

# Getting Down to Facts on Education Data Systems in California

May 10, 2019  
Sacramento, California



# What is Getting Down to Facts II?

National collaborative research project on California's PreK-12 education system including more than 100 researchers across the country.

- Sequel to the first GDTF released 10 years ago
- Input from multiple stakeholders: the public, teachers, principals, CBOs, superintendents (county and district), policy leaders
- 36 research studies, 19 research briefs and a summary paper

# Areas Covered



*Student  
Success*

*Governance*



*Personnel*

*Finance*



# Key Findings from Getting Down to Facts II

- California schools and students have been moving in the right direction.
- Great need remains for policies to address system weakness and build capacity.
- Specifically, areas for California to focus on:
  - Building on current reforms
  - Increasing funding and fixing systems
  - Addressing achievement gaps

# Evan White, JD, MPP

- Executive Director of the California Policy Lab at UC Berkeley
- Served as Senior Advisor to the the Consumer Financial Protection Bureau
- Previous Director of Fair Housing for Project Sentinel
- Clerked for the California Supreme Court, the Civil Rights Division of the Department of Justice, and the White House Office of Management and Budget



# Jesse Rothstein, PhD

- Professor of Public Policy and Economics, University of California, Berkeley
- Director, Institute for Research on Labor and Employment
- Director, California Policy Lab
- Co-Director, Opportunity Lab, University of California, Berkeley
- Research focuses on education policy and on the labor market



# Agenda

- Presentation, “Improving California’s Education Data Systems”
- Q&A

# Improving California's Education Data Systems

Jesse Rothstein

Evan White

May 2019



**Berkeley**  
UNIVERSITY OF CALIFORNIA

**UCLA**

**Building on:**

**Making California Data More Useful for Educational Improvement**

**Getting Down To Facts report**

Meredith Phillips, UCLA

Sarah Reber, UCLA

Jesse Rothstein, UC Berkeley



# Key findings

- California has improved its data systems, through CALPADS in particular.
- Data availability and usefulness still **fall short** of other states.
  - Data systems are not integrated.
- California has made a big investment in data systems, and with some additional investment, these **data could be made more widely available and more useful**.
- Developing a comprehensive statewide data system requires **leadership, commitment, and planning**.

**California lags behind.**

**States well ahead of us include:**

- Tennessee
- North Carolina
- Texas
- Florida

# California needs a statewide data system

## Linkages within and between sectors would improve service delivery and policymaking

### Evaluate state policies and programs

- Long term impact of early childhood policy initiatives (Dodge, Bai, Ladd, and Muschkin 2017)
- Impact of statewide curricular changes (Jacob, Dynarski, Frank, and Schneider 2017)

### Connect academic and human services data

- School staff can understand students' mental health/social welfare involvement
- Ensure that college students get CalFresh benefits they deserve.

Development of early warning and on-track indicators (Allensworth 2013) and identifying bright spots (Cannata, Smith, and Haynes 2017)

### Many possible users

- Teachers
- School and district leaders
- CDE
- Higher education institutions and segments
- Social service, employment, and law enforcement agencies
- Researchers inside and outside government

# Barriers to linked data are surmountable

## Technical

- Lack of a common identifier
- Privacy concerns
- Some agencies don't have the capacity to link data even if they want to

## Bureaucratic

- Each agency owns its own data
- Fragmented governance – no overarching authority

## Political

- The teacher & accountability wars
- Concerns about immigration enforcement
- Fear about what research will show

→ Building the data isn't enough – we also need to figure out ways to ensure that the research policymakers need gets done.

# Four near-term steps suggested in our GDTF report

1. Expand data linkages, built on the existing infrastructure
2. Champion integration through strong political and agency leadership, both by articulating the immense value of integration and convening stakeholders to build a collaborative system
3. Establish a governance structure to develop rules on access, use, privacy, security, and IT infrastructure
4. Build infrastructure to support linked data systems, and capacity to use them

# Current efforts to expand statewide data linking

- Governor's budget and trailer bill
  - \$10 million for planning and development of a longitudinal data system
  - Would include P20W plus health and human services
  - Shorter timelines
- Assembly Bill 1466 (Irwin) and Senate Bill 2 (Glazer)
  - Would create education-focused data systems
  - SB2 is linked to higher ed accountability organization
  - Longer timelines

# Open questions about design and implementation

- Where will it live?
  - New CPEC? A new office?
- How frequently will it be updated?
  - “Fresh” or “frozen”?
- What data would be included?
  - P20W? HHS? CJ? Build to grow?
- What governance?
  - Who has access? And for what uses?
  - Approval process
  - Dissemination review process
  - Costs to users, if any

# A Data Roadmap for California

## A Roadmap for Linking Administrative Data in California



### A long-run vision for state data

California is the home of Silicon Valley but the state's data systems lag surprisingly far behind. Forty-four states systematically link student-level postsecondary data to K-12 or workforce data, but California does not.

Recent calls to create a "P-20W" or "cradle-to-career" longitudinal data system in California have sparked interest in better connecting the state's administrative data. These efforts are heartening, given the value that such data could provide state policymakers. **But a narrow focus on educational data risks missing the bigger picture.** Education intersects with other domains, and student success is bolstered by healthy families, safe neighborhoods, and economic well-being. The state needs a long-term vision for making better use of all of the state's data for improving policy outcomes. Several recent reports by expert panels emphasize the critical role that linked data could play in other areas, like anti-poverty programs and criminal justice.<sup>4</sup>

California is the fifth largest economy in the world. It needs data systems that help the state understand its residents and serve them better. We propose here a roadmap toward that goal: (1) **create a new, independent office** with the mandate and expertise to link data across siloes, (2) **sequence the linkage process** by starting with education and expanding outward, and (3) **establish streamlined governance** that makes data available to improve state policies and programs.

### The case for linked data

Most complex policy problems cut across sectors. Homelessness, for example, intersects housing, health, and criminal justice. And educational achievement is driven not only by teacher and school quality, but also family well-being, neighborhood factors, and student health.

California does not have the tools necessary to address these complex problems because its data is siloed in ways that obscure good information about potential causes and solutions. **We cannot solve a Rubik's Cube when we see only one side.**

We currently cannot answer simple questions such as:

- Which California high schoolers are prepared for college?
- How many families are claiming the safety net benefits for which they're eligible, and what consequences does this have for their children's life outcomes?
- Which programs successfully help formerly incarcerated individuals reintegrate into society?

Linking California's data assets could help answer these questions and lift tens of thousands out of poverty, improve our schools, and save the state tens of millions of dollars through program improvements. These data could be used by researchers to inform policy, by households to inform school and neighborhood choice, and by policymakers to inform program improvements.

- **A new data linkage office must have all five vowels...**
  - **A**uthority to compel data
  - **E**xpertise in linking and curating data
  - **I**ndependence from data contributors
  - **O**pen source orientation
  - **U**nwavering commitment to privacy and security

# A Data Roadmap for California

## A Roadmap for Linking Administrative Data in California



### A long-run vision for state data

California is the home of Silicon Valley but the state's data systems lag surprisingly far behind. Forty-four states systematically link student-level postsecondary data to K-12 or workforce data, but California does not.

Recent calls to create a "P-20W" or "cradle-to-career" longitudinal data system in California have sparked interest in better connecting the state's administrative data. These efforts are heartening, given the value that such data could provide state policymakers. **But a narrow focus on educational data risks missing the bigger picture.** Education intersects with other domains, and student success is bolstered by healthy families, safe neighborhoods, and economic well-being. The state needs a long-term vision for making better use of all of the state's data for improving policy outcomes. Several recent reports by expert panels emphasize the critical role that linked data could play in other areas, like anti-poverty programs and criminal justice.<sup>4</sup>

California is the fifth largest economy in the world. It needs data systems that help the state understand its residents and serve them better. We propose here a roadmap toward that goal: (1) **create a new, independent office** with the mandate and expertise to link data across silos, (2) **sequence the linkage process** by starting with education and expanding outward, and (3) **establish streamlined governance** that makes data available to improve state policies and programs.

### The case for linked data

Most complex policy problems cut across sectors. Homelessness, for example, intersects housing, health, and criminal justice. And educational achievement is driven not only by teacher and school quality, but also family well-being, neighborhood factors, and student health.

California does not have the tools necessary to address these complex problems because its data is siloed in ways that obscure good information about potential causes and solutions. **We cannot solve a Rubik's Cube when we see only one side.**

We currently cannot answer simple questions such as:

- Which California high schoolers are prepared for college?
- How many families are claiming the safety net benefits for which they're eligible, and what consequences does this have for their children's life outcomes?
- Which programs successfully help formerly incarcerated individuals reintegrate into society?

Linking California's data assets could help answer these questions and lift tens of thousands out of poverty, improve our schools, and save the state tens of millions of dollars through program improvements. These data could be used by researchers to inform policy, by households to inform school and neighborhood choice, and by policymakers to inform program improvements.

## Education and workforce

- Higher ed segments
- Quarterly earnings
- K12 (CALPADS)
- Cal Grants & FAFSA

## Health and Human Services

- Birth and death records
- Medicaid claims
- Hospital discharge
- Safety net participation

## Sequence for success

- Start with education, but don't end there

## Backbone Data

- DMV
- Tax records

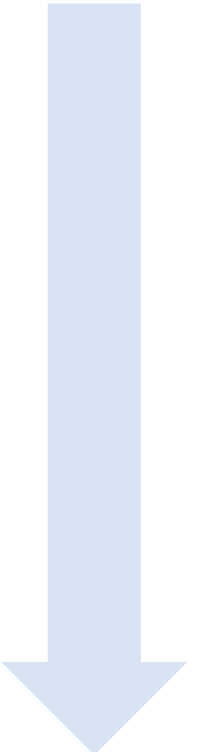
## Criminal Justice

- Criminal history
- Prison and parole

## Early Childhood

## Future Linkages

- Local and regional data
- Housing, transportation, environment





# A Data Roadmap for California

## A Roadmap for Linking Administrative Data in California



### A long-run vision for state data

California is the home of Silicon Valley but the state's data systems lag surprisingly far behind. Forty-four states systematically link student-level postsecondary data to K-12 or workforce data, but California does not.

Recent calls to create a "P-20W" or "cradle-to-career" longitudinal data system in California have sparked interest in better connecting the state's administrative data. These efforts are heartening, given the value that such data could provide state policymakers. **But a narrow focus on educational data risks missing the bigger picture.** Education intersects with other domains, and student success is bolstered by healthy families, safe neighborhoods, and economic well-being. The state needs a long-term vision for making better use of all of the state's data for improving policy outcomes. Several recent reports by expert panels emphasize the critical role that linked data could play in other areas, like anti-poverty programs and criminal justice.<sup>4</sup>

California is the fifth largest economy in the world. It needs data systems that help the state understand its residents and serve them better. We propose here a roadmap toward that goal: (1) **create a new, independent office** with the mandate and expertise to link data across silos, (2) **sequence the linkage process** by starting with education and expanding outward, and (3) **establish streamlined governance** that makes data available to improve state policies and programs.

### The case for linked data

Most complex policy problems cut across sectors. Homelessness, for example, intersects housing, health, and criminal justice. And educational achievement is driven not only by teacher and school quality, but also family well-being, neighborhood factors, and student health.

California does not have the tools necessary to address these complex problems because its data is siloed in ways that obscure good information about potential causes and solutions. **We cannot solve a Rubik's Cube when we see only one side.**

We currently cannot answer simple questions such as:

- Which California high schoolers are prepared for college?
- How many families are claiming the safety net benefits for which they're eligible, and what consequences does this have for their children's life outcomes?
- Which programs successfully help formerly incarcerated individuals reintegrate into society?

Linking California's data assets could help answer these questions and lift tens of thousands out of poverty, improve our schools, and save the state tens of millions of dollars through program improvements. These data could be used by researchers to inform policy, by households to inform school and neighborhood choice, and by policymakers to inform program improvements.

- **Depoliticize data governance**
  - Learn from others' mistakes
  - Limit vetoes
  - Presumption in favor of use
  - Empower subject matter experts
  - "Get to yes" on privacy and security

# Thanks!

---



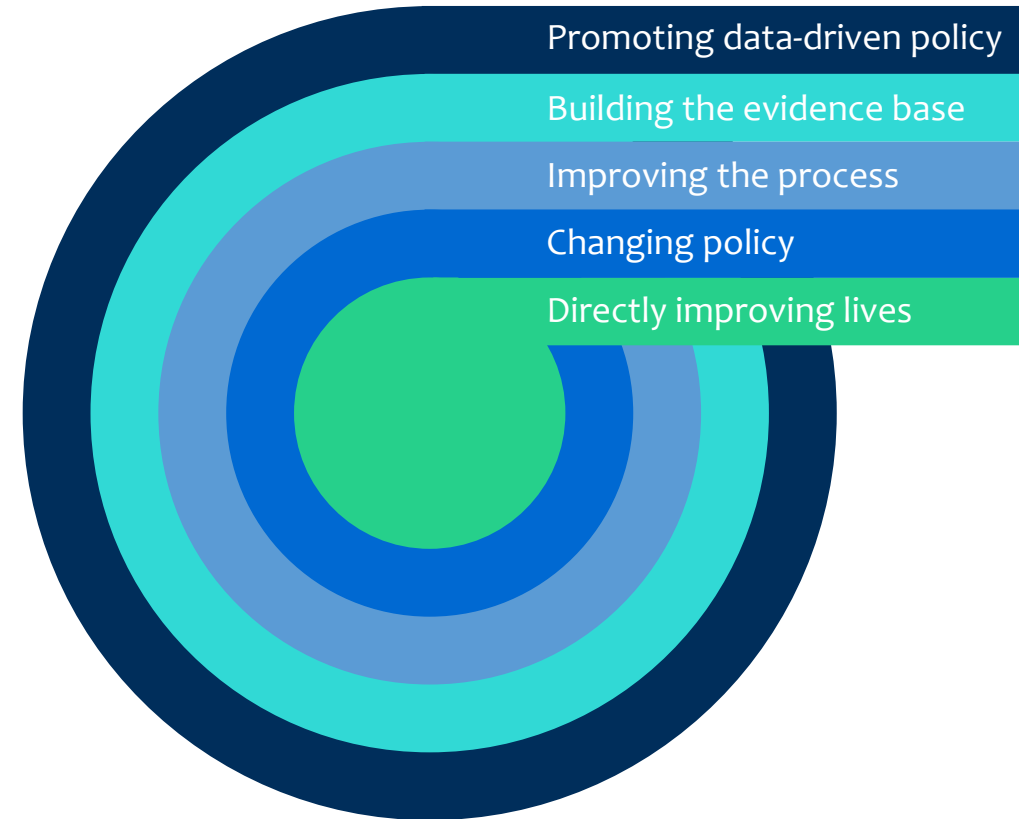
Contact us at [rothstein@berkeley.edu](mailto:rothstein@berkeley.edu) or [evan@capolicylab.org](mailto:evan@capolicylab.org)



Sign up for our newsletter at [www.capolicylab.org](http://www.capolicylab.org)



Donate to support our work at [www.capolicylab.org/donate](http://www.capolicylab.org/donate)



Questions?



# Upcoming PACE Seminar

## The Challenges of Employee and Retiree Health Benefit Costs for California Districts

Tuesday, June 11, 10:00-11:00am

Location: Library Galleria



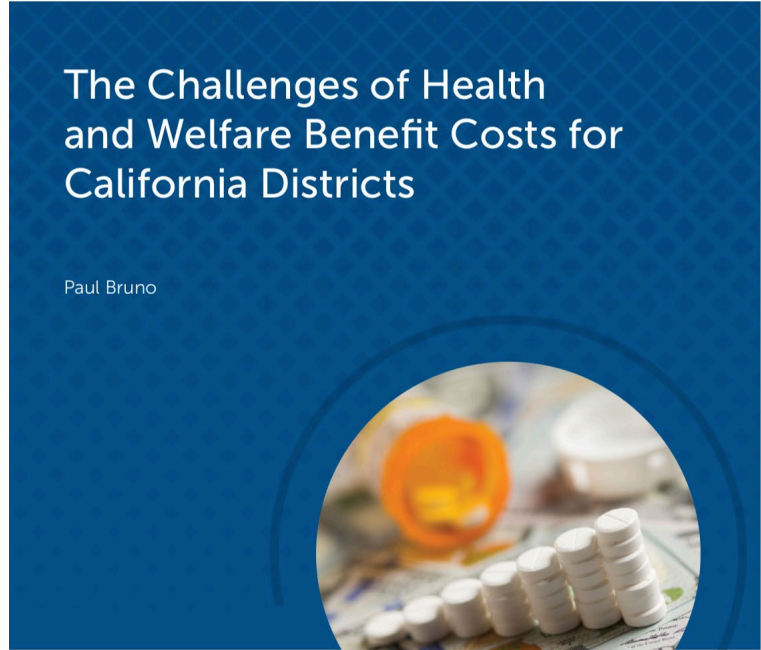
**Paul Bruno**  
University of Southern  
California Rossier  
School of Education



**Superintendent  
Jorge Aguilar**  
Sacramento City  
Unified School District



**Michael Fine**  
Fiscal Crisis and  
Management Assistance  
Team



May 2019

