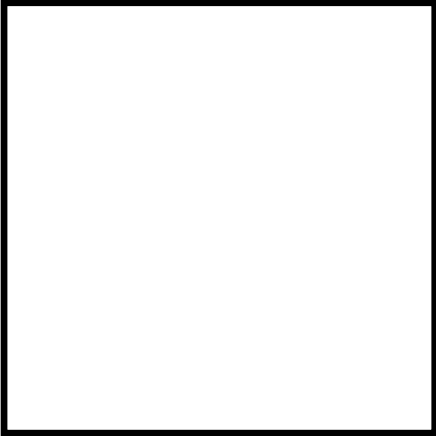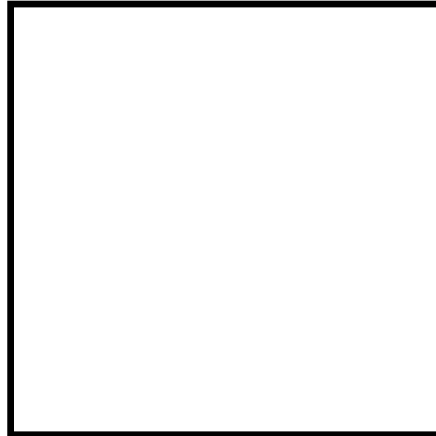
1982   **1983**   **2015**
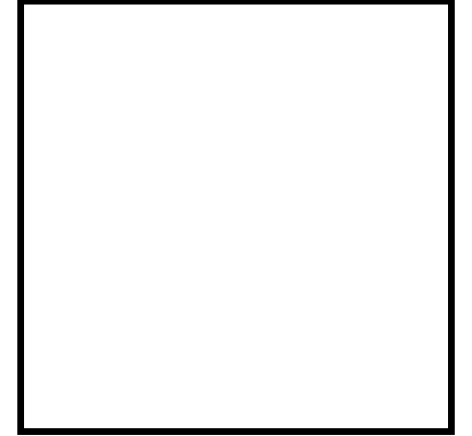
1. Pre-PSE
2. During PSE
3. Post-PSE
4. Tracked using master ID

PSE data complex + tax data complex =
doubly complex

Linking Post-PSE Outcomes to PSE: Master ID
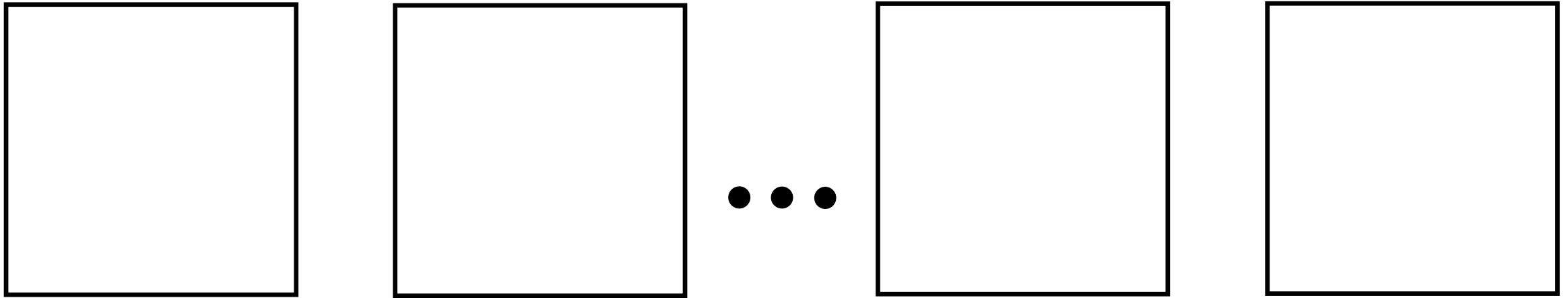
PSE as Above

Post-PSE

Customizing for Each Project

- Start with research question
- Identify structure of data required and build (unit of observation critical!)
- Populate with needed variables
  - PSE-based variables
  - Tax-based variables
  - Constructed variables (e.g., socio-economic status or labour market conditions from census or LFS)

Possible Structures and Mechanisms for Advancing & Carrying Out Research With the ELLP

What Is This All About? ("I Just Want to Do My Research")

Data/Research Context
- Complexity of the data: data preparation & manipulation
- Overlap of research questions & analytical approaches

Goals
- Reduce barriers to entry – hit the deck running!
- Maximize efficient use of the data

Sound familiar? (IRP, MESA Project...)

A Proposal: ELLP "Research Hubs"

- Highly skilled research team(s) focused on the ELLP

- Provide access for individual researchers/teams (but not a "gate-keeper" of the ELLP)

- Coordinated, synergistic research

Add Policy Focus – Government Projects? What a Research Hub Could Do

- Help turn policy questions into research projects
- Effectively manage projects – with specific "deliverables"
- Produce a range of research products
- Disseminate the findings across a range of stakeholders – including the general public
- For small or large projects
- Or *sets* of projects

Potential National/Regional System of ELLP Research Hubs

- Share ideas, experience & expertise: "capacity building"

- Facilitate coordination of research across hubs as appropriate

- Ensure both national and regional policy questions are addressed

What Will Make These ELLP Research Hubs Happen?

- Core, stable funding to maintain capacity

- Key role for government policy stakeholders

- Other stakeholders: PSE groups, employer organizations, others

- Universities and collective university organisations

Could Fit Into Existing Structures – Where They Exist

- Quebec well-suited for one or more ELLP Research Hubs?

- Build onto existing structures?

- Data will be in RDCs

Summary of ELLP Research Hubs Idea

- Context of complex data and overlapping projects → efficiencies of research undertakings

- Win-win-win…
  - Researchers (plus students who get training)
  - Policy makers and other PSE stakeholders
  - General public

"The Platform" (ELLP): Opportunities and Challenges – A Researcher's Perspective

Ross Finnie (rfinnie@uottawa.ca)

EPRI — Education Policy Research Initiative
Initiative de recherche sur les politiques de l'éducation

The Big Picture