

**Statistics on the Small Business Administration's  
Scale-Up America Program**

**by**

**C.J. Krizan  
U.S. Census Bureau**

**CES 19-11**

**March, 2019**

The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the U.S. Census Bureau. Many of these analyses take the form of CES research papers. The papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors.

To obtain information about the series, see [www.census.gov/ces](http://www.census.gov/ces) or contact Christopher Goetz, Editor, Discussion Papers, U.S. Census Bureau, Center for Economic Studies 5K028B, 4600 Silver Hill Road, Washington, DC 20233, [CES.Working.Papers@census.gov](mailto:CES.Working.Papers@census.gov). To subscribe to the series, please click [here](#).

## Abstract

This paper attempts to quantify the difference in performance, of “treated” (program participant) and “non-treated” (non-participant) firms in SBA’s Scale-Up initiative. I combine data from the SBA with administrative data housed at Census using a combination of numeric and name and address matching techniques. My results show that after controlling for available observable characteristics, a positive correlation exists between participation in the Scale-Up initiative and firm growth. However, publicly available survey results have shown that entrepreneurs have a variety of goals in-mind when they start their businesses. Two prominent, and potentially contradictory ones are work-life balance and greater income. That means that not all firms may want to grow and I am unable to completely control for owner motivations. Finally, I do not find a statistically significant relationship between participation in Scale-Up and firm survival once other business characteristics are accounted for.

---

\* The analysis, thoughts, opinions, and any errors presented here are solely those of the authors and do not necessarily reflect any official position of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The Disclosure Review Board release number DRB-B0011-CED-20181108. Republication in whole or part must be cleared with the author. C.J. Krizan is a Senior Economist in the Center for Economic Studies at the U.S. Census Bureau and the corresponding author on this paper, 4600 Silver Hill Road, Washington, DC 20233, [cornell.j.krizan@census.gov](mailto:cornell.j.krizan@census.gov)

## **I. Introduction:**

There is a great deal of interest across government agencies and academic institutions in linking program and administrative data, in part because linked administrative-program data can be used for a wide range of purposes, including improving the quality of both datasets, reducing respondent burden (repeatedly answering the same questions for multiple agencies), and also for program evaluation studies.<sup>1</sup> The U.S. Census Bureau houses a great deal of administrative data in support of its programs. One such dataset is the Business Register (BR). The BR contains establishments of almost all non-farm domestic businesses as well as organizational units of multi-establishment businesses. The Small Business Administration (SBA) by contrast is a major program agency that seeks to make small businesses more successful through its various programs. Examples of such programs include Scale-Up America, the Regional Innovation Clusters Initiative, and the Emerging Leaders Program. In this paper I report the results from a project to link the Scale-Up America data to the BR and examine the employment, payroll, and survival outcomes of the treated firms compared to a similar group of non-participant businesses.

This project helps both agencies understand the challenges and benefits from merging their respective data with a particular emphasis on the ability to measure the difference in performance, of “treated” (program participant) and “non-treated” (non-participant) firms. While the Scale-Up initiative lasted just two years (2015 and 2016) before being discontinued, it is still useful to understand how the treated firms performed relative to their counterparts. The lessons learned from this exercise can be applied both to currently existing programs as well as future projects that measure their performance.

---

<sup>1</sup> See for example the report of the Commission on Evidence-based Policymaking: <https://www.cep.gov/cep-final-report.html>, and the following inter-agency reports on data matching and program evaluation: [https://www.sba.gov/sites/default/files/aboutsbaarticle/Building\\_Smarter\\_Data1.pdf](https://www.sba.gov/sites/default/files/aboutsbaarticle/Building_Smarter_Data1.pdf) [https://www.cep.gov/content/dam/cep/comments/Employer-Data-Matching-Workgroup\\_White-Paper.pdf](https://www.cep.gov/content/dam/cep/comments/Employer-Data-Matching-Workgroup_White-Paper.pdf)

## II. Matching and Evaluation Literature:

This paper follows numerous other studies that have merged Census and program data to compute program data statistics. Jarmin (1999) looked at the correlation between firms' participation in the National Institute of Standards and Technology's Manufacturing Extension Program (MEP) and their productivity and sales growth. The goal of the MEP is to help small- and medium-sized manufacturers grow and become more productive through providing business and technical assistance. Because MEP program data and the Longitudinal Research Database (LRD) did not contain common numeric identifiers, Jarmin matched the two data sets using name and address. He found evidence that MEP assistance is positively correlated with productivity growth but did not find a significant relationship with sales growth.

Following up on Jarmin's work, Ordowich et al. (2012) match more recent MEP data to the Census Bureau's Longitudinal Business Database (LBD). They use lagged dependent variable and difference-in-difference (DID) techniques to look for correlations between participation in MEP programs and labor productivity. Overall the two models produced differing results and they were unable to draw clear conclusions about the existence or direction of the relationship between firm productivity growth and program participation. However, they did find positive and significant coefficients on program participation for small firms.

A related study was also recently published by the National Academies in which these same authors compared the MEP program to similar programs in other countries.<sup>2</sup> They sought to identify foreign practices that the MEP could adopt as well as to compare funding levels and outcomes across programs/nations. They found that MEP is a relatively low-cost program with evidence of producing results. Many countries sponsor technology adoption programs. The most successful ones have long-term stable funding, are staffed with well-qualified experts,

---

<sup>2</sup> <http://www.nap.edu/catalog/18448/21st-century-manufacturing-the-role-of-the-manufacturing-extension-partnership>

operate with high levels of independence and have strong ties to local industry clusters. Most recently, Lipscomb et al. (2017) found some evidence that MEP services had positive effects on productivity and sales while increasing the probability of establishment survival. The effects varied somewhat by establishment size and sector.

Krizan (2015) focused on the International Trade Administration's Global Markets program. Global Markets, like Scale-Up, is a counseling program. The Global Markets program assists and advocates for U.S. businesses in international markets in order to increase U.S. exports, and by extension, employment. It provides U.S. companies with counseling, market research, commercial diplomacy, trade events, and other services. Krizan (2015) finds that, although he can't rule-out several sources of bias, the model results show that participation in the Global Markets program is positively and significantly correlated with export growth. Furthermore, some – though not all - of the models also find a positive and significant relationship between employment growth and counseling, especially for small and medium-sized enterprises (500 employees or less).

The most recent work in this group of studies focused on the Small Business Administration (SBA) loan programs and was performed by Brown and Earle (2018). In the programs that Brown and Earle studied, the SBA provides loan guarantees to small businesses that may otherwise not be able to obtain funding. Their goal is to foster business growth. Brown and Earle use both fixed effects and matching models and find evidence of a strong correlation between program participation and employment growth of the firm.

### **III. SBA's Scale-Up America Initiative:**

SBA's Web site states: "The SBA's new Scale-Up America Initiative is designed to help small firms with high potential "scale up" and grow their businesses so that they will provide more jobs and have a greater economic impact, both locally and nationally. The SBA has structured this community-focused initiative with local entrepreneurial ecosystems in mind: a key

emphasis of the program is building and strengthening entrepreneurial networks *within a particular community*, so that firms can grow by leveraging and complimenting the existing resources and expertise in their areas” (emphasis added).<sup>3</sup>

Scale-Up provides counseling and enhanced access to capital (through networking, not direct loans from the SBA) to help existing businesses better integrate into the local business ecology. As the Office of Entrepreneurial Development’s 2015 “Year in Review” publication explains:<sup>4</sup>

*ScaleUp America, offers vital support and guidance to small businesses during the critical time period just beyond the start-up phase. To effectively cross into the next phase and “scale up,” businesses require special attention, expert advice and access to capital to nurture growth. The purpose of ScaleUp America is to fill a gap in comprehensive entrepreneurial development services for these companies so they transition into more stable organizations, create more jobs, and drive future growth and success. ScaleUp America participants are part of an intensive classroom program combined with one-on-one management consulting and networking support to connect participants to new business opportunities and growth capital.*

The goal therefore is to help them expand, and presumably, survive longer. This benefits both the firms and the regional economies. A main focus of this paper will be compare the performance of these treated firms with other similar businesses in their communities.

That Scale-Up was a new, small, short-lived initiative provides both benefits and challenges to the matching process. Had the program continued, matching program and administrative records early-on would have allowed researchers to refine matching techniques over time. It would have been possible to provide feedback to the program agency on what additional useful information could be gathered to aid the matching process. In this case a key variable of interest, the Employer Identification Number or EIN is missing for many firms. The EIN is

---

<sup>3</sup> <https://www.sba.gov/tools/local-assistance/suac>

<sup>4</sup> [https://www.sba.gov/sites/default/files/resources\\_articles/2015\\_OED\\_Year\\_In\\_Review.pdf](https://www.sba.gov/sites/default/files/resources_articles/2015_OED_Year_In_Review.pdf)

important for this exercise because it serves as a unit of identification as well as the main linking variable across data sets.

#### **IV. The Census Bureau's Business Register**

As described by DeSalvo, Limehouse, and Klimek (2016), the BR is the Census Bureau's internal master list of all non-agricultural U.S. Businesses. It is used as a sampling frame for censuses and surveys as well as a data source for statistical products such as the County Business Patterns (CBP)<sup>5</sup>, Non-employer Statistics<sup>6</sup>, and Statistics on U.S. Businesses<sup>7</sup>. These products produce estimates of the populations of: all U.S. employer businesses, all businesses without paid employees, and regional estimates of U.S. businesses respectively.

From a researcher's perspective, the BR is a set of master lists of U.S. businesses, either employer or non-employer, that can be linked to other economic surveys or programs in cross-section or linked to itself longitudinally. Typically researchers wishing to use the longitudinal nature of the BR prefer to use another data product based on it: the Longitudinal Business Database or LBD. The LBD links the BR over time and contains a number of improvements to the linkages through name and address matching and statistical retiming (Jarmin and Miranda (2002), Stinson, White, and Lawrence (2019)).

The improvements embodied in the LBD's require additional time for processing and analysis so its release typically lags that of the BR. The BR itself, in large part because it is also used as the basis of statistical products that require high-quality data, is released with a 1-2 year lag behind the current calendar year. The Scale-Up initiative covered 2015 & 2016 and the main variables of interest are growth and survival into 2016 and 2017, just a year or two behind the current

---

<sup>5</sup> <https://www.census.gov/programs-surveys/cbp.html>

<sup>6</sup> <https://www.census.gov/programs-surveys/nonemployer-statistics.html>

<sup>7</sup> <https://www.census.gov/programs-surveys/susb.html>

calendar year. Because the Census staff need time for processing and quality control checks, this project could not use the LBD. Instead, the Census data utilized here are from the BR for the years 2014-2017<sup>8</sup>.

As DeSalvo, Limehouse, and Klimek note, the BR is growing over time and it currently contains about 7.757 million establishments. An establishment is a physical location where the firm's business is conducted. A firm is a legal entity that may own one or more establishments. When a firm owns more than one establishment, it is called a "multi-unit" firm or MU. Firms with only one establishment are called "single-unit" firms or SUs. There is an extensive literature documenting the substantial differences in performance shown by these two groups of firms (Davis, Haltiwanger, and Shuh (1996), Foster, Haltiwanger, and Krizan (2006), Foster, Haltiwanger, Krizan, and Ohlmacher (2016), etc.).

The BR contains a several key variables for this study. As previously mentioned, the EIN is the main linking variable between the two data sources. However, because EIN is not populated for a large portion of the SBA data, name and address data were exploited to perform both hand and fuzzy/statistical matching on these variables. Once the data were linked, the BR also provided measures of employment, payroll, age, and survival.

## **V. Matching SBA Data to the BR:**

The initial phase of the project was to match the 2015 and 2016 Scale-Up data to the BR. As Table 1 highlights, we received about 650 firm observations from SBA.<sup>9</sup> Slightly more than half of the records arrived with EINs, the numeric identifier used to link businesses across the datasets. EINs have several advantages as identifiers for cross-agency data matching. First,

---

<sup>8</sup> The 2017 used in this study are still in preliminary format and may be subject edits before their final release.

<sup>9</sup> In accordance with Census policy, all counts and statistics are rounded to prevent disclosure. This may cause column and row totals to appear not to sum-up correctly but the pre-disclosure totals were checked and passed consistency requirements.

they are relatively accurate. EINs are numeric and therefore easier to use for matching purposes than firm name. They also have the advantage of broad coverage. All employer firms, most exporters and importers, and many firms without employees such as partnerships and S Corporations have them. Finally, they are relatively constant over time.

EINs also have drawbacks as identifiers however. Firms can apply for an unlimited number of EINs at no cost beyond the time needed to fill out the application. Also, many firms, particularly large multi-unit businesses, have multiple EINs and may not report the same number to different agencies. If a single firm reports different EINs to two different agencies, the agencies will not be able to use EINs to discern that they are dealing with the same firm. Even the Census Bureau, with its multi-decade long data on the universes of businesses has an incomplete list of all EINs in use by firms. This was demonstrated during the recent construction of the LFTTD. As documented by Bernard, Jensen, and Schott (2009), depending on the year of the data, roughly 18-28% of the EINs reported during foreign trade transactions cannot be matched to the BR.

To maximize the match rate to the BR, records that arrived without EINS – as well as those that arrived with EINs not found in the BR – were run through a combination of statistical name and address matching software (SAS Proc DQMATCH) and/or hand-matching techniques. That is, EINs were added to the records where the field was blank and alternative EINs were attached to those cases where an already existing EIN were not found in the BR. Statistical name and address matching was the primary tool for blank records while hand matching was primarily used in cases where statistical matching did not produce a match, when it produced multiple potential matches that had to be sorted through manually, or when the record already had an EIN that didn't match the BR and an alternative was needed.

Table 1 shows that after the hand and name & address matching runs were made the share of records without EINs drops from nearly one-half to 22% and the share of SBA records matched

to the BR is 64%. This total includes both matches obtained from adding EINs to blank records, matches on alternative EINs added to records whose original EIN wasn't found in the BR, and name and address matches.

The project's final match rate of 64% is not ideal but it is important to recall that the Scale-Up initiative focuses on young, small, dynamic businesses. The literature shows that it can take statistical agencies a little longer to find and keep track of young and small businesses such as these, particularly if they appear between Economic Census years (DeSalvo et. all (2016)). Furthermore, this match rate is very close that achieved by similar studies (Jarmin (1999), Krizan (2015)).<sup>10</sup> The final number of SBA records used in the study did decline further however because of restrictions on employment, payroll, industry, and age. Further details are provided in following sections and tables.

## VI. **Methodology**

### A) Program Participation Issues:

It's impossible to perfectly assess the effect (if any) that SBA counseling services has on firm performance, or what would have happened to SBA assisted businesses in the absence of the program (the "counterfactual"). However, both random control trials and other methods can be used that allow for reasonable assessments of program effects. The two main approaches are random control trial (RCT) and a quasi-experimental design (QED). RCT's rely on developing a group of businesses that did not receive assistance (the "control" group) and comparing their outcomes to the program participants ("treated" group) while QEDs look for the effects of counseling while controlling for as many other factors relevant to performance as possible. This study takes the latter approach.

---

<sup>10</sup> The characteristics of the matched and non-matched SBA records were similar in terms of age, size, and state. No obvious source of bias was found.

Perhaps the biggest drawback to QEDs is that not all of the factors affecting firm outcomes are observable to the econometrician. It is possible to control for general firm characteristics such as industry, size, and age. However, firms receiving counseling may differ systematically from non-treated firms in ways that these standard variables may not capture. For example, it may be that only more competent, motivated firms seek out aide from SBA. Alternatively, it may be that the firms most likely to get assistance are those experiencing hardship and/ or in need of new markets for their goods making SBA's promise to integrate them into the local economy very attractive.

The nature of the Scale-Up initiative accentuates the selection bias issue. Scale-Up's goal is to assist firms that are eager to expand their size. Recent research using administrative and survey data has shown that a large share of small businesses do not grow over time and that expansion is not a universally held goal among entrepreneurs. For example, Haltiwanger, Jarmin, and Miranda (2013) used the Census Bureau's LBD to investigate the widely-held perception that job growth is fueled almost entirely by small businesses. Their findings show a more nuanced picture of what's happening. While small firms create a large share of jobs, firm age is an important factor as well. In fact, once they control for firm age, there is not a clear relationship between businesses' size and growth rate. Although they do not focus on policy recommendations, they note that programs aimed at small businesses should recognize the importance of firm age. Specifically, they note the key role played by new firms which, on average, grow faster than more mature firms and that there is a strong "up-or-out" phenomena where small young firms tend to either exit or grow rapidly. Most firms don't survive but those that do either expand immediately or are very stable. That is, if a firm grows, it does so from an early stage.

While some share of the growth rate dispersion among young firms is due to factors such as financing, productivity, product quality, luck, etc., recent survey results from the U.S. Census

Bureau's Annual Survey of Entrepreneurs shows that there is also dramatic, sometimes contradictory heterogeneity among owners' motivations and goals. For example, Table 2 highlights two key goals business owners had when they started their businesses: balancing work and family on one hand and earning greater income on the other.

Roughly half of the respondents report that both balancing work and family as well as earning greater income are "Very important" illustrating the diverse, sometimes contradictory goals of small business owners. In particular, not all business owners may want to expand their operations if it would require that disturb their balance between work and family life. Given that the Scale-Up initiative provides only counseling (not loans or grants) aimed at generating firm growth, and that this is not "Very important" to about 1/2 of the population of U.S. entrepreneurs, program participants are likely a select population.

Unfortunately, it is impossible for me to observe the motivations of the program participants or the owners of the control group firms and I cannot control for them fully. As discussed below, I will attempt to use firm age, size, and prior growth patterns, factors highlighted by Haltiwanger, Jarmin, and Miranda (2013), to control for these differences to the extent possible. Clearly however, these are second-best measures of the control group owners' desire for business growth.

Another issue that must be considered is that firms being treated by SBA may participate in more than one program. For example, firms in the SBA program could be getting assistance from government agencies like the Economic Development Agency (EDA). This additional assistance can affect the accuracy of our results if the changes in performance are attributed to participation in the SBA program, when in fact it is the EDA helped the firm change.

Unfortunately, I have no clear way to control for participation in other programs and couch the results accordingly.

Another generic issue in calculating program-level statistics is that I may under-estimate the effects of counseling if there are substantial spillover effects across firms. That is, if firms participating in SBA programs are able to pass along their knowledge of exporting to nearby firms, I will systematically under-estimate the correlation between SBA and changes in firms' performance. This is a particular concern since the Scale-Up initiative is focused on local communities.

Finally, I only have a measure of participation in the program. I do not have data on the intensity of help provided by SBA. If some of the firms participated more intensively in the program than others, it will be a source of measurement error in my variable of interest and bias my results downward.

#### B) Control Group Selection:

Recent applied (Brown and Earle (2017), Amaral et al. (2018)) and theoretical work (Iacus, King, and Porro (2012)) has highlighted a relatively new technique for improving the measurable balance between treatment and control group covariates called coarse exact matching. It has several advantages over other techniques such as propensity score matching, particularly in this context where I have relatively few treatment firms and literally millions of potential control observations. Iacus et al. argue that "Coarsened exact matching is faster, is easier to use and understand, requires fewer assumptions, is more easily automated, and possesses more attractive statistical properties for many applications than do existing matching methods."

The technique amounts to requiring the treatment and control groups' key variables to match either within narrow bands or exactly. For example, Brown and Earle (2017) require candidate control firms to match the SBA treatment firms on: year, single-unit status, age-group, size group (or exact size for small firms), and 4-digit industry. Once they obtain a set of

candidate control firms, they further refine the matches via propensity score weights based on employment history and wage.

I largely follow Brown and Earle's methodology while adapting it to the unique circumstances of the Scale-Up initiative. I require the control firms to match exactly on SU status, 4-digit industry, year, and state (location). Employment measures are, like Brown and Earle, based on the year prior to receiving SBA assistance and matched firm pairs are required to have positive employment. Additionally, they are also required to be within 1 (for 10 employee firms or smaller) or 2 (10 employees or more) workers of each other. Similarly, treated and control firms must be within 1 (firm age  $\leq 10$  years) or 3 years of age.

I also follow, to the extent possible, Brown and Earle's requirement that control firms' employment history is similar to that of the treatment firms. In particular I require that the absolute value of the lagged employment change ( $emp_{t-1} - emp_{t-2}$ ) of firms where both (t-1) and (t-2) employment is positive is within  $abs(5)$  of their paired SBA firm. This requirement seems particularly important given that the Scale-Up initiative solicits participation by entrepreneurs wishing to increase the size of their business. Matching on prior year changes is a step towards controlling for the growth patterns among the two groups of firms prior to contact with SBA.

Recall also that the Scale-Up initiative aims to foster growth by better integrating their client firms into the local community. Given the importance of the regional environment to the program, I further require that control firms be within 20 miles of their paired treated firm.<sup>11</sup> While an imperfect measure of sharing a local community, it should help ensure that the control firms are within a reasonable distance to the operating environment of the treatment firms and share at least some of the same potential resources.

---

<sup>11</sup> Distance is measured using the Haversine formula and the zip code centroids of the respective firms as inputs. See: [https://en.wikipedia.org/wiki/Haversine\\_formula](https://en.wikipedia.org/wiki/Haversine_formula) for details.

Table 3 illustrates the effects of these restrictions on the set of potential control firms. Row 1 gives the means of the initial BR sample covering the time-period 2014-2017. The middle row shows the characteristics of the final BR matched sample, and the final row shows the characteristics of the SBA firms that were matched to the BR.

The full BR sample spans the length of time needed for the prior-year employment as well as growth-rate calculations. The total number of observations is roughly 28 million establishments before any restrictions beyond being active (positive payroll, employment and populated industry, zip, and age fields) is imposed. This is roughly 7 million establishments per year of data and it includes establishments from both multi and single unit firms. It can be thought of as representative of a “typical business” in the U.S. during this time period. Comparing it to the bottom row of data, we see that a firm, on average, is larger (35.7 vs 5.6 employees), and older (10.8 vs. 7.0 years) than the typical Scale-Up firm, as can be expected given the nature of the initiative.

The next column shows the typical distance between a U.S. firm and a Scale-Up client. The typical U.S. enterprise is located about 140 miles from a Scale-Up firm, well outside the immediate economy. The next column, “# Inds” shows that the BR covers a much wider spectrum of industries (450) than do the Scale-Up firms (50). Most businesses are not likely to match the Scale-Up firms’ industry classification. Overall then, the typical U.S. business is bigger and older than Scale-Up firms, likely to be in a different industry, and are located far outside the Scale-Up client’s local economy.

The picture is far different after the matching is performed. The matched BR firms are slightly smaller (3.7 vs 5.6 employees), very similar in age (7.7 vs. 7.0 years), and very close-by (within 10 miles or so) their matched SBA counterpart. Recall also that, although not shown in the table, the matching process also required a perfect match on SU status and state and that a tolerance for prior-period employment change of 5 employees was also imposed.

Overall then the control firms look more similar to the SBA firms than does the general population of businesses. However, I can neither fully measure nor adjust for unobservable characteristics such as the owners' motivations and goals or any other aide they may be receiving from private or government sources. These unobserved and un-controlled-for characteristics could be powerful drivers of differences in firm growth and survival.

## VII. Results

### A) Employment and Payroll Change or Growth:

Although the coarse exact matching technique employed here helps control for observable differences between Scale-Up firms and the general population, the employment and payroll growth/change regressions attempt to further control for remaining differences in key characteristics that have been shown to affect growth rates. Specifically, the growth/change regressions take the following form:<sup>12</sup>

$$\begin{aligned} Change_{it} = & \beta_0 + \beta_1 ScaleUp_{it-1} + \beta_2 Age_{it-1} + \beta_3 Size_{it-1} + \beta_4 Distance_{it-1} \\ & + \beta_5 State_{it-1} + \beta_6 Industry_{it-1} + \beta_7 Year_{it} + \beta_8 Round_{it-1} + \varepsilon_{it} \end{aligned} \quad (1)$$

Where  $i$  indexes the firm and  $t$  the year. *Change* can be defined as either ( $Employment_t - Employment_{t-1}$ ) or ( $Payroll_t - Payroll_{t-1}$ ). *Scale-Up* is a dummy variable for whether or not the firm participated in the program and is the main variable of interest. *Age* and *Size* are continuous variables that control for these important firm characteristics. Following Davis, Haltiwanger, and Scuh (1996) and Haltiwanger, Jarmin, and Miranda (2013), size is defined as the average employment over the two time periods. *Distance* is a measure of the firm's

---

<sup>12</sup> All regressions are weighted by the inverse of the number of paired matches. That is, if a SBA firm is paired with 5 matched BR firms, each of the BR control firms receives a weight of 0.20. Treatment firms get a weight of one.

distance, in miles, from its paired SBA firms. The coefficients for these terms are reported in Table 4.

The remaining coefficients are not reported due to space and disclosure constraints. They include *State* which controls for the geographic location of the firm. *Industry* is a measure of the businesses' 4-digit NAICS code. *Year* is the year the change was observed (2015-2016, 2016-2017) and *Round* is the year/round during which the firm participated in the Scale-Up initiative (2015 or 2016). Firms receiving counseling in 2015 can appear more than once in the sample, as can their controls.

Table 4 provides the coefficients and standard errors for the employment and payroll growth and change regressions. The main variable of interest, Scale-Up is positive and significant in all four panels. Taking employment change as an example, the results indicate that participation in the program results in employment growth of about 0.7 jobs at a typical firm. This is less than the 3 jobs per firm that Brown and Earle (2017) estimated result from participation in an SBA-sponsored loan program but more than the jobs estimated to result from export counseling in Krizan (2015).

These results show that controlling for all observable characteristics, a positive correlation exists between participation in the Scale-Up initiative and firm growth - but this does not prove causation. There are numerous unobservable factors that are correlated with program participation that may be biasing the coefficient. Perhaps the most troubling, particularly given the positive correlation, is owner motivations. To some extent this may be controlled for by requiring that matched firms have similar prior-period growth histories. However, this is an imperfect proxy for owners' goals and not all firms in the sample qualified for this restriction to be applied.

The other variables that are consistently significant are size, and for the regressions of payroll and employment change, age.<sup>13</sup> These are the two variables often found to be most important for explaining employment growth in the literature (Davis, Haltiwanger, and Shuh (1996), Haltiwanger, Jarmin, and Miranda (2013)). Their significance here aligns with those findings. The measure of distance to the SBA firm is not statistically significant in any of the models.

## B) Firm Survival

While Scale-Up does not specifically mention increasing the likelihood of firm survival in its program descriptions, it does specifically target “...established, high-potential small businesses and entrepreneurs that are primed for growth beyond the start-up or early stages.”<sup>14</sup> Both the entrepreneurship and trade literatures have shown that there is a strong “up or out” pattern in business growth (Eaton, Eslava, Jinkns, Krizan, and Tybout (2015), Haltiwanger, Jarmin, and Miranda (2013)). That is, once businesses survive either domestically or internationally for a few periods, they are very likely to grow. Most exit occurs early-on. It is therefore reasonable to expect that if Scale-Up increases the likelihood of growth it would also increase the odds of business survival.

To investigate the correlation between Scale-Up participation and firm survival, I run Model #2 below with both a linear probability and logit estimator.

$$Survival_{it} = \beta_0 + \beta_1 ScaleUp_{it-1} + \beta_2 Age_{it-1} + \beta_3 Size_{it-1} + \beta_4 Distance_{it-1} + \beta_5 State_{it-1} + \beta_6 Industry_{it-1} + \beta_7 Year_{it} + \beta_8 Round_{it-1} + \varepsilon_{it} \quad (2)$$

---

<sup>13</sup> Although the literature generally employs categorical measures of size and age in regressions such as these, I chose to use continuous measures because of the small sample sizes and limited amount of variation in the data, particularly after exact matching was performed.

<sup>14</sup> <https://www.sba.gov/about-sba/sba-newsroom/press-releases-media-advisoris/sba-seeks-proposals-Scale-Up-america-entrepreneurship-training-program>

The model is essentially the same as those run for growth and change above except that the left hand side variable is now a dummy variable to indicate whether or not the firm survived. The coefficients are reported in Table 5.

As with the growth models, size and age are consistently positive and statistically significant. However, there is not a statistically significant relationship between participation in Scale-Up and firm survival once other observable characteristics are accounted for. The unobservable characteristics of Scale-Up clients would likely have a mixed impact on firm survival and largely be contingent on firm growth. That is, if the clients' goal for their business is growth and not just lifestyle, they may be more likely to close (stay open) if the business doesn't (does) grow.

#### **VIII. Conclusion:**

As noted in a recent White Paper from an OMB-convened working group: "Matching and reusing data on employers across Federal government agencies can have multiple and significant benefits, but it is currently very difficult to do.... There is substantial potential to achieve efficiencies in matching U.S. employer data across Federal data sets for data analysis, evaluations, and statistical activities, based on common needs across agencies. The greatest barrier to matching data on employers across data sets is the lack of a common, or universal, business identifier."<sup>15</sup>

This paper is one of several recent research studies that have addressed the challenges and benefits from merging data across government agencies to generate useful statistics. The most common cross-agency numeric identifier is the EIN and this paper used data from the SBA that was partially populated with EINs and combined it with data from the BR which contains an incomplete universe of EINs. When possible, the non-matched EINs and observations with

---

<sup>15</sup> [https://www.cep.gov/content/dam/cep/comments/Employer-Data-Matching-Workgroup\\_White-Paper.pdf](https://www.cep.gov/content/dam/cep/comments/Employer-Data-Matching-Workgroup_White-Paper.pdf)

blank EINs were either matched by hand or by using name and address matching algorithms. The final match rate was roughly 65%, though a substantial number of the matched SBA-BR cases had to be dropped because of data quality constraints.

I used the combined data to measure the difference between the performance of “treated” (program participant) and “non-treated” (non-participant) firms. The results show that controlling for all observable characteristics, a positive correlation exists between participation in the Scale-Up initiative and firm growth and employment/payroll change. However, this does not prove causation. There are numerous unobservable factors that are correlated with program participation that may be biasing the coefficient. Controlling for prior employment growth patterns may help but it is an imperfect proxy, imperfectly applied.

I did not find a statistically significant relationship between participation in Scale-Up and firm survival once other observable characteristics are accounted for. The unobservable characteristics of Scale-Up clients would likely have a mixed impact on firm survival and largely be contingent on firm growth. That is, if the clients’ goal for their business is growth and not just lifestyle, they may be more likely to close (stay open) if the business doesn’t (does) grow.

While the Scale-Up initiative lasted just two years (2015 and 2016) before being discontinued, it is still useful to understand how the treated firms performed relative to their counterparts. The lessons learned from this exercise can be applied both to currently existing programs as well as future projects that measure their performance.

## Bibliography

Amaral, M., Herrin, W.E., Saxe, J.M., Jacobson, L.E., and Courtney D. Jensen (2018). "The use of coarsened exact matching to evaluate treatment mode in the rib fracture patient" *Journal of Surgical Research*, 223, 259-260.

Bernard, A.J., Jensen, J.B., and Schott, P.K., (2009). "Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods," in T. Dunne, J.B. Jensen, M.J. Roberts, eds., *Producer Dynamics: New Evidence from Micro Data*, University of Chicago Press.

Brown, J. D. and Earle, J.S., (2017). "Do SBA Loans Create Jobs?," *Journal of Finance American Finance Association*, vol. 72(3), pages 1039-1080, June.

Davis, S., Haltiwanger, J., Schuh, S., (1996). *Job Creation and Destruction*, Cambridge, MA: MIT Press.

DeSalvo, B., Limehouse, F., and Shawn D. Klimek, (2016). "Documenting the Business Register and Related Economic Business Data," Working Papers 16-17, Center for Economic Studies, U.S. Census Bureau.

Eaton, J., Eslava, M., Jinkins, D., Krizan, C.J. and Tybout, J.R., (2015). "A Search and Learning Model of Export Dynamics," mimeo.

Foster, L., Haltiwanger, J. and C. J. Krizan, (2006). "Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s," *The Review of Economics and Statistics*, MIT Press, vol. 88(4), pages 748-758, November

Foster, L., Haltiwanger, J., Krizan, C.J., and Scott Ohlmacher, (2016). "The Evolution of National Retail Chains: How We Got Here." Handbook on the Economics of Retailing and Distribution Emek Basker Editor, Edward Elgar Publishing, Northampton MA USA.

Haltiwanger, J.C., Jarmin, R.S. Miranda, J., (2013). "Who Creates Jobs? Small versus Large versus Young," *Review of Economics and Statistics*, 95(2), 347-361.

Iacus, S.M., King, G., and Giuseppe Porro (2012). "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis*, 20(1), 1-24.

Jarmin, R.S., (1999). "Evaluating the Impact of the Manufacturing Extension Partnership on Productivity Growth," *Journal of Policy Analysis and Management*, 18(1), 99-119.

Jarmin R.S., and Miranda, J., (2002). "The Longitudinal Business Database," Center for Economic Studies Working Papers, 02-17.

Krizan, C.J. (2015), "Statistics on the International Trade Administration's Global Markets Program." Working Papers 15-17, Center for Economic Studies, U.S. Census Bureau

Lipscomb, Clifford & Youtie, Jan & Shapira, Philip & Arora, Sanjay & Krause, Andy. (2017). "Evaluating the Impact of Manufacturing Extension Services on Establishment Performance." *Economic Development Quarterly*. 32. 089124241774405. 10.1177/0891242417744050.

Ordowich C., Cheney, D., Youtie, J., Fernández-Ribas, A., and Shapira, P., (2012). "Evaluating the Impact of MEP Services on Establishment Performance: A Preliminary Empirical Investigation," Center for Economic Studies Working Papers, 12-15.

Stinson, M., White, K., and Lawrence, J. (2019). "Upcoming Improvements to the Longitudinal Business Database and the Business Dynamics Statistics." Mimeo.

**Table 1: Scale-Up-BR Match Rates:**

	Total Obs	%Match BR	% Don't Match BR	% Without EIN
Scale-Up with Name Matches Added	650	0.64	0.16	0.22

**Table 2: Entrepreneurs' Reported Motivation for Starting Their Business:**

	% Owners
Balance work and family: Not important	16.7
Balance work and family: Somewhat important	34
Balance work and family: Very important	48.9
Greater income: Not important	11.6
Greater income: Somewhat important	32.7
Greater income: Very important	55.5
Source: 2016 Annual Survey of Entrepreneurs ( <a href="https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk#">https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk#</a> )	

**Table 3: Means of Population and Matched Pairs:**

	Log Payroll (\$)	Employees	Age (years)	Distance (miles)	# Inds	N
Business Register Population	5.1	35.7	10.8	138.2	450	27,540,000
BR Matched Sample	4.3	3.7	7.7	10.1	50	3,700
SBA Sample	4.6	5.6	7.0		50	250

**Table 4: Growth and Change Regression Results (N=3918 all models)**

Employment Growth			Payroll Growth		
	Estimate	SE		Estimate	SE
Intercept	-0.008	0.148	Intercept	0.111	0.135
Scale-Up	<b>0.157</b>	0.040	Scale-Up	<b>0.170</b>	0.036
Age	-0.005	0.004	Age	-0.005	0.004
Size	<b>0.022</b>	0.003	Size	<b>0.020</b>	0.003
Distance	0.006	0.004	Distance	0.004	0.004
R <sup>2</sup>	0.119		R <sup>2</sup>	0.142	
Employment Change			Payroll Change		
	Estimate	SE		Estimate	SE
Intercept	0.094	0.608	Intercept	<b>46.370</b>	20.090
Scale-Up	<b>0.720</b>	0.163	Scale-Up	<b>24.780</b>	5.360
Age	<b>-0.064</b>	0.015	Age	<b>-1.705</b>	0.491
Size	<b>0.244</b>	0.010	Size	<b>4.907</b>	0.319
Distance	<b>0.032</b>	0.014	Distance	0.410	0.458
R <sup>2</sup>	0.246		R <sup>2</sup>	0.199	

*Unreported Controls Include:* 4-digit industry, U.S. state, year, ScaleUp round

**Table 5: Survival Results: (N=3918 all models)**

Linear Probability Model			Logit		
	Estimate	SE		Estimate	SE
<b>Intercept</b>	<b>0.956</b>	0.053	<b>Intercept</b>	6.052	597.700
<b>Scale-Up</b>	0.025	0.014	<b>Scale-Up</b>	-0.024	0.375
<b>Age</b>	0.003	0.002	<b>Age</b>	0.013	0.082
<b>Size</b>	<b>0.011</b>	0.001	<b>Size</b>	<b>0.763</b>	0.174
<b>Distance</b>	-0.001	0.002	<b>Distance</b>	-0.010	0.063
R <sup>2</sup>	0.165		Likelihood Ratio		
			P>ChiSq	0.047	

*Unreported Controls Include:* 4-digit industry, U.S. state, year, Scaleup round