

Longitudinal Employer-Household Dynamics (LEHD) Program

Samuel R. Bondurant

Dallas-Fort Worth Federal Statistical Research Data Center
Center for Enterprise Dissemination
US Bureau of the Census

Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau.

Goals of this Presentation

- Give a comprehensive review of the LEHD and its various components
- Provide useful resources to access more info about the LEHD

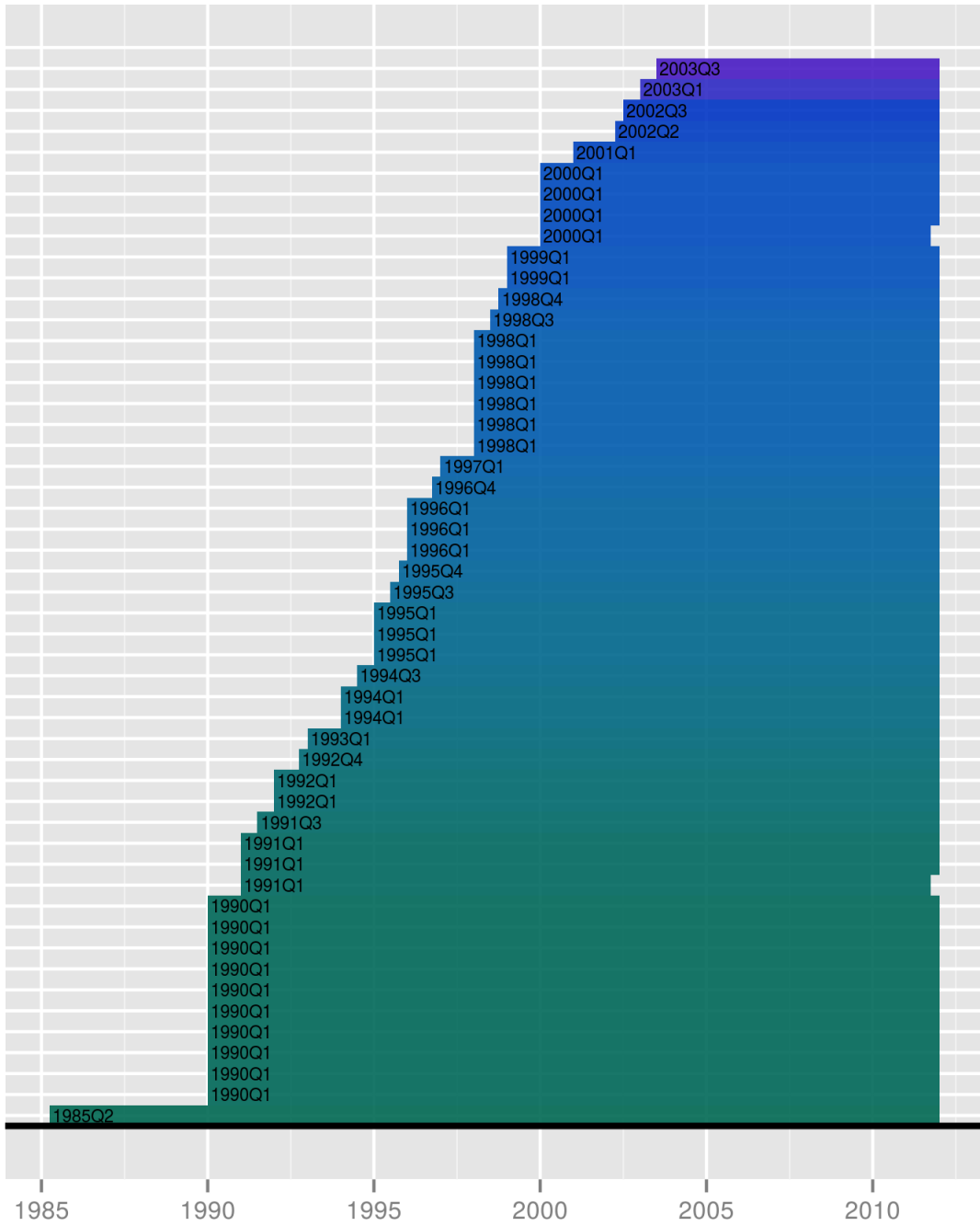
Census Working Papers on LEHD

- CES Discussion Paper 18-27R
 - Lars Vilhuber (2018)
 - <https://ideas.repec.org/p/cen/wpaper/18-27r.html>
- The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators
 - Abowd et al. (2006, 2009)
 - <https://www.nber.org/chapters/c0485.pdf>

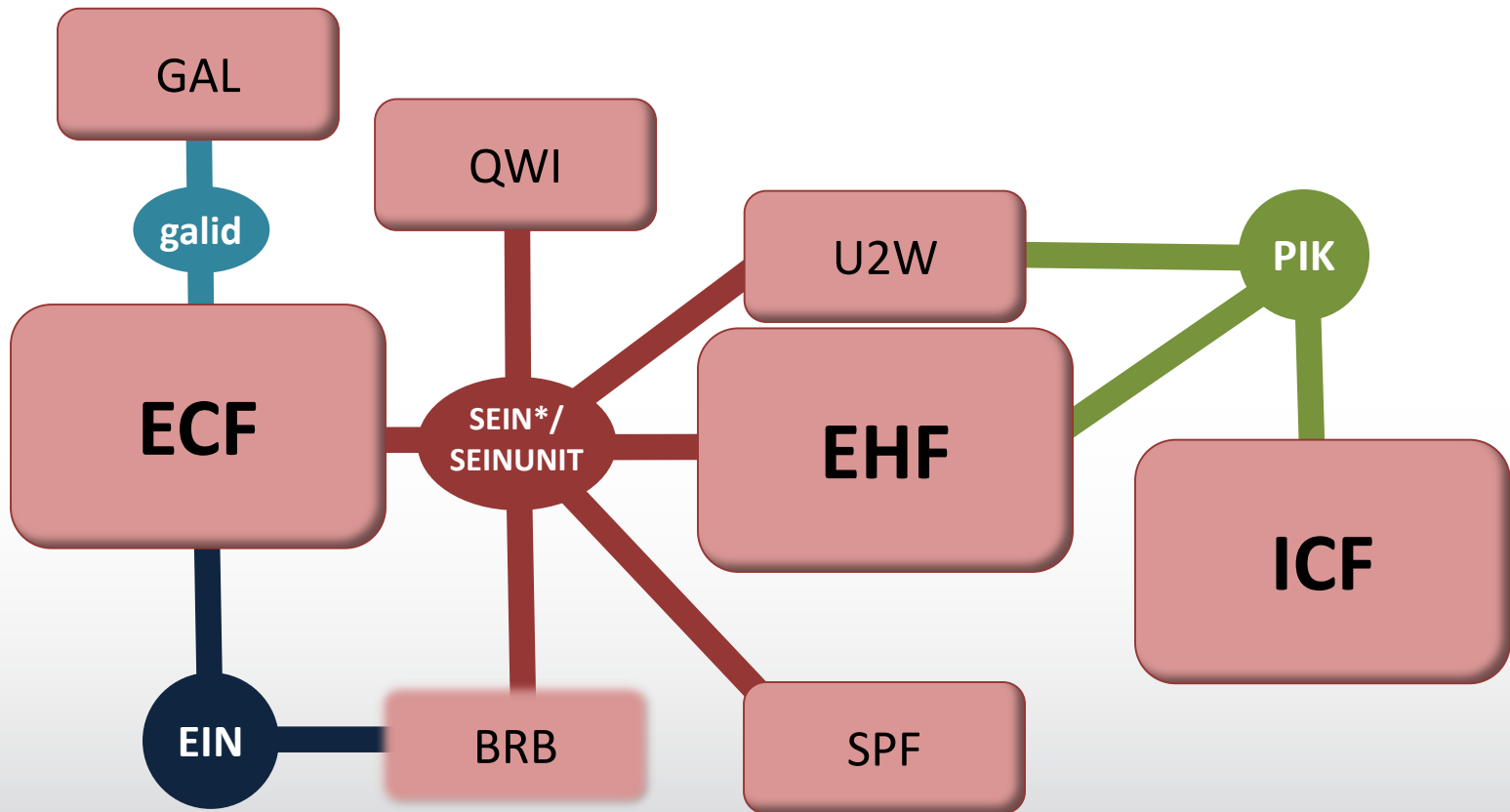
Short description of the LEHD

- Data system connecting workers to employers for any business that pays unemployment insurance
- Data: Unemployment Insurance (UI) records and ES-202 (QCEW) establishment records
 - State labor offices, Census, IRS, BLS, SSA
- 96% of all private-sector jobs are covered
 - Stevens (2007) survey's coverage for a subset of LEHD states

Mississippi
 New Hampshire
 Arkansas
 District of Columbia
 Alabama
 Vermont
 Oklahoma
 Ohio
 OPM
 Utah
 Nebraska
 Iowa
 Delaware
 Virginia
 Tennessee
 South Carolina
 Nevada
 North Dakota
 Michigan
 West Virginia
 Kentucky
 New Jersey
 Maine
 Connecticut
 Hawaii
 New Mexico
 Texas
 Rhode Island
 New York
 Minnesota
 South Dakota
 Georgia
 Montana
 Florida
 Wyoming
 Arizona
 California
 Pennsylvania
 Oregon
 North Carolina
 Wisconsin
 Washington
 Missouri
 Louisiana
 Kansas
 Indiana
 Illinois
 Idaho
 Colorado
 Alaska
 Maryland



LEHD System



*Majority are matched using just the SEIN and not the SEIN+SEINUNIT pairing

- ECF - Employer Characteristics File
 - Employer file that contains one record for every quarter-year an establishment/firm is present in either the ES-202 or the UI.
 - Two data files: establishment level and firm level
 - Each contain the same information, establishments were aggregated to create the firm level file
 - Contents: size, wages, location (lat/lon), industry
 - Obs. level: estab/firm-quarter-(month)
 - Data is in wide format (3 variables for monthly emp)
 - Source: ES-202 and the UI
 - FTI version includes: IRS firm ID, firm age, firm size

- EHF - Employment History File
 - Jobs file containing one record for each employee-employer (job) combination for each individual appearing in UI wage records (earning at least \$1 in a quarter) and for each firm that appears in the QCEW records with positive employment.
 - Contents: quarterly earnings
 - Obs. level: job[PIK-SEIN]-year-(quarter)
 - Data is in wide format (4 variables for quarter earnings)
 - Source: ES-202 and the UI
 - FTI version includes: N/A
 - Notes:
 - Data quality issues for some states in the 90's refer to Vilhuber (2018) for exact state-quarter-years
 - Top-coded max earnings vary across states over time

- ICF - Individual Characteristics File
 - Demographic info for every person ever employed in any LEHD state over the time period spanned by state's UI records, conditional on participation in the LEHD program.
 - Contents: DOB, sex, POB (country), race, ethnicity, and education
 - Obs. Level: individual (PIK)
 - Source: Person Characteristics File and short/long-form 2000 Decennial Census
 - FTI version includes: county-year and lat/lon of residence
 - Notes: Lot of imputation for some variables. Education is 88%

- GAL – Geocoded Address List
 - State’s unique commercial and residential addresses with Census geocodes down to the block level and geo coordinates (lat/lon)
 - Contents: addresses
 - Obs. Level: unique address (galid)
 - Source: ES202 (90+), SSEL (90-01), BR (02-10), MAF (06-14), ACS-POW (01-07), and AHS (02)
 - Addresses are restricted depending on the source
 - GAL core and GAL_T26 (FTI version) available for external projects
 - Table 4.1 from Vilhuber (2018)

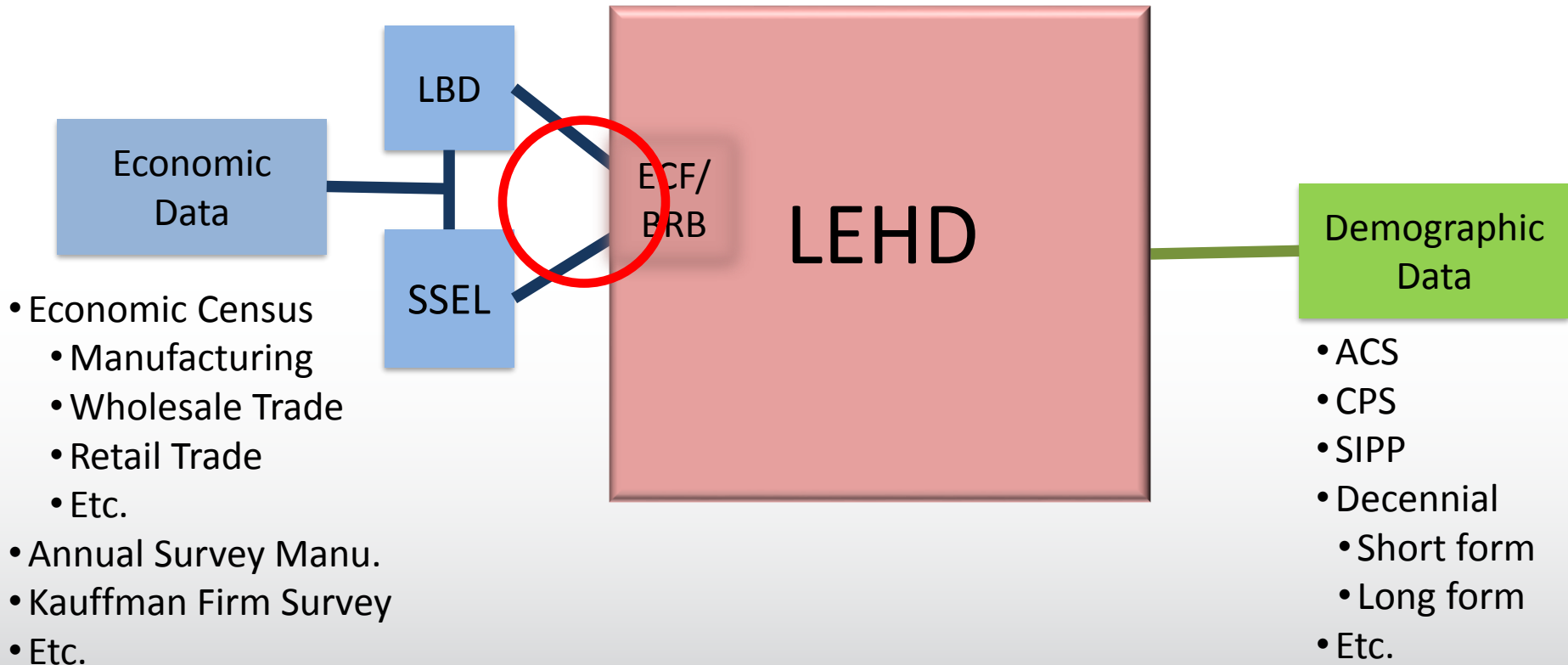
- QWI – Quarterly Workforce Indicators
 - Estab. level file containing quarterly measures of workforce composition and worker turnover.
 - Contents: aggregate tabulations by sex*agegroup, sex*educgroup, race*ethnicity, or no demographic groups
 - Obs. Level: estab/firm-quarter
 - Source: ECF, EHF, and ICF
 - FTI version includes: N/A
 - Notes: Public-use QWI file is available w/o additional permissions and is provided upon request

- **SPF – Successor-Predecessor File**
 - A suite of files providing intertemporal flow-based links for firms that changed their SEIN.
 - SEINs change for many reasons, if an employer changes SEINs but makes no other changes, then workers would appear to have left their original firm even though their employment status remains unchanged.
 - Contents: successor-predecessor links, ratio of shared jobs between successor and predecessor
 - Obs. Level: firm-quarter
 - Source: EHF
 - FTI version includes: N/A
 - Notes: not fully documented, use with caution

- U2W – Unit-to-Worker Impute
 - Imputes establishments for each employee of a multi-unit employer
 - UI records do not provide establishment identifiers, nor industry or geographic details of the establishment. They only provide a firm identifier.
 - Needed for 30 to 40 percent of workers
 - Contents: 10 imputed establishment identifiers
 - Obs. Level: job [PIK-SEIN]
 - Source: ECF, GAL, EHF, ICF, and SPF
 - FTI version includes: N/A

- BRB – Business Register Bridge
 - A link file between the ECF and the Business Register firm and establishment microdata that provides a crosswalk at various business-unit aggregations.
 - Contents: SEINUNIT to EIN “links”
 - Obs. Level: “establishment”-quarter
 - [firm EIN—state—4-digit SIC—county]
 - Source: ECF and BR
 - Notes:
 - Only available in the 2008 LEHD snapshot
 - Current LEHD snapshot include 6-digit NAICS which are also in the BR

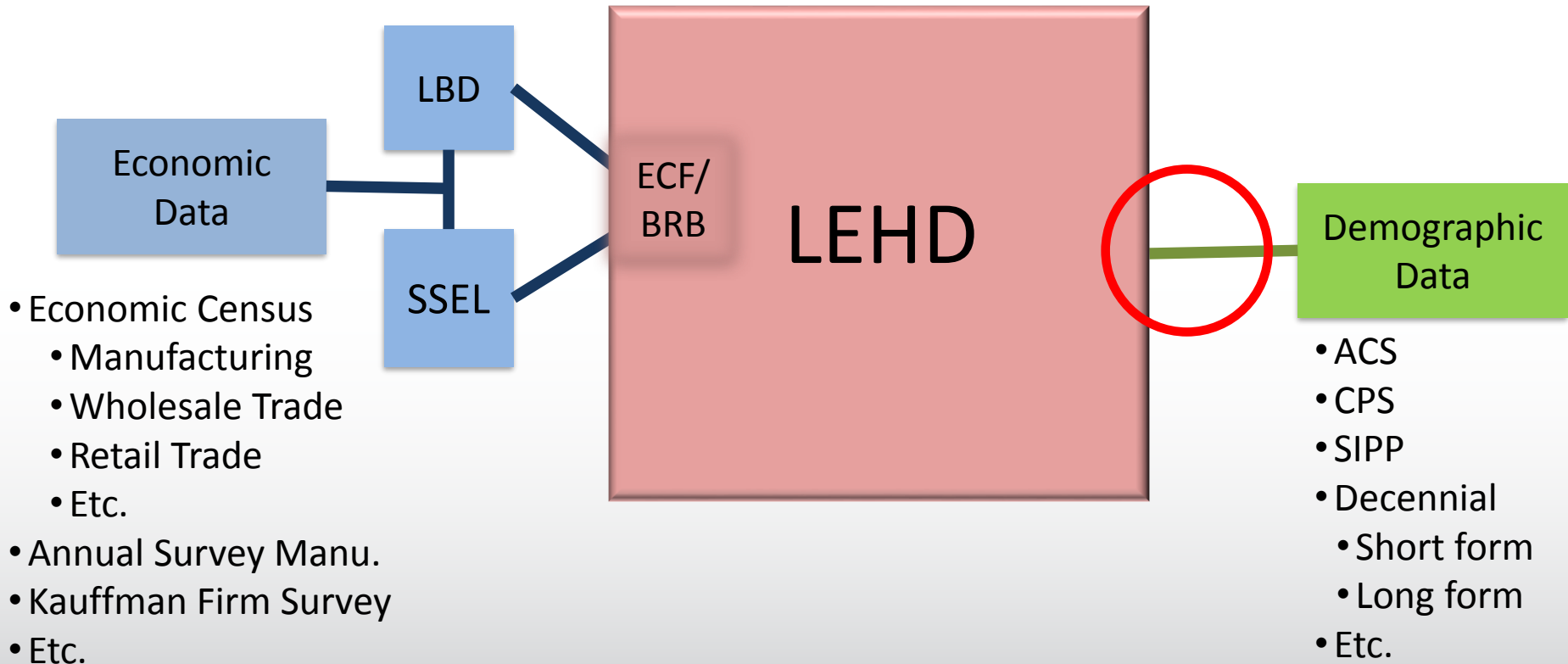
Longitudinal Employer-Household Dynamics (LEHD)



Linking Firms/Establishments LEHD – Economic Census Data

- Linking firms from the LEHD to economic Census data is straightforward for single-unit firms
- Linking establishments from the LEHD to economic Census data is tricky for multi-unit firms
 - There does not exist a crosswalk for the SEINUNIT to Census establishment identifiers
 - Matches for establishments can be made using common variables across datasets
 - Federal EIN, NAICS codes, geography, employment size, etc.

Longitudinal Employer-Household Dynamics (LEHD)



Protected Identification Keys (PIKs)

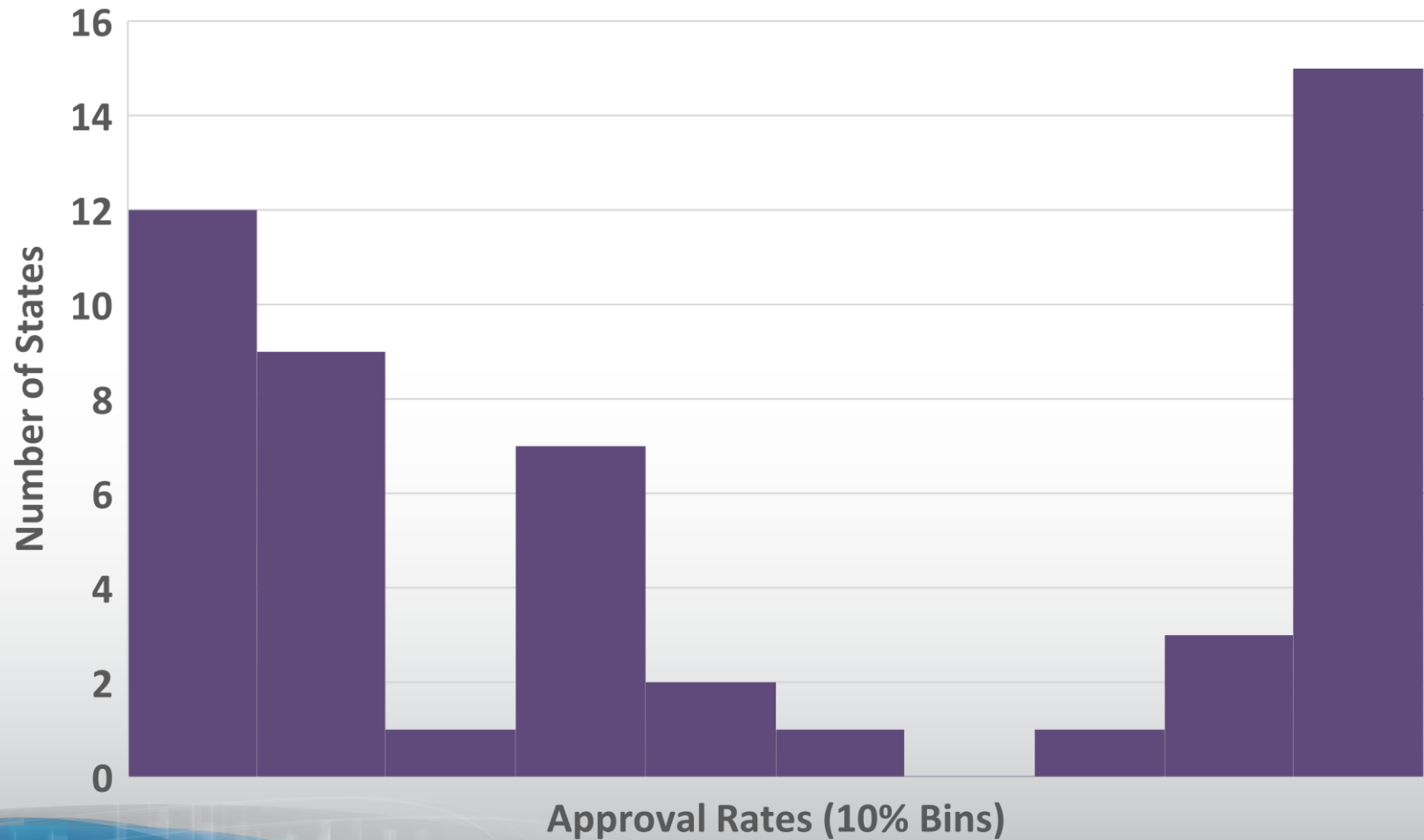
- Identifier corresponding to a unique individual
 - Assigned independently and randomly by the Person Identification Validation System (PVS)
 - Uses: Name, DOB, SSN, Address
- External individual data can be assigned a PIK
 - Costs ~\$25k or more
- Crosswalks to household surveys (ACS, AHS, CPS, SIPP, etc.) are available and must be requested separately

LEHD Restrictions

Proposal process:

- Census Approval
 - LEHD Benefit (appeals to the states more)
- Individual states approve project proposals
 - Some states defer to Census approval
 - Independent state review, default-reject if no response
 - Around 23 states approve on average
- IRS approval (if requesting FTI)
 - Version of the LEHD without FTI exists
- SSA approval for the ICF

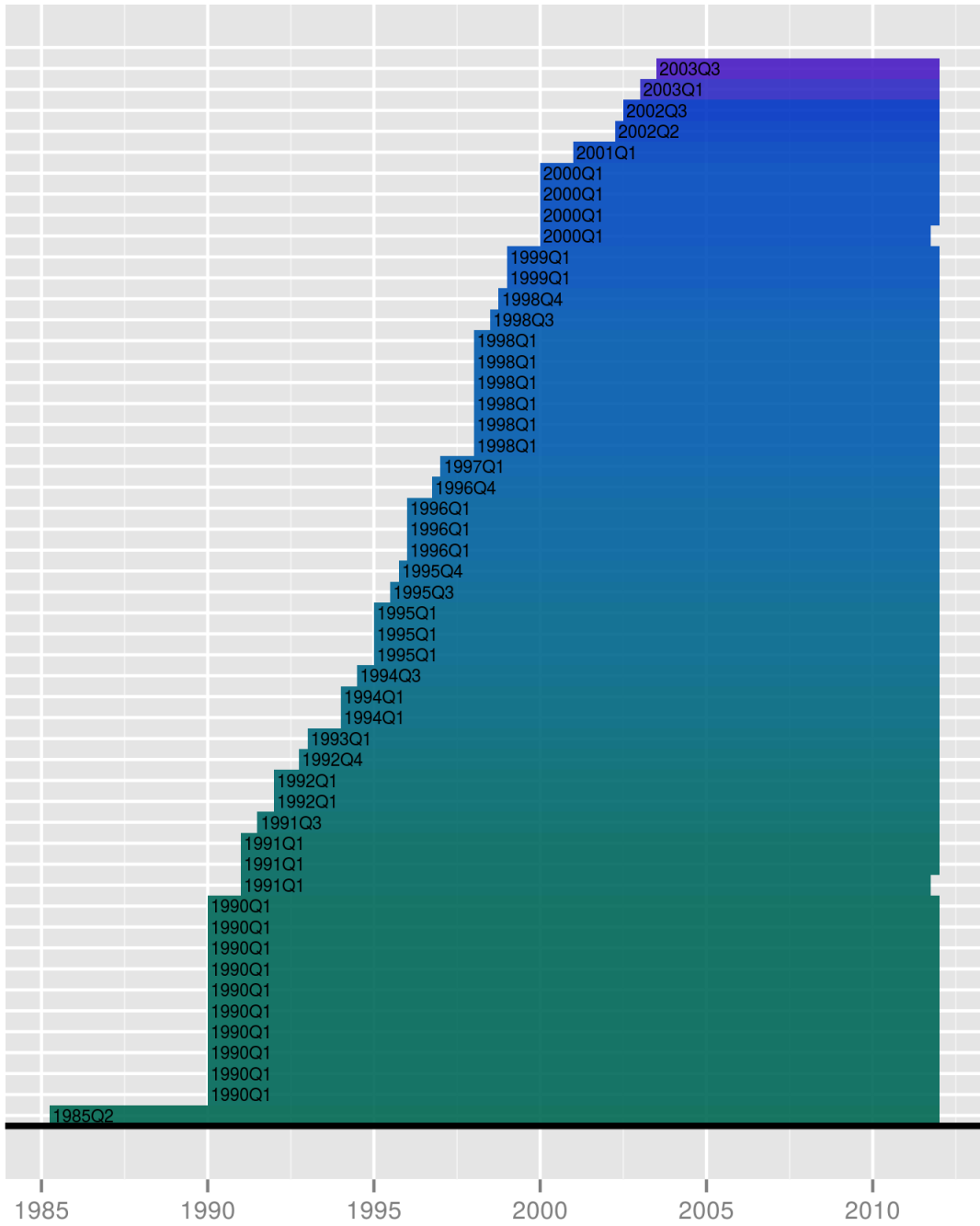
Histogram of state approval rates since 2015



LEHD Restrictions

- Disclosure rules
 - Individual states cannot be singled out in results
 - Standard output must include at least three states with one state not comprising more than 50% of the sample.
 - State/sub-state output require state LMI approval
- LEHD does not have firm/establishment names
- Data is extremely large, SAS is recommended (almost needed) to handle the computation
 - Work with representative subsample first

Mississippi
 New Hampshire
 Arkansas
 District of Columbia
 Alabama
 Vermont
 Oklahoma
 Ohio
 OPM
 Utah
 Nebraska
 Iowa
 Delaware
 Virginia
 Tennessee
 South Carolina
 Nevada
 North Dakota
 Michigan
 West Virginia
 Kentucky
 New Jersey
 Maine
 Connecticut
 Hawaii
 New Mexico
 Texas
 Rhode Island
 New York
 Minnesota
 South Dakota
 Georgia
 Montana
 Florida
 Wyoming
 Arizona
 California
 Pennsylvania
 Oregon
 North Carolina
 Wisconsin
 Washington
 Missouri
 Louisiana
 Kansas
 Indiana
 Illinois
 Idaho
 Colorado
 Alaska
 Maryland



It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States

- Barth, Bryson, Davis, Freeman (2016)
 - Journal of Labor Economics
 - www.journals.uchicago.edu/doi/abs/10.1086/684045
- Much of the 1970s-2010s increase in earnings inequality results from increased dispersion of the earnings among the establishments where individuals work
- LEHD, Economic Censuses, Decennial Censuses, CPS, etc.

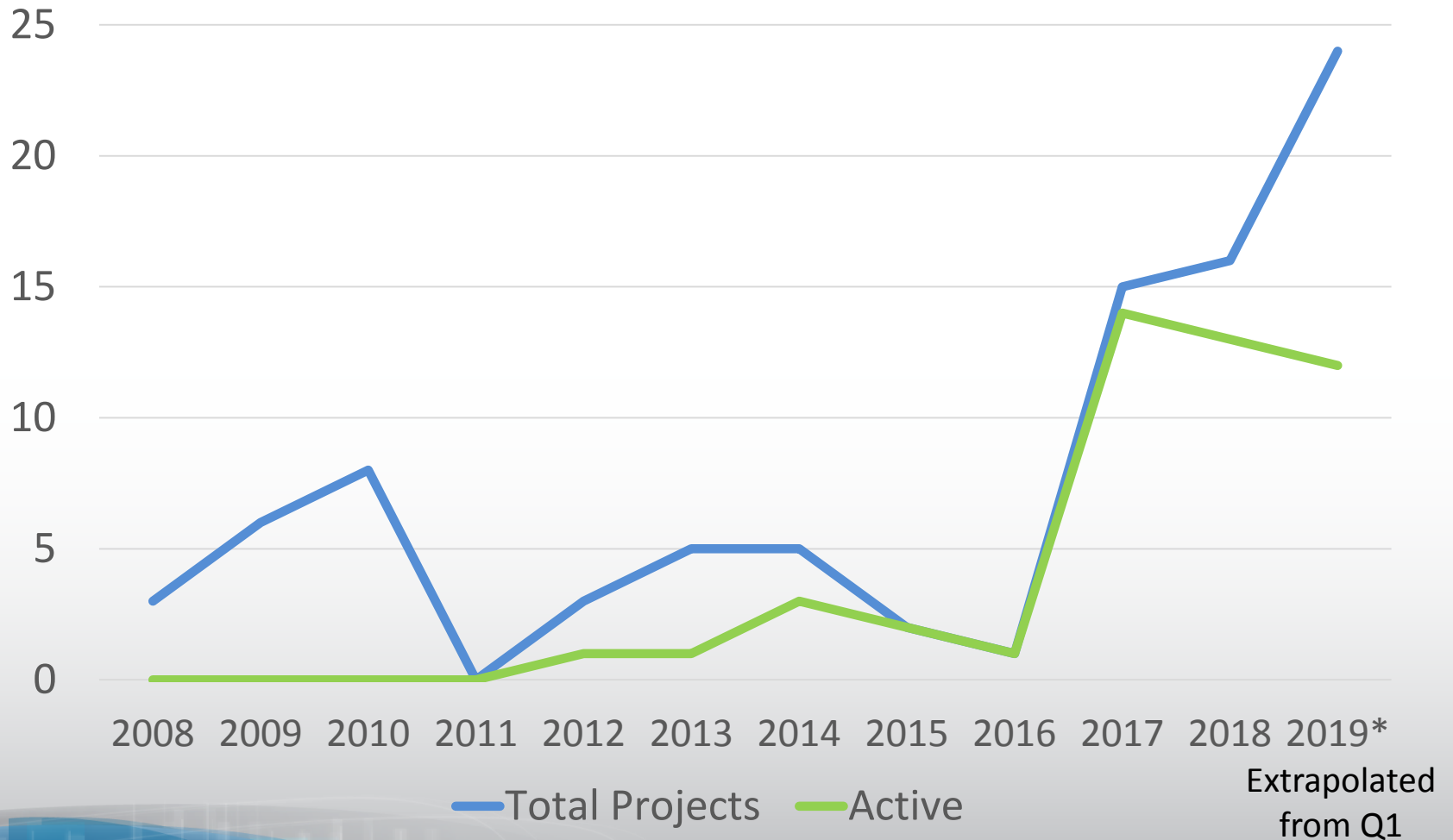
Firm Performance and the Volatility of Workers Earnings

- Juhn, McCue, Monti, Pierce (2018)
 - Journal of Labor Economics
 - www.journals.uchicago.edu/doi/pdfplus/10.1086/694167
- Shocks to firm revenues do transmit to the earnings of continuing employees but the effect is small. Firms insulate workers from idiosyncratic shocks.
- LEHD, ARTS, ASM, SAS, Economic Censuses

Employee Costs of Corporate Bankruptcy

- Graham, Kim, Li, Qiu (2016)
 - *R&R Journal of Finance*
 - ideas.repec.org/p/cen/wpaper/13-37.html
 - papers.ssrn.com/sol3/papers.cfm?abstract_id=2276753
- Annual wages decrease by 30% the year after a firm files for bankruptcy and remain below pre-bankruptcy earnings for five years. Affected employees are likely to work fewer hours and leave the firm, industry, and local labor market.
- LEHD, LBD, External data: Bankruptcy, Compustat

Approved LEHD Projects



Census Working Papers on LEHD

- CES Discussion Paper 18-27R
 - Lars Vilhuber (2018)
 - <https://ideas.repec.org/p/cen/wpaper/18-27r.html>
- The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators
 - Abowd et al. (2006, 2009)
 - <https://www.nber.org/chapters/c0485.pdf>

Publications by RDC Researchers and CES Staff: 2017 and Forthcoming

Econ Journals (by rank)	RDC		Total
	Researchers	CES staff	
AAA (1-5)	11	0	11
AA (6-20)	11	7	18
A (21-102)	22	18	40
B (103-258)	7	3	10
C (259-562)	5	1	6
D (562-1,202)	0	0	0
<hr/>			
Journals outside of Econ	16	4	20
Book Chapters	9	12	21
TOTAL	81	45	126

Contact Info

- Email: samuel.r.bondurant@census.gov
 - samuel.bondurant@dal.frb.org
- Phone: 214-922-6074
- Website: www.dfwrdc.org
 - Proposals tab

ACS-POW	American Community Survey-Place of Work	NAICS	North American Industry Classification System
AHS	American Housing Survey	PIK	Protected Identification Key
ARTS	Annual Retail Trade Survey	POB	Place of Birth
AWTS	Annual Wholesale Trade Survey	QCEW	Quarterly Census of Employment and Wages
BLS	Bureau of Labor Statistics	QWI	Quarterly Workforce Indicators
BR	Business Register	SAS	Data software
BRB	Business Register Bridge	SAS	Service Annual Survey
CES	Center for Economic Studies	SEIN	State EIN
CPS	Current Population Survey	SIC	Standard Industrial Classification
DOB	Date of Birth	SIPP	Survey of Income and Program Participation
ECF	Employer Characteristics File	SPF	Successor-Predecessor File
EHF	Employment History File	SSA	Social Security Administration
EIN	Employment Identification Number	SSN	Social Security Number
ES-202	see QCEW	SSEL	Standard Statistical Establishment Listing (see BR)
FTI	Federal Tax Information (IRS data)	T26	Title 26 (IRS data)
GAL	Geocoded Address List	U2W	Unit-to-Worker Impute
ICF	Individual Characteristics File	UI	Unemployment Insurance
IRS	Internal Revenue Service		
LAT	Latitude		
LBD	Longitudinal Business Database		
LMI	Labor Market Information (offices)		
LON	Longitude		
MAF	Master Address File		