

# Using the Kauffman Firm Survey (KFS) to Study Gender and Entrepreneurship

UN Technical Meeting on Measuring Entrepreneurship from a Gender Perspective, December 6<sup>th</sup>, 2013

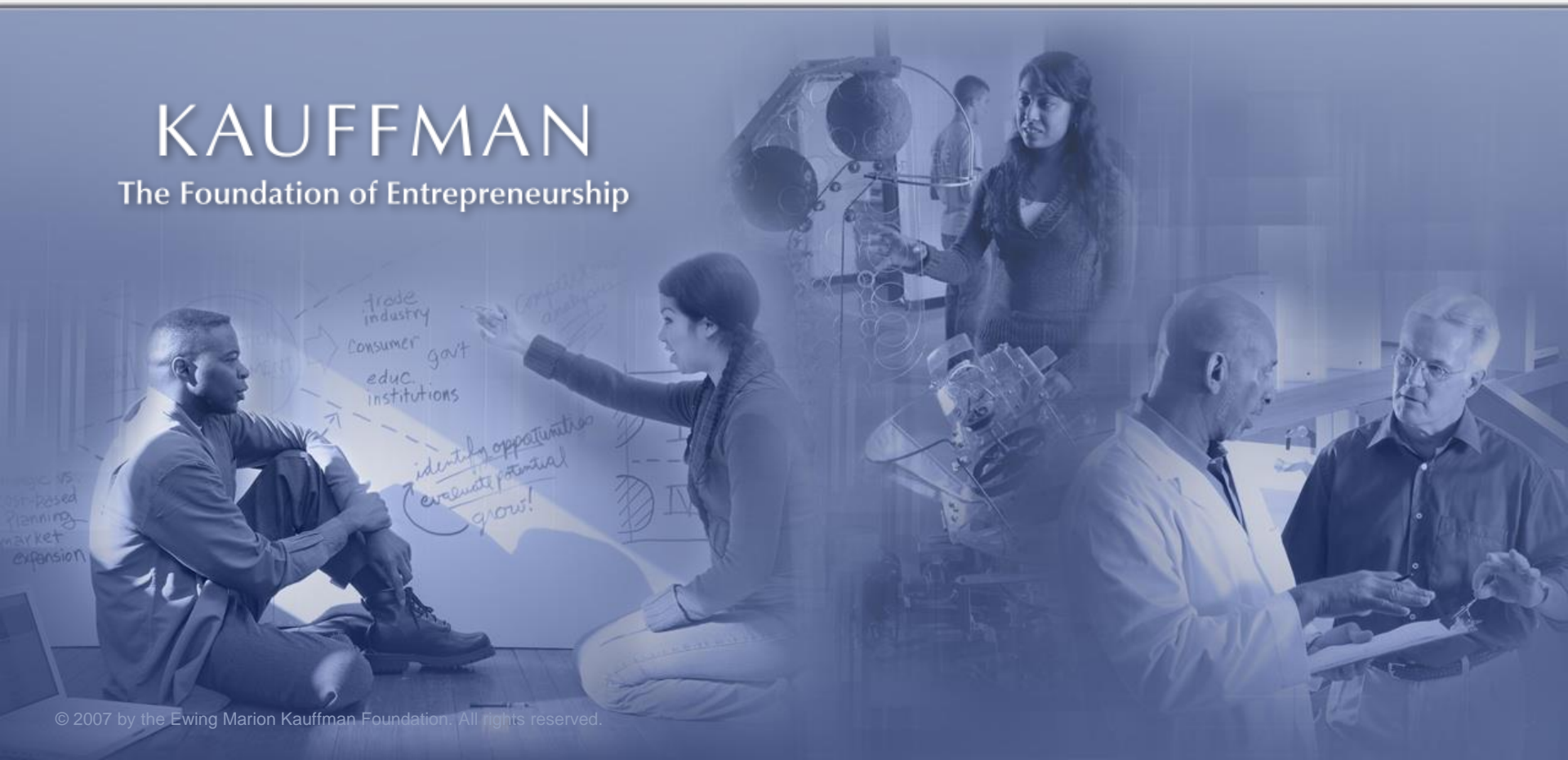
Alicia Robb, Kauffman Foundation

“Every individual that we can inspire, that we can guide, that we can help to start a new company, is vital to the future of our economic welfare.”

— Ewing Kauffman

## KAUFFMAN

The Foundation of Entrepreneurship



# Census Data: Survey of Business Owners

## U.S. Nonfarm Firms by Gender and Ownership, 1997, 2002, and 2007

As Published		All Firms		Firms with paid Employees			
		Firms (Number)	Receipts (Millions of dollars)	Firms (number)	Receipts (millions of dollars)	Employees (number)	Annual payroll (millions of dollars)
<b>Women-owned firms</b>							
	2007 <sup>1</sup>	7,793,425	1,192,781	911,285	1,010,470	7,587,020	218,136
	2002 <sup>1</sup>	6,489,483	940,775	916,768	804,097	7,146,229	173,709
	1997 <sup>2</sup>	5,417,034	818,669	846,780	717,764	7,076,081	149,116
	growth 2002-2007	20.1%	26.8%	-0.6%	25.7%	6.2%	25.6%
	growth 1997-2002	19.8%	14.9%	8.3%	12.0%	1.0%	16.5%
	growth 1997-2007	43.9%	45.7%	7.6%	40.8%	7.2%	46.3%
<b>All U.S. firms</b>							
	2007 <sup>1</sup>	27,110,362	30,181,461	5,752,975	29,208,766	118,668,699	4,886,977
	2002 <sup>1</sup>	22,974,685	22,627,167	5,524,813	21,859,758	110,786,416	3,813,488
	1997 <sup>2</sup>	20,821,934	18,553,243	5,295,151	17,907,940	103,359,815	2,936,493
	growth 2002-2007	18.0%	33.4%	4.1%	33.6%	7.1%	28.1%
	growth 1997-2002	10.3%	22.0%	4.3%	22.1%	7.2%	29.9%
	growth 1997-2007	30.2%	62.7%	8.6%	63.1%	14.8%	66.4%
<b>Percent of total U.S. firms</b>							
	Women-owned (2007)	28.7%	4.0%	15.8%	3.5%	6.4%	4.5%
	Women-owned (2002)	28.2%	4.2%	16.6%	3.7%	6.5%	4.6%

# Survey of Business Owners (2007)

	Average receipts	Average Employment	Average payroll
Female-owned	\$ 153,567	8.3	\$ 28,547
Male-owned	\$ 609,918	12.7	\$ 36,794
Equally male-/female-owned	\$ 276,967	7.7	\$ 26,747
Publicly held and other firms	\$ 23,910,915	111.2	\$ 47,496
All	\$ 1,108,464	20.5	\$ 41,112

# 2012 Survey of Business Owners Tentative Release Schedule

## Data Products

## Tentative Release Dates

■ Women-Owned Businesses	June 2015
■ Company Summary	December 2015
■ Characteristics of Businesses	December 2015
■ Characteristics of Business Owners	December 2015

- VERY DATED BY THE TIME DATA ARE RELEASED
- REPEATED CROSS SECTIONS, SURVIVAL BIAS
- DETAILED MICRODATA DIFFICULT TO ACCESS

# Motivation...Panel Study of Entrepreneurial Dynamics


- PSED is the only source of data on business formation for a sample of firm founders representative of an underlying population that has information on individual and firm-level variables that match theories used by entrepreneurship scholars

<http://www.psed.isr.umich.edu/>

- PSED is great, but...
  - A surprisingly small portion of people “in the process of starting a business” manage to create a business
  - Potentially high performing businesses are not evenly distributed across industries
  - Financing activities are an important part of firm formation, but we don't have a lot of data on them


# The Approach: Eligibility Screening

The Dun and Bradstreet (D&B) database was chosen as the sampling frame source [251,282].




## Include

Businesses that were started as independent business, or by the purchase of an existing business, or by the purchase of a franchise in the 2004 calendar year



## Exclude

Businesses that were started as a branch or a subsidiary owned by an existing business, that were inherited, or that were created as a not-for-profit organization in the 2004 calendar year



## Include

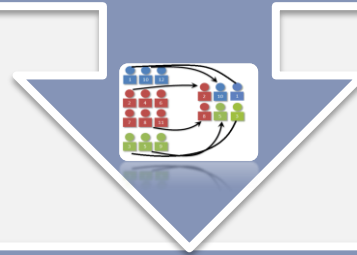
Businesses that have a valid business legal status (sole proprietorship, limited liability company, subchapter S corporation, C-corporation, general partnership, or limited partnership) in 2004

# The Approach: Eligibility Screening

## Include

Businesses that have at least one of the following activities:

1. Acquired employer identification number during the 2004 calendar year;
2. Organized as sole proprietorships reporting that 2004 was the first year they used Schedule C or Schedule C-EZ to report business income on a personal income tax return;
3. Reported that 2004 was the first year they made state unemployment insurance payments;  
or
4. Reported that 2004 was the first year they made federal insurance contribution act payments.



Stratified sample of 32,469 businesses was selected. Oversampling on technology status :Three technology strata: high-tech, medium-tech, and non-tech businesses

4,928 businesses as the final sample of eligible businesses

**Table 2. Estimated Count and Sample for the Seventh Follow-Up Survey from the D&B Database: Businesses That Had a 2004 Start Date and Completed the Baseline Survey**

Technology Sampling Strata	Woman-Ownership Sampling Strata	Estimated Number of New Businesses in 2004 <sup>a</sup>	Sample for Seventh Follow-Up Survey <sup>b</sup>
<b>Total</b>		<b>73,278</b>	<b>4,928</b>
High-technology		1,313	705
	Yes	190	103
	No	1,123	602
Medium-technology		9,675	1,329
	Yes	2,026	271
	No	7,649	1,058
Non-technology		62,290	2,894
	Yes	14,366	513
	No	47,924	2,381



# KFS Milestones

- 2004 Survey development began
- 2005 Large pilot test
- 2005-06 Baseline Survey conducted
- 2006-12 Annual Follow-Up Surveys conducted
- Complete survey covers calendar years 2004-2011 (8 year panel)

# KFS Milestones

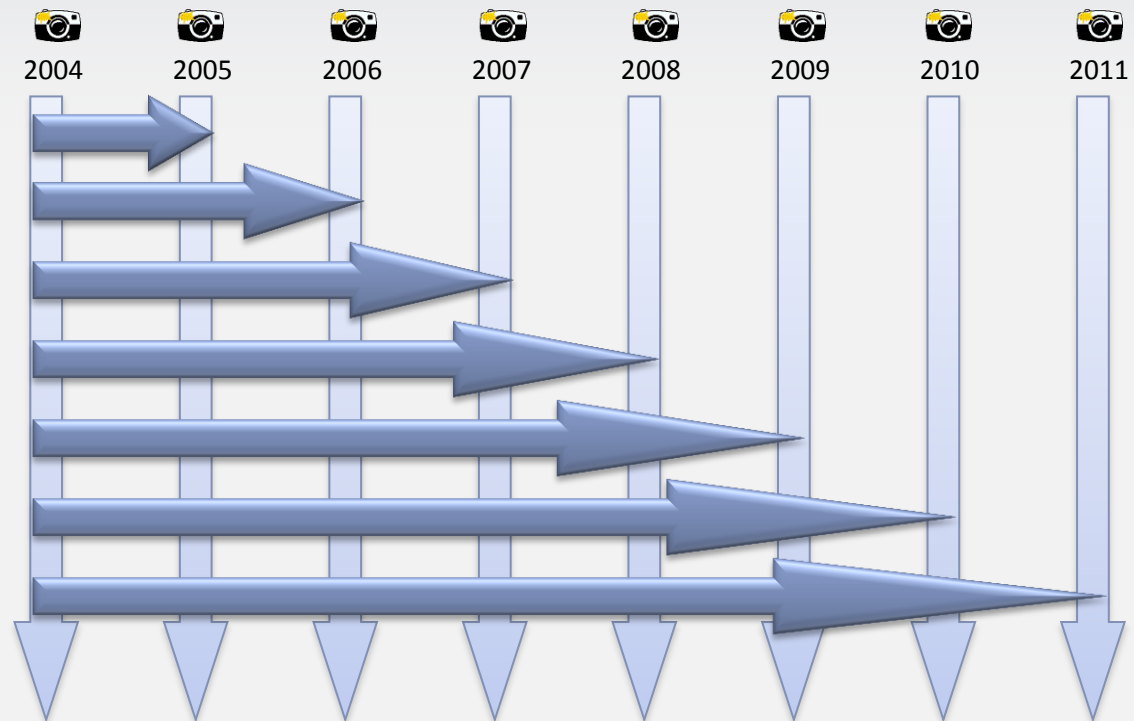
## Cross Sectional

Status / Survey year	2004	2005	2006	2007	2008	2009	2010	2011
Refusal		561	743	900	865	794	816	676
Stopped operation or sold or merged in any of previous follow-ups			303	671	1,015	1,399	1,685	1,941
Out of Business		260	321	299	344	250	218	234
Merged Or Sold		43	47	45	40	36	38	40
Temporarily Stopped		66	124	98	58	41	45	30
Complete	4,928	3,998	3,390	2,915	2,606	2,408	2,126	2,007
Total	4,928	4,928	4,928	4,928	4,928	4,928	4,928	4,928

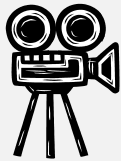
# KFS Milestones

The KFS longitudinal data can be analyzed, either as a cross-section or longitudinally.

## Static : Cross-section



Change :  
Longitudinal



# Detailed Data: Baseline & Over Time

## Business Characteristics

- Information about business legal status, business's primary industry, number of full- and part-time employees, number of owners, number of owner-operators, and primary location of the business

## Strategy and Innovation

- Product/Service Offerings, Intellectual Property, Licensing In and Licensing Out, R&D, Intangible Asset spending, comparative advantage, sources of comparative advantage. Product/Service Offerings, Intellectual Property, Licensing In and Licensing Out, R&D, Intangible Asset spending, comparative advantage, sources of comparative advantage, percentage of sales to individuals, government, and other businesses

## Business Organization and Human Resource Benefits

- Types of Benefits Offered (health insurance, retirement plan, stock options, bonus plan, tuition reimbursement, paid vacation, paid sick days, flex time and other benefits ) , Task/Work Structure (human resources, sales, marketing, executive, R&D, production, general administration and financial administration)

# Detailed Data: Baseline & Over Time

## Business Finances: The most comprehensive part of the survey

- At the business level:
  - ▶ Sources and amount of business debt financing (credit cards, loans from a commercial bank, line of credit, loans from a non-bank financial institution, loans from any family or friends, and loans from any other individuals)
  - ▶ Income Statement Info (Revenue, Expenses, Profits)
  - ▶ Balance Sheet Info (Assets, Liabilities, Equity),
  - ▶ Credit Market Experiences, Discouraged borrowers,
  - ▶ Trade financing, R&D spending
  - ▶ The purchase of new or used machinery or equipment, Rental or lease of buildings or machinery or equipment.

# Detailed Data: Baseline & Over Time

## Business Finances: The most comprehensive part of the survey

- At the owner level (up to 10 owners):
  - ▶ Sources and amount of equity financing (owners, spouses, parents, venture capitalists, companies, government and other sources)
  - ▶ Sources and amount of personal debt financing (personal credit cards, personal loans from a bank, business credit cards, personal loans from any family or friends,)

## Work Behaviors and Demographics of Owners

- Education, Age, Race, Ethnicity, Gender, Citizenship, Immigrant Status, Hours Worked, Previous Years of Work Experience, Previous Start-up Experience (same/different industry as this firm)

# Percentage of respondents who were female (2011)

	Sample Size	Estimate	Standard Error
All	1,892	30.2	0.99
High-tech	289	13.5	1.54
Medium-tech	589	22.3	1.25
Non-tech	1,014	32.2	1.19

Oversampled high- and medium-tech firms

# Accessing the Data: Three Avenues

- Public-use Microdata
  - Available on the web
  - Identifying features omitted
  - Can be used to get familiarity with data and some research
  - [www.kauffman.org/kfs](http://www.kauffman.org/kfs)
- Data Enclave
  - Secure server available via remote access
  - More identifying variables available
  - <http://dataenclave.norc.org/>
- Census RDCs
  - Onsite access to KFS data matched to Census data
  - Only available to scholars completing the Census procedures for researcher access
  - <http://www.census.gov/ces/>

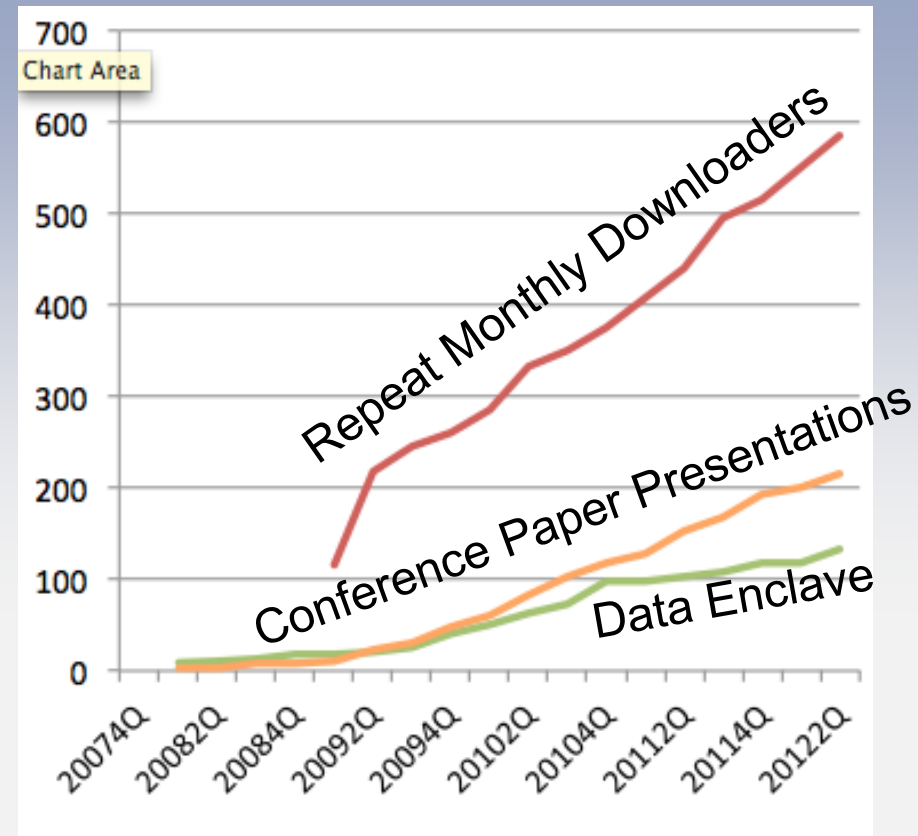


# KFS Impact : By the Numbers

Measures of Research Productivity (August 2013)	
Conference Paper Presentations	252
Research Reports/Working Papers	59
Peer-reviewed Journal Articles	30
Book Chapters	10
Books	3
Dissertations and Theses	13
Current Researchers in the Data Enclave	160
Unique Downloaders of Public Data	2554
Repeat Downloaders of Public Data	602
Conference Paper Presentations	252
Research Reports/Working Papers	59
Peer-reviewed Journal Articles	30

# KFS Project Management Looking Forward

- Support growing community
- Ensure research quality
- Summarize findings for broader audiences through other publications
- Qualitative interviews with highest growth businesses looking more deeply inside their experiences, matched to like firms that were not high growth
- Merging In LinkedIn Data, Website Data, Detailed Closure data (from an exit survey)



# Data Wiki

- The KFS Data Wiki was established by researchers for researchers as a means of building a more comprehensive understanding of how the KFS data can be used and what it means.
- <http://www.kauffman.org/kfs/KFSWiki/default.aspx>
- The Wiki is a place where the research community can post and share their own research in the [Examples](#) pages, pose questions to the community in the [Research Forum](#), or dig deeper into specific variables in the [Data Dictionary](#) page.
- Listing of upcoming events and submission deadlines

# Funding Possibilities for KFS Researchers:

- Travel Grants
- Research Grants
- Promising Paper Awards
- Enclave Seat Sponsorship

# What's Next: New data collection efforts

- Early stages of exploring KFS 2.0
- Timeline: present final plans next year to Board
- Initial thoughts
  - Multiple Cohorts (three cohorts followed over four years each)
  - Young high performing or high growth potential firms
  - Focus on financing, strategy, innovation, jobs
  - Barriers, constraints, challenges to growth
  - Women and minorities
  - Regional variations
  - Clusters

# Questions?

- [arobb@kauffman.org](mailto:arobb@kauffman.org)
- [www.kauffman.org](http://www.kauffman.org)
- [www.kauffman.org/kfs](http://www.kauffman.org/kfs)