

Response Rates and Response Patterns Among New Employers: Results from the Kauffman Firm Survey

Frank Potter, Zhanyun Zhao, Yuhong Zheng, David DesRoches, and Janice Ballou
Mathematica Policy Research (MPR), Inc., Princeton, New Jersey 08543-2393

Abstract

Economic growth is related to continued formation of new enterprises. The Ewing Marion Kauffman Foundation sponsors research into entrepreneurship and new business formation. The Kauffman Firm Survey (KFS) is a national-level longitudinal survey of new businesses conducted for the Foundation by Mathematica Policy Research (MPR). The KFS consists of an initial survey of 5,000 enterprises formed in 2004 with annual follow-up surveys to collect data on the new business characteristics and financing needed to create and sustain them. Survey data will be available to researchers to improve the understanding of the dynamics of new businesses. MPR selected the sample using D&B data files and is using web-based and CATI data collection procedures to capture the information. The purpose of this paper is to describe the characteristics and determinants of response among this elusive population.

Key Words: Nonresponse; Establishment Surveys; Respondent Characteristics, Web-Based, CATI

1. Overview

The Ewing Marion Kauffman Foundation sponsors research into entrepreneurship and new business formation. As part of this on-going effort, the Foundation sponsored the Kauffman Firm Survey, a national-level longitudinal survey of new businesses. The Foundation contracted with Mathematica Policy Research, Inc. (MPR) to design and implement the KFS. The KFS collected data from new businesses to better understand business creation and maturation, and the characteristics of the founders of new businesses. It will also investigate the longitudinal development and sustainability of firms. For the survey, the sampling frame was based on those firms listed in the Dun & Bradstreet (D&B) database with a starting date in 2004. The sampling units are “enterprise” units or the “ultimate parent” units as identified in the D&B data base. Branches or subsidiary units are excluded from the sampling frame and, if found in the sample, ineligible for the survey. The survey will establish a cohort of new

businesses and MPR will follow the cohort longitudinally over a four-year period.

Survey data will be available to researchers to improve the understanding of the dynamics of entrepreneurship and new businesses. MPR selected the sample using D&B data files and used web-based and CATI data collection procedures. The purpose of this paper is to describe the characteristics and determinants of response among this elusive population.

After an initial pilot test (Pilot Test 1) that explored business eligibility criteria, the Foundation agreed that a more extensive pilot was required to test components of the study design. The second pilot study in 2004 evaluated the instrument length and structure, the effectiveness of incentives, and the collection of data through a Web survey option. The pilot sample was drawn from the D&B database and included businesses that had been started in 2003, with an oversample of businesses in high-technology industries. Pilot Test 2 ended in April 2005.

Lessons learned during the two pilots were incorporated into the Baseline Survey, which began in July 2005 using a self-administered Web survey and Computer-Assisted Telephone Interviewing (CATI). The Baseline Survey goal was to interview 5,000 owners of businesses started in 2004 and data collection ended in June 2006 with 4,931 interviews.

1.1 Sampling Design

The sampling frame for the KFS is based on the Dun and Bradstreet (D&B) database and restricted to firms (or enterprise units) that started in 2004. Establishments that were branches or intermediary units within an enterprise were excluded from the sampling frame. The D&B database was partitioned into six sampling strata defined by industrial technology categories (based on industry designation) and gender of the owner or CEO of the firm (based on the D&B data element and additional information). The high technology strata were defined based on categorization developed by Hadlock et al. (1991). Specifically, the sampling strategy called for 3,000 interviews to be completed

in two categories of high-technology businesses and 2,000 interviews to be completed in all other industrial classifications. The industries in the technology strata are shown in Table 1.

Table 1: Technology Strata Definitions		
Technology Stratum	SIC Code	Industry
1	28	Chemicals and allied products
	35	Industrial machinery and equipment
	36	Electrical and electronic equipment
	38	Instruments and related products
		Crude Petroleum and natural gas operations
2	131	gas operations
	211	Cigarettes
	291	Petroleum refining
	299	Miscellaneous petroleum and coal products
	335	Nonferrous rolling and drawing
	371	Motor vehicles and equipment
	372	Aircraft and parts
	376	Guided missiles, space vehicles, parts
	737	Computer and data processing services
	871	Engineering and architectural services
	873	Research and testing services
	874	Management and public relations
	899	Services, not elsewhere classified
	229	Miscellaneous textile goods
	261	Pulp mills
267	Miscellaneous converted paper products	
348	Ordnance and accessories, not elsewhere classified	
379	Miscellaneous transportation equipment	
3		All other industries

The D&B data base is a compilation of data from various sources including credit bureaus, state offices where some new firms are registered and companies (such as credit card companies and shipping companies) that are likely to be used by all businesses. However, compiling information on

newly formed business enterprises is particularly difficult because there is no single business registry of new enterprises and the time between establishing the firm and the firm showing up in one of D&B sources may be 6 or more months. To capture as complete a picture of firms starting in 2004 as possible, we arranged with D&B to provide multiple files at different time points during 2005. We obtained a file in June 2005 and then a new file in November 2005. As shown in Table 2, D&B provided MPR with a file of 188,000 enterprises with the starting year of 2004 in June 2005. This number was approximately 30 percent lower than a similar file received in June 2004 of firms starting in 2003. We investigated the lower number and could find no clear changes in operations by D&B and no evidence use available from federal sources to verify or disprove this count.

Since we were going to obtain a second file in November from D&B, we needed to release a sample sufficiently large to accommodate the expected response and eligibility rates, but we wanted to balance the sample size between the two files to reduce unequal sampling weights. However, because the high technology firms numbered only 2,500 (again fewer than expected) and we wanted a large pool of these firms for the longitudinal panel, we decided to include all of these firms in the sample in order to obtain an adequate count of these firms. For the other strata, we were somewhat conservative, but still released relatively large samples.

The November D&B file included 63,000 firms with start date in 2004, resulting a total pool of 251,000 firms from the combined June and November files (Table 2). However, 13,000 firms from June file (7 percent) were not in the November file (see Table 3) and were presumed to be no longer in business (which we called deceased firms) and left only 238,000 firms in the combined data base less the deceased firms.

Because of the interest in the subsequent analysis of the high technology firms, we again decided to include all of these firms in the sample in order to obtain an adequate count of these firms for the longitudinal panel. For the other two strata, we attempted the balance the final sample across the two files and the sampling strata. The final sample size and sampling rates are in Table 4.

1.2 Data Collection Methods

To be eligible for the KFS survey, the owner of the new business surveyed had to indicate that their new business was not a new business, branch or

Table 2: Sampling Frame Of Firms In D&B Data Base: Firms With Start Date Of 2004

Technology Strata	Woman Owned	June File	Deceased In November ¹	New In November	Total All (June And November)	Operating Total
Total		188,292	13,439	62,990	251,282	237,843
High-Tech		2,593	144	1,276	3,869	3,725
	Yes	361	21	166	527	506
	No	2,232	123	1,110	3,342	3,219
Medium-Tech		22,544	926	7,117	29,661	28,735
	Yes	4,332	153	1,215	5,547	5,394
	No	18,212	773	5,902	24,114	23,341
Non Tech		163,155	12,369	54,597	217,752	205,383
	Yes	32,016	2,177	9,951	41,967	39,790
	No	131,139	10,192	44,646	175,785	165,593

¹Deceased in November is the count of enterprises in the D&B data base in June 2005 that were not in the data base in November 2005.

Table 3: Sampling Frame Losses of Firms In D&B Data Base: June 2005 To November 2005, With 2004 Start Date

High-Tech	Woman Owned	June File	Deceased	% Deceased
Total		188,292	13,439	7.1%
High-Tech	Total	2,593	144	5.6%
	Yes	361	21	5.8%
	No	2,232	123	5.5%
Medium-Tech	Total	22,544	926	4.1%
	Yes	4,332	153	3.5%
	No	18,212	773	4.2%
Non Tech	Total	163,155	12,369	7.6%
	Yes	32,016	2,177	6.8%
	No	131,139	10,192	7.8%

¹Deceased in November is the count of enterprises in the D&B data base in June 2005 that were not in the data base in November 2005.

Table 4. Samples From June And November Sampling Frame From D&B Data Base Firms With 2004 Start Date

High-Tech	Woman Owned	June Sample	June Frame Percent	Nov Sample	Nov Frame Percent	Total Sample
Total Sample		23,942	12.7	8,527	13.5	32,469
High-Tech	Total	2,593		1,276		3,869
	Yes	361	100.0	166	100.0	527
	No	2,232	100.0	1,110	100.0	3,342
Medium-Tech	Total	5,769		1,805		7,574
	Yes	1,029	23.8	237	19.5	1,266
	No	4,740	26.0	1,568	26.6	6,308
Non Tech	Total	15,580		5,446		21,026
	Yes	2,090	6.5	670	6.7	2,760
	No	13,490	10.3	4,776	10.7	18,266

subsidiary owned by an existing business, a business inherited from someone else, but was a new, independent business created by a single person or a team of people, the purchase of an existing business, or the purchase of a franchise. Organizations that were designed for social and charitable objectives and legally establish as a “not-for-profit” were also excluded. Businesses were excluded as having been established in a year prior to 2004 if they had an employer identification number, schedule C income, a legal form, or had paid state unemployment insurance or federal social security taxes prior to 2004.

The sample that were released for data collection were contacted through an advance letter sent to the business principals. The letter provided a brief overview of the purpose of the study and included personalized user name and password information for each business to access the web survey. The letter did not mention completing the survey over the telephone, so that the Web-based mode was initially imposed on the full sample. The advance letter package included instructions for completing the survey on the Web and a set of Frequently Asked Questions (FAQ). The Web instructions page provided steps sample members could follow to log on and complete the Web survey, and the FAQ page provided more detailed information about the Kauffman Foundation, the survey firm, and the survey objectives.

One week after the advance mailing, a reminder postcard was sent to all sample members not completing the Web survey, which repeated the key points of the letter and reminded them to visit the website to complete the questionnaire. Principals of new businesses not responding to the Web survey after the advance letter and postcard received a telephone call to complete the screening, and eligible respondents were asked to complete the interview.

Prior to the start of telephone interviewing, the list of Web non-respondents was sent to a telephone matching service of current business telephone numbers. Once CATI interviewing began, cases with a bad telephone number were sent to locators to attempt to update the information for the business, or to determine that the business was closed. The locating effort consisted of calls to directory assistance, yellow pages searches, online searches to locate the principal of the business or a business website, and calls to entities listed at the same address.

2. Sample

2.1 Data Collection Results

The data collection on the Baseline survey was completed on July 29, 2006, after one full year of data collection. A total of 32,469 pieces of sample were fielded in the six releases, and approximately 383,000 calls were made to sampled businesses not completing the survey through the Web during the Baseline data collection. Among the businesses we contacted and screened, 34.4 percent were eligible to participate in the study. Extensive refusal conversion efforts were needed to screen and interview a sufficient number of businesses, and roughly 8,000 were finalized as refusals at the end of the data collection period.

We completed interviews with the principals of 4,931 businesses who started operations in 2004. A total of 2,895 interviews were completed among businesses in the two categories of high technology businesses and 2,036 interviews were completed among businesses of all other industrial classifications. The decision to conclude the Baseline data collection effort short of the goal of 5,000 completed interviews was made in collaboration with the Principal Investigator and the Foundation. It was generally felt that the amount of resources that would be required to reach 5,000 completes would be better utilized for the follow-up surveys.

2.2 Locating and Response

The weighted location and response rates for the full sample is given in Table 5 and by technology stratum in Tables 6A, 6B and 6C. For the nonresponse adjustments to the sampling weights, we used weighted logistic propensity modeling by technology stratum and separately for location and response among the located sample. The analysis of the location and response patterns showed considerable interactive effects. Three-order and four-order interactions were common and were difficult to interpret.

In general, high tech and medium tech firms responded at a higher rate than non tech firms. Firms in the Midwest census region responded at a higher rate for all types of firms, but high tech firms in the Northeast also had a high response rate. Woman-owned firms and firms outside of MSAs responded at a higher rate, but the differential was smaller among high tech firms for both of these. Firms with no employees or only 1 employee had higher response rate than larger firms, but again the differential was smaller in the high tech firms.

References

Hadlock, Paul, Daniel Hecker, and Joseph Gannon. “High Technology Employment: Another View.” *Monthly Labor Review*, July 1991, pp. 26-30.

Table 5: Response and Location Weighted Rate Among Enterprises in the Kauffman Firm Survey: Firms With Start Date Of 2004

	Sample Count	Located	Weighted Percentage Located	Completes	Weighted Percentage Complete/Located	Weighted Percentage Completes
All	32,469	29,526	90.7	123,394	54.2	49.1
High Tech						
1: High Tech	3,869	3,640	94.1	2,051	56.3	53.0
2: Medium Tech	7,574	6,839	90.3	15,743	58.8	53.1
3: Non Tech	21,026	19,047	90.7	105,600	53.5	48.5
Woman Ownership						
1: Woman Owned	4,553	4,150	91.4	26,205	59.7	54.6
2: Not Woman Owned	27,916	25,376	90.5	97,189	52.8	47.8
Employee Size						
1: 0-1	13,590	12,164	89.4	52,672	56.4	50.4
2: 2-4	13,424	12,241	91.0	53,595	53.7	48.9
3: 5+	5,455	5,121	93.5	17,127	49.5	46.3
MSA						
Not In MSA	4,445	4,119	92.5	20,725	61.1	56.6
In MSA	28,024	25,407	90.4	102,669	52.9	47.8
Census Division						
Midwest	6,587	6,090	92.0	27,311	58.1	53.5
Northeast	5,272	4,822	91.0	18,801	51.7	47.0
South	12,288	11,033	89.6	46,828	53.7	48.1
West	8,322	7,581	91.0	30,454	53.1	48.4
Urbanicity Code						
Metro Areas With At Least 1 Million Residents	19,433	17,504	89.7	66,663	50.8	45.6
Metro Areas With Fewer Than 1 Million Residents	8,810	8,085	91.5	35,250	55.4	50.7
Remaining	4,226	3,937	93.2	21,482	65.1	60.7
Woman Ownership By Employee Count						
1: 0-1	2,136	1,911	89.2	12,180	61.9	55.2
2: 2-4	1,735	1,608	93.2	10,600	59.0	55.0
3: 5+	682	631	93.2	3,425	54.8	51.1
1: 0-1	11,454	10,253	89.4	40,493	54.9	49.1
2: 2-4	11,689	10,633	90.5	42,995	52.5	47.6
3: 5+	4,773	4,490	93.5	13,702	48.3	45.2

Table 6A: Response and Location Weighted Rate Among Enterprises
in the Kauffman Firm Survey: Firms With Start Date Of 2004
High Tech Firms

	Sample Count	Located	Weighted Percentage Located	Completes	Weighted Percentage Complete/Located	Weighted Percentage Completes
All	3,869	3,640	94.1	2,051	56.3	53.0
Census Division						
Midwest	841	797	94.8	473	59.3	56.2
Northeast	655	623	95.1	375	60.2	57.3
South	1,240	1,167	94.1	646	55.4	52.1
West	1,133	1,053	92.9	557	52.9	49.2
Woman Ownership						
1: Woman Owned	527	491	93.2	287	58.5	54.5
2: Not Woman Owned	3,342	3,149	94.2	1,764	56.0	52.8
MSA						
Not In MSA	538	514	95.5	310	60.3	57.6
In MSA	3,331	3,126	93.8	1,741	55.7	52.3
Population (County)						
Less than 100, 000	696	661	95.0	405	61.3	58.2
100,000-500,000	1,121	1,075	95.9	630	58.6	56.2
500,000 or more	2,052	1,904	92.8	1,016	53.4	49.5
Employees						
1: 0-1	1,047	952	90.9	536	56.3	51.2
2: 2-4	1,552	1,462	94.2	848	58.0	54.6
3: 5+	1,270	1,226	96.5	667	54.4	52.5
Legal Status						
Unknown / Unavailable	1,101	1,012	91.9	555	54.8	50.4
Corporation	1,542	1,481	96.0	849	57.3	55.1
Partnership	588	570	96.9	319	56.0	54.3
Proprietorship	638	577	90.4	328	56.8	51.4
Sales Volume						
Less than \$100,000	1,012	916	90.5	524	57.2	51.8
\$100,000 - \$199,999	880	827	94.0	450	54.4	51.1
\$200,000 - \$999,999	1,198	1,140	95.2	692	60.7	57.8
\$1,000,000 or more	779	757	97.2	385	50.9	49.4

Table 6B: Response and Location Weighted Rate Among Enterprises
in the Kauffman Firm Survey: Firms With Start Date Of 2004
Medium Tech Firms

	Sample Count	Located	Weighted Percentage Located	Completes	Weighted Percentage Complete/Located	Weighted Percentage Completes
All	7,574	6,839	90.3	15,743	58.8	53.1
Census Division						
Midwest	1,444	1,334	92.4	3,265	62.6	57.8
Northeast	1,275	1,171	91.8	2,654	58.1	53.3
South	2,859	2,538	88.8	5,694	57.1	50.7
West	1,996	1,796	89.9	4,131	58.8	52.9
Woman Ownership						
1: Woman Owned	1,266	1,132	89.6	3,167	63.8	57.2
2: Not Woman Owned	6,308	5,707	90.5	12,576	57.6	52.1
MSA						
Not In MSA	722	656	90.9	1,728	67.2	61.1
In MSA	6,852	6,183	90.3	14,015	57.9	52.2
Population (County)						
Less than 100, 000	983	898	91.4	2,328	66.2	60.5
100,000-500,000	2,373	2,164	91.2	5,110	60.2	54.9
500,000 or more	4,218	3,777	89.6	8,306	56.2	50.3
Employees						
1: 0-1	4,222	3,777	89.5	9,118	61.6	55.1
2: 2-4	2,196	2,002	91.1	4,468	57.0	52.0
3: 5+	1,156	1,060	91.7	2,157	52.1	47.8
Legal Status						
Unknown / Unavailable	3,510	3,126	89.1	7,232	59.2	52.7
Corporation	2,158	1,974	91.5	4,266	55.1	50.4
Partnership	697	646	92.6	1,430	56.8	52.6
Proprietorship	1,209	1,093	90.4	2,815	65.3	59.1
Sales Volume						
Less than \$50,000	1,735	1,507	86.9	3,563	60.3	52.4
\$50,000 - \$99,999	2,571	2,333	90.8	5,730	62.6	56.8
\$100,000 - \$199,999	1,431	1,312	91.7	2,886	56.2	51.5
\$200,000 or more	1,837	1,687	91.8	3,564	54.1	49.7

Table 6C: Response and Location Weighted Rate Among Enterprises
in the Kauffman Firm Survey: Firms With Start Date Of 2004
Non Tech Firms

	Sample Count	Located	Weighted Percentage Located	Completes	Weighted Percentage Complete/Located	Weighted Percentage Completes
All	21,026	19,047	90.7	105,600	53.5	48.5
Census Division						
Midwest	4,302	3,959	92.0	23,573	58.0	53.0
Northeast	3,342	3,028	90.8	15,772	50.6	45.9
South	8,189	7,328	89.6	40,489	53.2	47.7
West	5,193	4,732	91.2	25,766	52.3	47.7
Woman Ownership						
1: Woman Owned	2,760	2,527	91.6	22,751	59.2	54.2
2: Not Woman Owned	18,266	16,520	90.4	82,849	52.1	47.1
MSA						
Not In MSA	3,185	2,949	92.6	18,687	60.6	56.2
In MSA	17,841	16,098	90.3	86,913	52.2	47.1
Population (County)						
Less than 100, 000	4,127	3,809	92.4	23,837	59.8	55.3
100,000-500,000	6,672	6,118	91.7	34,940	54.9	50.4
500,000 or more	10,227	9,120	89.3	46,823	49.8	44.5
Employees						
1: 0-1	8,321	7,435	89.3	43,018	55.4	49.5
2: 2	9,676	8,777	90.9	48,279	53.4	48.5
3: 3+	3,029	2,835	93.6	14,303	48.9	45.8
Legal Status						
Unknown / Unavailable	11,215	10,008	89.3	54,322	53.1	47.5
Corporation	4,890	4,511	92.3	25,496	53.9	49.8
Partnership	1,729	1,624	94.0	8,393	50.0	47.0
Proprietorship	3,192	2,904	91.0	17,388	55.9	50.8
Sales Volume						
Less than \$50,000	4,800	4,300	89.6	25,659	56.1	50.3
\$50,000 - \$99,999	5,479	4,869	89.0	27,820	54.8	48.8
\$100,000 - \$199,999	5,152	4,694	91.2	25,388	53.0	48.4
\$200,000 or more	5,595	5,184	92.7	26,733	50.4	46.7