Evidence-Based Policy and Its Limits

April 16, 2015 | Child and Family Research Partnership
The University of Texas at Austin | LBJ School of Public Affairs

CHILD & FAMILY
RESEARCH PARTNERSHIP
THE UNIVERSITY OF TEXAS AT AUSTIN | LBJ SCHOOL OF PUBLIC AFFAIRS

Dr. Ron Haskins

Brookings Institution

Co-Director, Center on Children and Families, Budgeting for National Priorities
Senior Fellow, Economic Studies
The Cabot Family Chair

BROOKINGS | QUALITY. INDEPENDENCE. IMPACT.

The Obama Evidence-Based Revolution

Ron Haskins
Cabot Family Chair & Co-Director, Center on Children & Families
The Brookings Institution

LBJ School of Public Affairs
University of Texas at Austin
April 16, 2015

Two Goals of Obama's Evidence-Based Strategy for Grant Making

- Spend most federal grant dollars on evidence-based programs (some spending on innovative programs)
- Continuous evaluation

Two Parts of Obama's Evidence-Based Approach

- Federal Agencies (PART)
- Federal Grants

How to Identify Evidence-Based Programs

- Identified by federal agencies
- Identified by grant applicants

Type of Evidence Required in Six Obama Evidence-Based Initiatives

Initiative	*Statutory requirement*	*Agency evidence review or funding opportunity announcement requirement*
Home Visiting	"Significant" "positive outcomes" from "randomized controlled" or "quasi-experimental" research designs	At least one high- or moderate-quality impact evaluation
Teen Pregnancy Prevention	"Proven effective through rigorous evaluation"	At least one high-or moderate-quality impact evaluation
Investing in Innovation (i3)	Language about effects that programs must have demonstrated in order to be considered eligible	Evidence from randomized controlled trials or quasi-experimental designs
Social Innovation Fund (SIF)	Evidence from "rigorous evaluations of program effectiveness"	Must have evidence of impact; definitions of evidence are consistent with those for i3
Workforce Innovation Fund (WIF)	"Evidence-based strategies," but type of study not defined	Three tiers of grants, each tied to specific level of evidence
Trade Adjustment Assistance Community College and Career Training (TAACCCT)	No language on evidence	Strong evidence, defined as a "study or multiple studies whose designs can support strong, causal conclusions"

Overview of Six Evidence-Based Initiatives

Evidence-based initiative	*Initial funding*	*Administering agency*	*Date of first awards*
Teen Pregnancy Prevention (TPP)*	$110 million	HHS	September 2010
Maternal, Infant, and Early Childhood Home Visiting*	$1.5 billion	HHS	July 2010
Investing in Innovation (i3)	$650 million	Department of Education	August 2010
Social Innovation Fund (SIF)	$50 million	Corporation for National and Community Service	July 2010
Trade Adjustment Assistance Community College and Career Training (TAACCCT)	$2 billion	Department of Labor	September 2011
Workforce Innovation Fund (WIF)	$125 million	Department of Labor	June 2012

*HHS identified evidence-based programs

Obama Evidence-Based Initiatives: Number of Local Projects

Initiative	Number of projects
Teen Pregnancy	102
Home Visiting	774
Investing in Innovation	117
Social Innovation Fund	221
Workforce Innovation Fund	26
TAACCCT	185
Total	**1,425**

Issues with RCTs

- Causal density
- Modest impacts
- Expense: time & money
- Are RCTs ethical?

Why the Obama Evidence-Based Initiative Worked

- Stellar leadership
- Relentless focus on using evidence
- Clever and persistent legislative strategies
- Competitive (not formula) grants
- Decent review panels

Signs of Success

- Initiatives survive the 2015 and 2017 transitions
- Culture of evidence is established
- Early positive outcomes are achieved
- In case of failure, programs improve based on evidence

Major Issues

- Role of RCTs
- When a program fails, what's next
- Innovation vs. evidence
- Implementation
- Do we have good model programs?

Dr. Cynthia Osborne

LBJ School of Public Affairs | University of Texas at Austin

Associate Professor of Public Affairs
Director of the Child and Family Research Partnership

Potential Risks of Evidence-Based Policy Making: A Case Study from Home Visiting

#evidencebased

CHILD & FAMILY
RESEARCH PARTNERSHIP
THE UNIVERSITY OF TEXAS AT AUSTIN | LBJ SCHOOL OF PUBLIC AFFAIRS

Evidence-based policymaking is the right approach

- Social scientists have been waiting for this!
- Home visiting has embraced this approach more than other programs
- Evidence base of home visiting is stronger and more rigorous than most areas of social policy or early childhood interventions
- **BUT….potential risks threaten its success**

Potential Risk #1

We believe ALL programs work for ALL families on ALL outcomes

- Not all programs have measured or shown impact on all outcomes

	EHS-HB	HIPPY	NFP	PAT
Maternal and Newborn Health	No Effect	Not Measured	Favorable	No Effect
Prevention of Child Injuries, Child Abuse, Neglect, or Maltreatment and Reduction of Emergency Department Visits	Not Measured	Not Measured	Favorable	Favorable
Improvement in School Readiness and Achievement	Favorable	Favorable	Favorable	Favorable
Reduction in Crime or Domestic Violence	Not Measured	Not Measured	Favorable	Not Measured
Improvements in Family Economic Self-Sufficiency	Favorable	Not Measured	Favorable	Favorable
Improvements in the Coordination and Referrals for Other Community Resources and Supports	Favorable	Not Measured	No Effect	Not Measured

Source: US Department of Health and Human Services (2015). Home Visiting Evidence of Effectiveness (HomVee).

#evidencebased

Potential Risk #1

We believe ALL programs work for ALL families on ALL outcomes

- Not all programs have measured or shown impact on all outcomes
- Choose a program that has demonstrated impacts for the outcomes you aim to change

Potential Risk #2

We set unrealistic expectations for the programs

- Home visiting programs DO work - and *most* of the evidence is rigorous
- But the results are modest and many results apply only to the hardest-to-serve populations

Evidence-based home visiting programs have "been critical in improving maternal and child health outcomes in the early years, leaving long-lasting, positive impacts on parenting skills, children's cognitive, language, and social-emotional development and school readiness."

- President Obama's Plan for Early Education for All Americans (White House, 2013)

"States should invest in quality evidence-based home visiting programs,...Fostering positive parenting skills and family responsibility and health today sows the seeds for safer, healthier children who are better prepared to learn tomorrow. And tax payers reap the benefit when many of our nation's costliest social problems—school failure, child abuse, and welfare dependence—are prevented."

- Pew Charitable Trusts,
The Case for Home Visiting (2010)

"I think it's fantastic, that kind of home visitation, it really does work. It's very conservative; it does things that make differences in what is going to happen in our public over a long period of time. That we are making a difference—a big difference maker for the state of Texas, for the future. So that 20 years from now and 40 years from now we will have a population that is even less likely and less prone to drop out of school. They'll be much more educated, much more highly educated, because of the programs they are doing."

- Honorable Texas State Representative Jerry Madden (R) in The Case for Home Visiting (Pew Charitable Trusts,2011)

"The visits have been studied extensively through randomized controlled trials – the gold standard of evidence – *and are stunningly effective*...we have an anti-poverty program that is cheap, is backed by rigorous evidence, and *pays for itself several times over* in reduced costs later on."

- New York Times columnist, Nicholas Kristoff, *The Way to Beat Poverty* (Sept. 2014)

Prenatal Care

Program Model Evaluated	Measure	Comparison	Program
EHS	Percentage receiving prenatal care services during their pregnancy	N/A	95%
EHS	Percentage receiving prenatal care during the first trimester	N/A	82%
NFP	**Attending childbirth classes during pregnancy**	**54%**	**70%**
NFP	**Knowledge on the number of prenatal care services**	**4.9**	**5.5**
NFP	Number of prenatal visits	10.5	10.5
HFA	Received any prenatal care	N/A	94%
HFA	Received early prenatal care	N/A	75%

Note. Findings in bold and highlighted in blue represent a statistically significant difference between the program and comparison groups

Breastfeeding

Program Model Evaluated	Measure	Comparison	Program
EHS	Ever-breastfed rate	N/A	59%
NFP	**Ever-breastfed rate**	**16%**	**26%**
HFA	Ever-breastfed rate	45%	46%
HFA	Length of breastfeeding (months)	1.04	1.01

Note. Findings in bold and highlighted in blue represent a statistically significant difference between the program and comparison groups

Well-Child Visits/Immunizations

Program Model Evaluated	Measure	Comparison	Program
EHS	Receiving any immunizations at age two	98.2%	98.2%
EHS	Receiving any immunizations at age three	98.5%	99.2%
NFP	Number of well-child visits	4.8	4.6
NFP	Percentage of children who had current immunizations	68%	70%
HFA	Adequate well-child visits at age two	8%	4%
HFA	Number of well-child visits at age one	4.54	4.61
HFA	Up-to-date immunizations at age one	82%	82%
HFA	Up-to-date immunizations at age two (Hawaii trial)	87%	85%
HFA	Up-to-date immunizations at age two (Alaska trial)	27%	27%
HFA	Number of well-child visits at age three	1.9	2.4
HFA	Up-to-date immunizations at age three	82.4%	84%
PAT	Percentage of children who had current immunizations	65%	56%
PAT	**Fully-immunized for his/her age (3-year follow-up)**	**8%**	**40%**

Note. Findings in bold and highlighted in blue represent a statistically significant difference between the program and comparison groups

#evidencebased

Learning Support

Program Model Evaluated	Measure	Comparison	Program
EHS	The HOME Inventory (Kindergarten)	35.2	33.7
EHS	Teaching activities (Kindergarten)	10.8	11.3
EHS	Reading daily (Kindergarten)	27.3%	35.1%
EHS	The HOME Inventory: Language & Literacy (age two)	10.1	10.3
EHS	The HOME Inventory: Language & Literacy (age three)	10.7	10.9
EHS	The HOME Inventory: Language & Literacy (age five)	10.6	11.2
NFP	The HOME Inventory (age two)	30.9	32.3
NFP	The HOME Inventory Provisions of Appropriate Play Materials subscale (among the most disadvantaged at 10 months)	5.94	7.35
PAT	Reading aloud to child to child (4-point scale) at 1-year assessment (among the very low income group)	2.5	3.0
PAT	Tells stories, says nursery rhymes, sings with child (4-point scale) at 2-year assessment (among the very low income group)	2.9	3.4
HFA	The HOME Inventory (age one)	35.2	35.2
HFA	The HOME Inventory (age two)	34.1	34.6
HFA	Self-reported estimate of the time spent reading to the child on a weekly basis	2.72	2.46
HFA	The Nursing Child Assessment Satellite Teaching (NCAST) scale	11.9	11.8

Note. Findings in bold and highlighted in blue represent a statistically significant difference between the program and comparison groups

Child Maltreatment

Program Model Evaluated	Measure	Comparison	Program
EHS	Likelihood of an encounter with child welfare	OR = 0.64	
EHS	Number of encounters with child welfare btwn ages 5 & 9	B = -2.50	
EHS	Substantiated report of physical or sexual abuse	B = -1.24	
NFP	Substantiated abuse or neglect (2-year follow-up)	10%	5%
NFP	Rates of child abuse and neglect among the most disadvantaged group at age 2	.19	.04
NFP	Substantiated reports of child abuse and neglect (incidence; 15-year follow-up)	.54	.29
PAT	Opened cases of child abuse and neglect	2.4	0
HFA	Extreme physical abuse at 3-year follow-up (CTS-PC)	2%	4%
HFA	Minor physical assault at 3-year follow-up (CTS-PC)	86%	86%
HFA	Neglect at 3-year follow-up (CTS-PC)	27%	22%
HFA	Percent with a confirmed abuse or neglect (HPO subgroup) by age 7	19.3%	9.9%
HFA	Rate of confirmed CPS report for any abuse or neglect (RRO subgroup) by age 7	60.4%	41.5%
HFA	Number of total confirmed reports for mothers as confirmed subject (RRO subgroup) by age 7	1.6	0.8

Note. Findings in bold and highlighted in blue represent a statistically significant difference between the program and comparison groups

Potential Risk #2

We set unrealistic expectations for the programs

- Home visiting programs DO work - and *most* of the evidence is rigorous
- But the results are modest and many results apply only to the hardest-to-serve populations
- Home visiting is ***not*** a silver bullet. It is one important element of services to support families and children

Potential Risk #3

We forget that RCTs are a GUIDE, not a GUARANTEE

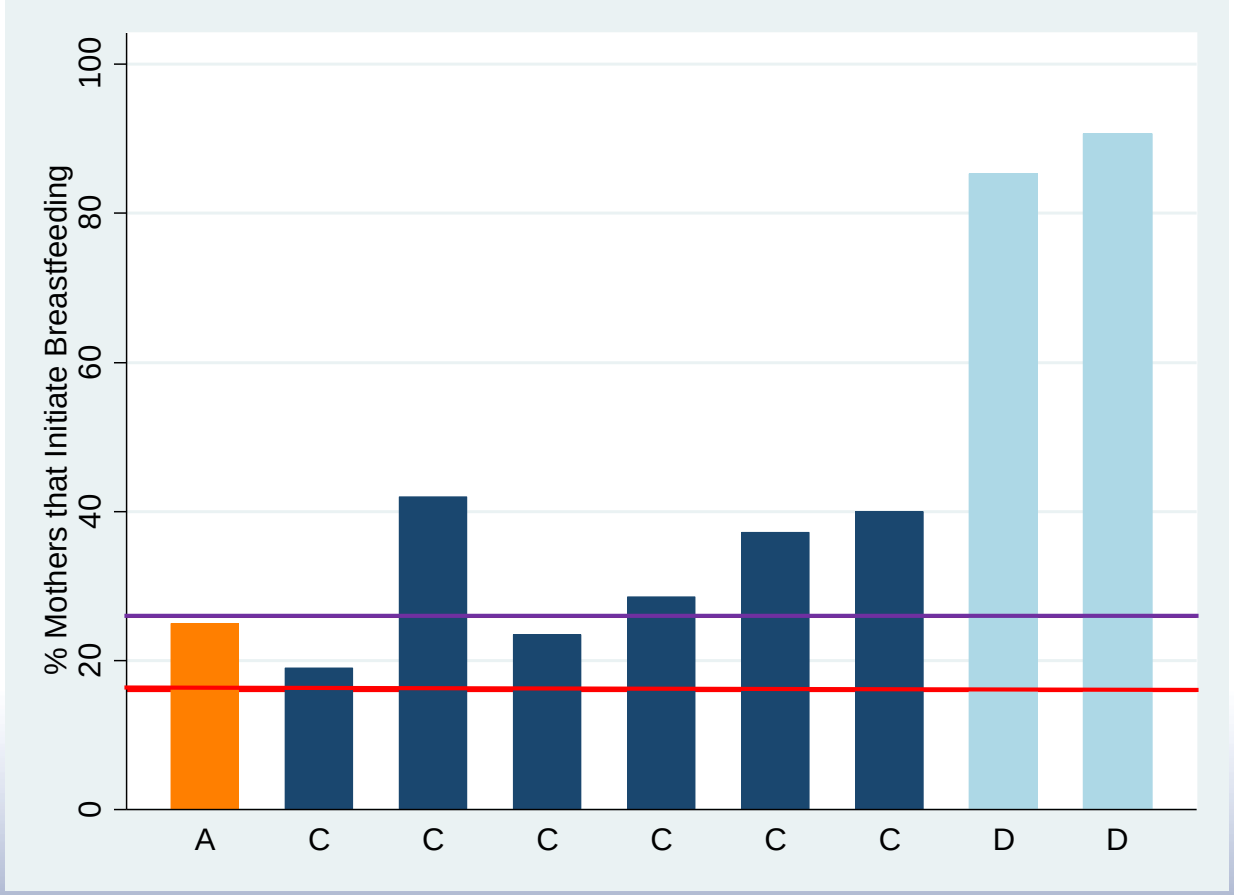
- Results among RCTs vary widely (even within a program model)
- Taking programs to scale will reveal even greater variation in outcomes

Variation is expected, given differences in:

- Motivation and capacity to implement a program
- Population characteristics of participants (e.g. geography, race/ethnicity, socioeconomic status)
- Fidelity to the model
- Data quality and measures

Program Model	Population Served	Race/ Ethnicity
EHS 1996-2001	1385 families in urban and rural areas in US	39% W 28% B 29% H
NFP 1978-1980	400 1st time mothers in Elmira, NY	90% W
NFP 1990-1991	1139 1st time mothers in Memphis, TN	92% B
NFP 1994-1995	735 1st time mothers in Denver, CO	46% H 36% W 15% B
PAT 1991-1993	717 teens in Southern CA	56% H 21% W 20% B
PAT 1992-1996	375 mothers in Northern CA	80% H

Percent Mothers Who Initiate Breastfeeding



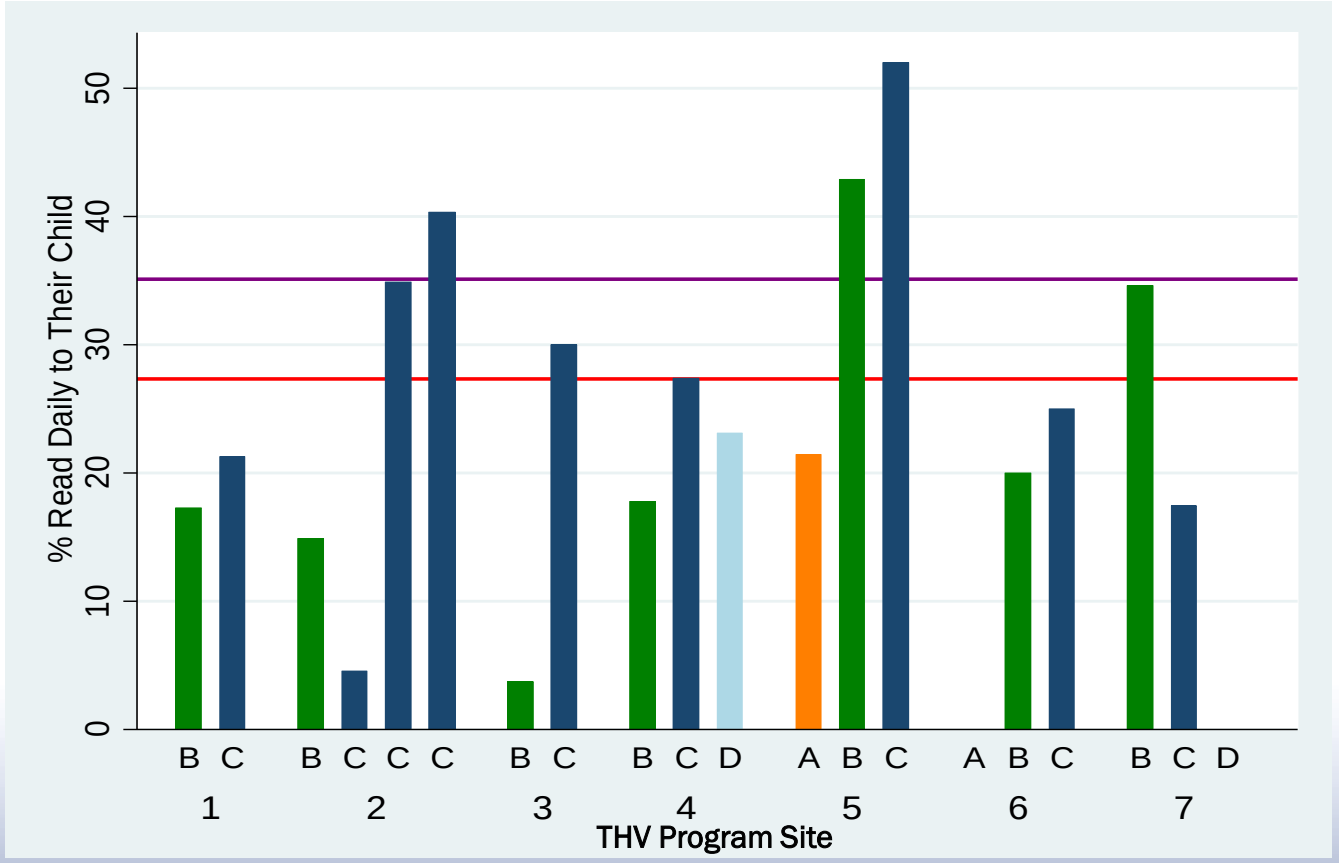
NFP Treatment Group: 26%

NFP Control Group: 16%

Sources: NFP Memphis RCT (Kitzman et al., 1997); THV Data Collection System. Notes: Data from THV programs restricted to sites where the denominator > 10

#evidencebased

Percent Mothers Reading Daily to Their Child



EHS Treatment Group: 35.1%

EHS Control Group: 27.3%

Sources: Early Head Start Research and Evaluation Project (Jones Harden, et al., 2012); THV Data Collection. System. Notes: Data from THV programs restricted to sites where the denominator > 10

#evidencebased

Potential Risk #3

We forget that RCTs are a GUIDE, not a GUARANTEE

- Results among RCTs vary widely (even within a program model)
- Taking programs to scale will reveal even greater variation in outcomes
- Learn more about **WHY** programs work to assist in broader implementation

Potential Risk #4

We neglect to innovate

- The programs on the approved list are not the only programs that work – they are the ones we have tested
- New challenges will arise (e.g. prenatal smoking v. obesity)
- Continue to evaluate the evidence-based programs –
  - Whether they work
  - Why they work
  - For whom do they work best
  - How can we improve

Potential Risks Are Worth It

Moving forward:

- Target interventions toward problems they can solve
- Set realistic expectations, and know the limitations
- Prepare for variation in outcomes, and learn why a program model works to enhance fidelity
- Innovate and improve

Dr. Paul von Hippel

LBJ School of Public Affairs | University of Texas at Austin

Assistant Professor of Public Affairs

Evidence-Based Policy: A Comparison of Government and Business

#evidencebased

CHILD & FAMILY
RESEARCH PARTNERSHIP
THE UNIVERSITY OF TEXAS AT AUSTIN | LBJ SCHOOL OF PUBLIC AFFAIRS

My Background: Business

- 2006-10
  - Data scientist
    - CheckFree (internet banking services)
    - JPMorgan Chase (checking and savings accounts)
  - Projects
    - Automated fraud detection
      - Internet billpay
      - Debit cards
      - Checks
    - Overdraft policies

My Background: Government Policy

- 2010-
  - Assistant Professor, LBJ School of Public Affairs
  - Program evaluations
    - Year-round school calendars
    - PEEQ
      - Ranking Teacher Preparation Programs (TPPs) effects on K-12 test scores
    - Texas Fitness Now
      - PE grants for high-poverty middle schools

Contrast #1: Willingness to Generalize

- Businesses are reluctant to generalize
  - "It worked at Bank of America."
    - "That doesn't mean it will work for us."
  - "It works in our established markets."
    - "That doesn't mean will work in new markets."
  - "It back-tested well in our 2013 data."
    - "That doesn't mean it will work today."
- Evidence is continually collected
  - Monitoring of established policies
  - Champion-challenger testing (A/B testing) of new ideas

Contrast #1: Willingness to Generalize

- Governments generalize all too willingly
  - The Perry Pre-School had great effects (123 children, Ypsilanti MI, 1960s)
    - Therefore Head Start is effective? (1 million children, nationwide, today)
  - CATCH reduced obesity in a 2005 nonrandomized trial in El Paso
    - but not in a 1994 randomized trial in Texas, California, Lousiana, and Minnesota
    - Today CATCH is the most popular anti-obesity curriculum in Texas
      - And other popular programs are similar
    - Existing sites should be evaluated

Contrast #2: Integration into Decision Making

- Businesses use evidence in decisions
  - Back-test policies in archival data
  - Prospective tests of policies in small random samples of the portfolio
  - Decisions
    - Adopt new policy, buy new service
    - Continue/discontinue old one
    - Fire someone

Contrast #2: Integration into Decision Making

- Governments also use evidence in decisions
- But often decision-making is uncoordinated from evidence gathering
  - Decision dates: annual budgets, biennial legislative sessions
  - Evaluation dates: program calendars, grant cycles, journal review decisions

Contrast #2: Integration into Decision Making

- Example 1: Texas Fitness Now
  - Texas Legislature authorized in 2007, 2009, terminated in 2011
    - $37 million
  - von Hippel & Bradbury evaluated in 2014
    - $25,000
    - Results under embargo till journal publication in 2015/16

- Example 2: Head Start Impact Study
  - Randomized evaluation
    - 1998: recommended
    - 2002: launched
    - 2006, 2008: 1st, 3rd grade data collection completed
    - 2008: draft report completed
    - 2010, 2012: public report released
  - Disappointing results. Little change to the program.

Contrast # 3: Time Horizon

- Businesses focused on short time horizon
  - Don't wait for default
    - Monitor delinquencies (late, missed payments)
  - Can be a weakness as well as a strength
    - Cuts to R&D
    - Mortgage balloon payments that postpone default
- Government, academia increasingly interested in long time horizon
  - E.g., preschool and kindergarten effects on adult wages
  - Interesting & important
    - But can't affect decisions today
  - How common are long-term "sleeper" effects when there are no short-term effects?

Dr. Carolyn Heinrich

LBJ School of Public Affairs | University of Texas at Austin

Sid Richardson Professor of Public Affairs
Director of the Center for Health and Social Policy

Value of Evidence

- **Having rigorous (experimental) evidence is critical, but we also need to have a healthy appreciation for and cultivate the generation and use of a variety of types of evidence**
  - *"We always viewed the random assignment design as the skeleton on which to build an evaluation that used multiple techniques to address a broader range of questions—including those raised by practitioners and managers, and by researchers who used different methods to diagnose the problems and understand why people behave as they do and how social programs work in practice… They [program funders and government partners] cared about why program participants and managers behaved as they did, why programs were or were not successfully implemented, and what could be done to improve the results." (Judy Gueron, former MDRC President, and Howard Rolston)*

Two-Way Research Partnerships

- Having an in-depth understanding of the context, issues and challenges that our practitioner partners face is critical not only to ensuring that the research we undertake will generate valuable knowledge that can be applied, but also to making sure we are asking the right or most important questions and grasping new opportunities for producing cutting-edge knowledge that can be truly transformational.
  - Minneapolis Public Schools Research and Evaluation Director, commenting on our research-to-practice collaborative: *"We are just going to eat up everything you can give to us!"*

CHILD & FAMILY RESEARCH PARTNERSHIP

The Child and Family Research Partnership (CFRP) is an independent, nonpartisan research group at the LBJ School of Public Affairs at The University of Texas at Austin, specializing in issues related to young children, teens, and their parents. We engage in rigorous research and evaluation work aimed at strengthening families and enhancing public policy.

childandfamilyresearch.org

@cfrplbj

facebook.com/cfrplbj

bit.ly/cfrplinkedin

Email List • bit.ly/cfrpnews

Jobs List • bit.ly/jobsatcfrp

#evidencebased