

RESEARCH WORKING PAPER SERIES

EARLY-STAGE ENTREPRENEURSHIP:
SOME KEY INDICATORS AND A SUMMARY INDEX

Robert Fairlie^I and Sameeksha Desai^{II}

June 2020

ABSTRACT



Entrepreneurship is a process and a series of dynamic steps, rather than a binary or static outcome. To provide more granular insights into the early stages of entrepreneurship, we elaborate on four indicators and a summary index capturing different dimensions of entrepreneurial activity within the population and within new businesses. The purpose of these indicators is to provide simple, interpretable, and comparable insight to technical and non-technical user audiences.

Keywords: entrepreneurship, indicators, early-stage, rate of new entrepreneurs, first year survival, jobs, opportunity share

Acknowledgements: We thank AJ Herrmann, Travis Howe, Hayden Murray, and Derek Ozkal for their feedback.

Research Working Papers have not necessarily been peer reviewed, and are made available by the Ewing Marion Kauffman Foundation to share research and encourage discussion. The views and findings expressed herein are those of the authors and do not reflect the official views of the Kauffman Foundation.

^I Department of Economics, University of California, Santa Cruz; ^{II} Ewing Marion Kauffman Foundation

INTRODUCTION

Substantial interest in entrepreneurship among policymakers at the federal, state, and local levels; program managers at nonprofits and community organizations; investors and banks; and others highlights the need for accessible and interpretable measures of entrepreneurship.



However, measuring entrepreneurship is a challenge for several reasons. First, whereas many approaches treat entrepreneurship as a binary, static activity – such as start / don't start a business, or own / don't own a business – it is actually a dynamic process or a “journey” (Bennett and Chatterji, 2019) that may be characterized by different needs and challenges at different stages of the process (Looze and Desai, 2020). We therefore narrow our approach in measuring early-stage entrepreneurship activity to focus within the first year of this journey.

Second, “entrepreneurship” is actually a catch-all descriptor of new business activity, which is the phenomenon of interest. New business entities are an important, but not the only, manifestation of entrepreneurship. For example, many entrepreneurs operate as unincorporated self-employed, which means that any entity-based measure would leave them out. Some such entrepreneurs may operate this way but decide to incorporate the business in the future – and some may remain unincorporated self-employed entrepreneurs for the length of their activity. At the same time, many dynamics of interest are, by nature, related to entities, such as job creation. Our approach, therefore, is to take a broad view of entrepreneurial activity and to consider entrepreneurship both as something that occurs at the level of the person and the level of a business entity.

Third, creating indicators for broad use by audiences that include non-technical users requires considering user needs. We took several considerations into account based on feedback from a wide range of users interested in applying data to inform their understanding of entrepreneurship and their considerations for decision making.¹ A driving factor is the need to create “modular” style indicators that allow users to hone in on a specific metric that can stand alone, and to examine them together for a big picture view.

We, therefore, prioritized several goals when creating indicators. To provide value for users tasked with making decisions about entrepreneurship, a series of indicators should focus narrowly on a specific dimension. This increases the level of granularity and nuance visible to the user. We, therefore, focus on metrics within the first year to capture an early part of the entrepreneurship journey. At the same time, a series of indicators should appropriately allow for the realistic heterogeneity of entrepreneurship (Dilli et al., 2018). While no single indicator can provide a complete picture of all types of entrepreneurial activity at any given time, our approach considers that entrepreneurship is both about people (who decide to start a business) and businesses. We also prioritized selection of indicators for inclusion in the series that also be viewed collectively in a summary index for a big picture overview. Finally, we use underlying data that have timely release where possible, instead of lags, in order to enhance immediacy of application for decision making.

¹ These indicators were informed by feedback from nonprofits and community organizations, policymakers, researchers, internal staff, and the media. Two of the indicators – the rate of new entrepreneurs (year) and opportunity share of new entrepreneurs (year) – were created by Fairlie and continue as part of this series. Two new indicators – startup early job creation and startup early survival rate – were developed to provide entity-based metrics to round out this series.

This paper lays out the approach used to create a series of four standalone indicators on early-stage entrepreneurship as well as a comprehensive summary index. These measures form the Kauffman Indicators of Early-Stage Entrepreneurship.²

DATA SOURCES

Two data sources are the primary sources for the indicators. A special panel dataset created from the Current Population Survey (CPS) is used to construct the indicators that aggregate individual- or person-level entrepreneurial activity. This is the source for the rate of new entrepreneurs and opportunity share of new entrepreneurs. The CPS is a monthly survey of approximately 60,000 households and is the official source used to calculate the household-based measure of the unemployment rate by the U.S. Bureau of Labor Statistics. These surveys, conducted monthly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics, represent the entire U.S. population and contain observations for more than 130,000 people each month. The survey primarily asks questions focused on the employment status of household members, including whether they are unemployed, out of labor force, wage/salary worker, or a business owner.³



The special panel dataset matches the basic monthly files of the CPS over time. By linking the CPS files over time, longitudinal data are created, allowing for the examination of month-to-month changes in business creation. Combining the monthly files creates a sample size of roughly 700,000 adults ages 20 to 64 each year. The CPS microdata capture all business owners, including those who own incorporated or unincorporated businesses, and those who are employers or non-employers, and can identify changes to a respondent's main job. This method of creating panel data takes advantage of the household surveying strategies used for the CPS. Households in the CPS are interviewed each month over a four-month period. Eight months later, they are re-interviewed in each month of a second four-month period. Thus, individuals who are interviewed in January, February, March, and April of one year are interviewed again in January, February, March, and April of the following year. The CPS rotation pattern makes it possible to match information on individuals monthly and, therefore, to create two-month panel data for up to 75 percent of all CPS respondents. To match these data, the household and individual identifiers provided by the CPS are used. False matches are removed by comparing race, sex, and age codes from the two months of data. After removing all non-unique matches, the underlying CPS data are checked extensively for coding errors and other problems.

Monthly match rates are generally between 94 percent and 96 percent. Household moves are the primary reason for non-matching. A somewhat non-random sample (mainly geographic movers) will, therefore, be lost due to the matching routine. Moves do not appear to create a serious problem for month-to-month matches, however, because the observable characteristics of the original sample and the matched sample are very similar.

² The early-stage indicators, summary reports and data downloads are available on indicators.kauffman.org.

³ See <https://www.census.gov/programs-surveys/cps.html>.



The CPS sample was designed to produce national and state estimates of the unemployment rate and additional labor force characteristics of the civilian, non-institutional population ages 16 and older.⁴ The total national sample size is drawn to ensure a high level of precision for the monthly national unemployment rate. For each of the 50 states and the District of Columbia, the sample also is designed to guarantee precise estimates of average annual unemployment rates, resulting in varying sample rates by state.⁵ Sampling weights provided by the CPS, which also adjust for non-response and post-stratification raking, are used for all national and state-level estimates.

The Business Employment Dynamics (BED) series, from the U.S. Department of Labor, is the source for the startup early job creation and startup early survival rate indicators. The BED is derived from the Quarterly Census of Employment and Wages (QCEW), or ES-202, program. The data include all establishments subject to state unemployment insurance (UI) laws and federal agencies subject to the Unemployment Compensation for Federal Employees program. It covers all employer establishments in the United States – approximately 7.5 million.

The BED data include numbers of businesses tabulated by firm age, establishment age, employment size, and geography (national and state). Firm age information is used to identify and measure the number of startups, defined as employer businesses younger than one year old.

Because the BED is based on underlying administrative data that covers the universe of employer establishments, sampling concerns like standard errors and confidence intervals are irrelevant. Nonetheless, non-sampling errors still could occur. These could be caused, for example, by data entry issues or by businesses submitting incorrect employment data.

We create four standalone indicators and a summary index that, when viewed as a series, provide insights into dimensions of early-stage entrepreneurship. The indicators are the rate of new entrepreneurs, opportunity share of new entrepreneurs, startup early job creation, and startup early survival rate. The summary early-stage entrepreneurship index evenly weights contributions from the four indicators.

Each of the indicators is based on either a nationally representative sample of more than a half-million observations each year (using CPS data), or the universe of new employer businesses or new employer establishments in the United States (using BED data) (roughly 400,000 new employer businesses and 650,000 new employer establishments each year). These datasets allow for an examination of entrepreneurs and the early-stage startups that they create. For all the indicators, the underlying definitions and

⁴ The civilian non-institutional population is defined as persons 16 years of age and older residing in the 50 states and the District of Columbia, who are not inmates of institutions (e.g., penal and mental facilities, homes for the aged), and who are not on active duty in the Armed Forces. This number is reported regularly by the Federal Reserve and is available here: <https://fred.stlouisfed.org/series/CNP16OV>.

⁵ See Polivka (2000).

methodology are the same for national and state estimates, with appropriate adjustments for geography and population size by state.

EARLY-STAGE INDICATORS

Rate of new entrepreneurs

The rate of new entrepreneurs captures the percentage of the adult, non-business owner population that starts a business each month. New business owners are defined here as those individuals who work an average of 15 or more hours per week in their businesses in the preceding month.⁶ To create the rate of new entrepreneurs all individuals who do not own a business as their main job are identified in the first survey month. By matching monthly CPS files, it is then determined if these individuals own a business as their main job with 15 or more hours worked per usual week in the following survey month. Changes to respondents' main jobs are measured accurately because CPS survey takers ask whether the individual has the same main job that they reported in the previous month. If yes, the interviewer carries forward job information, including business ownership, from the previous month's survey. If no, the respondent is asked the full series of job-related questions. Survey-takers ask this question at the beginning of the job section to save time during the interview process and improve consistency in reporting.

The main job is defined as the job with the most hours worked. Individuals who start side businesses will, therefore, not be counted if they are working more hours on a wage/salary job. The requirement that business owners work 15 or more hours per week in the second month is imposed to rule out part-time business owners and very small business activities. The rate of new entrepreneurs may, therefore, underestimate or overestimate the percent of individuals creating any type of business.

The rate of new entrepreneurs provides a broad measure of entrepreneurship, capturing all new business owners, including those who own incorporated or unincorporated businesses, and those who are employers or non-employers.⁷

The rate of new entrepreneurs excludes individuals who owned a business and worked fewer than 15 hours in the first survey month. Thus, it does not capture business owners who increased their hours from less than 15 per week in one month to 15 or more hours

⁶ Note that estimates of annual business creation rates would be approximately six to eight times higher, not twelve times higher than monthly rates, because individuals potentially can enter and leave business ownership at multiple times within the same year. For example, an individual with a sole proprietorship might work more than 15 hours a week during one month, showing up as a new entrepreneur, then be unable to find a new project for that business for several months, taking a seasonal position as an employee of another business during that time. Later in the year, the individual may find a new project which enables them to activate the business and work more than 15 hours in a subsequent month. This individual would show up twice in the data, though the business is the same from an ownership point of view. Yearly figures reported for the rate of new entrepreneurs are averages of the monthly rate.

⁷ Most self-employed business owners are non-employers, but about 1 million self-employed business owners are classified as employer businesses. See <https://www.census.gov/epcd/nonemployer/view/define.html>.

per week in the second month. It also does not capture when they changed from being non-business owners to business owners with less than 15 hours worked. These individuals are excluded from the sample but may actually have been at the earliest stages of starting a business. At the same time, the rate of new entrepreneurs may overstate entrepreneurship because of how individuals report their work status. Longstanding business owners who are also salaried in the business may, for example, not report that business ownership is their main job if their wage/salary jobs had more hours in that particular month. If these individuals later report having worked more hours in business ownership in a subsequent month, it would appear that a new business had been created.

For the rate of new entrepreneurs calculations presented in this report, all observations from the CPS with allocated labor force status, class of worker, and hours worked variables are excluded. The rate of new entrepreneurs is substantially higher for allocated or imputed observations.



Opportunity share of new entrepreneurs

The opportunity share of new entrepreneurs reflects the percent of the total number of new entrepreneurs who were not unemployed and not looking for a job as they started the new business. Using the same data for the rate of new entrepreneurs in the CPS panel, it considers individuals' initial labor market status in the first survey month.

The opportunity share of new entrepreneurs is rooted in the idea that some individuals will be opportunity entrepreneurs that choose entrepreneurship, including those coming out of work, school, or other labor market status, and some individuals will be necessity entrepreneurs to avoid unemployment. The distinction between opportunity versus necessity has been discussed extensively in the entrepreneurship literature (Desai, 2017; Fairlie and Fossen, 2017). It is conceptually useful because the motivations for starting a business could influence the type, nature, and future direction of the business; it is also meaningful because it reflects to some extent the landscape of economic opportunity for entrepreneurs. Although there is some convergence about the theoretical distinction between the two motivations for business creation, a clean distinction is difficult to make with empirical data. It is important to note that although the motivations for starting businesses can differ (Kauffman Foundation, 2020) and can be in the context of weak economic conditions and high unemployment rates, necessity businesses could eventually become very successful (Fairlie et al., 2019; Fairlie, 2011; Block and Sandner, 2009; Hinz and Junhauer-Gans, 2010; Caliendo and Kritikos, 2010).

Distinguishing between opportunity and necessity entrepreneurship using prior labor market status offers broad insight into the influence of economic conditions on overall business creation among new entrepreneurs.

Over the past two decades, the opportunity share of new entrepreneurs appears to have increased when economic conditions were improving and decreased when economic conditions were worsening. The largest opportunity share of new entrepreneurs occurred

in the late 1990s, and the smallest share was observed in 2009, at the end of the Great Recession. The opportunity share of new entrepreneurs also decreased in the recession of the early 2000s and increased in the growth period that followed in the mid-2000s.

Startup early job creation

Startup early job creation uses BED data to capture early-stage job creation among startup cohorts each year. It reflects total employment created by new employer firms in their first year for every 1,000 people, meaning that it is a measure of average first year job creation by a startup, and it is comparable across time and by geography.

This indicator is an annual measure and uses startups defined as new employer establishments that are younger than one year old in a given year. The total employment generated by these startups in their first year is normalized by the population to create the per capita startup early job creation measure. To focus on early-stage business success, a one-year window is used to measure job creation. Population data is obtained from the U.S. Census Bureau.



Focusing on only the quantity of employer startups or the average number of jobs created per startup alone would not capture the potential of startups for early job creation. Startup early job creation reflects job creation power of a typical startup within the first year. This is also appropriate to understand entrepreneurial job creation because the majority of new businesses that ever become employers will do so within the first year (Bayard et al., 2018). This indicator does not reflect dimensions of job quality, such as compensation or longevity, which can be important characteristics (Burton et al., 2016), and it doesn't reflect industry.

Startup early survival rate

As with the previous indicator, the startup early survival rate uses BED data to measure the percentage of new employer establishments that survive their first year of operation. It is calculated as the percentage of new employer establishments that are still active after one year of operation.

This indicator measures the early survival rates of new establishments rather than new firms. A firm is an establishment (when it is a single entity that is independent) or can be a combination of several establishments (Sadeghi et al, 2016). For example, a new location of a service-oriented business (such as a restaurant or gas station) would count as a new establishment but not as a new firm. Historically, the establishment first-year survival rate has been very similar to the firm first-year survival rate (U.S. Bureau of Labor Statistics, 2015).

Startup early survival rate is cut off at one year and is, thus, a measure of immediate survival. It does not reflect the long-term survival of a startup.



Summary index

The four indicators described above are standalone metrics that each provide insight on a specific dimension of early-stage entrepreneurship. Using these indicators, we also calculate a summary early-stage entrepreneurship index to provide a measure of overall trends in early-stage entrepreneurship over time. The index is equally weighted and uses the four normalized indicators.

Each of the measures is normalized by creating a Z-score for the variable, using national annual estimates from the first two decades of available data (1996 to 2015) to calculate the mean and standard deviation for each component.⁸ The time period is used because it covers two business cycles (starting at the midpoint of the 1990s growth period). The national unemployment rate was 5.4 percent in 1996, and after two recessions and an intermediate growth period, returned to 5.3 percent in 2015. The same normalization method, which is based on national data, is used for both geographic levels – national and state – for comparability and consistency over time. The index is centered at 0. Thus, values greater (lower) than 0 indicate values greater (lower) than the first two-decade average of the index. In some cases, an index score may be driven by one very high or low indicator.

CONCLUSION

The early-stage entrepreneurship indicators described in this paper are a set of measures that represents new business creation in the United States, integrating several high-quality, timely sources of information.

REFERENCES

- Bennett, V. and Chatterji, A. (2019) The entrepreneurial process: evidence from a nationally representative survey, *Strategic Management Journal*, forthcoming.
- Block, J. and Sandner, P. (2009) Necessity and opportunity entrepreneurs and their duration in self-employment: evidence from German micro data, *Journal of Industry, Competition and Trade*, 9(2): 117-137.
- Burton, D., Dahl, M. and Sorenson, O. (2016) Do startups create good jobs? Working paper.
- Caliendo, M. and Kritikos, A. 2010. Startups by the unemployed: characteristics, survival and direct employment effects, *Small Business Economics*, 35(1): 71-92.
- Desai, S. (2017) Measuring Entrepreneurship: Type, motivation, and growth." *IZA World of Labor* 2017: 327 doi: 10.15185/izawol.327.

⁸ Note that this calculation uses 1996-2015 to calculate the mean. We fix the time period focusing on a time period that covers two business cycles. This is an update to Fairlie et al. (2018, 2019), which uses all years to calculate the mean (in other words, the index for the year 2018 would have used 1996 to 2018). This update changes the base for the index by three out of twenty years; estimates using the full time period from 1995 to 2019 to calculate the means for the Z-scores are similar, as expected.



- Dilli, S., Elert, N. and Herrmann, A. (2018) Varieties of entrepreneurship: exploring the institutional foundations of different entrepreneurship types through varieties of capitalism arguments, *Small Business Economics*, 51: 293-320.
- Fairlie, R. (2014) Kauffman Index of Entrepreneurial Activity, 1996–2013, Kansas City: Ewing Marion Kauffman Foundation.
- Fairlie, R. (2011) Entrepreneurship, economic conditions, and the Great Recession, *Journal of Economics and Management Strategy*, 22(2): 207–231.
- Fairlie, R. and Fossen, F. (2017) The two components of business creation: opportunity versus necessity entrepreneurship, Stanford Institute for Economic Policy Research Discussion Paper No. 17-014.
- Fairlie, R., Desai, S. and Herrmann, A. (2019) 2018 National Report on Early-Stage Entrepreneurship, Ewing Marion Kauffman Foundation.
- Fairlie, R., Desai, S. and Herrmann, A. (2018) 2017 National Report on Early-Stage Entrepreneurship, Ewing Marion Kauffman Foundation.
- Hinz, T. and Junbauer-Gans, M. (2010). Starting a business after unemployment: characteristics and chances of success, *Entrepreneurship and Regional Development*, 11(4): 371-333.
- Looze, J. and Desai, S. (2020) Challenges along the entrepreneurial journey: considerations for entrepreneurship supporters, Ewing Marion Kauffman Foundation.
- Polivka, A. (2002) Using Earnings Data from the Monthly Current Population Survey, Washington, D.C.: Bureau of Labor Statistics.
- Reedy, E. and Litan, R. (2011) Starting Smaller; Staying Smaller: America's Slow Leak in Job Creation, Ewing Marion Kauffman Foundation.
- Ruef, M. (2005) Origins of organizations: the entrepreneurial process, pp 63-100, in Keister, L. (ed) *Entrepreneurship Research in the Sociology of Work*, Volume 15, Elsevier.
- Sadeghi, A., Talan, D. and Clayton, R. (2016) Establishment, firm, or enterprise: does the unit of analysis matter? *Monthly Labor Review*, U.S. Bureau of Labor Statistics, November.
- U.S. Bureau of Labor Statistics 2015. Legacy BDS Establishment Characteristics Data Tables 1977-2014; <https://www.census.gov/programs-surveys/bds/data/data-tables/legacy-establishment-characteristics-tables-1977-2014.html>.
- U.S. Bureau of Labor Statistics, Civilian Noninstitutional Population [CNP16OV], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/CNP16OV>.
- U.S. Bureau of Labor Statistics. 2017. Labor Force Statistics from the Current Population Survey (CPS), <http://www.bls.gov/cps/>.
- U.S. Census Bureau (2017). Business Dynamics Statistics (BDS), <https://www.census.gov/ces/dataproducts/bds/>.