THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Federal Sources of Entrepreneurship Data: A Compendium

Andrew Reamer, Research Professor
December 16, 2019

Prepared for the Ewing Marion Kauffman Foundation

Introduction

The E.M. Kauffman Foundation has asked the George Washington Institute of Public Policy (GWIPP) to prepare a compendium of federal sources of data on self-employment, entrepreneurship, and small business development.

The availability of useful, reliable federal data on these topics enables robust descriptions and explanations of entrepreneurship trends in the United States and so helps guide the development of effective entrepreneurship policies. Achieving these ends first requires the identification and detailed description of available federal datasets, as provided in this compendium. Its contents include:

- An overview and discussion of 18 datasets from four federal agencies, organized by two categories and five subcategories.
- Tables providing information on each dataset, including:
  - scope of coverage of self-employed, entrepreneurs, and businesses;
  - data collection methods (nature of data source, periodicity, sampling frame, sample size);
  - dataset variables (owner characteristics, business characteristics and operations, geographic areas);
  - data release schedule; and
  - data access by format (including fixed tables, interactive tools, API, FTP download, public use microdata samples [PUMS], and confidential microdata).
- For each dataset, examples of studies, if any, that use the data source to describe and explain trends in entrepreneurship.

The author's aim is for the compendium to facilitate an assessment of the strengths and weaknesses of currently available federal datasets, discussion about how data availability and value can be improved, and implementation of desired improvements.

Overview of Findings

Four federal agencies collect and publish 18 datasets useful for describing and analyzing entrepreneurship trends:

- Census Bureau, U.S. Department of Commerce (9)
- Bureau of Labor Statistics (BLS), U.S. Department of Labor (6)
- Bureau of Economic Analysis (BEA), U.S. Department of Commerce (2)
- Statistics of Income Division (SOI), Internal Revenue Service (IRS), U.S. Department of the Treasury (1)

These 18 datasets can be organized into two categories. The 13 datasets in the first group measure the number of self-employed persons, entrepreneurs, and small businesses.

- Five datasets disaggregate the count by business variables (e.g., industry, earnings), but not owner characteristics.
- Five datasets disaggregate the count by business variables and owner characteristics (e.g., sex, age, race, ethnicity).
- Three datasets go into substantial detail regarding owners' personal history in relation to business development.

The five datasets in the second group measure business dynamics—patterns in the rise and decline of individual businesses.

- One focuses on business applications and formations.
- Four measure establishment and/or firm openings, expansions, contractions, and closures.

The 18 datasets are organized by category and subcategory in the box below.

Federal Sources of Entrepreneurship Data

I. The Self-Employed, Entrepreneurs, and Small Businesses – Counts and Characteristics

A. By Business Characteristics
- Proprietor Employment and Income (BEA)
- Business Tax Returns (IRS)
- Non-Employer Statistics (Census)
- Self-Employed Persons by Industry (BEA)
- Employment Projections (BLS)

B. By Proprietor Demographic Characteristics and Business Characteristics
- Annual Business Survey (Census)
- American Community Survey (Census)
- Current Population Survey Basic Labor Force (BLS)
- CPS Annual Social and Economic Supplement (Census)
- Geographic Profile of Employment and Unemployment (BLS)

C. Personal Histories -- Context for Self-Employment and Entrepreneurship
- National Longitudinal Survey of Youth 1979 (BLS)
- National Longitudinal Survey of Youth 1997 (BLS)
- Survey of Income and Program Participation (Census)

II. Business Dynamics

A. Business Applications and Formations
- Business Formation Statistics (Census)

B. Openings, Expansions, Contractions, and Closings
- Business Employment Dynamics (BLS)
- Business Dynamics Statistics (Census)
- Statistics of US Businesses (Census)
- Quarterly Workforce Indicators, Longitudinal Employer-Household Dynamics (Census)

Federal Sources of Entrepreneurship Data

Table 1: Coverage, Sources, Methods, Characteristics, Geography

Dataset Name	Agency	Nature of Data Source	Universe	Coverage of Entrepreneurs, Self-Employed, and Small Businesses	Periodicity	Release Date	Most Recent Nationwide Count: Date	Most Recent Nationwide Count: Seasonally Adjusted Status	Most Recent Nationwide Count: Nonfarm/Farm or Public/Private	Most Recent Nationwide Count	Owner Characteristics	Business Characteristics	Business Operations	Geography	Sampling Frame	Sample Size
I. The Self-Employed, Entrepreneurs, and Small Businesses – Counts and Characteristics																
A. By Business Characteristics									Nonfarm	Farm						
Proprietor Employment and Income	BEA	Admin records	Proprietorships and partnerships (from tax returns)		Count -- annually	First Nov after data year	2018		44,581,000	1,790,000	NA	Farm, nonfarm	Income	US, state, county, CBSA	Schedules C and F, Form 1040; Form 1065; USDA farm income estimates	46,371,000
					Earnings -- US - monthly; states - quarterly, metro/county - annually	US -- first month after data month; state - 3 mos after qtr; metro/county -- first Nov after data yr										
Business Tax Returns	IRS	Admin records	Nonfarm sole proprietorships filing tax returns		Annually	Second Sept after data year	2017			26,426,406	Sole proprietor	Industry (2-digit)	Business receipts, selected deductions, payroll, and net income; income statements; profits and losses; business expenses	US		26,426,406 Schedule Cs
			Business partnerships filing tax returns		Annually	Multiple dates	2017		3,905,335 partnerships; 27,500,929 partners		General, limited, limited liability, and foreign partnerships	Industry (1-digit)	Total assets, trade or business income and deductions, portfolio income, rental income and total net income; balance sheet items; profit status; foreign transactions	US	Forms 1065 and 1065-B	3,905,335 partnerships
			S corporations filing tax returns		Annually	Fourth June after data year	2014		4,380,125 firms; 7,215,740 shareholders		S corporation	Industry (2-digit), # of shareholders	Assets, liabilities, receipts, deductions, income	US	Form 1120S	4,380,125 S corporations
Non-Employer Statistics (NES)	Census Bureau	Admin records	For-profit businesses with tax return	No employees, receipts => $1K	Annually	Second June after data year	2017			25,701,671	Legal form of business (C, S, sole, partnership)	Industry (5-digit)	Sales, value of shipments, or revenue	US, state, county, CBSA	Census Business Register	25,701,671 businesses
Self-Employed Persons by Industry	BEA	Household survey	Active proprietors or partners who devote a majority of their working hours to their unincorporated businesses		Annually	First March after data year	2018		8,941,000	766,000	NA	Industry (2-digit)	NA	US	Current Population Survey, BLS	162,000 households annually
Employment Projections	BLS	Household survey	Employed workers	Self-employed (nonfarm, farm)	Annually	First Sept after data year	2018		8,924,000	762,800	Occupation (6 digit SOC)	NA	NA	US	National Employment Matrix	for self-employed -- 12 months of CPS interviews
							2028		9,453,400	733,400					NA	NA
B. By Proprietor Demographic Characteristics and Business Characteristics									Private	Public						
Annual Business Survey (ABS)	Census Bureau and NCSES	Business survey	Businesses filing tax returns, by legal form	Nonfarm employer businesses filing IRS tax forms as sole proprietorships, partnerships, or any other type of corporation, with receipts of $1,000 or more	Annually	Second December after data year	2016 (As ASE. First ABS year is 2017.)		5,333,444	268,314	Sex, ethnicity, race, veteran status, age, educational attainment, citizenship status, prior business experience (private owners only)	Private/publicly-held; family-owned; spouse involvement; home-based; industry; firm age; firm size (employees, payroll, receipts); employee benefits	Employees; payroll; receipts; e-commerce; R&D (1-9 employees only); innovation; technology; IP; financing and profitability; globalization; business structure	For years ending in 2,7 -- US, state, MSA, county, place; for other years -- US, state, select MSAs	IRS tax returns -- Schedule C and Forms 1065, 941, 944, 1120	850,000 nonfarm employer business in years ending in 2, 7; 300,000 in other years
Note: The ABS is the new incarnation of the Survey of Business Owners (SBO) and the Annual Survey of Entrepreneurs (ASE). The SBO was conducted in years 2 and 7, the ASE in all other years.																
									Nonfarm/Farm							
American Community Survey (ACS)	Census Bureau	Household survey	Labor force participants by class of worker	Unincorporated self-employed and unpaid family workers	Annually	1-yr averages - first September after data year. 5-yr averages, first December after fifth data year	2018			9,605,440	Sex, occupation	Industry (2-digit), FT year-round	Median earnings (all, FT year-round)	US, Census region, state, county, census tract, place, congressional district, school district, Urban Area, CBSA, ZCTA	Sampling strata; MAF	3,540,000 addresses annually
				Incorporated self-employed						5,674,572						
									Nonfarm	Farm						
Current Population Survey (CPS) Basic Labor Force	BLS	Household survey	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Monthly	First Friday of next month	Oct-19	SA	8,802,000	796,000	Race, ethnicity (A,Q), Age, sex, FT/PT status, employment status (A, M), certification and licensing status (A)	NA	NA	US	Primary Sampling Units (PSUs); Master Address File (MAF)	54,000 households interviewed
								NSA	8,926,000	786,000						
					Quarterly	First Friday of next quarter	3Q19	SA	8,893,000	779,000						81,000 unique households per
								NSA	8,916,000	806,000						
					Annually	First January after data year	2018		8,736,000	766,000						162,000 unique households annually
CPS Annual Social and Economic Supplement (ASEC)	Census Bureau	Household survey	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Annually	First Sept after data year	2018	Longest job	14,135,000	750,000	Gender, race, ethnicity	Industry (1-digit)	Earnings	US	PSUs, MAF	99,000 households annually
								Longest job of year-round, full-time workers	9,042,000	495,000						
Geographic Profile of Employment and Unemployment (GP)	BLS	Household survey	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Annually	First October after data year	2018		8,916,000	764,000	Gender, race, ethnicity	NA	NA	US, Census region, state	PSUs, MAF	162,000 unique households annually

C. Personal Histories -- Context for Self-Employment and Entrepreneurship									Nonfarm/Farm						
National Longitudinal Survey of Youth 1979 (NLSY79)	BLS	Household survey	Persons 14-22 years old when first surveyed in 1979	Self-employed persons	Biennially	Two-three years after interviews	2016		Not published -- must be calculated	Employment history, education and training, age, gender, race, ethnicity, health and health care, income, attitudes, unsocial behavior, family history, personal history	Industry	Earnings	Census region, metro/non-metro, urban/rural	102 PSUs selected - cross-sectional sample with oversample of blacks, Hispanics	sample of 12,686 young people 14-22 years old when first surveyed in 1979
				Business owners					Not published -- must be calculated	Above plus family involvement, share, family history in business, patents, entrepreneurship status, prior businesses	Industry, legal form, year established, acquisition (year, funds amount and source), end (year, method)	Number employees, number locations, revenue, patents			
National Longitudinal Survey of Youth 1997 (NLSY97)	BLS	Household survey	Persons 12-16 years old as of December 31, 1996	Self-employed persons	Biennially	Two-three years after interviews	2015-16		Not published -- must be calculated	Employment history, education and training, age, gender, race, ethnicity, health and health care, income, attitudes, unsocial behavior, family history, personal history, citizenship	Industry	Earnings	Census region, metro/non-metro, urban/rural	147 PSUs selected - cross-sectional sample with oversample of blacks, Hispanics	sample of 8,984 12-16 years old as of December 31, 1996
				Business owners					Not published -- must be calculated		Industry, legal form	Number employees, income			
Survey of Income and Program Participation (SIPP)	Census Bureau	Household survey	Labor force participants by class of worker	Self-employed persons	Annually	In third calendar year after base year	2016		Not published -- must be calculated	Age, sex, race, ethnicity, educational attainment, employment status and history, health care, program participation, assets, adult well-being	Legal form, industry	Earnings	Selected by researcher	820 PSUs, MAF	29,825 households interviewed Feb-June 2014

II. Business Dynamics A. Business Applications and Formations									Nonfarm/Farm					**Records Frames**	**Universe Size**
Business Formation Statistics (BFS)	Census Bureau	Admin records	Employers with EINs	Business applications (BA)	Quarterly	US -- Two weeks after data quarter. States on delayed basis.	3Q2019		860,125	Legal status	Propensity of turning into businesses with payroll		US, state	IRS SS-4 filings	650,000 to 1,000,000 per quarter
Note: See appendix worksheet for 12 BFS measures				Business formations after application (BF)								Business formation within 4 quarters of application, within 8 quarters of application		New business operating records	100,000 per quarter
				within 4 quarters (BF4)					78,012						
				within 8 quarters (BF8)					94,972						

B. Openings, Expansions, Contractions, and Closings									Nonfarm/Farm						
Business Employment Dynamics (BED)	BLS	Admin records	Establishments in state UI systems	Private establishments with employment (all sizes and ages)	Quarterly and annually (March-to-March change)	Seven months after data quarter	1Q2019	SA	8,229,541	NA	Industry (3-digit), establishment size	Establishment births, expansions, contractions, deaths (gross job gains/losses, number of establishments); size of change (# jobs)	US, state	Quarterly Census of Employment and Unemployment (QCEW)	8,208,709 establishments
								NSA	8,208,709						
			Firms in state UI systems	Private firms with employment	Quarterly				5,282,000	NA	Firm size by # employees	Gross job gains (expanding firms, opening firms) and gross job losses (contracting firms, closing firms) by firm size class	US		5,282,800 firms
Business Dynamics Statistics (BDS)	Census Bureau	Admin records	Private establishments	Same as County Business Patterns (private establishments with employees)	Annually	2017 will be released in Spring 2020	2016		6,886,453	NA	Sector (1-digit), establishment size, initial establishment size, establishment age	Establishment entry, exit; job creation (new establishments, continuing establishments); job destruction (exiting establishments, continuing establishments)	US, state	County Business Patterns establishments in Longitudinal Business Database	6,886,453 establishments
			Private firms						5,165,983	NA	Sector (1-digit), firm size, initial firm size, firm age	Job creation (new establishments, continuing establishments); job destruction (exiting establishments, continuing establishments); firm deaths	US, state, metro/nonmetro		5,165,983 firms
Statistics of US Businesses (SUSB)	Census Bureau	Admin records	Private firms and establishments	Same as County Business Patterns (private establishments with employees)	Count - Annually	Second December after data year	2016		5,954,684 firms and 7,757,807 establishments	NA	Firm size, industry (6-digit), legal form of organization	Employment, payroll (and receipts for years ending in 2 and 7)	US, state, CBSA, county, congressional district	County Business Patterns establishments in Longitudinal Business Database	5,954,684 firms and 7,757,807 establishments
			Private establishments		Year-to-year change - annually	Third March after data year	2015-2016		6,872,350	NA	Firm size, industry	Establishment births, deaths, expansions, contractions -- number of establishments, change in employment	US, state, CBSA, county		6,872,350 establishments
Quarterly Workforce Indicators (QWI), Longitudinal Employer-Household Dynamics (LEHD)	Census Bureau	Admin records	Employers participating in state UI systems		Quarterly	Nine mos after data quarter	1Q2019		NA (Totals not in dataset)	NA	Industry (4-digit), firm age, firm size	Employment, hires, separations, turnover, firm employment change (gains, losses), average monthly earnings, worker characteristics (sex, age, education, race, ethnicity)	US, state, county, CBSA, WIA	QCEW	10,203,000 establishments

Prepared by Andrew Reamer, George Washington University, for the E.M. Kauffman Foundation
December 16, 2019

Federal Sources of Entrepreneurship Data

Table 2: Options for Data Access

Dataset Name	Agency	Universe	Coverage of Entrepreneurs, Self-Employed, and Small Businesses	Periodicity	Data Access: Fixed Tables	Data Access: Interactive Table Creator	Data Access: Application Programming Interface (API)	Data Access: File Transfer Protocol (FTP)	Data Access: Public Use Microdata Samples (PUMS)	Data Access: Confidential Microdata
I. The Self-Employed, Entrepreneurs, and Small Businesses – Counts and Characteristics										
A. By Business Characteristics										
Proprietor Employment and Income	BEA	Proprietorships and partnerships (from tax returns)		Count -- annually Earnings -- US - monthly; states - quarterly, metro/county - annually	https://www.bea.gov/data/income-saving	https://apps.bea.gov/itable/index.cfm	https://apps.bea.gov/API/signup/index.cfm			Federal Statistical Research Data Centers
Business Tax Returns	IRS	Nonfarm sole proprietorships filing tax returns		Annually	https://www.irs.gov/					
		Business partnerships filing tax returns		Annually	https://www.irs.gov/					
		S corporations filing tax returns		Annually	https://www.irs.gov/					
Non-Employer Statistics (NES)	Census Bureau	For-profit businesses with tax return	No employees, receipts => $1K	Annually	https://www.census.gov/programs-surveys/nonemployer-	https://data.census.gov/	https://www.census.gov/data/developers/data-sets/cbp-			Federal Statistical Research Data Centers
Self-Employed Persons by Industry	BEA	Active proprietors or partners who devote a majority of their working hours to their unincorporated businesses		Annually	BEA Interactive Tables (Table 6.7D)	BEA Interactive Tables (Table 6.7D)				Federal Statistical Research Data Centers
Employment Projections	BLS	Employed workers	Self-employed (nonfarm, farm)	Annually	https://www.bls.gov/emp/tables/employm	https://www.bls.gov/emp/data.htm	https://www.bls.gov/developers/			
B. By Proprietor Demographic Characteristics and Business Characteristics										
Annual Business Survey (ABS)	Census Bureau and NCSES	Businesses filing tax returns, by legal form	Nonfarm employer businesses filing IRS tax forms as sole proprietorships, partnerships, or other type of corporation, with receipts of $1,000 or more	Annually	https://www.census.gov/programs-surveys/abs/data/tables.html	data.census.gov	https://www.census.gov/data/developers/data-sets/ase.html			Federal Statistical Research Data Centers
Note: The ABS is the new incarnation of the Survey of Business Owners (SBO) and the Annual Survey of Entrepreneurs (ASE). The SBO was conducted in years 2 and 7, the ASE in all other years.							https://www.census.gov/data/developers/			
American Community Survey (ACS)	Census Bureau	Labor force participants by class of worker	Unincorporated self-employed and unpaid family workers	Annually	data.census.gov	data.census.gov	https://www.census.gov/data/developers/data-sets/acs-		https://www.census.gov/programs-surveys/acs/data/pu	Federal Statistical Research Data Centers
			Incorporated self-employed							
Current Population Survey (CPS) Basic Labor Force	BLS	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Monthly	https://www.bls.gov/web/empsit/cpseea0 https://www.bls.gov/web/empsit/cpseea2		https://www.bls.gov/developers/		https://www2.census.gov/programs-surveys/cps/methodology/PublicUseDocum	Federal Statistical Research Data Centers

				Quarterly	https://www.bls.gov/web/empsit/cpsee_e https://www.bls.gov/web/empsit/cpsee_e				entation_final.pdf	
				Annually	https://www.bls.gov/cps/cpsaat12.htm					
CPS Annual Social and Economic Supplement (ASEC)	Census Bureau	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Annually	https://www.census.gov/data/tables/time-series/demo/income-	https://www.census.gov/cps/data/cpstablecreator.html			https://www.census.gov/topics/population/foreign-	Federal Statistical Research Data Centers
Geographic Profile of Employment and Unemployment (GP)	BLS	Labor force participants by class of worker	Unincorporated self-employed (farm, nonfarm)	Annually	https://www.bls.gov/opub/geographic-profile/home.htm		https://www.bls.gov/developers/			Federal Statistical Research Data Centers

C. Personal Histories -- Context for Self-Employment and Entrepreneurship

National Longitudinal Survey of Youth 1979 (NLSY79)	BLS	Persons 14-22 years old when first surveyed in 1979	Self-employed persons and business owners	Biennially		https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY79			https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY79	https://www.bls.gov/nls/nlsfaqs.htm#anch26
National Longitudinal Survey of Youth 1997 (NLSY97)	BLS	Persons 12-16 years old as of Dec 31, 1996	Self-employed persons and business owners	Biennially		https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY9			https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY9	https://www.bls.gov/nls/nlsfaqs.htm#anch26
Survey of Income and Program Participation (SIPP)	Census Bureau	Labor force participants by class of worker	Self-employed persons	Annually		https://www.census.gov/programs-surveys/sipp/data/da		https://thedataweb.rm.census.gov/ftp/sipp_ftp.html	SIPP Synthetic Beta Data	Federal Statistical Research Data Centers

II. Business Dynamics

A. Business Applications and Formations

Business Formation Statistics (BFS)	Census Bureau	Employers with EINs	Business applications (BA) Business formations (BF)	Quarterly	https://www.census.gov/econ/bfs/index.h	https://www.census.gov/econ/bfs/data.ht				

B. Openings, Expansions, Contractions, and Closings

Business Employment Dynamics (BED)	BLS	Establishments in state UI systems	Private establishments with employment (all sizes and ages)	Quarterly and annually (March-to-March change)	Quarterly tables of job gains and losses Q1 Distribution of	https://www.bls.gov/bdm/#data	https://www.bls.gov/developers/	https://download.bls.gov/pub/time.series/bd/		Federal Statistical Research Data Centers
Business Dynamics Statistics (BDS)	Census Bureau	Private establishments and firms	Same as County Business Patterns (private establishments with employees)	Annually	https://www.census.gov/ces/dataproducts/bds/data.html		https://www.census.gov/data/developers/data-sets/business-			Federal Statistical Research Data Centers
Statistics of US Businesses (SUSB)	Census Bureau	Private firms and establishments Private establishments	Same as County Business Patterns (private establishments with employees)	Count - Annually Year-to-year change - annually	https://www.census.gov/programs-surveys/susb/data/ta https://www.census.gov/data/tables/2016			https://www.census.gov/programs-surveys/susb/data/datasets.html		Federal Statistical Research Data Centers

Quarterly Workforce Indicators (QWI), Longitudinal Employer-Household Dynamics (LEHD)	Census Bureau	Employers participating in state UI systems	Quarterly	LED Extraction Tool	QWI Explorer	https://www.census.gov/data/developers/data-sets/qwi.html	https://lehd.ces.census.gov/data/#qwi		Federal Statistical Research Data Centers

Prepared by Andrew Reamer, George Washington University, for the E.M. Kauffman Foundation
December 16, 2019

Business Formation Statistics Variables

The Business Application Series contains 4 filtered series of EIN applications, measuring selected groupings based on a variety of factors.

BA	Business Applications: The core business applications series that corresponds to a subset of all EIN applications. BA includes all applications for an EIN, except for applications for tax liens, estates, trusts, or certain financial filings, applications with no state-county geocodes, applications from certain agricultural, public entities, and applications in certain industries (e.g. private households, civic and social organizations)
HBA	High-Propensity Business Applications: A subset of BA that contains all applications with a high-propensity of turning into a business with a payroll, based on various factors
WBA	Business Applications with Planned Wages: A subset of HBA that contains all applications that indicate a planned date for paying wages
CBA	Business Applications from Corporations: A subset of HBA that contains all applications that come from a corporation or a personal service corporation

The Business Formation Series measures new employer firm births, the point when an application is realized and a business begins to operate. The Business Formations Series is available in two subsets, measuring from application to realization of the business paying wages, within 4 and 8 quarters.

BF4Q	Business Formations within 4 Quarters: The number of employer businesses that originate from Business Applications within four quarters from the quarter of application
PBF4Q	Projected Business Formations within 4 Quarters: The projected number of employer businesses that originate from Business Applications within four quarters from the quarter of application
SBF4Q	Spliced Business Formations within 4 Quarters: Combines BF4Q and PBF4Q to provide the entire time series for the actual and projected business formations within four quarters
BF8Q	Business Formations within 8 Quarters: The number of employer businesses that originate from Business Applications (BA) within eight quarters from the quarter of application
PBF8Q	Projected Business Formations within 8 Quarters: The projected number of employer businesses that originate from Business Applications (BA) within eight quarters from the quarter of application
SBF8Q	Spliced Business Formations within 8 Quarters: Combines BF8Q and PBF8Q to provide the entire time series for the actual and projected business formations within eight quarters
DUR4Q	Average Duration (in Quarters) from Business Application to Formation within 4 Quarters: A measure of delay between business application and formation, conditional on business formation within four quarters
DUR8Q	Average Duration (in Quarters) from Business Application for Formation within 8 Quarters: A measure of delay between business application and formation, conditional on business formation within eight quarters

Economic Research based on Federal Sources of Entrepreneurship Data: Examples

Prepared by Andrew Reamer
December 16, 2019

Table of Contents

I. Counts and Characteristics of the Self-Employed, Entrepreneurs, and Small Businesses

A. By Business Characteristics

Proprietorships (BEA)

Keith Graham Debbage and Shaylee Bowen, Non-farm proprietorship employment by US metropolitan area, *Journal of Enterprising Communities: People and Places in the Global Economy*, Vol. 12 No. 2, pp. 139-157, May 2018

Summary

The entrepreneurial process is a result of an interaction between an individual entrepreneur and the surrounding entrepreneurial ecosystem. The purpose of this paper is to determine whether US metropolitan areas with disproportionately high shares of entrepreneurs are systematically linked to particular attributes of the entrepreneurial support system?

The relative share of NFP employment by metropolitan area exhibited a strong positive relationship with percentage of employment in finance, insurance and real estate, median age, percentage of Hispanic population and median home value. It is argued that the combination of significant predictors captures both out-of-necessity self-employment (e.g. low-skilled Hispanic and aging populations) and a self-employment of opportunity (e.g. access to capital).

Use of BEA Proprietorships to Reach Conclusions

In this paper, non-farm proprietorship (NFP) employment data from the US Bureau of Economic Analysis is used as a dependent variable proxy for entrepreneurship. NFP data are widely used in the entrepreneurship literature. Data on all independent variables were obtained from the Census Bureau's American Community Survey and the Bureau of Labor Statistics by metropolitan area and subject to a stepwise linear regression analysis.

Xiaowen Liu, "Tax Policy and Entrepreneurship," University of Tennessee, Knoxville, 8-2014.

Summary

Small businesses and the entrepreneurial spirit are among the driving forces in economic growth and development in the United States. The US governments (both federal and state) have long been aware of the importance of entrepreneurship, and many policies are directed toward helping small businesses. However, whether such policies give rise to expected behavioral responses from small businesses remains inconclusive. This dissertation looks into the behavioral response of self-employed filers to individual income tax and the impact of state

and federal tax policies on entrepreneurship. In the first chapter, we examine taxpayers' behavioral response to the Alternative Minimum Tax (AMT). We find strong evidence that taxpayers, especially self-employed individuals, appear to manipulate their incomes to avoid the AMT. We also find suggestive evidence that the notch created by the AMT generates both a real response and an evasion response. These results have important policy implications for the AMT design and for the evaluation of the welfare loss from taxation of small businesses. The second chapter examines the effect of state tax policies on entrepreneurial activity. This paper contributes to the literature in several important ways: first, we explore dynamic specifications to capture inherent time trends among entrepreneurial performance. Second, we consider a number of intensive-margin measures of state nonfarm proprietors' success. Third, we extend the earlier research by including a longer panel (1978-2009) of state data. Despite these innovations, our empirical results echo the recent studies in this area and suggest that most of the highly-visible state tax policies do not have statistically significant impacts on entrepreneurial performance. The last chapter uses time series analysis to explore the effect of federal tax policies on entrepreneurial performance and whether the effect is heterogeneous across different stages of the business cycle. We do not find that tax policy affects the small businesses sector differently between economic ups and downs.

Use of BEA Proprietorships to Reach Conclusions

Our paper is the first to use nonfarm proprietors' income as a direct measure of entrepreneurial success at the state level. We investigate several measures of small business performance derived from nonfarm proprietors' income and employment data.

Business Tax Returns (IRS)

Anmol Bhandari and Ellen R. McGrattan, Sweat Equity in U.S. Private Business, NBER Working Paper No. 24520, April 2018.

Summary

In this paper, we first provide evidence that existing measures of business incomes and valuations based on widely-used surveys such as the Survey of Consumer Finances are mismeasured. We then develop a theory disciplined by U.S. national accounts and business census data to measure net incomes and private business sweat equity—which is the value of time to build customer bases, client lists, and other intangible assets. We estimate an aggregate sweat equity value of 0.65 times GDP, with little cross-sectional dispersion in valuations when compared to business net incomes and large cross-sectional dispersion in rates of return. Our estimate of sweat equity is close to the estimate of marketable fixed assets used in production by private businesses, implying a high ratio of intangible to total assets. We use the model to evaluate the impact of greater tax compliance of private businesses and lower tax rates on the net income of both privately held and publicly traded businesses. We find larger sectoral and aggregate effects from the tax policy experiments relative to studies that abstract from private business and, in particular, the accumulation of sweat capital. Finally, we show that our results are robust to including nonpecuniary benefits of business ownership.

We restrict attention to U.S. private businesses in three legal organization categories, namely, S corporations, sole proprietorships, and partnerships, and we report the ratios by industry, age, and different measures of business size.

Use of IRS Tax Returns to Reach Conclusions

Key parameters of our baseline model are chosen to ensure that model income and product shares are consistent with U.S. national account data, model taxable income distributions are consistent with IRS data, and model business age profiles and hours are consistent with U.S. Census data. For information on business incomes, we use data from the IRS, and for information on business owners, we use data from the U.S. Census Bureau's Survey of Business Owners (SBO).

David Summers, "The Economic Impact of Entrepreneurship: Setting Realistic Expectations," *Academy of Entrepreneurship Journal*, 2015, vol. 21, n. 2, pp. 88-107.

Summary

Many involved in economic development are convinced encouraging entrepreneurship and small business activity is key to economic growth. Often, however, the positive economic impact of entrepreneurship and small business is exaggerated because of faulty expectations

based on hype rather than a critical analysis of current data. Realistic expectations are especially important to guide economic development as the nation's economy rebounds from recession. This paper addresses faulty expectations by evaluating current data and trends in new business startups, job creation, and the family income and net worth of families that report business ownership to establish a realistic assessment of the impact of entrepreneurship and small business activity. Overall entrepreneurial startup activity and business dynamism, which measures firm births, deaths, expansions and retractions, are in decline in recent years. The expectations established by analysis of the numbers show that economic development activities may best be focused on supporting the small number of high-growth employer firms rather than on encouraging mass business startups.

In 2011, the latest data available, the Internal Revenue Service (IRS) stated the average reported annual pretax net income for sole proprietorships was $14,285 which is a slight decrease from the 2007 level of $14,490 (Internal Revenue Service, 2014). This means that in 2011 an average of slightly less than $1,200 per month of pretax net income was generated. This may not seem like significant income, but for a struggling family, it could mean the difference in financially meeting the family's needs each month. From another viewpoint, a 10 year monthly cash flow of $1,200 at a 5% discount rate would have a present value of $113,138 representing additional family financial wealth.

Use of IRS Tax Returns to Reach Conclusions

The author uses IRS tax data to ascertain the economic impacts of entrepreneurship: “In 2011, the latest data available, the Internal Revenue Service (IRS) stated the average reported annual pretax net income for sole proprietorships was $14,285 which is a slight decrease from the 2007 level of $14,490 (Internal Revenue Service, 2014). This means that in 2011 an average of slightly less than $1,200 per month of pretax net income was generated.”

Non-Employer Statistics (Census Bureau)

Abraham, Katharine G., John Haltiwanger, Kristin Sandusky, and James Spletzer, The Rise of the Gig Economy: Fact or Fiction?, AEA Papers and Proceedings, 109: 357-61, 2019.

Summary

Gig work mediated through online platforms has received much recent attention. We find only one sector—the transportation services sector—in which there is unambiguous evidence of substantial and rapidly growing gig activity. A challenge for tracking and understanding the rise in gig activity is that core household surveys are missing the recent overall rise in self-employment that is apparent in administrative and private sector transactions data. We show that this limitation of available household survey data is evident even in the transportation services sector, where the growth in self-employment activity since 2013 has been exponential.

Based on published Census Bureau statistics on nonemployer businesses, most of which are Schedule C sole proprietorships, the self-employment rate rose from 13 percent to 15 percent between 2004 and 2016. This may be indicative of a rise in the gig economy.

Use of NES to Reach Conclusions

The authors use NES data in combination with other federal data sources on self-employment, primarily to set the context for their analysis.

Zoltan J. Acs, Brian Headd and Hezekiah Agwara, Nonemployer Start-Up Puzzle, SSRN, 2010

Summary

Understanding the dynamics of nonemployer firms has been an unsolved puzzle because of a dearth of data. Nonemployer firms represent three-quarters of the businesses in the United States but only 3 percent of business receipts. While they represent a relatively small share of economic activity, nonemployer firms are important as a gateway to becoming employer firms, providing flexible work opportunities and a path to economic prosperity. We find nonemployers have a startup rate nearly three times the rate of employer firms, 35 percent and 13 percent, respectively. Analysis of the determinants of nonemployer start-up rates across states and industries indicates that they seem countercyclical with respect to the labor market, while employer start-up rates move in line with the overall economic cycles. The determinants of start-ups have disparate effects on employer and nonemployer firms.

Use of NES to Reach Conclusions

The authors use special tabulations from the U.S. Census Bureau's annual nonemployer firm program to generate their conclusions.

Employment Projections (BLS)

These examples are from BLS and describe employment projects for the self-employed. I did not find research papers that relied on EP data.

Dubina, Kevin; Morisi, Teri; Rieley, Michael; and Wagoner, Andrea, Projections Overview and Highlights, 2018–28, *Monthly Labor Review*, October 2019.

Summary

Provides projections of agricultural and non-agricultural self-employed persons in 2028.

T. Alan Lacey, Mitra Toossi, Kevin S. Dubina, and Andrea B. Gensler, Projections overview and highlights, 2016–26, *Monthly Labor Review*, October 2017.

Summary

Provides projection of agricultural and non-agricultural self-employed persons in 2026.

Richard Henderson, Industry employment and output projections to 2024, *Monthly Labor Review*, December 2015.

Summary

Nonagricultural self-employed jobs are projected to increase from almost 8.6 million in 2014 to nearly 9.2 million in 2024. The increase of 579,300 jobs, occurring at an annual rate of 0.7 percent, is smaller than the decline of 883,400 in self-employment from 2004 to 2014.

III. The Self-Employed, Entrepreneurs, and Small Businesses – Counts and Characteristics

B. By Proprietor Demographic Characteristics and Business Characteristics

Annual Business Survey (Census Bureau and NCSES)

J. David Brown, John S. Earle, Mee Jung Kima, and Kyung Min Lee, Immigrant Entrepreneurs and Innovation in the U.S. High-Tech Sector, NBER, February 2019.

Summary

We estimate differences in innovation behavior between foreign versus U.S.-born entrepreneurs in high-tech industries. Our data come from the Annual Survey of Entrepreneurs, a random sample of firms with detailed information on owner characteristics and innovation activities. We find uniformly higher rates of innovation in immigrant-owned firms for 15 of 16 different innovation measures; the only exception is for copyright/trademark. The immigrant advantage holds for older firms as well as for recent start-ups and for every level of the entrepreneur's education. The size of the estimated immigrant-native differences in product and process innovation activities rises with detailed controls for demographic and human capital characteristics but falls for R&D and patenting. Controlling for finance, motivations, and industry reduces all coefficients, but for most measures and specifications immigrants are estimated to have a sizable advantage in innovation.

Use of ASE to Reach Conclusions

We exploit new confidential microdata from the Census Bureau's 2014 Annual Survey of Entrepreneurs (ASE). The ASE is an annual survey that supplements the Survey of Business Owners (SBO), conducted every five years, providing detailed demographic characteristics on business owners and their motivations to start a business, as well as economic characteristics of their firms. Of particular importance for this paper, it includes a rich set of innovation measures, which are the main outcome variables in our study.

The ASE sample contains non-farm businesses with at least one paid employee and receipts of $1,000 or more. Using the Census Business Register (BR) as the sampling frame, the ASE sample is stratified by the 50 most populous Metropolitan Statistical Areas (MSAs), state, and the firm's number of years in business. The ASE sample is randomly selected, except for large companies in each stratum, which are selected with certainty based on volume of sales, payroll, or number of paid employees. The initial 2014 ASE sample was about 290,000 employer firms, and the response rate was 74 percent.

For this paper, we restrict the full ASE sample to firms in the high-tech sector as defined by the share of Science, Technology, Engineering, and Mathematics (STEM) employment in the industry. This represents about 5.31 percent of firm-owner observations in the ASE. We also exclude businesses where no individual owns at least 10 percent of the equity, because detailed owner information is not provided for such businesses. We drop owners who choose the same answers for every motivation question (all very important, all somewhat important, or all not important), because those answering patterns may not reflect the true intensity for each question, as well as firm-owner observations that have missing values for any of the variables used in the regressions. Our final sample consists of about 11,000 owners of 7,400 firms. We weight each owner by their ownership equity share, adjusting them to sum up to one within each firm, and we weight each firm by ASE survey weights to make the sample representative for the U.S. economy.

Our main variable of interest is an indicator for whether the owner is an immigrant, defined in the ASE as a noncitizen at birth.

Anmol Bhandari and Ellen R. McGrattan, "Sweat Equity in U.S. Private Business," NBER Working Paper No. 24520, April 2018.

Summary

In this paper, we first provide evidence that existing measures of business incomes and valuations based on widely-used surveys such as the Survey of Consumer Finances are mismeasured. We then develop a theory disciplined by U.S. national accounts and business census data to measure net incomes and private business sweat equity—which is the value of time to build customer bases, client lists, and other intangible assets. We estimate an aggregate sweat equity value of 0.65 times GDP, with little cross-sectional dispersion in valuations when compared to business net incomes and large cross-sectional dispersion in rates of return. Our estimate of sweat equity is close to the estimate of marketable fixed assets used in production by private businesses, implying a high ratio of intangible to total assets. We use the model to evaluate the impact of greater tax compliance of private businesses and lower tax rates on the net income of both privately held and publicly traded businesses. We find larger sectoral and aggregate effects from the tax policy experiments relative to studies that abstract from private business and, in particular, the accumulation of sweat capital. Finally, we show that our results are robust to including non-pecuniary benefits of business ownership.

Use of SBO to Reach Conclusions

For information on business owners, we use data from the U.S. Census Bureau's Survey of Business Owners (SBO). The SBO provides information on turnover rates of business, time allocation to business operations, and financing requirements for business start-ups.

The Census data do not include business valuations but do include information about businesses and owners that, along with theory, can be used to infer sweat equity valuations. More specifically, to discipline our model, we use information from the 2007 SBO public use microdata sample (PUMS) on the year of the business acquisition, the hours spent working in the business, and capital sources and requirements for business start-ups.

Arnobio Morelix, Victor Hwang, Inara S. Tareque, Zero Barriers: Three Mega Trends Shaping the Future of Entrepreneurship, State of Entrepreneurship 2017, Kauffman Foundation, 2017.

Summary

After a long Great Recession hangover, entrepreneurship is finally rebounding in the United States. Entrepreneurs are driving a resurgence of business activity in America—in new business creation, local small business activity, and the growth of small firms into larger businesses.

But underneath this reassuring surface, turbulent shifts are shaping the future of entrepreneurship to be dramatically different than what it is today, or was in the past. We posit that three mega trends will be defining forces shaping the future of entrepreneurship for decades to come. These three trends reflect the changing demographics, map, and nature of American entrepreneurship.

(Multiple findings are based on analysis of the ASE.)

Use of ASE to Reach Conclusions

The authors provide an extensive series of graphs describing ASE findings by various variables, particularly funding sources by race and ethnicity.

Mark D. Heileman, Timothy L. Pett, and Adrian J. Mayer, "Small-To-Medium-Sized Enterprise Characteristics and Performance: An Exploratory Examination Of The Census Bureau's Survey Of Business Owners," *Journal of Business and Entrepreneurship*, Fall 2016.

Summary

The purpose of this paper is to explore the relationship of small-to-medium-sized enterprise (SME) characteristics and performance. The results provide some interesting differences across the SME firm size groups. In particular, smaller SMEs have different characteristics regarding owner demographics, business acquisition methods, and business performance metrics compared to larger SMEs. The study concludes with some future research directions.

Use of SBO to Reach Conclusions

The study draws on the Census Bureau's Survey of Business Owners (SBO) Public Use Microdata Sample (PUMS). SBO response variables regarding owner demographics, business acquisition, and business performance are the outcome variables of interest in this study. A one-way analysis of variance (ANOVA) is applied to test whether or not there are differences across SMEs based on firm size.

American Community Survey (Census Bureau)

Cheridan Christnacht, Adam Smith, and Rebecca Chenevert, Measuring Entrepreneurship in the American Community Survey: A Demographic and Occupational Profile of Self-Employed Workers, Working Paper Number SEHSD-WP2018-28, November 2018.

Summary

Entrepreneurs help drive the economy forward by increasing competition, introducing new ideas, and fueling job creation. A wide range of data have been used in an attempt to study this population. However, few studies have used the American Community Survey (ACS). In this paper we explore whether the ACS can be used to describe the demographic characteristics and economic outcomes of entrepreneurs.

Perhaps the most difficult part of studying this population is the lack of clear definition of the "entrepreneur," which results in studies interchanging the terms "entrepreneur," "self-employed," and "business owner". As seen in a large portion of research on entrepreneurs, this paper uses self-employment as a proxy for entrepreneurship.

Specifically, using self-employment data from the ACS we compare characteristics of the self-employed with their wage-and-salary counterparts, with special attention paid to industry, occupation, income, nativity, race, and sex. We also compare self-employment rates across demographic characteristics and highlight demographic and economic differences among the self-employed by occupation and incorporation status. In addition, using ACS industry and occupation write-in response fields, we analyze the specific job titles, primary job duties, and industries most associated with self-employed workers. Overall, we find that demographic trends and economic outcomes of self-employed workers are similar to what prior research has found.

Use of ACS to Reach Conclusions

Using self-employment data from the ACS, we compare characteristics of the self-employed with their wage-and-salary counterparts, with special attention paid to industry, occupation, income, nativity, race, and sex. We also compare self-employment rates across demographic characteristics and highlight demographic and economic differences among the self-employed by occupation and incorporation status. In addition, using ACS industry and occupation write-in response fields, we analyze the specific job titles, primary job duties, and industries most associated with self-employed workers.

Robert W. Fairlie and Magnus Lofstrom, Immigration and Entrepreneurship, IZA Discussion Papers, No. 7669, Institute for the Study of Labor (IZA), Bonn, 2013.

Summary

Immigrants are widely perceived as being highly entrepreneurial and important for economic growth and innovation. This is reflected in immigration policies and many developed countries have created special visas and entry requirements in an attempt to attract immigrant entrepreneurs. Not surprisingly, a large body of research on immigrant entrepreneurship has developed over the years. In this chapter we provide an overview of the economics literature with respect to some of the most fundamental immigrant entrepreneurship issues as well as the empirical methods and data used. The main themes we address are immigrant entrepreneurs' contributions to the economy, entrepreneurship differences across groups and group differences in entrepreneurial success.

Immigrants make substantial contributions to business ownership, business income and employment in the United States. In this section, we present some estimates from the American Community Survey documenting just how large these contributions are to the U.S. economy. Estimates from the 2006-10 ACS indicate that there are 2.4 million immigrant business owners, representing 18.2 percent of all business owners (the ACS data are discussed in detail in Appendix A).

Use of ACS to Reach Conclusions

Immigrants make substantial contributions to business ownership, business income and employment in the United States. We present estimates from the American Community Survey documenting just how large these contributions are to the U.S. economy. We then return to the ACS data to explore the characteristics of immigrant entrepreneurs and their businesses. (The authors also use long-form data from the 1990 and 2000 censuses, which were the precursor to the ACS.)

Current Population Survey Basic Labor Force (BLS)

Carlianne Patrick, Heather Stephens, and Amanda Weinstein, Where are all the self-employed women? Push and pull factors influencing female labor market decisions, *Small Business Economics*, March 2016, Volume 46, Issue 3, pp 365–390.

Summary

Previous research focuses on factors that influence self-employment participation, in part because entrepreneurship has been associated with economic growth. This literature has tended to focus only on men or the comparison of women to men, while ignoring substantial heterogeneity in employment decisions among women. By investigating the impact of individual, household, and local economic and cultural characteristics on the labor market outcomes of different groups of women, we get a more comprehensive picture of their self-employment decision. Recognizing self-employment as one of multiple labor market choices, we use multinomial logit and two confidential, geocoded micro-level datasets to study women`s career choices in urban areas. We find that the effects of various push and pull factors differ between married and unmarried women. In particular, more progressive gender attitudes pull married women into self-employment, while household burdens associated with children push them into self-employment. For unmarried women, the local business climate and individual characteristics have the strongest influence. In both cases, the motivations for women are quite different than men.

Over time, women's share of self-employment and their self-employment numbers have been steadily increasing. Between 1975 and 1995, women's total self-employment numbers more than doubled, increasing from 1.5 million to over 3.4 million (US BLS CPS). Although self-employed men still outnumber self-employed women, women gained ground and even surpassed men in terms of their self-employment growth rates. Women's self-employment rates grew by over 75 % between 1975 and 1995, narrowing the gender gap in self-employment over this time. However, since 1995, both men's and women's overall self-employment growth rates have been fairly stagnant. It is not until we break out women by marital status that we see that unmarried women's self-employment growth path has diverged from that of married women (and men) and has not been stagnant over this period but has continued to grow.

Use of CPS to Reach Conclusions

The authors use CPS data to describe aggregate trends in female self-employment, by marital status, between 1976 and 2014. They then analyze confidential microdata from the National Longitudinal Survey of Youth 1979 (NLSY79) to describe and explain in much greater detail the labor market decisions of women in U.S. urban areas from 1994 through 2008.

Carl Kogut, Donna Luse, and Larry Short, Female Self-employment Over Time, *Academy of Business Research Journal*; Gulfport Vol. 2 (2016): 19-31.

Summary

The women's movement of the 1960s and 1970s has significantly changed the role of women in the business world. The purpose of this study is to determine if similar changes have occurred with women's involvement in the field of entrepreneurship.

Analysis of U.S. Census Bureau, Current Population Survey (CPS) data over the past 40 years indicates that significant changes have taken place in both the involvement of women in entrepreneurship and also the nature of the women themselves who are entrepreneurs. Not only have women become more involved in entrepreneurship, they now represent almost 35% of the self-employed in the USA in 2015 in contrast to 24% in 1975. Self-employed women have become more racially diverse, more ethnically diverse, and are better educated than their counterparts of 40 years ago and of the general population today. Some of the more surprising findings are that Black women, Hispanic women and single women have not entered the field of entrepreneurship as aggressively as other women.

Use of CPS to Reach Conclusions

The authors analyze CPS self-employment data from 1975 through 2015, disaggregated by gender, age, race, ethnicity, educational attainment, and marital status.

CPS Annual Social and Economic Supplement (Census Bureau)

Areendam Chanda and Bulent Unel, Do Attitudes Toward Risk Taking Affect Entrepreneurship? Evidence from Second-generation Americans, October 2019.

Summary

This paper empirically investigates the impact of willingness to take risks on the likelihood of being an entrepreneur. We use twenty-four years of data on second-generation Americans from Current Population Surveys in conjunction with country level measures of willingness to take risks from Global Preference Surveys. The average level of risk taking in the country of origin is found to have a positive and significant impact on the likelihood of being an entrepreneur. A one-standard deviation increase in risk taking increases the probability of being entrepreneur by almost 15 percent. We also examine other preferences and cultural measures including trust, patience, and individualism. We find some evidence that these measures are associated with entrepreneurship. However, they tend to lose their explanatory power once one controls for risk taking.

Use of ASEC to Reach Conclusions

The individual level data used to examine occupational choices are drawn from the Annual Social and Economic (ASEC) files of the Current Population Survey (CPS). Surveys are publicly available at the Integrated Public Use Micro Samples (IPUMS) website (Ruggles et al., 2019). Our analysis uses repeated cross-section data that cover 50 states and D.C. from 1995 to 2019. The ASEC survey includes information about individuals' gender, race, age, education, nativity (including their parents) as well as their current and prior year worker class for their major job, industry where they work/worked, etc. The survey classifies individuals as wage and salary workers or self-employed, and the latter are further classified as incorporated and unincorporated.

As discussed, our analysis uses two samples: immigrants and second-generation Americans (i.e., U.S.-born individuals whose parents are foreign-born). For the former, we only include immigrants whose birthplace is the same as their parents. Thus, immigrants whose birthplace is different from their parents and immigrants whose parents are from different countries are excluded from this sample. The second sample includes U.S.-born individuals whose parents are from the same foreign country, and thus excludes the second-generation Americans whose parents are from different countries.

Kyung Min Lee, Health Insurance and the Supply of Entrepreneurs: Evidence from the ACA Medicaid Expansion, October 2019.

Summary

I examine whether the expansion of Medicaid eligibility under the Affordable Care Act increases the supply of entrepreneurs as measured by self-employment. Using the 2003–2017 Current Population Survey and focusing on childless adults in low-income households, I apply difference-in-differences, propensity score weighting, and instrumental variable (IV) methods. I find that expanding Medicaid eligibility raises the self-employment rate by 0.8 to 1.6 percentage points, without increasing self-employment exit. IV estimates imply that covered individuals have 8 to 11 percentage points higher probability to become self-employed. In the analysis of policy heterogeneity, I find evidence that the underlying mechanism of the effect was through the reduction of entrepreneurship lock. The results suggest that limited access to health insurance may be a barrier to entrepreneurship.

Use of ASEC to Reach Conclusions

Exploiting the geographic and time variation created by state policy implementation, I compare self-employment outcomes between expansion and non-expansion states, before and after the states' adoption of the ACA Medicaid expansion. I focus on low-income childless adults—the group experiencing the largest expansion in the Medicaid eligibility. My data are from the Annual Social and Economic Supplement of the Current Population Survey (CPS) from 2003 to 2017. My main outcome variable is the probability of self-employment, but I also analyze self-employment entry and exit, which may be useful in reflecting potential negative income effects. Exploiting the sample rotation design of the CPS, I link individuals across years and create two-year panels with a large number of observations to capture self-employment transitions.

Geographic Profile of Employment and Unemployment (BLS)

Mary Ellen Benedict and Irinah Hakobyan, Regional Self-Employment: The Effect of State Push and Pull Factors, *Politics and Policy*, vol. 36, n. 2, April 2008.

Summary

This study investigates the effects of business cycle, tax, and self-employment policy on the self-employment rate at the state level. Using 1999-2002 data, the results of the analysis suggest that the recession-push hypothesis of self-employment is not supported for the United States. In fact, it appears as though the prosperity-pull hypothesis is a more likely candidate for explaining self-employment and business cycle changes. Further, the Self-Employment Assistance Program, based on the recession-push hypothesis, appears to have a negative effect on the self-employment rate. These results suggest that current programs to promote self-employment have incorrectly focused on the unemployed, and/or have not prepared participants for self-employment. The study also examines the tax effect on self-employment and finds only weak support for the tax avoidance hypothesis.

Use of GP to Reach Conclusions

To measure the business cycle and tax effects on self-employment, we use panel data for U.S. states. Our data cover the period of 1999-2002. We selected this time frame because of the availability of detailed state data necessary for the analysis. Two potential dependent variables were considered as measures of self-employment: the self-employment rate (defined as the number of self-employed people over the civilian labor force) and the self-employment ratio (the number of self-employed divided by the population). We focus ultimately on the self-employment rate rather than ratio, because the regressions with both rate and ratio give similar results, and the rate measures the number of self-employed as a fraction of the economically active part of the population, which can be more informative when observing self-employment as a function of business cycle variables. The number of self-employed as well as the labor force and population statistics is provided by the BLS (Geographic Profile of Employment and Unemployment). We also concentrate on the nonagricultural sector because, compared with workers in other industries, those in agriculture face a unique economic environment and decision-making process concerning whether or not to be self-employed.

I. The Self-Employed, Entrepreneurs, and Small Businesses – Counts and Characteristics

C. Personal Histories -- Context for Self-Employment and Entrepreneurship

National Longitudinal Survey of Youth 1979 (BLS)

Robert W. Fairlie, "Self-employment, entrepreneurship, and the NLSY79," *Monthly Labor Review*, February 2005.

Summary

In the sections that follow, this article presents estimates of self-employment from the NLSY79, reviews findings from previous studies that used the survey, and discusses some of the merits of the data sets making up the survey.

Self-employment rates increase rapidly as the NLSY79 cohort ages. The self-employment rate is defined as the fraction of workers that is self-employed. At age 22, only 5.1 percent of men and 2.6 percent of women are self-employed. By age 42, however, 12.1 percent of men and 9.8 percent of women are self-employed.

Use of NLSY79 to Reach Conclusions

The NLSY79 is an excellent source of data for conducting research on self-employment and entrepreneurship. The wealth of information available in the survey allows one to build rich empirical models of the entrepreneurial process. Measures of previous wage and salary, self-employment, and unemployment experience can be created, and the NLSY79 contains several uncommon variables, such as those associated with detailed asset categories, family background information, data on criminal activities, Armed Forces Qualification Test (AFQT) scores, and psychological characteristics. Furthermore, a plethora of measures of the dynamics of self-employment may be extracted from the longitudinal data in the survey. For example, measures of transitions to and from self-employment, number of years of self-employment, and whether an individual ever tries self-employment can easily be created. Finally, the returns to self-employment, measured as earnings, job satisfaction, net worth, or other outcomes, can be estimated. Changes over time in labor market status can be used to identify the effects of self-employment, potentially removing biases created by unobserved heterogeneity across individuals.

Yasuyo Abe, Hannah Betesh, and A. Rupa Datta, "A Longitudinal Analysis of Early Self-employment in the NLSYs," for the Office of Advocacy, Small Business Administration, August 2010

Summary

While the existing literature on self-employment offers a wealth of information on the characteristics of self-employed workers at a single point in time, to date few studies have taken workers' patterns of self-employment as their unit of analysis. Few studies describe how involvement in self-employment is changing for the new generation of workers. The purpose of this research is to provide policy-relevant analysis of the characteristics and career paths of those Americans who have chosen self-employment. Specifically, this study will (a) provide new empirical findings regarding the dynamics of self-employment that underpin individual entrepreneur-ship during early adult work life; and (b) document generational changes in self-employment patterns in early adult work life between two cohorts born in the second half of the 20th century.

Use of NLSY79 to Reach Conclusions

To address these research issues, this study utilizes two National Longitudinal Surveys of Youth, the 1979 Cohort (NLSY79) and the 1997 Cohort (NLSY97), which offer extensive information on economic activity, as well as data on personal and family backgrounds, and allow detailed longitudinal investigation of self-employment activities.

Gustavo Manso, "Experimentation and the Returns to Entrepreneurship," March 2016.

Summary

Using longitudinal data, I find patterns that are consistent with entrepreneurship as experimentation: entrepreneurship spells are short; the probability of abandoning entrepreneurship is higher after bad performance; and failed entrepreneurs are not punished when they return to the salaried workforce.

Lifetime earnings computed from longitudinal data incorporate the value of the options embedded in entrepreneurship. Once the value of these options are taken into account returns to entrepreneurship are more attractive than suggested by previous research. Successful entrepreneurs earn significantly more than salaried workers with similar characteristics, while failed entrepreneurs are able to quickly move back to the salaried workforce limiting their losses. The option to abandon entrepreneurship increases the return and reduces the risk faced by entrepreneurs.

This view of entrepreneurship as experimentation and real options flips the interpretations of some of the previous findings. High variance in cross-sectional self-employed earnings, as found in previous research, is actually valuable for entrepreneurs since this variance increases the

value of their real options. Failed entrepreneurs will quickly abandon entrepreneurship and variance in cross-sectional earnings will not be reflected in lifetime earnings.

Use of NLSY79 to Reach Conclusions

To test the predictions of the model, I use the National Longitudinal Survey of Youth-1979 (NLSY79). From the NLSY79, I obtain information on demographics, educational attainment, labor market outcomes, and prelabor market traits. The main advantage of the NLSY79 is that it follows individuals over time, allowing one to compute the lifetime returns to self-employed and salaried workers.

Deepak Hegde and Justin Tumlinson, "Asymmetric Information and Entrepreneurship," October 2018.

Summary

Why do individuals become entrepreneurs? We argue that information asymmetries and the quest to maximize pecuniary returns produce entrepreneurs. In our model, individuals signal their hidden ability to employers (e.g., via educational qualifications). However, signals are imperfect and individuals with greater ability than their signals convey to employers become entrepreneurs. Empirical analysis of two longitudinal samples of U.S. and U.K. residents supports the model's predictions that (i) entrepreneurs have higher ability than employees with comparable signals, (ii) employees have better signals than equally able entrepreneurs, and (iii) entrepreneurs' earnings are higher and exhibit greater variance than employees with similar signals.

Use of NLSY79 to Reach Conclusions

We test these theoretical predictions using data drawn from the nationally representative National Longitudinal Survey of Youth (NLSY), first administered to 12,686 individuals born between 1957 and 1964, and resident in the U.S. in 1979. The NLSY provides a detailed record of their education and work histories to the present. Analyzing this sample, we find that those who become self-employed (or entrepreneurs) scored higher on cognitive ability tests administered to them as adolescents than employees with similar educational credentials, our proxy for observable signals. Despite their higher ability scores, the self-employed have lower academic credentials. In fact, the larger the gap between an individual's own ability and the median ability of individuals with his same academic credentials, the more likely he is to choose entrepreneurship. The median self-employed worker earns 7.3 percent more than the comparably educated wage-employee, and entrepreneurial earnings have higher variance. These empirical differences between the self-employed and wage-employed prevail for both self-employed workers who incorporate their businesses as well as those who do not, with the results on income and wealth differences being particularly stark for incorporated entrepreneurs—those most likely to be residual claimants of high growth enterprises.

National Longitudinal Survey of Youth 1997 (BLS)

Yasuyo Abe, Hannah Betesh, and A. Rupa Datta, "A Longitudinal Analysis of Early Self-employment in the NLSYs," for the Office of Advocacy, Small Business Administration, August 2010

Summary

While the existing literature on self-employment offers a wealth of information on the characteristics of self-employed workers at a single point in time, to date few studies have taken workers' patterns of self-employment as their unit of analysis. Few studies describe how involvement in self-employment is changing for the new generation of workers. The purpose of this research is to provide policy-relevant analysis of the characteristics and career paths of those Americans who have chosen self-employment. Specifically, this study will (a) provide new empirical findings regarding the dynamics of self-employment that underpin individual entrepreneur-ship during early adult work life; and (b) document generational changes in self-employment patterns in early adult work life between two cohorts born in the second half of the 20th century.

Use of NLSY97 to Reach Conclusions

To address these research issues, this study utilizes two National Longitudinal Surveys of Youth, the 1979 Cohort (NLSY79) and the 1997 Cohort (NLSY97), which offer extensive information on economic activity, as well as data on personal and family backgrounds, and allow detailed longitudinal investigation of self-employment activities.

Viktoriya Nikolov and Michael S. Bargar, "Determinants of Self-Employment in the United States," *Undergraduate Economic Review*, vol. 6, issue 1, 2010.

Summary

The prominence entrepreneurs have occupied in the popular imagination belies their relative neglect in formal economic theory. This paper adds to the growing body of work on entrepreneurs by examining the characteristics of self-employed individuals in the National Longitudinal Survey of Youth 1997. We believe our article to be the first that uses this fresh body of data for this purpose. Employing the standard binomial probit model with a list of potentially significant variables drawn from existing literature, we discovered that women are significantly less likely to be self-employed than men.

Use of NLSY97 to Reach Conclusions

We contribute to the available literature by being the first to look at individual data from a source new to this line of research, the National Longitudinal Survey of Youth 1997. Here we examine the difference in levels of self-employment between men and women and how a set

of variables such as the respondent's level of education, presence of children and region of residence, affect the choice of whether to become self-employed.

Survey of Income and Program Participation (Census)

Julia Beckhusen, "Employment Transitions among the Self-Employed during the Great Recession," U.S. Census Bureau, August 2014.

Summary

Entrepreneurs base their decision to start a business on a range of factors, from age, education and assets to macroeconomic conditions. While the majority of these factors have a well-understood impact on entering and exiting self-employment, the effect of macroeconomic conditions is less clear. During periods of recession, self-employment may increase due to its attractiveness as an alternative to unemployment. However, the difficulty of maintaining a business through the downturn can lead to a decrease in the self-employed. Understanding the transitions in and out of self-employment would help us better appreciate how entrepreneurs experience recessions.

The research results suggest that the probability of entering self-employment depends on characteristics of the individual while movements out of self-employment are contingent on characteristics of the business. Furthermore, transitions from unemployment to self-employment increased during the recession months and transitions from self-employment to wage-work increased in the post-recession months.

Use of SIPP to Reach Conclusions

We use a robust set of longitudinal data from the Survey of Income and Program Participation (SIPP) to analyze the movements between self-employment, unemployment and wage-work during the Great Recession.

Magnus Lofstrom, Timothy Bates, and Simon Parker, "Why Are Some People More Likely to Become Small-Businesses Owners than Others: Entrepreneurship Entry and Industry-Specific Barriers," *The Journal of Business Venturing*, 29:2 (2014).

Summary

We classify industries using measures of entry barriers and proceed to investigate how determinants of entry vary in high- as opposed to low-barrier fields. The wealth and educational background characteristics potential entrepreneurs possess, we found, predispose them to make distinctly different industry choices, both because of the differing rewards available to them and the very different entry barriers they face. The characteristics of potential entrants, in other words, draw them toward some industries and away from others.

The top 2 quintiles of the personal wealth distribution consistently predict entry into high-barrier lines of business. Thus, across a wide range of the distribution, wealth appears to alleviate borrowing constraints, facilitating entry into high-barrier lines of business. College-education level strongly predicts entry into high-barrier industries, yet the opposite outcome

describes low-barrier fields. College graduates positively select into industries where entry barriers elevate the expected earnings of firm ownership, while steering clear of low-remuneration fields. We conclude that industry context heavily shapes impacts of entrant resource endowments on entrepreneurial entry choices.

Use of SIPP to Reach Conclusions

This study utilizes data drawn from the 1996 and 2001 panels of the U.S. Census Bureau's Survey of Income and Program Participation (SIPP). Entrepreneurship is operationalized as small business ownership. Our sample includes adults aged between 25 and 59, irrespective of gender and labour force status. The age restriction is imposed to minimise confounding issues relating to schooling and retirement decisions. We also restrict our sample to individuals for whom wealth information is available, who are observed in the SIPP at least twice in consecutive years, and whose original state was not self employment. These restrictions are necessary because our analysis of transitions into small-firm ownership is based on changes in year-over-year labour market states.

Caroline Bruckner and Thomas L. Hungerford, "Failure to Contribute: An Estimate of the Consequences of Non- and Underpayment of Self-Employment Taxes by Independent Contractors and On-Demand Workers on Social Security," Boston College, January 2019.

Summary

While existing academic and government research has focused on the size, growth trajectory, and labor and tax law implications of independent contractors, freelancers, and workers selling goods and services online and through app-based platforms (the "on-demand" economy), less work has been devoted to quantifying the Social Security implications for the on-demand economy and its workers. Although it is known that self-employed workers have tax compliance and reporting issues, the existing reporting rules applicable to most workers earning income in the on-demand economy substantially increase the likelihood that these taxpayers are failing to contribute to Social Security and Medicare through payment of the self-employment tax (SE tax). As such, this paper sheds light on the Social Security implications of current federal tax rules for independent contractors generally and, in particular, workers earning income through occupations occurring in the on-demand economy.

Uses of SIPP to Reach Conclusions

By analyzing 2014 SIPP data, we identify a population of self-employed, non-employer respondents working outside of a traditional employment relationship ("independent contractors"), as well as individuals working in occupations in the on-demand economy ("on-demand workers"). SIPP data have the potential to capture workers who earn income using on-demand platforms to connect with customers and process payments ("on-platform work"), as well as workers who earn income in occupations occurring in the on-demand economy who do not use on-demand platforms ("off-platform work").

Additionally, with SIPP data, we are able to estimate the income that independent contractors and on-demand workers earned in these employment relationships in 2014. In addition, using SIPP, we were able to provide supplemental demographic data on independent contractors and the on-demand platform workforce.

II. Business Development Flows

A. Business Applications and Formations

Business Formation Statistics (Census Bureau)

In light of the newness of this data source, no published reports or articles were found that use the BFS to highlight trends in entrepreneurship.

II. Business Development Flows

B. Openings, Expansions, Contractions, and Closing

Business Employment Dynamics (BLS)

ME Canon, Y Liu, Firm Size and Employment Dynamics, *Economic Synopses*, St. Louis Federal Reserve Bank, 2015

Summary

Small employers (firms with fewer than 50 employees) have been fundamental to employment growth in the United States, particularly during recessions. In their 2012 study, Moscarini and Postel-Vinay use 1978-2009 data and find that job flows at large employers are more cyclical. In particular, large firms, on net, destroy proportionally more jobs than small employers when unemployment is above trend (i.e., late in a recession and immediately after a recession) and create more jobs when unemployment is below trend (i.e., late in a typical expansion). The authors find that this disparity is due to different patterns of entry and exit for large and small employers and that it also holds for employers of different ages.

Use of BED to Reach Conclusions

The authors analyze 1993-2013 BED job gains and losses data by firm size.

John Haltiwanger, Top ten signs of declining business dynamism and entrepreneurship in the US, University of Maryland, 2015

Summary

The U.S. has exhibited a substantial and pervasive decline in measures of business dynamism, entrepreneurship and labor market fluidity in the last several decades. We have learned this through the relatively recent development of comprehensive longitudinal business databases tracking the private, non-farm sector of the U.S. Numerous studies have documented the decline and explored its causes and consequences. In this short synopsis, the basic facts of this decline are summarized by highlighting the top ten signs of the decline.

Use of BED to Reach Conclusions

The authors analyze 1992-2015 BED data to identify multiple trends in business dynamism (startups and early growth). They also analyze the Census Bureau's Business Dynamics Statistics (BDS) for the same purpose.

Business Dynamic Statistics (Census Bureau)

N Goldschlag, E Perlman, Business Dynamic Statistics of Innovative Firms, 2017.

Summary

A key driver of economic growth is the reallocation of resources from low to high productivity activities. Innovation plays an important role in this regard by introducing new products, services, and business methods that ultimately lead to increased productivity and rising living standards. Traditional measures of innovation, particularly those based on aggregate inputs, are increasingly unable to capture the breadth and depth of innovation in modern economies.

Use of BDS to Reach Conclusions

In this paper, we describe an effort at the US Census Bureau, the Business Dynamics Statistics of Innovative Firms (BDS-IF) project, which aims to address these challenges by extending the Business Dynamics Statistics data to include new measures of innovative activity. The BDS-IF project will produce measures of firm, establishment, and employment flows by firm age, firm size, and industry for the subset of firms engaged in activities related to innovation. These activities include patenting and trademarking, the employment of STEM workers, and R&D expenditures. The exibility of the underlying data infrastructure allows this measurement agenda to be extended to include copyright activity, management practices, and high growth firms.

J Haltiwanger, RS Jarmin, J Miranda, Business dynamics statistics briefing: Where have all the young firms gone?, SSRN 2049909, 2012.

Summary

The Census Bureau's Business Dynamics Statistics (BDS) provides data on business dynamics for U.S. firms and establishments with paid employees. This briefing highlights some key features of the most recent BDS update, which now has data through 2010. As the most complete public-use dataset allowing for the analysis of business dynamics in the United States, the BDS is a key source of knowledge about the changing state, as well as the national, economy.

The new BDS data release shows that, in 2010, 394,000 startups created 2.3 million jobs (these were not simply establishment openings but new firms whose establishments also were new to the economy). This reflects substantial job creation in a time of anemic overall economic activity. Over the same period from March 2009 to March 2010, the net job creation from all U.S. private sector firms was - 1.8 million jobs. Without the contribution of business startups, the net employment loss would have been substantially greater.

Use of BDS to Reach Conclusions

A potentially troubling trend identified from earlier BDS releases is that the pace of business startups has exhibited a long-run decline that dates back to the 1980s. The newly released BDS shows that this trend has continued through 2010.

Statistics of US Businesses (Census Bureau)

JB Bailey, DW Thomas, Regulating away competition: The effect of regulation on entrepreneurship and employment, *Journal of Regulatory Economics*, 2017.

Summary

Many scholars have worried that regulation deters entrepreneurship because it increases the cost of entry, reduces innovation in the regulated industry, and benefits large firms because they can overcome the costs of complying with regulations more easily than smaller firms. Using novel data on the extent of US federal regulations by industry and data on firm births and employment from the Statistics of US Businesses, we run fixed effects regressions to show that more-regulated industries experienced fewer new firm births and slower employment growth in the period 1998–2011. Large firms may even successfully lobby government officials to increase regulations to raise their smaller rivals' costs. We also find that regulations inhibit employment growth in all firms and that large firms are less likely to exit a heavily regulated industry than small firms.

Use of SUSB to Reach Conclusions

To quantify the effect of regulation on firms of different sizes and employment growth, we use industry-level data on firms from the Statistics of US Businesses (SUSB), together with RegData's index of regulatory intensity and several control variables. Our sample contains data from 215 industries for 1997–2011.

The SUSB is compiled annually by the US Census Bureau using data on the full population of US firms—it is not simply a sample subject to sampling error. We use the dynamic version of the SUSB maintained by the Office of Advocacy of the US Small Business Administration. While several datasets such as the static SUSB and the Quarterly Census of Employment and wages provide counts of companies and employment by industry at a point in time, the dynamic SUSB provides information on how these variables are changing each year. It tracks the number of new firms in each industry (firm births), the number of firms exiting each industry (firm deaths), and the number of employees hired and fired for each industry. Key variables from the SUSB are summarized in Table 1. The data identify industries down to four-digit North American Industry Classification System (NAICS) codes. The NAICS breaks down industries to progressively greater levels of detail, starting with the two-digit level, such as 31 (manufacturing). Three-digit codes dig deeper, with industry classifications such as 311 (food manufacturing). Four-digit codes provide still greater detail, with industry classifications such as 3111 (animal food manufacturing) and 3112 (grain and oil seed milling). The SUSB describes 290 four-digit industries.

B Headd, B Kirchhoff, The growth, decline and survival of small businesses: An exploratory study of life cycles, *Journal of Small Business Management*, September 2009.

Summary

Explorative findings showed growing firms as generally a constant share of the economy with a minor business cycle effect, growing firms outnumbering decliners, new firms not growing much, and fast growers in a given year tending to revert to the mean in the following year. The findings are presented as an opening statement, but are far from final salvo into the discussion of small-firm growth.

Use of SUSB to Reach Conclusions

Using U.S. Census Bureau special tabulations, we follow a cohort of small (single establishment) firms formed in 1992 and a cohort of fast growers to 2002 to track their employment changes.

Longitudinal Employer-Household Dynamics (Census Bureau)

T Babina, P Ouimet, R Zarutskie, Going entrepreneurial? IPOs and new firm creation, SSRN, 2017.

Summary

Using matched employee-employer US Census data [from LEHD], we examine the effect of a successful initial public offering (IPO) on employee departures to startups. Accounting for the endogeneity of a firm's choice to go public, we find strong evidence that going public induces employees to leave for start-ups. Moreover, we document that the increase in turnover following an IPO is driven by employees departing to start-ups; we find no change in the rate of employee departures for established firms. We present evidence that, following an IPO, many employees who received stock grants experience a positive shock to their wealth which allows them to better tolerate the risks associated with joining a startup or to obtain funding. Our results suggest that the recent declines in IPO activity and new firm creation in the US may be causally linked. The recent decline in IPOs means fewer workers may move to startups, decreasing overall new firm creation in the economy.

Use of LEHD to Reach Conclusions

We add worker-level data using the Longitudinal Employer-Household Dynamics (LEHD) data, This database tracks employers, employees, and their earnings on a quarterly basis. The LEHD data also allow us to observe the age, gender, race, and place of birth of each employee. We link the LEHD to firm identifiers in the LBD using the employer identification number (EIN). The LEHD data are collected from the unemployment insurance records of states participating in the program.4 Data start in 1990 for several states and the number of states included increases over time. The data coverage ends in 2008. Our project has access to data from 31 states.

C Goetz, H Hyatt, E McEntarfer, K Sandusky, The promise and potential of linked employer-employee data for entrepreneurship research, NBER, 2015

Summary

In this paper, we highlight the potential for linked employer-employee data to be used in entrepreneurship research, describing new data on business start-ups, their founders and early employees, and providing examples of how they can be used in entrepreneurship research. Linked employer-employee data provides a unique perspective on new business creation by combining information on the business, workforce, and individual. By combining data on both workers and firms, linked data can investigate many questions that owner-level or firm-level data cannot easily answer alone - such as composition of the workforce at start-ups and their role in explaining business dynamics, the flow of workers across new and established firms, and the employment paths of the business owners themselves.

Use of LEHD to Reach Conclusions

Specifically, our goals in this paper are threefold:

(1) To familiarize researchers with the U.S. linked employer-employee data and how it can be used in entrepreneurship research;
(2) To describe newly available public use statistics derived from linked employer-employee data and provide examples of how it can be used to study entrepreneurship; and
(3) To outline future plans to expand the set of available data to study entrepreneurship by linking in new administrative data sources on self-employment and partnerships, as well as identifying the employment history and human capital formation of entrepreneurs themselves.