Statewide Longitudinal Data Systems

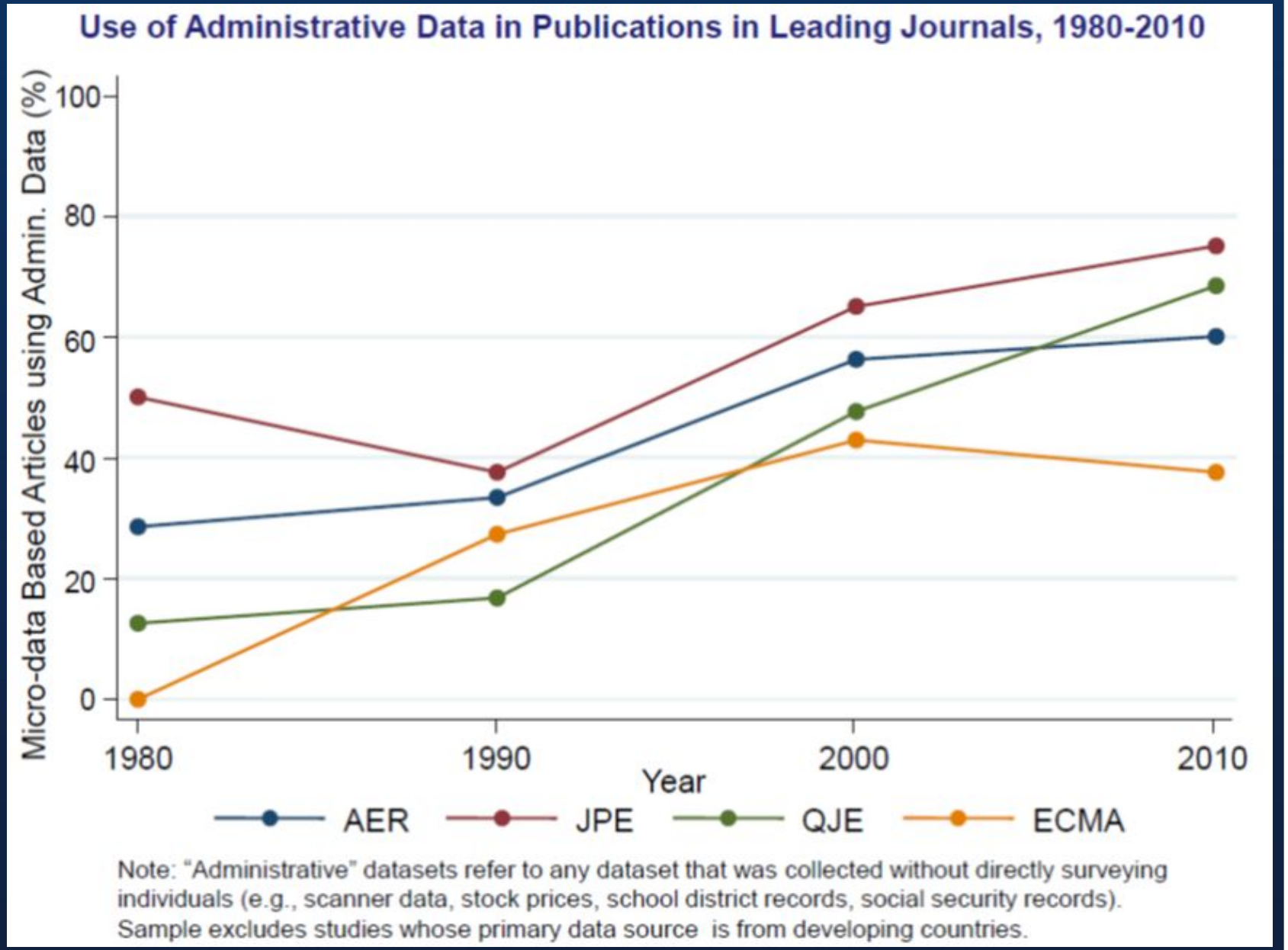
What constitutes the Michigan Education Data Center and how can it be used to further STEM Education research

Kevin Stange, Professor of Public Policy
Kyle Kwaiser, Data Architect

MICHIGAN EDUCATION DATA CENTER
UNIVERSITY OF MICHIGAN

Overview

- What are State Longitudinal Data Systems (SLDS) in Education?
- Michigan Education Data Center (MEDC) holdings deep dive
- Example projects
  - K12
  - Postsecondary
- Brief mention about funding
- Q&A



Source: Chetty, Raj. 2012. "Time Trends in the Use of Administrative Data for Empirical Research" presentation at NBER Summer Institute, July 2012.

Context: Administrative data revolution in social science

Growth in Use of Admin Higher Ed Data

1993 NBER Volume

2020 NBER Volume

Chapter focus		
Theory/conceptual	2	1
Survey data	4	1
Admin data	2	7

Particularly in economic analysis of higher education (my field)

Administrative data is often used where survey sources historically have been

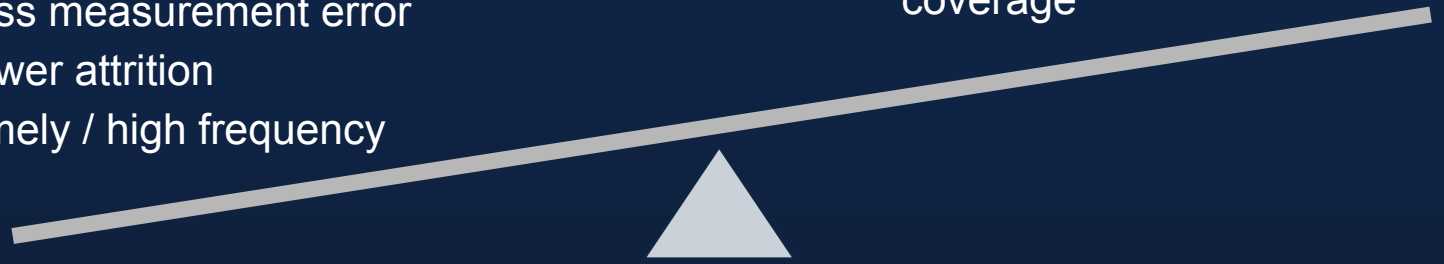
- NCES longitudinal studies (many)
- NSF Recent College Graduates survey
- NSF Survey of Earned Doctorates
- Census/American Community Survey

Benefits

- Large samples
  → Subsamples
  → Better research strategies
- Lower cost of collection
- Less measurement error
- Lower attrition
- Timely / high frequency

Drawbacks

- Domain-specific
  - Limited variables
  - Matching issues
- Access
- Specific context (e.g. state)
- Incomplete geographic coverage



Admin data has strengths (and some drawbacks)

Statewide Longitudinal Data System

Often contains data from four domains

- Early learning
- K12
- Postsecondary
- Workforce

Each domain includes varying sub-domains (e.g. K12 enrollment, test scores, demographics, attendance, discipline, teachers)

Sometimes other domains: criminal justice, health and human services

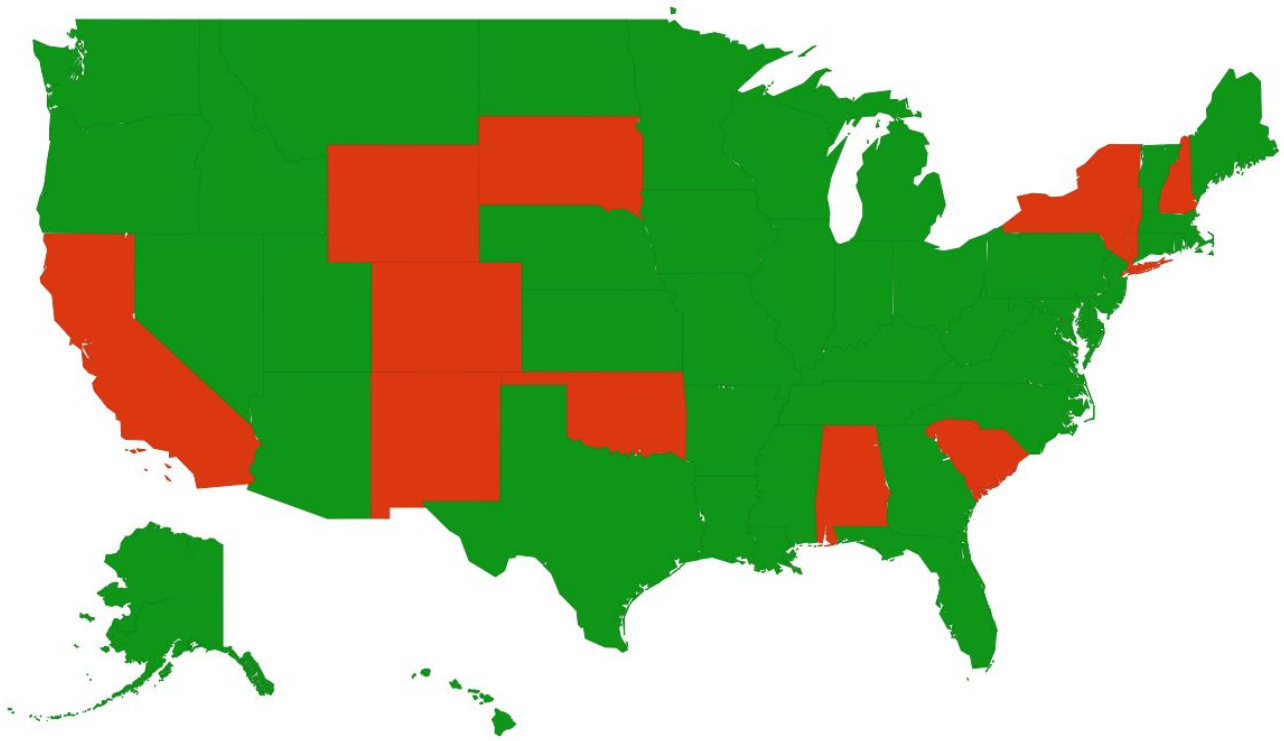
De-identified, but linked across domains via unique identifier

Largely funded by US Dept of Education (a legacy of focus on test-based accountability in NCLB)

What is an SLDS?

- Major source of administrative data used in education research

States with Statewide Longitudinal Data Systems, 2021

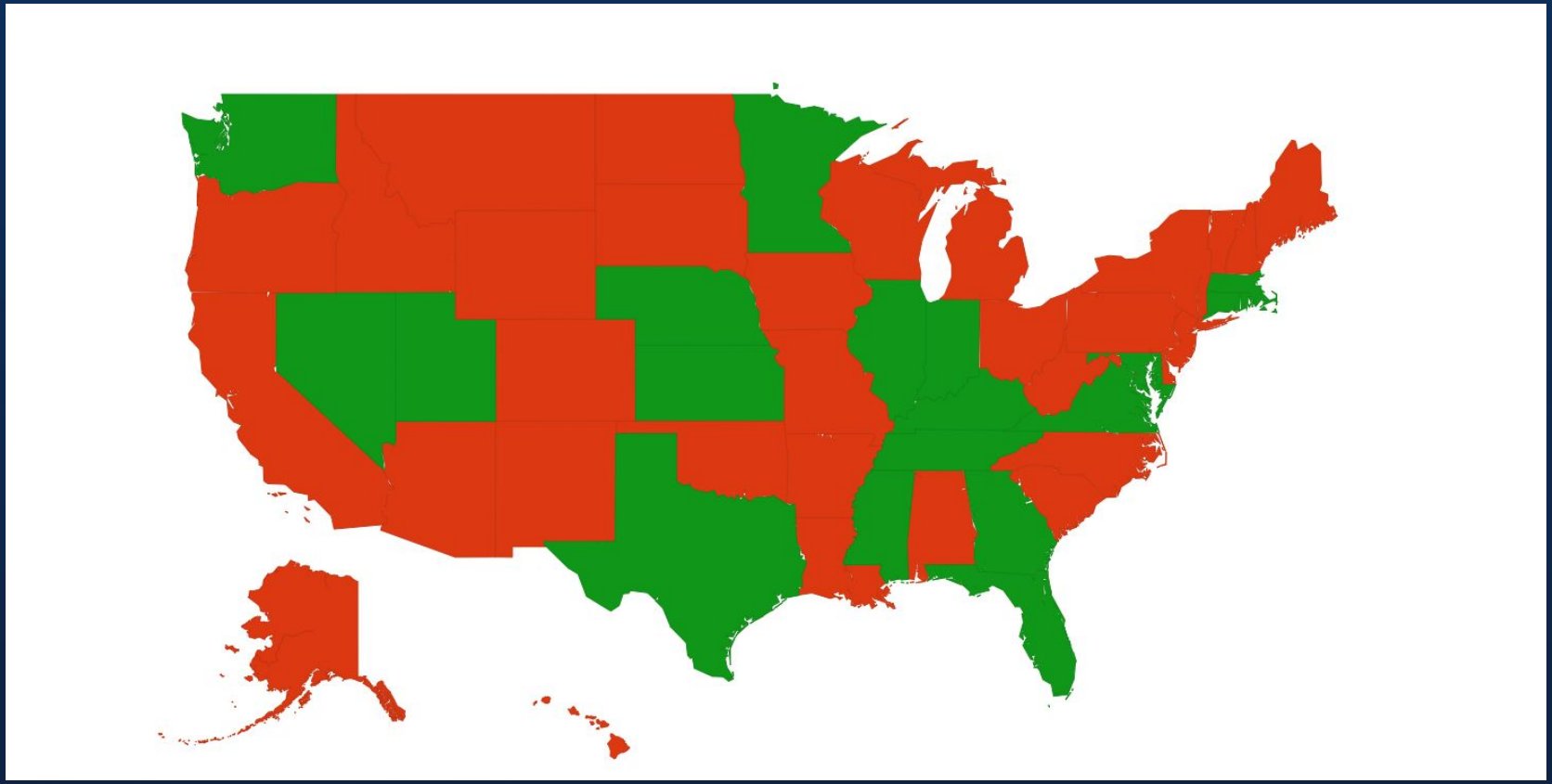


Source: Education Commission of the States, *50-State Comparison: Statewide Longitudinal Data Systems, December 2021. https://www.ecs.org/state-longitudinal-data-systems/*

40 states have state-wide systems

- With at least 2 of the domains
- Even states without SLDS will often have sub-state systems (e.g. CSU in CA, CUNY in NY)

States with SLDS containing all four domains, 2021



Source: Education Commission of the States, *50-State Comparison: Statewide Longitudinal Data Systems, December 2021. https://www.ecs.org/state-longitudinal-data-systems/*

19 states have extensive SLDS

Containing matched records from

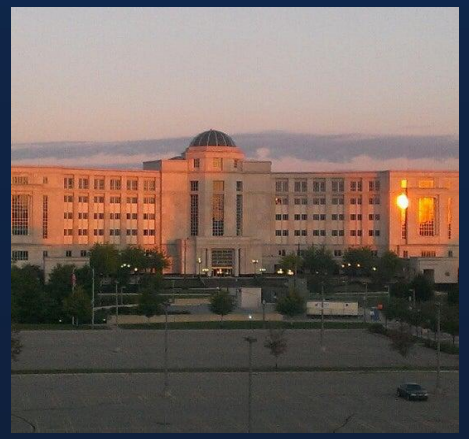
- early learning
- K12
- postsecondary
- workforce

MICHIGAN EDUCATION DATA CENTER

- Founded in 2018, the Michigan Education Data Center (MEDC) was born out of a new partnership with the State of Michigan (MDE, CEPI), University of Michigan, and Michigan State

- MEDC's goals are to:
  1. Enhance and increase the amount of robust education research happening in Michigan
  2. Assist MDE in fulfilling its research agenda
  3. Create consistency across Michigan's education research data sets

- MDE and CEPI have final sign off on all research and data sharing



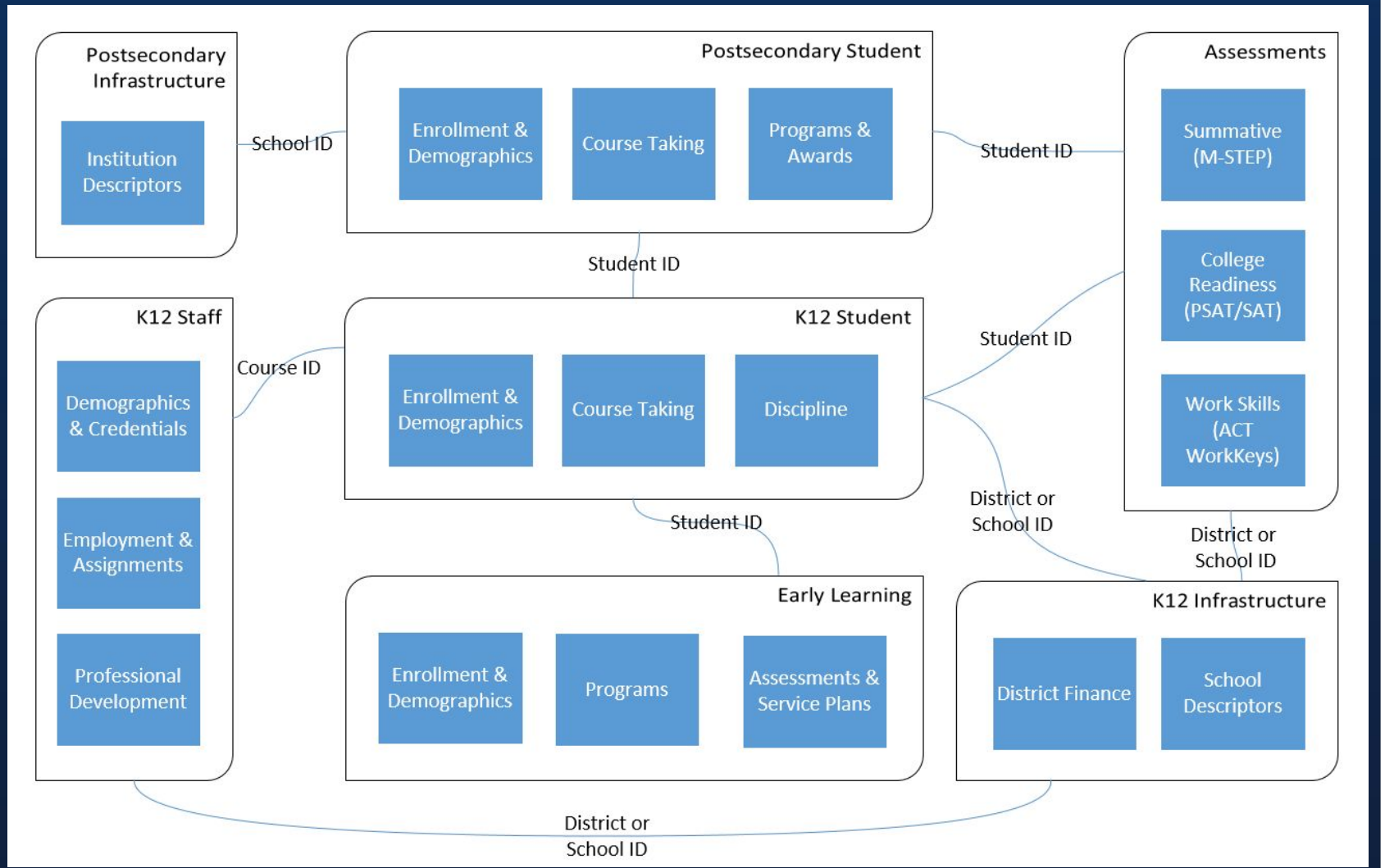
Michigan Department of Education
Center for Educational Performance & Information

The Education Policy Initiative
University of Michigan

Education Policy Innovation Collaborative
Michigan State University

Bringing a researcher's lens to the ways in which administrative education data can be used to answer tough questions facing educators, school leaders, parents, and policy makers.

MICHIGAN'S EDUCATION DATA



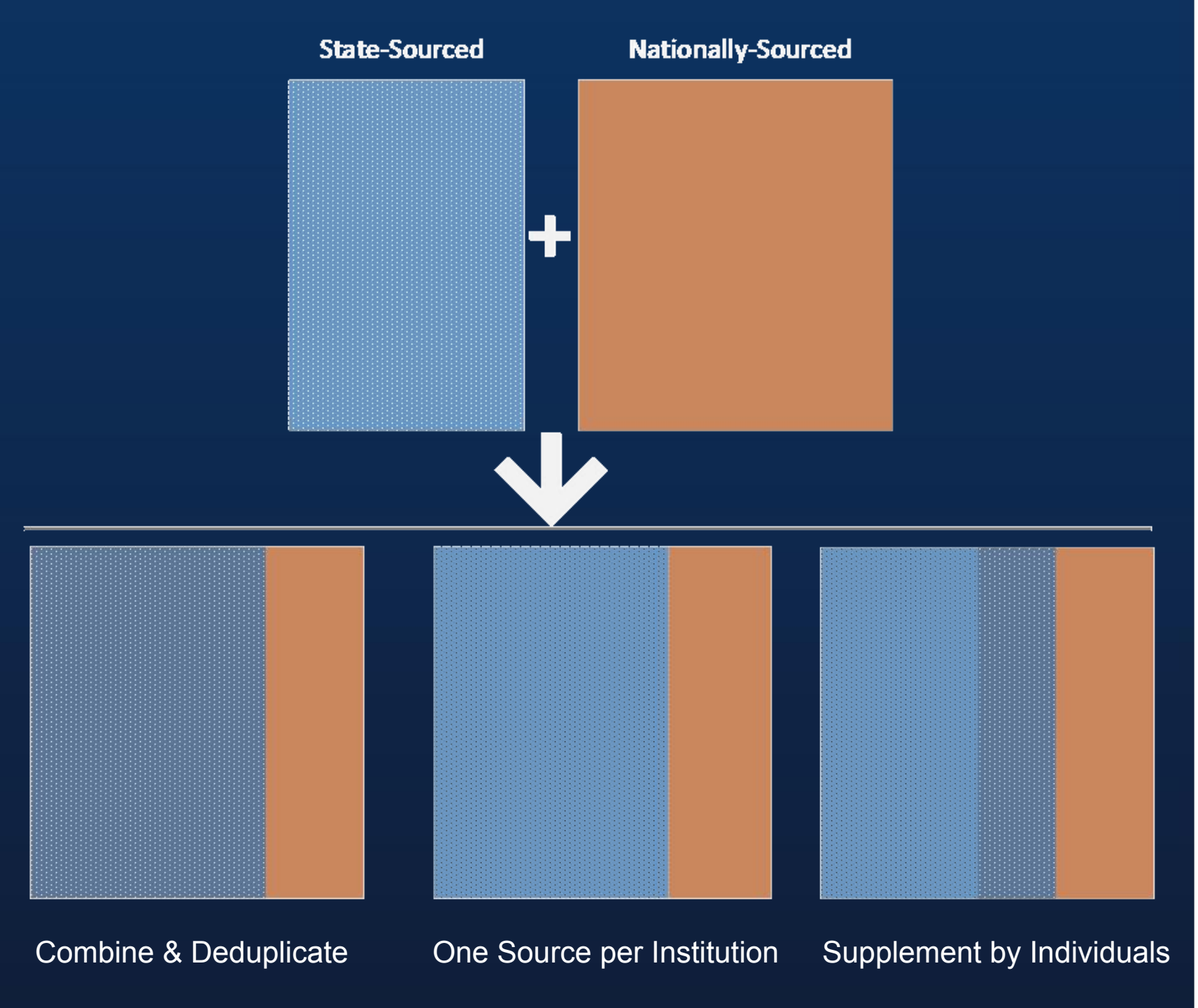
True P-20 data system with persistent identifiers and variable naming standardized from 10+ collection systems.

MICHIGAN'S EDUCATION DATA

Topics	Start Year	Delivery
Early Learning	2013	Aug
K12 Student	2003	Aug
Postsecondary	2010	Feb
Assessment	2003	Nov
Staff	2004	Sep
Infrastructure	2004	Jul

Generally, a 20 year panel with a number of exceptions.

https://medc.miedresearch.org/resources

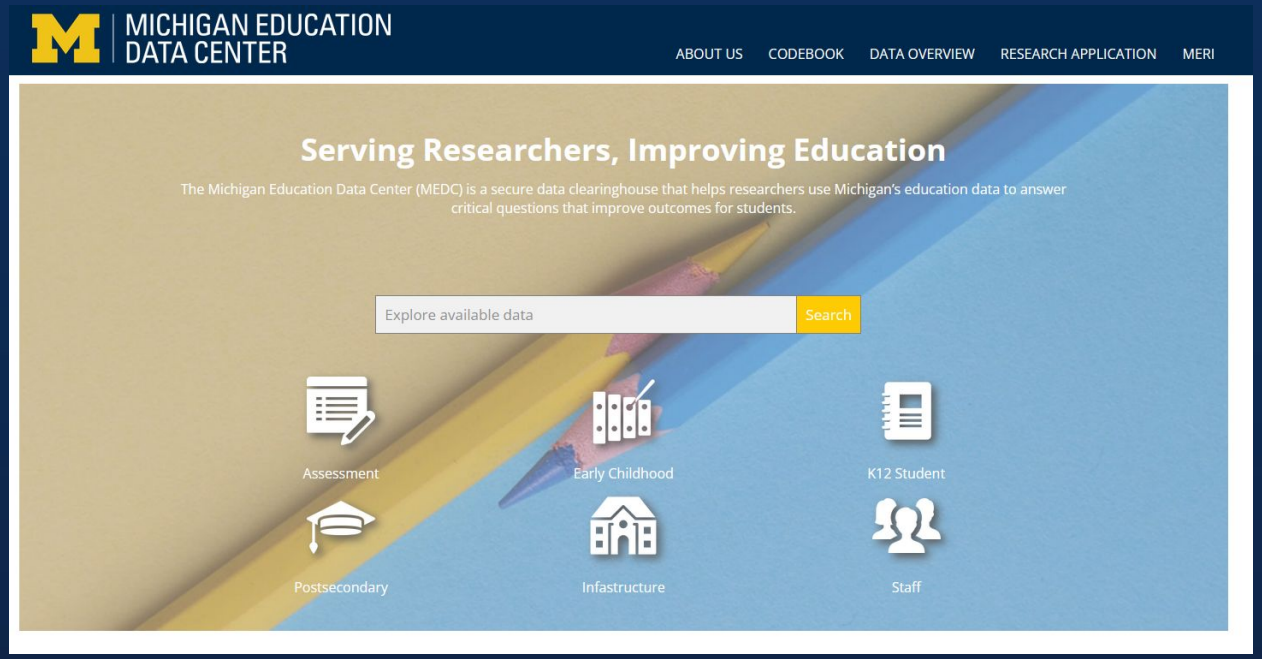


Combining Post-Secondary

Balancing coverage, utility and accuracy when combining large databases.

Rich Metadata

- Dataset overviews and codebooks medc.miedresearch.org



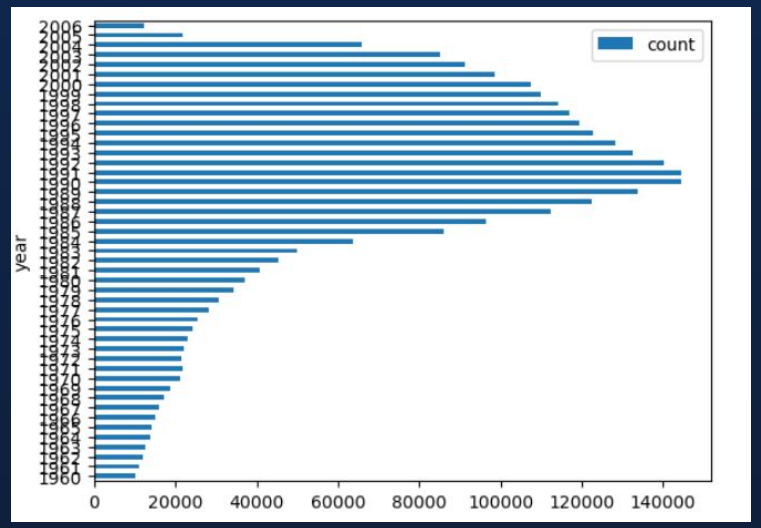
- Detailed technical notes generated by subject matter experts
- Monthly question submissions to state partners

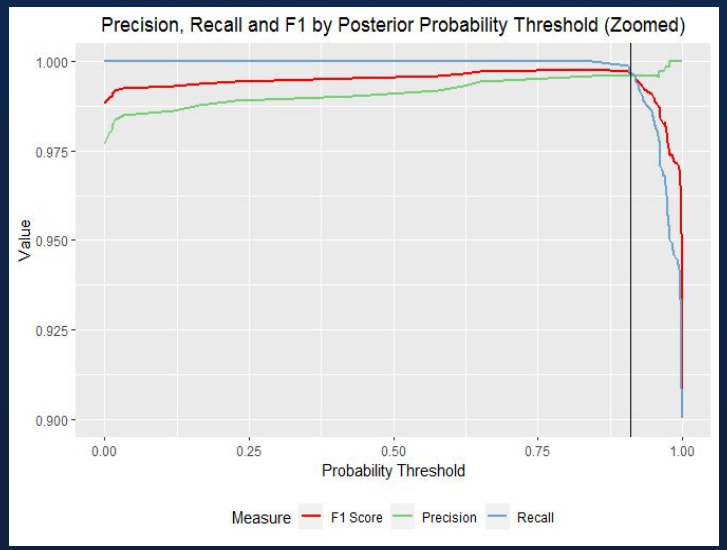
Standardized format covering everything from option sets through to data provenance.

- Enabled by comprehensive PII

	Year of Birth	Number Individuals	Address
Pre-12	1984-2017	4,874,716	Yes
Postsecondary	1984-2004	3,020,563	No
K12 Staff	1942-1999	1,160,075	No

- R's Fastlink implementation of Fellegi-Sunter model



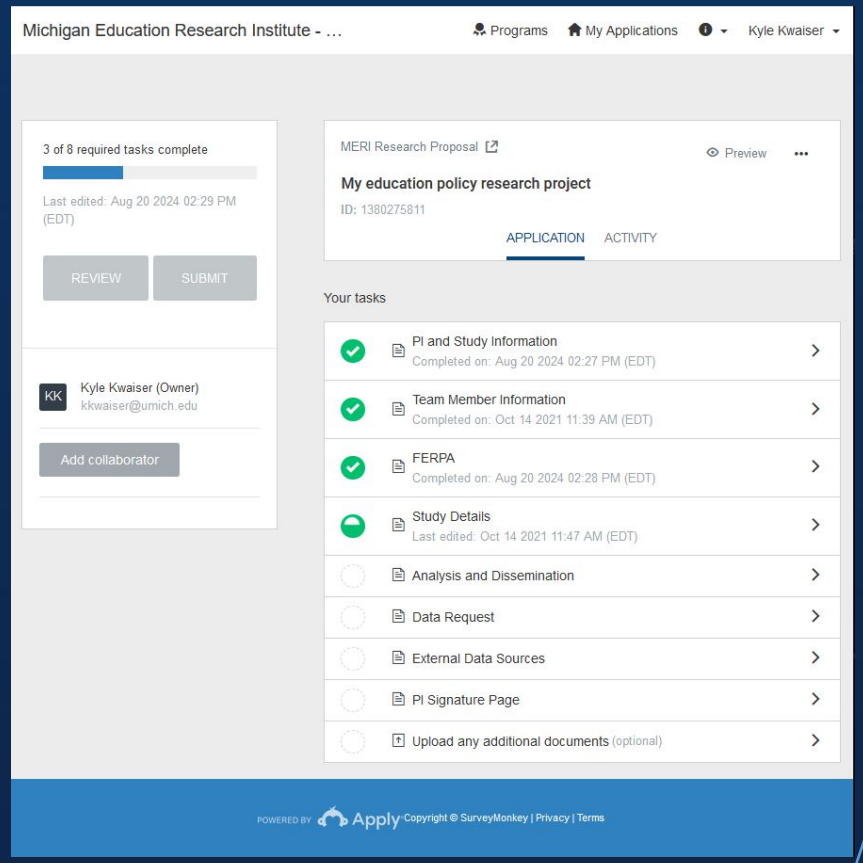


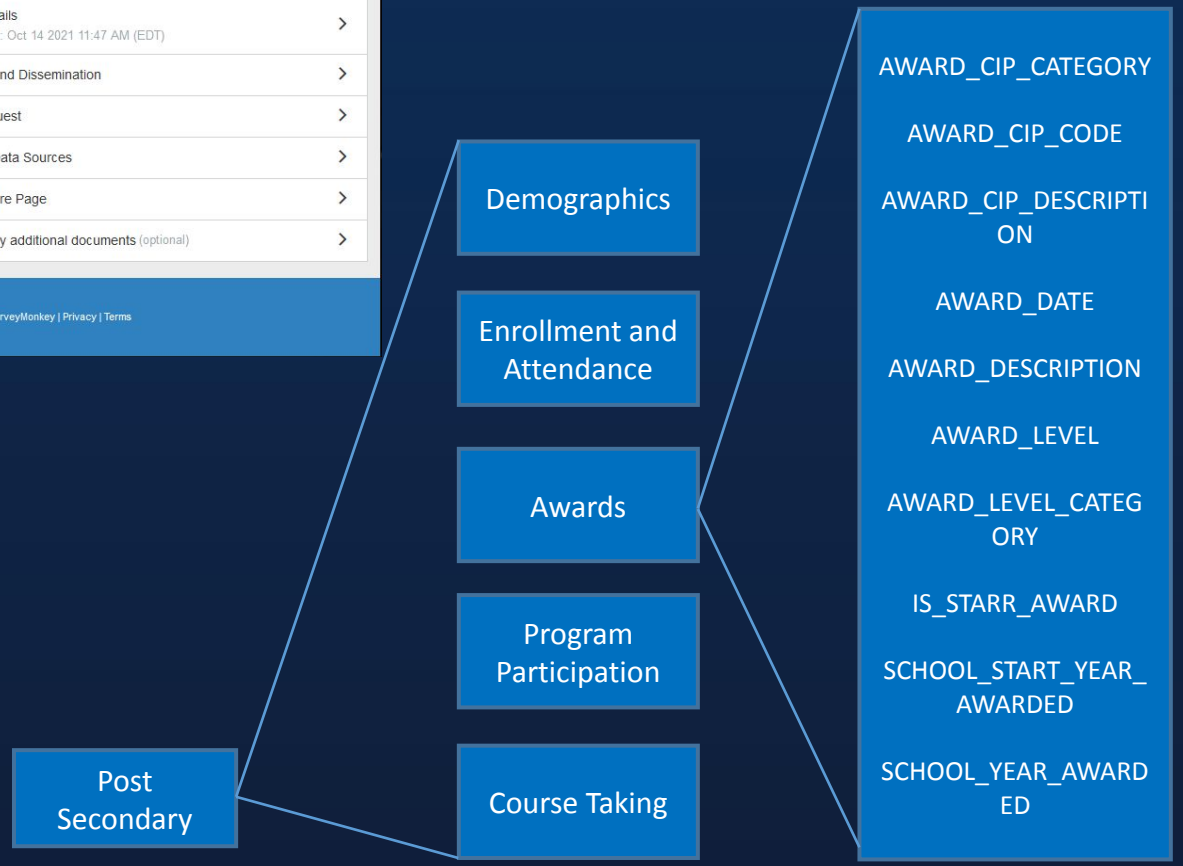
Postsec PII: Individuals by birth year

Robust Matching Model

Have capability to match on external data (with permission)

Sophisticated probabilistic entity resolution enabling matches leveraging name, date of birth and/or address fields.





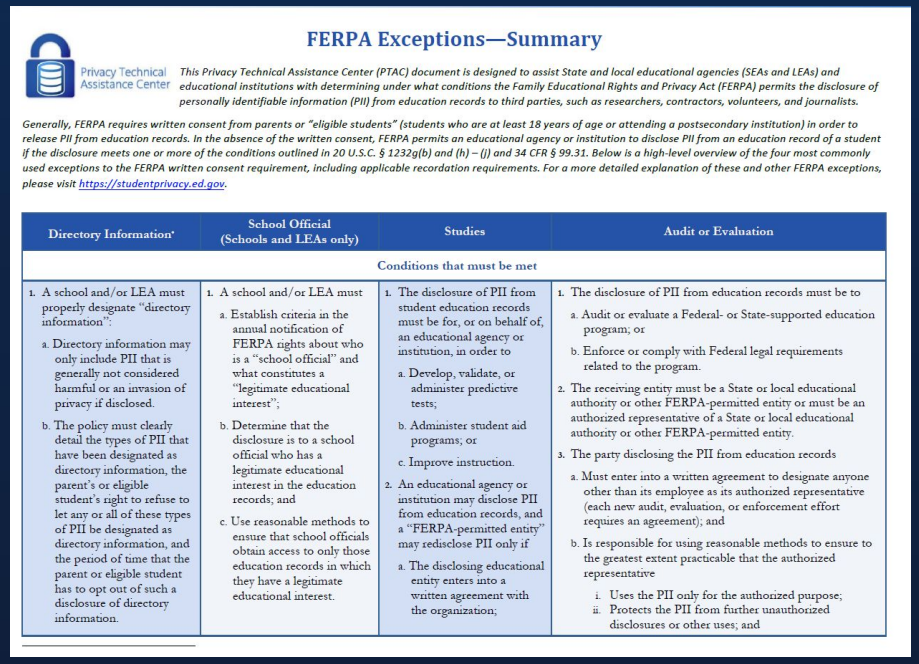
Research Application & Review

We provide the infrastructure for submitting research proposals and support before and during the application review process.

- Data are FERPA-protected, must meet exceptions
  - Studies exception
  - Audit or Evaluation

- Address Michigan's Top 10 Strategic Education Plan
  - https://miedresearch.org/agenda/

- Data security



FERPA Exceptions—Summary

Privacy Technical Assistance Center

This Privacy Technical Assistance Center (PTAC) document is designed to assist State and local educational agencies (SEAs and LEAs) and educational institutions with determining under what conditions the Family Educational Rights and Privacy Act (FERPA) permits the disclosure of personally identifiable information (PII) from education records to third parties, such as researchers, contractors, volunteers, and journalists.

Generally, FERPA requires written consent from parents or "eligible students" (students who are at least 18 years of age or attending a postsecondary institution) in order to release PII from education records. In the absence of the written consent, FERPA permits an educational agency or institution to disclose PII from an education record of a student if the disclosure meets one or more of the conditions outlined in 20 U.S.C. § 1232g(b) and (h) – (j) and 34 CFR § 99.31. Below is a high-level overview of the four most commonly used exceptions to the FERPA written consent requirement, including applicable recordation requirements. For a more detailed explanation of these and other FERPA exceptions, please visit https://studentprivacy.ed.gov.

Directory Information*	School Official (Schools and LEAs only)	Studies	Audit or Evaluation
Conditions that must be met			
1. A school and/or LEA must properly designate "directory information": a. Directory information may only include PII that is generally not considered harmful or an invasion of privacy if disclosed. b. The policy must clearly detail the types of PII that have been designated as directory information, the parent's or eligible student's right to refuse to let any or all of these types of PII be designated as directory information, and the period of time that the parent or eligible student has to opt out of such a disclosure of directory information.	1. A school and/or LEA must a. Establish criteria in the annual notification of FERPA rights about who is a "school official" and what constitutes a "legitimate educational interest"; b. Determine that the disclosure is to a school official who has a legitimate educational interest in the education records; and c. Use reasonable methods to ensure that school officials obtain access to only those education records in which they have a legitimate educational interest.	1. The disclosure of PII from student education records must be for, or on behalf of, an educational agency or institution, in order to a. Develop, validate, or administer predictive tests; b. Administer student aid programs; or c. Improve instruction. 2. An educational agency or institution may disclose PII from education records, and a "FERPA-permitted entity" may redisclose PII only if a. The disclosing educational entity enters into a written agreement with the organization;	1. The disclosure of PII from education records must be to a. Audit or evaluate a Federal- or State-supported education program; or b. Enforce or comply with Federal legal requirements related to the program. 2. The receiving entity must be a State or local educational authority or other FERPA-permitted entity or must be an authorized representative of a State or local educational authority or other FERPA-permitted entity. 3. The party disclosing the PII from education records a. Must enter into a written agreement to designate anyone other than its employee as its authorized representative (each new audit, evaluation, or enforcement effort requires an agreement); and b. Is responsible for using reasonable methods to ensure to the greatest extent practicable that the authorized representative i. Uses the PII only for the authorized purpose; ii. Protects the PII from further unauthorized disclosures or other uses; and

Michigan's Top 10 Strategic Education Plan

Research Application Tips

The process is detailed but MEDC staff are available for consultation and involved throughout.

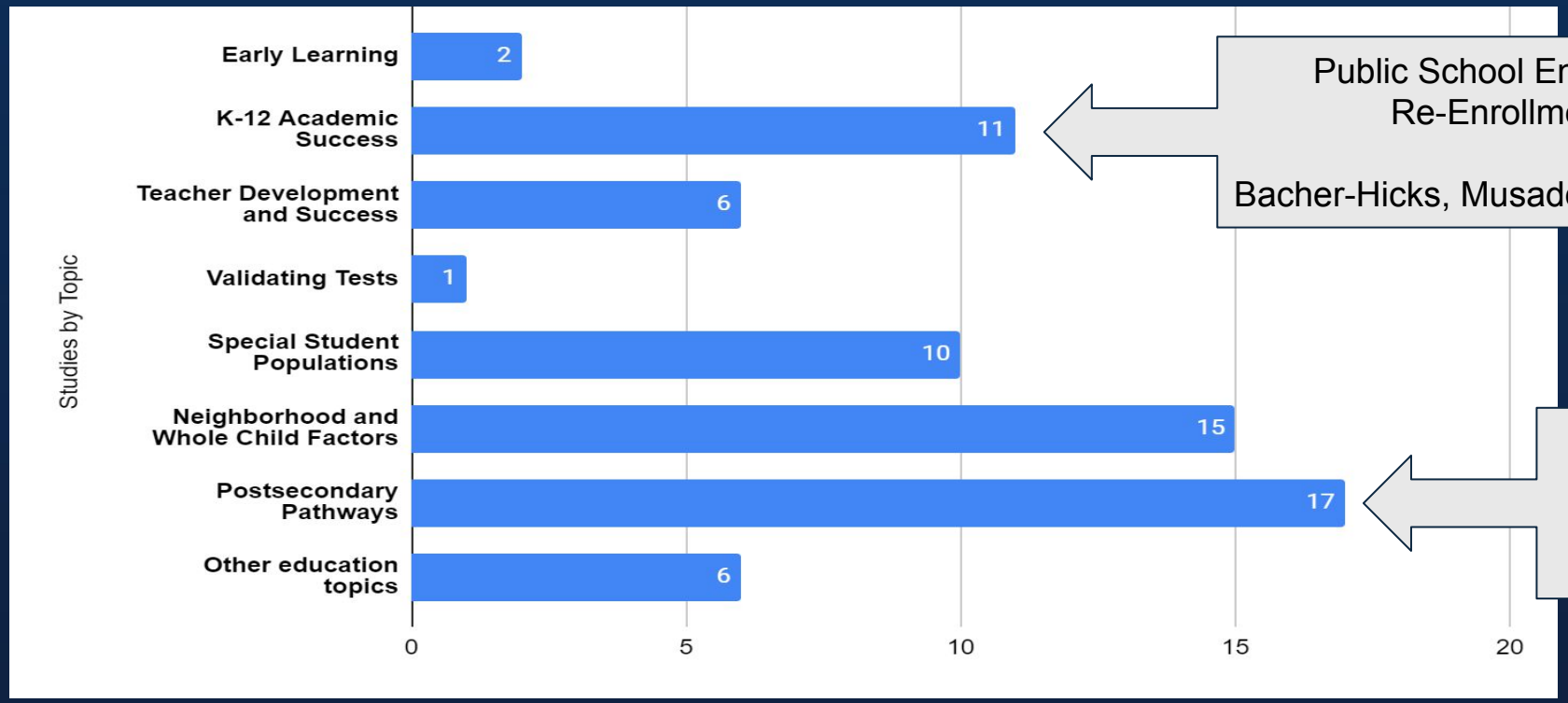
- Research applications reviewed ~every 2 months
  - Friday, January 26, 2024
  - Friday, March 29, 2024
- https://medc.miedresearch.org/application
- Cost recovery for staff and IT infrastructure

MEDC Service	FY24 fees for U-M researchers	FY24 fees for all other researchers
Fee for new projects	$4,830	$6,230
Fee for data updates	$605	$780
Probabilistic matches	Dependent on match complexity	Dependent on match complexity

Apply For Access!

Apply early as all projects are vetted, approved and sponsored by State of Michigan staff (who are quite busy).

Active MEDC Projects, by Topic



Many applications

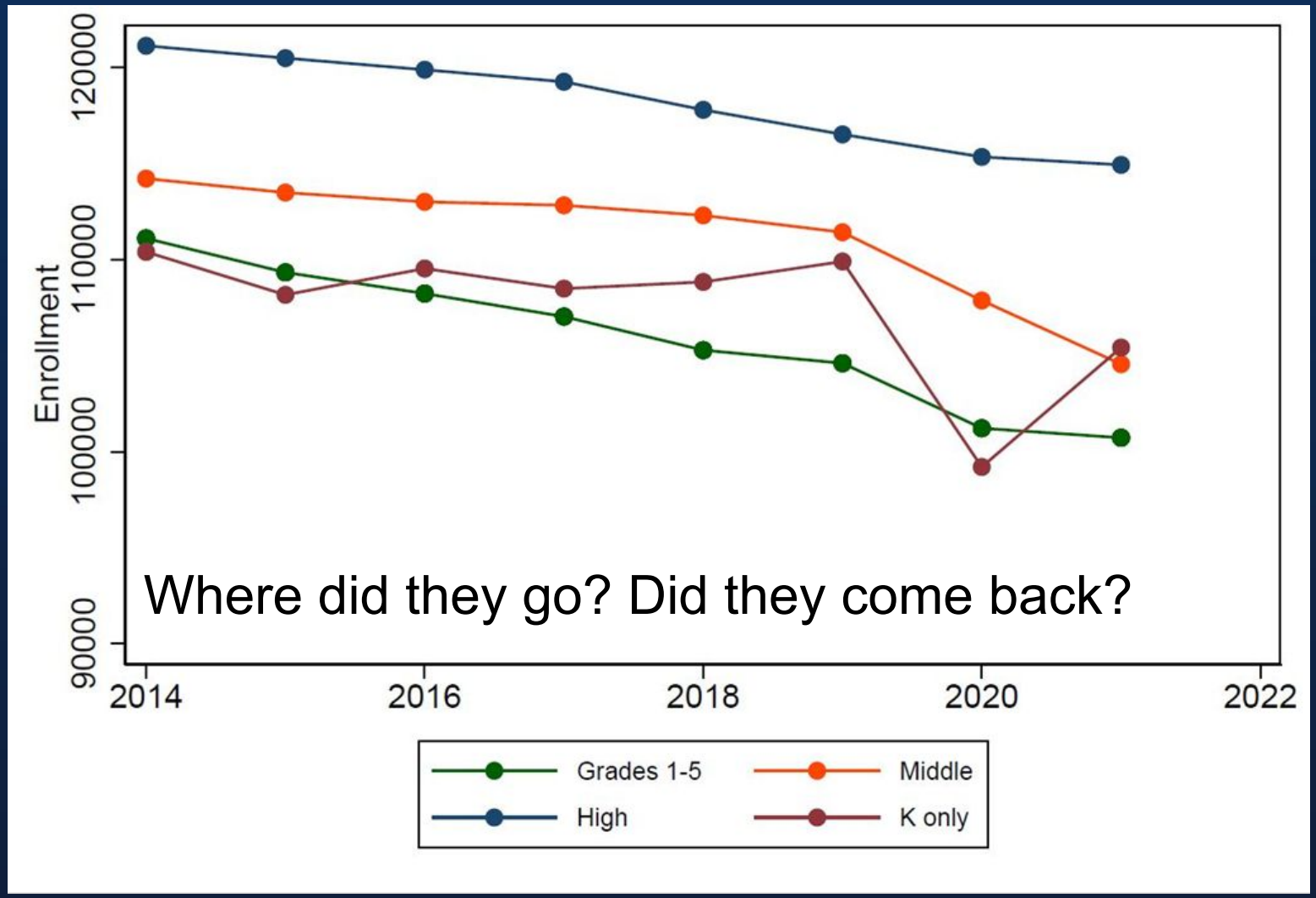
Public School Enrollment, Dis-Enrollment, and Re-Enrollment During the Pandemic

Bacher-Hicks, Musaddiq, Goodman & Stange (2023)

Does “Marginal Price” Impact Student Course-taking and Time-to-degree?

Hemelt & Stange (2015)

Public School Enrollment per Grade in Michigan

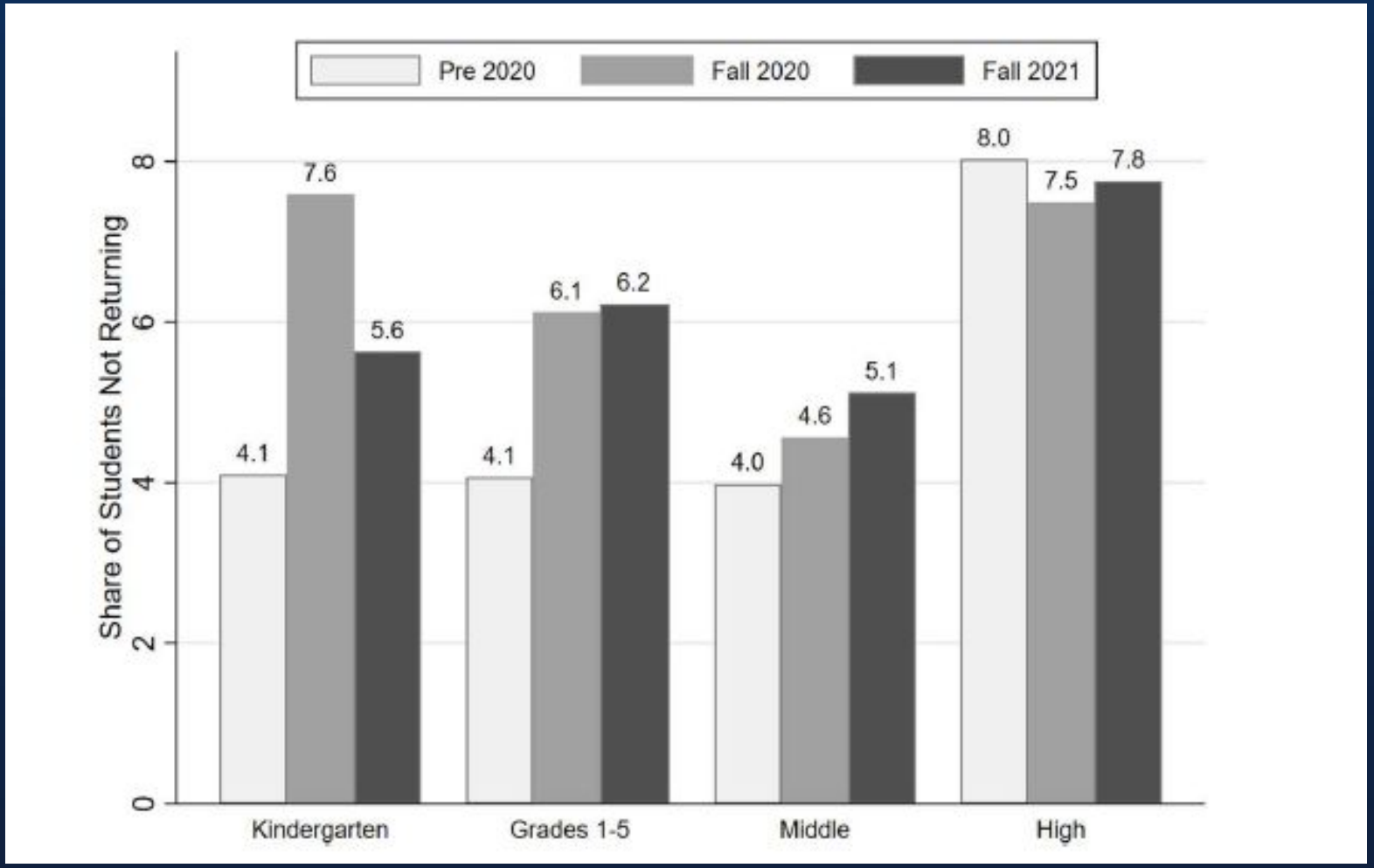


Overall enrollment in Michigan public schools

- Clear drop in enrollment (2-3% overall) in Fall 2020
- Largest drops in kindergarten, but also in other elementary and middle school grades.
- Enrollment still below pre-pandemic trends in Fall 2021, but partially bounces back.
- Could be created with aggregate/public data too

Source: Bacher-Hicks, Musaddiq, Goodman & Stange (2023)

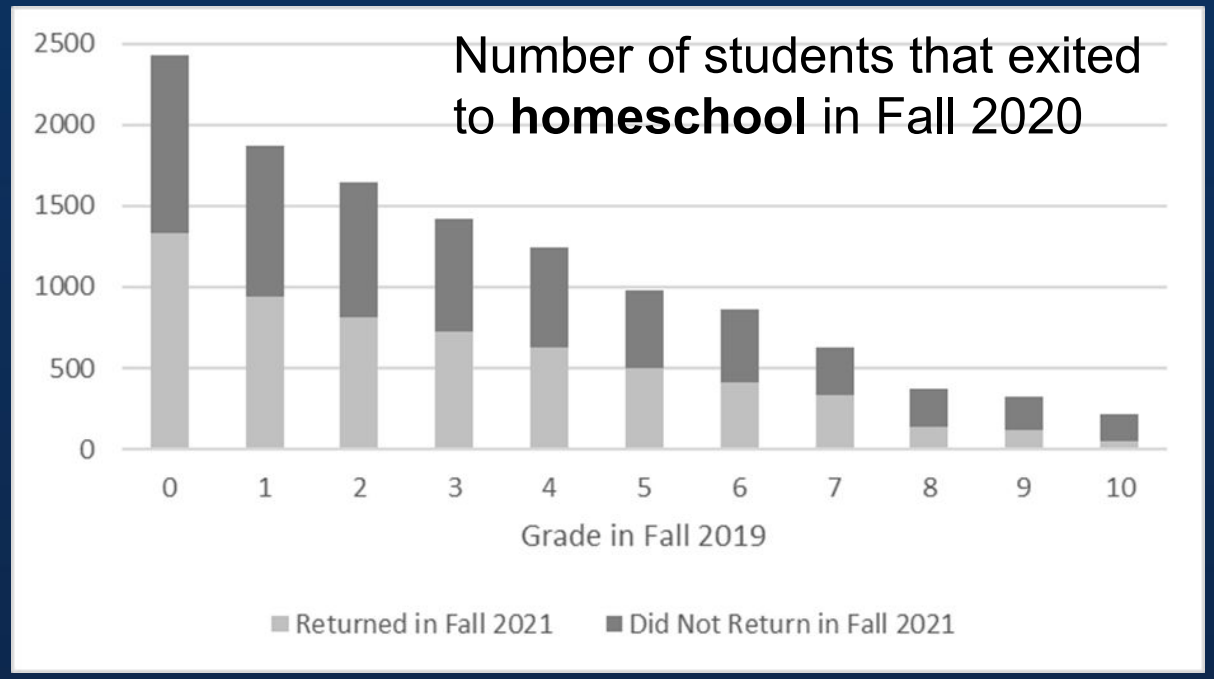
Exit Rate from Michigan Public Schools

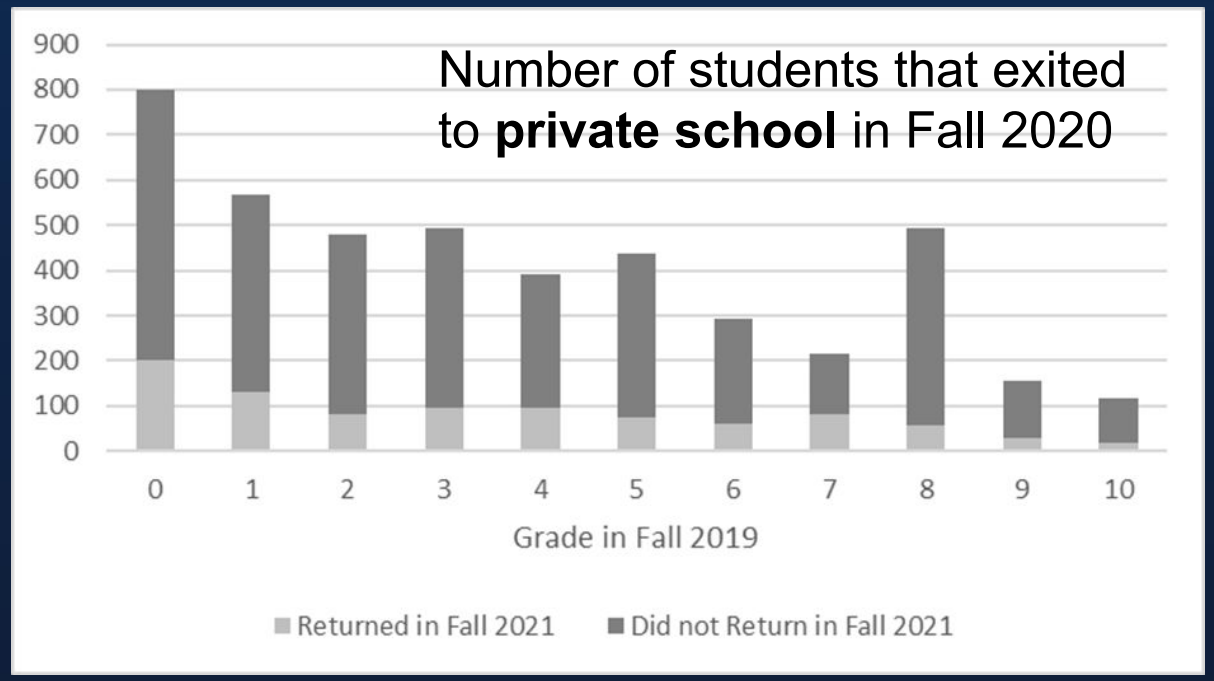


Exit rates increased substantially in younger grades

- Two-thirds of the increase in exits is explained by homeschooling and private schooling.
- We see big increases in both types of exits, but especially to homeschooling.
- Analysis relies on longitudinal data linking students over time

Source: Bacher-Hicks, Musaddiq, Goodman & Stange (2023)



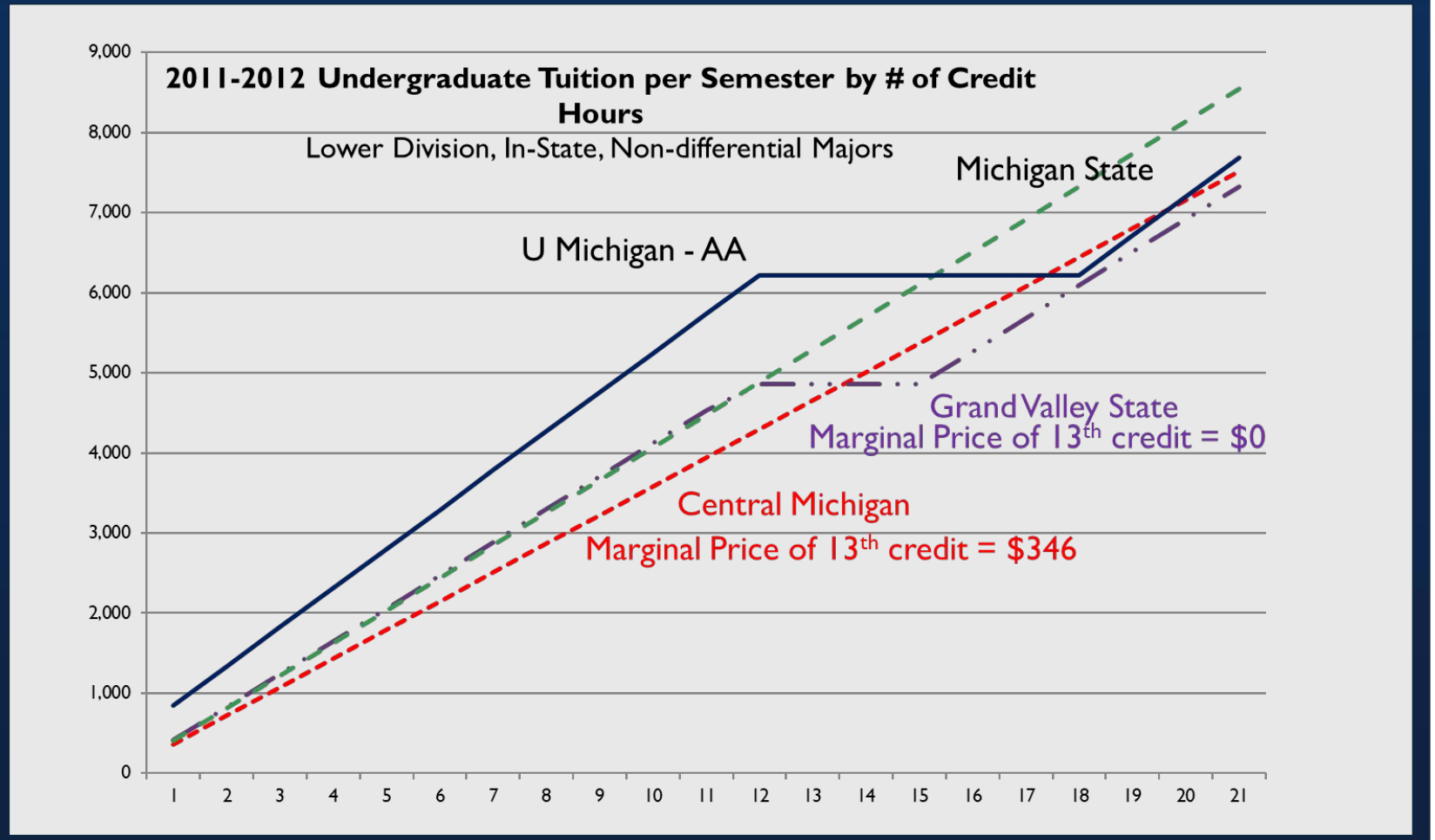


Reentry in Fall 2021 differed by destination

- Among those who exited for homeschooling in 2020, half returned in 2021
- Among those who exited for private schooling in 2020, most did not return in 2021
- Pandemic-driven school enrollment declines may persist, particularly among higher-income families

Source: Bacher-Hicks, Musaddiq, Goodman & Stange (2023)

“Flat Pricing” Alters Tuition Schedule



Half of Michigan’s 15 4-year Colleges have “flat” pricing

- No additional tuition for credits taken above 12
- The rest charge per credit
- Creates incentive to take more credits per term

- Does “flat” pricing increase credits attempted & earned?
- …speed up degree completion?
- … alter mix of courses taken?

Source: Hemelt & Stange (2015)

Final sample

- Limit main analysis to full-time students (i.e., >= 12 credits):
  - N = 107,633 students and 194,375 student*semester observations across all four cohorts
  - Equals about 1.8 semesters per student

Information used

- Demographics (gender, race, ethnicity, Free and Reduced Price Lunch status (marker for income), limited English, special education)
- 11th grade achievement (math, reading, science, social studies, and writing scores) and ACT (composite and subject)
- Matched to college enrollment spells from National Student Clearinghouse (NSC)
- Full historical transcripts (course-level data) and credit accumulation from Student Transcript and Academic Record Repository (STARR)

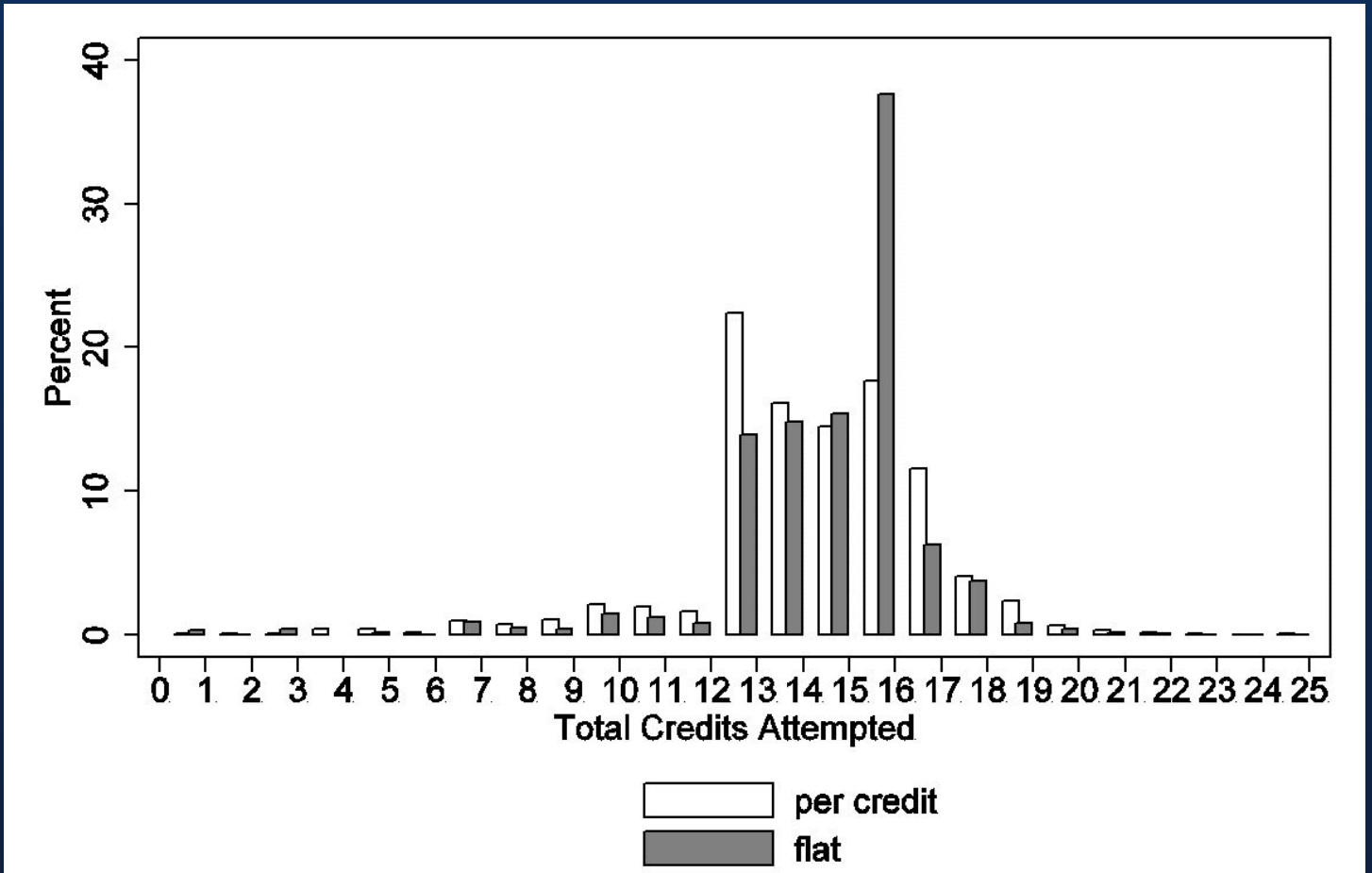
Data sources and Sample

- All MI public high school students who graduated in 2008, 2009, 2010, or 2011

- Keep students who attended a Michigan public 4-year university during 2011/12 academic year (have transcript data)

Source: Hemelt & Stange (2015)

Distribution of Credits Attempted, by Pricing Schedule



Notes: Data are student transcripts from State of Michigan Student Transcript and Academic Record Repository (STARR), for academic year 2011-2012.

Flat pricing shifts distribution

- Flat pricing associated with much higher share of students taking 15 credits
- Per-credit pricing associated with higher share of students taking 12 credits
- Worry about differences in students attending these two types of schools

Source: Hemelt & Stange (2015)

$$Y_{ijct} = \alpha + \beta_1 Flat_j + \beta_2 X_{ijt} + \beta_3 Z_j + \delta_t + \theta_c + \varepsilon_{isct}$$

(i = student, j = school, t = term, c = cohort)

	Outcome = Attempt 13 or more credits			
	(1)	(2)	(3)	(4)
All cohorts, all Fall and Spring terms (max n = 447,216)	0.093** (0.034)	0.073*** (0.023)	0.055** (0.026)	0.067** (0.026)
All cohorts, only 2011-2012 AY (max n = 195,829)	0.101*** (0.034)	0.083*** (0.023)	0.064** (0.025)	0.074** (0.026)
Only class of 2011, only 2011-2012 AY (max n = 52,193)	0.092** (0.037)	0.075** (0.028)	0.058* (0.028)	0.073** (0.028)
Only class of 2010, only 2011-2012 AY (max n = 48,069)		0.075*** (0.025)	0.056** (0.026)	0.065** (0.027)
Only class of 2009, only 2011-2012 AY (max n = 48,803)		0.094*** (0.024)	0.073** (0.024)	0.081*** (0.025)
Only class of 2008, only 2011-2012 AY (max n = 46,758)		0.089*** (0.027)	0.071** (0.031)	0.075** (0.033)
Student controls?	No	Yes	Yes	Yes
Institution controls?	None	None	ACT composite	ACT composite
Sample	All schools	All schools	All schools	Exclude UM-AA

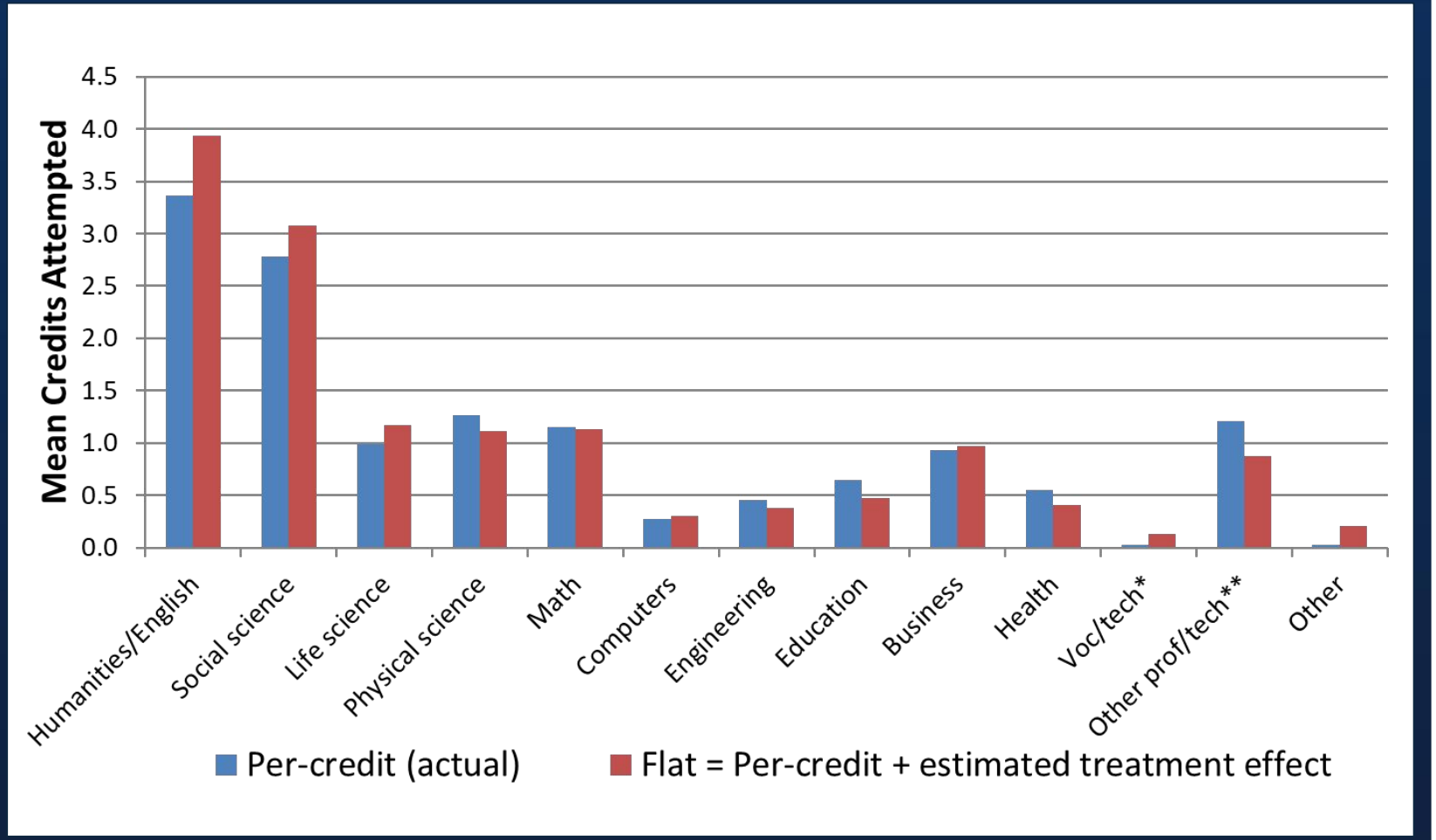
Notes: Each cell reports the coefficient on indicator for "Flat Pricing" from a separate regression. Sample is restricted to full-time students. All models include indicators for each unique term (e.g. Fall 2011). Individual controls include dummies for female, black, Hispanic, other race, LEP and FARM and composite ACT score. Specifications that pool multiple cohorts also include cohort fixed effects. Robust standard errors clustered at the college level appear in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Regression-based Estimates are Similar

- Compare credits attempted/earned (Y_ijct) by students attending "flat" pricing schools (marginal price = 0) to students attending per-credit pricing schools using OLS
- Control for rich measures of achievement and demographics (X_ijt), semester FE (δ_t), cohort FE (θ_c) and limited fixed institution characteristics (Z_j)
- Flat pricing is associated with 6-8 percentage pt increase in taking more than 12 credits

Source: Hemelt & Stange (2015)

Effects on Mean Credits Attempted by Subject



Notes: Per-credit mean is for all cohorts during 2011-2012 academic year. Flat (counterfactual) mean is per-credit mean plus estimated effect of flat pricing on average credits taken in subject. Model includes indicators for each unique term (e.g. Fall 2011), individual controls, institution-level ACT score, and cohort fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Additional credits primarily in Humanities and Social sciences

- Differences are not statistically significant
- May suggest that flat pricing is not effective way to get more students to take STEM courses

Source: Hemelt & Stange (2015)

Potential challenges for users

Funding

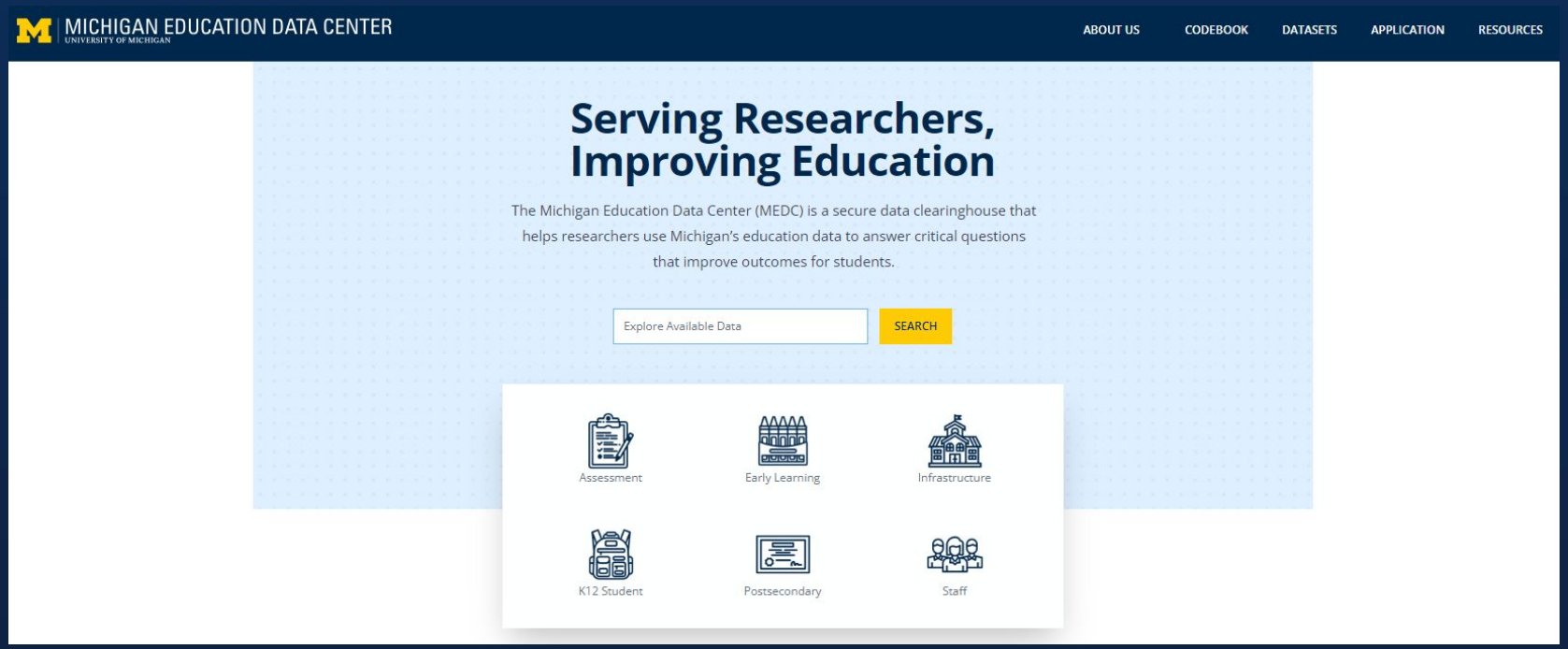
AERA has research grant for SLDS: https://www.aera.net/Professional-Opportunities-Funding/AERA-Funding-Opportunities/AERA-NSF-Grants-Program (due May 30th last year)

IRB

Data security requirements

Output review (30 day)

https://medc.miedresearch.org/



Questions?