Measuring the Relationship between Startup Success and Job Creation

Joshua Tryon

April 27, 2021

ECON 3161 – Dr. Shatakshee Dhongde

Abstract: Understanding the role that start-ups play in labor market dynamics can help economists expedite labor market recovery post-covid-19. This paper runs a regression using OLS least squares estimation to better identify the relationship between job growth rates and start-up survival rates. The regression provide evidence supporting the hypothesis that high rates of start-up survival will correspond with higher levels of job growth rates. Policymakers and entrepreneurs should use these results to inform their priorities as they lead the labor market recovery.

I. Introduction

As the global economy looks to recover from the novel-coronavirus-19 pandemic, they face a labor market comeback of historic proportions. The unemployment level in the United States has never been as high as it was at the outset of the pandemic in April 2020. Although the labor market improved through the end of the year, ending, current unemployment levels are still elevated above historic norms. The pandemic-driven labor market upheaval occurred at a time when start-ups are facing unprecedented questions. Rising level of regulation add to the bureaucratic burden, increase capital requirements, and extend time horizons for investors to realize returns. Regrettably, it is becoming increasingly difficult for startups to outlast vicious competition and simultaneously manage an environment of overt regulation. This comes at a cost to the American workforce.

Research indicates that start-ups are a key component of structural change in the labor market. In other words, they serve as an incredibly effective mechanism for re-allocating talent across labor markets. But, despite high levels of research, the effect of start-up driven structural transformation is still an area of relative obscurity. Specifically, little research has been performed to analyze the importance of survival rates in states with high levels of entrepreneurship. This means that we are unable to quantify the moment at which states begin to realize the labor market benefits of their investments in entrepreneurs. This study seeks to remedy the existing gap in research by examining the impact of startup survival rates on job growth rates in US states. This paper theorizes that a high level of start-up survival will exhibit a positive relationship with a state's job-openings rate when controlled for levels of startup activity and other factors influencing job growth. This conclusion infers that startup survival is a much more important predictor for the labor market than startup activity in general, and that states reap the rewards when more of their startups make it past infancy. The results of this analysis will have an application to the Beveridge (or UV curve,) which illustrates the negative relationship between the job opening rate and unemployment. If start-up survival displays a positive relationship with the job openings rate, the Beveridge curve would predict a negative relationship between start-up survival and unemployment. To provide support for these claims, we treat job openings data from the Bureau of Labor Statistics as the dependent variable and a measure of start-up survival from the Kauffman Indicators of Entrepreneurship as the primary independent variable.

This study seeks to inform regulators by examining how startup survival contributes to the health of labor markets. Analyzing the relationship between early-firm realization and available unfilled employment opportunities allows us to better understand their influence they have in our recovery. Post-Covid-19, it is more important than ever that we investigate outstanding questions and orient our regulations to optimize start-up contributions to our workforce.

II. Literature review

There is no shortage of research on the topic of startups. Economists have long recognized that startups play a critical role in the forces that alter the configuration of labor markets. Nevertheless, analysis of existing literature reveals consequential gaps in the existing understanding of the issue. Fonseca et al. (2001) is a relatively dated piece of research, but it provides important econometric research on the topic of entrepreneurship that serves as a baseline for this study. The study examined the contribution that startups make to the labor market in the context of interactions between shocks and institutions. The study examines the interaction between administrative barriers to entry for startups and shocks and focuses on the transition from a manufacturing to a service-based economy. Fonseca et al. (2001) hypothesizes that OECD (Organization for Economic Co-operation and Development) countries which impose greater barriers to entry struggled to absorb the movement of workers from the manufacturing sector in the late 20th century. The analysis predicts increase in unemployment as a function of shocks and national institutions. Unsurprisingly, the research indicates strong support for the hypothesis—countries with unfriendly restrictions on startups struggled to contain unemployment and create jobs at a pace rapid enough to accommodate changing economic conditions. The study suggests that startups play a crucial role in labor mobility and expansion. Moreover, it implies that governments ought to be cautious about imposing barriers for entrepreneurs to face. Essentially, forcing startups to meet lofty criteria will dull the ability of labor markets to effectively allocate productivity. Although informative, Fonseca et al. (2001) suffers from serious limitations. The research utilizes OECD level information. National data limits the granularity of the model and inhibits its usefulness in correcting for local market conditions. Further, Fonseca et al. (2001) fails to connect her findings to labor markets that are not facing high levels of stress. The study only identifies a relationship between unemployment and startup activity in the face of severe shocks. Without the unprecedented transfer of talent created by the technological explosion of the late 20th century, the research is inconclusive on the impact of startup activity.

Decker et al. (2014), studies the impact that older startups have in the US labor market. The authors use access to new data from the Business Dynamics Statistics and the Longitudinal Business Database maintained by the US Census Bureau to delineate between startups, and firms that are "new" as a result of a merger or acquisition. By better isolating startups from the larger majority of employers, the authors look to more effectively identify their relationship with labor markets. The study hypothesizes that that the entry, exit, expansion, and contraction of firms are closely tied to labor productivity (drawing on previous research,) and posit that startups play a role in expediting the business cycle and optimizing labor market efficiency. To support this position, the study identifies a plethora of supporting research for the notion that startups play a key role in job creation. They predict that an observed reduction in high growth startups will correlate with falling job creation. To accomplish this analysis, the authors track firm dynamics post-entry to identify a given regions rate of high growth ventures. They identify a strong relationship between job creation and growing startups. The study attributes this observation to the role that startups play to optimize business dynamics. Without startups, the rapid pace of reallocation for resources and productivity across the US economy would slow. This would negatively impact labor markets, which look to optimize the allocation of human productivity. The study demonstrates that established startups play a disproportionate role in job creation. The authors attribute this phenomenon to the high pace of labor reallocation that occurs among young firms. While informative, Decker et al. (2014) also has critical limitations. The study clearly identifies the role that startup productivity plays in the labor market but fails to clearly apply its results to firms that are not entirely established. Well established, productive firms indubitably play a role in creating jobs, but that information has little application to the way we regulate new ventures.

Perhaps the most recent and comprehensive research performed on the topic, Carlino and Drautzberg (2017) analyzes the relationship between startups and local labor markets. Their research focuses on the unique interaction between start-up shocks and labor markets. The authors hypothesize that a significant proportion of local job growth is generated by the interaction between startups and shocks, when controlling for external variables like local labor demand. The study crafts several important models. The model for startups predicts existing local startup activity based on a number of crucial variables such as profits, local wage rate, and borrowing habits. The study also creates a proxy for startup shocks by predicting job growth rate in the absence of startup activity. In order to do this, the study holds a Metropolitan Statistical Area's industrial composition of employment constant for a time. Using the model for startup activity, the study then predicts the lost job growth from an absence of startup activity and a reliance on the previous MSA industrial labor composition. The research analyzes the impact of startup activity on 354 MSAs and finds an interesting relationship between startup shocks and local labor markets. The research predicts a small but lasing increase in overall employment. Local labor demand rises sharply when new startups enter the market, but declines to a slightly higher than usual, but enduring levels. Carlino and Drautzberg (2017) provides informative insights on the value of startups for labor markets, but their conclusions still leave gaps in current knowledge of the topic. The analysis focuses on shocks occurring to previously existing startups. The research also focuses on Metropolitan Statistical Areas, which provides a picture of how existing startups impact smaller labor markets but fails to give a picture of their impact on larger, regional issues.

An overview of existing literature reveals two primary gaps in the current knowledge of the topic. First, existing research fails to identify the immediate versus long term impact of startup activity. Current studies provide analysis focused on the response of startups to labor market shocks and the role of more established startups, but this fails to provide information about when a states' economy reaps the labor market rewards of its investment in entrepreneurs. Second, previous literature lacks a state-level examination of startup activity. While local, federal, and international studies are beneficial, an observant economist would note that most entry barriers come as state-level requirements (e.g., state corporate tax rates.) In full, an analysis of current economic literature on the subject of startups and job growth demonstrates that current research does not address the distinction between infant startups and mature ventures on a state level. This study will address the current gap in research by examining the impact of start-up survival rates up on job growth rates in a state, when controlled for a general measure of new startup activity and other factors.

III. Data

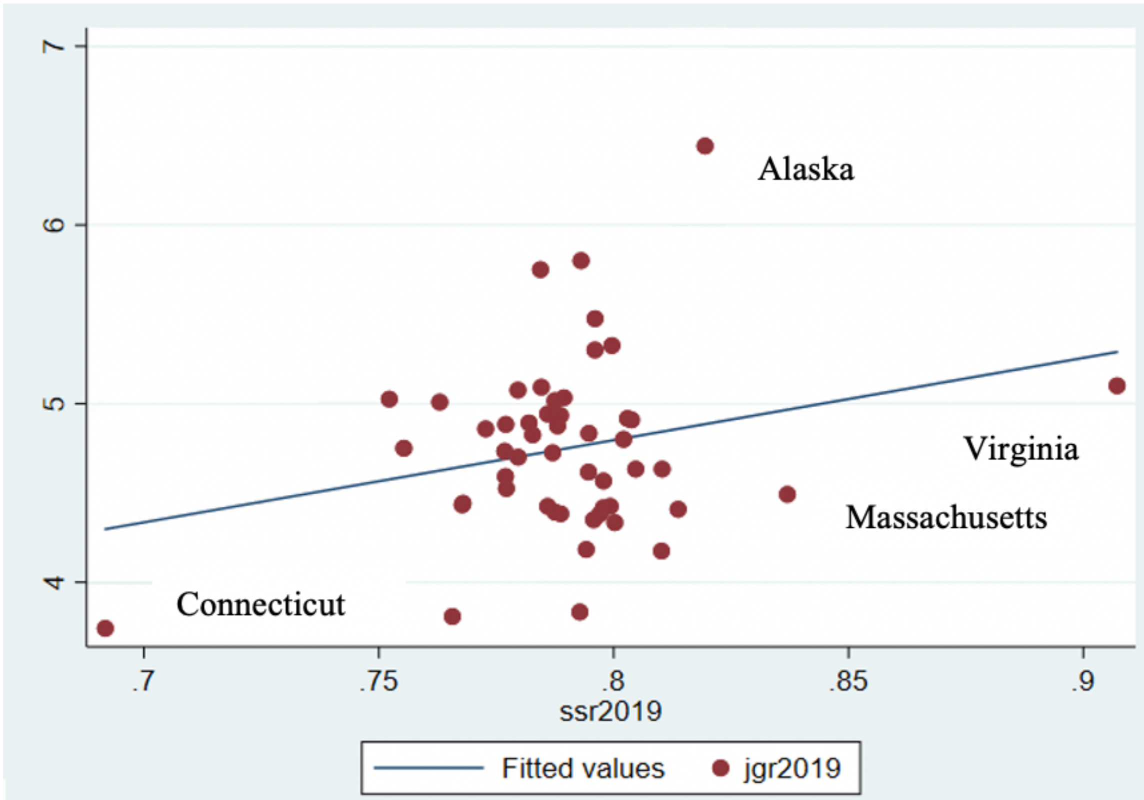
The data section of this analysis is divided into two sections. In section A, we will discuss the variables and data included in the model. In section B, we will revisit the data to discuss the all-important Gauss-Markov assumptions, and set the table to perform a regression analysis.

A. Endogenous variables

Job growth is influenced by a set of complex and often inter-related factors. Quantifying the impact of a states' startup-survival rate on its labor market requires us to control for a number of these factors. Among these factors are startup activity, aggregate demand, education levels, financing, taxation, and economic and financial openness. The dependent variable in this study is the job growth rate of the 50 states and the District of Columbia. To predict job growth rates, this study sets start-up survival rates as the primary independent variable. The non-farm job growth rate is a statistic published annually by the Bureau of Labor Statistics in the JOLTs experimental data. The data is from 2019 and no adjustments were made. The author considered using the log of job growth rates but found it unnecessary.

The primary independent variable in the sample is start up survival rates. Start-up survival rates is an annual statistic published by the Kauffman Indicators of Entrepreneurship. It is a measure of the percent of new establishments that are still active after one year of operation. This paper's goal in choosing job growth rates as the primary regressor is to understand the relationship that survival has with job growth rates and to provide a timeline for states to expect a return on investment. Figure one provides a visualization of the initial relationship between the dependent and primary independent variable.

Figure 1 – Relationship between start-up survival rates and job-growth



A number of variables were included in the model to control for factors that might also influence job growth rates. The first is a general measure of start-up activity. The metric come from the Kauffman Indicators of Entrepreneurship and measures the percentage of employers in a state that are new business owners. The model controls for this because we suspect that the quality of start-ups (as measured by the survival rate) is more important than the quantity of start-ups within a state, and because it is important to account for observed outliers (as discussed below.) By holding the number of start-ups constant, we can better understand the long-term effect that survival rates have on a state's labor market.

A GDP Growth by state is made endogenous because many studies have demonstrated the correlation between generic economic growth and job growth. This is to be expected. A measure of high school completion was chosen over a measure of college degree completion because a recent CNBC survey found that a majority of entrepreneurs did not complete a post-secondary degree. A measure of venture capital measures the ability of start-ups to get financing. Taxation is included due to its outsized impact on the business atmosphere within a state, and openness is a measure of the degree of economic and financial openness. Studies have demonstrated that openness plays a key role in the ability of an economy to fund and integrate innovate products and solutions, which can play a critical role in a start-ups ability to survive. Openness was measured as trade volume relative to GDP. The inclusion of openness in the model represents an attempt to quantify a state's sensitivity to competition. If a state is more willing or able to import better goods and services from other states or countries, it is indicative of a more competitive environment for entrepreneurship. Alternative measures for openness are described by Gräbner et al (2020) as "mostly based on sub-components and variations of the Trade/GDP approach." Gräbner et al (2020) suggests using imports/GDP when the primary goal of a metric is to measure the internal openness of an economy, so that was the calculation performed to quantify openness.

Figure two provides a detailed description of each variable included in the model.

Figure 2 – Variables Table

Variable ***(type)*** ***(name)***	***Description***	***Source***	***Date***
Job Growth Rates *(Dep.)* *(jgr2019)* *51 obs.*	Job growth rates are a calculation of the growth of available and unfilled employment opportunities within a state.	Bureau of Labor Statistics	2019
Start-up Survival Rates *(Primary ind.)* *(ssr2019)* *51 obs.*	Start-up survival rates measures the percentage of startups still active after a year of operation.	Kauffman Indicators of Entrepreneurship	2019
Start-up Activity *(Ind.)* *(sact2019)* *51 obs.*	A measure of startup activity within a state focused on new business applications. It measures the percentage of residents starting a business. KIE also titles "Rate of New Entrepreneurs."	Kauffman Indicators of Entrepreneurship	2019
Aggregate Demand Growth *(Ind.)* *(gdpgrowth)* *51 obs.*	Measures the growth of aggregate demand/spending in a state.	Bureau of Economic Analysis	2019

Pct. HS Completion *(Ind.)* *(pctHS)* *51 obs.*	Represents the percentage of a state's population that completed a basic high-school education. (GED included.)	The US Census Bureau	2018
Venture Capital Measure *(Ind.)* *ventCap* *51. Obs.*	Measures the level of venture capital deals per 100,000 members of a state's population. Represents the availability and acquisition of funding for young ventures.	24/7 Wall Street	2018
Taxation *(Ind.)* *taxClimate* *51. Obs.*	Measures the tax climate for businesses incorporated within a state. Index ranks from 1-51, the lower the index the better the tax climate for startups.	The Tax Foundation Business Tax Climate Index	2020
Openness *(Ind.)* *openness* *51. Obs.*	A measure of internal openness to competitive products and services. Measured as state imports/GDP. See Gräbner et al (2019) for a discussion of its calculation.	The US Census Bureau.	2020

Here, it is valuable to pause and recognize that certain states exhibit extreme deviation from the fitted value. Massachusetts likely exhibits this behavior because their tax structure successfully de-links their observed start-up success rates from the state's business environment. LLCs that incorporate in Massachusetts do not pay income taxes, so many start-ups geographically located outside of Massachusetts still file with the state. Because of this, there is a large population of startups whose success or failure is attributed to Massachusetts despite their lack of proximity. Alaska exhibits unique behavior due to their extreme economic reliance on traditional venues for job growth (manufacturing, agriculture,) and the lack of innovation. A number of variables (such as taxation and general start-up activity) were added to the model to account for the observed outliers.

Attached below, figure three provides a statistical summary of the variables included in the model.

Figure 3 – Summary Statistics*

Variable	Obs.	Mean	Std. Dev.	Min.	Max
jgr2019	51	4.75	0.49	3.74	6.44
ssr2019	51	0.79	0.03	.69	0.91
sact2019	51	0.11	0.77	.002	5.48
gdpgrowth	51	3.78	12.79	-0.10	4.20
pcths	51	88.22	9.10	27.4	93.60
ventCap	50	1.88	1.85	0.10	9.60
taxClimate	51	25.90	14.71	1	50
openness	51	0.12	0.15	0	0.90

**Appendix 1.1 displays the Stata command used and the results returned for Figure 3*

In order to run a successful regression analysis, out data must meet the Gauss-Markov/Classical Linear Model assumptions for a simple linear regression and a multiple linear regression. We will discuss these assumptions in section B.

B. Gauss-Markov/Classical Linear Model Assumptions

SLR.1 – Linear in Parameters

The first of the Gauss-Markov assumptions requires that the regression parameters are linear. The simple linear regression equation is defined as the equation below.

$$y = \beta_0 + \beta_1(x) + u$$

The model meets the requirements laid out in assumption one because there is a linear relationship between the dependent and primary independent variable.

SLR.2 – Random Sampling

The model meets the second assumption. The model's data is compiled from high quality sources such as the Census bureau or the Kauffman Indicators index. Moreover, the data does not exclude a state, and only excludes the territory of Puerto Rico (which many sources did not track data for.) It is the largest possible sample size to conduct a cross-state analysis.

SLR.3 – Sample Variation in the Explanatory Variable

This is possibly the easiest assumption to verify. SLR.3 requires that the dependent variable demonstrates some level of variation based on the given explanatory variable. The equation is written as follows:

$$\sum (x_i - x)^2 > 0$$

The data satisfies this assumption.

SLR.4 – Zero Conditional Mean

Per usual, the data does not meet this requirement. Zero conditional mean stipulates that the value of the dependent variable cannot carry information about the value of the unobserved factor u. In other words, the expected value of u, given x, is equal to 0. It is written below.

$$E(u_i|x_i) = 0$$

The unobserved variable does contain information about the value of the unobserved variable u. It would be difficult to devise a model that was able to account for each and every factor influencing job growth.

Because the model does not meet the first four assumptions, these are not unbiased estimators. This means that although the model will provide a prediction for the relationship between job growth and startup survival rates, it will over or under-predict this relationship.

SLR.5 – Homoskedasticity

The final assumption is homoskedasticity, which stipulates that the variance of the unobserved factors cannot change based on the given dependent variable.

$$Var(u_i|x_i) = \sigma^2$$

It is likely that the values of the variance are constant for unobserved factors, so the model satisfies the requirements of SLR.5.

The multiple regression model must also satisfy the Gauss-Markov assumptions. SLR.1 and SLR.2 are the same as MLR.1 and MLR.2. MLR.3 states that the model must not have perfect collinearity, which figure four below verifies.

Figure 4 – Collinearity Test Results

	jgr2019	***ssr2019***	***sact2019***	***gdpgrowth***	***pcths***	***ventCap***	***taxClimate***	***openness***
jgr2019	*1.00*							
ssr2019	*0.23*	*1.00*						
sact2019	*0.22*	*0.03*	*1.00*					
gdpgrowth	*0.21*	*0.03*	*0.99*	*1.00*				
pcths	*-0.11*	*-0.05*	*-0.96*	*-0.39*	*1.00*			
ventCap	*-0.30*	*0.12*	*-0.08*	*0.20*	*0.08*	*1.00*		
taxClimate	*-0.51*	*-0.03*	*-0.24*	*-0.24*	*0.18*	*0.20*	*1.00*	
openness	*0.17*	*0.01*	*0.73*	*0.51*	*-0.64*	*0.31*	*-0.56*	*1.00*

**Appendix 1.2 displays the Stata command used and the results returned for Figure 4*

It should be noted that no variable is perfectly correlated with another, although sact2019 and gdpgrowth come very close. This will impact robustness tests.

MLR.4 is similar to SLR.4. It is re-written to reflect a multiple regression:

$$E(u_i|x_{i1}, x_{i2}, x_{i3}, \dots, x_{ik}) = 0$$

The additional independent variables are likely to reduce the information contained in the unobserved factor, but it is unlikely to eliminate it. Thus, the model still fails MLR.4, and is not unbiased, although the estimates are slightly more accurate than those given by a simple regression.

The model also satisfies MLR.5, which assumes homoskedasticity. Similar to the SLR, it is likely that the values of the variance are constant for unobserved factors.

The final assumption is relevant to satisfy the Classical Linear Model assumptions. Normality implies that the population error term is independent of the explanatory variables and features a normal distribution with constant variance. It is likely that the population does not satisfy the assumption, because it fails MLR.4. In order to undertake statistical inference, this paper imposes the assumption of normality upon the model. Because of this, we interpret our statistical inference very carefully.

The importance if these assumptions is to better interpret the results of our estimations. Because the model fails the CLM and Gauss-Markov assumptions, we can expect that the estimators will be imperfect, and either under or over-estimate the fitted value. This does not mean that the regression carries no value, but it requires interpretation with caution.

IV. Results

After reviewing the data and addressing the Gauss-Markov assumptions, a number of regression models are run to test the hypothesis.

Model 1: Simple Linear Regression

The first model tests the relationship between the job growth rate and the start-up survival rate. The simple linear regression is written as follows.

Model 1: $jgr2019 = \beta_0 + \beta_1(ssr2019) + u$

The model features 50 states and the District of Columbia. Using the statistical software package Stata, we estimate the simple regression equation as:

$$\textbf{Estimation 1}: jgr2019 = 1.11 + 4.6(ssr2019)$$

Figure 5 – Summary of Simple Linear Regression

y	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$
jgr2019	*1.11*	*4.60*

**Appendix 1.3 displays the Stata command used and the results returned for Figure 5 and estimated equation 1.*

The model has an R-squared value of .0630. This indicates a very mild correlation between job growth rates and start up survival rates. The positive coefficient on ssr2019 tells us that the relationship is positive. This means that a higher rate of start-up survival is related to higher job growth rates. Because the model employs a level-level approach to regression, out results can be interpreted to mean that the increase in job growth rates is equal to 1.11 plus 4.6 times the increase in startup survival rates.

Intepretation: $\Delta jgr2019 = 4.6 * \Delta ssr2019$

A simple regression model is informative but lacks the ability to infer any causality into the relationship. This is because the error term *u* prevents the model from drawing conclusions as to the ceteris paribus effect of *ssr2019* on *jgr2019*. The next model, a multiple linear regression, allows us to draw ceteris paribus conclusions of x on y by making other factors (formerly contained in u) endogenous in the equation. The first equation is consistent with our hypothesis.

Model 2: Multiple Linear Regression #2

The second regression is a multiple linear regression that features all of the variables contained in the model.

Model 2: $jgr2019 = \beta_0 + \beta_1(ssr2019) + \beta_2(sact2019) + \beta_3(gdpgrowth) + \beta_4(pcths) + \beta_5(ventCap) + \beta_6(taxClimate) + \beta_6(openness) + u$

This model contains similar features to the first model. It features 50 states and the District of Columbia. The estimation (using Stata statistical software) is as follows:

Estimation 2: $jgr2019 = -2.52 + 3.52(ssr2019) + 1.03(sact2019) - .05(gdpgrowth) + .06(pcths) + .02(ventCap) - .02(taxClimate) - 2.10(openness) + u$

Figure six provides a summary of the second model.

Figure 6 – Summary of Multiple Linear Regression

y	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
jgr2019	*-2.52*	*3.52*	*1.03*	*-.05*	*.06*	*.02*	*-.02*	*-2.10*

**Appendix 1.4 displays the Stata command used and the results returned for Figure 6 and estimated equation 2.*

There are a number of 50 total observations in this model. Therefore, we interpret this model with caution, understanding that a central defect might be micro-numerosity. This model features an R-squared value of 0.51, indicating a strong relationship between job growth rates and a states start-up survival rate. This is substantially higher than the R-squared value from the initial estimate, demonstrating the benefits of a ceteris paribus analysis. The intercept on this coefficient is much lower than the intercept on the first equation, which indicates that model one likely over-estimates the impact of start-up survival on job growth rates, because of the omitted variable bias. Given this model, a perfect success rate from a states start-up population can be expected to increase job growth rates by 3.52%, with all else constant. It is also worth noting that our second regression still lends support to our hypothesis.

Model 3: Multiple Linear Regression #2

The third model removes sact2019, gdpgrowth, and ventCap. These variables are removed because they lack individual significance, and for reasons discussed later in the robustness test.

Model 3: $jgr2019 = \beta_0 + \beta_1(ssr2019) + \beta_2(pcths) + \beta_3(taxClimate) + \beta_4(openness) + u$

The estimation for model three is as follows:

Estimation 3: $jgr2019 = 3.06 + 3.90(ssr2019) - .008(pcths) - .02(taxClimate) - .87(openness) + u$

The results are displayed below.

Figure 7 - Summary of Multiple Linear Regression #2

y	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$
jgr2019	*3.06*	*3.90*	*-.008*	*-0.02*	*0.87*

**Appendix 1.5 displays the Stata command used and the results returned for Figure 7 and estimated equation 3.*

This model demonstrates behavior that is still consistent with our hypothesis, the coefficient on ssr2019 still shows a positive relationship with job growth rate. The R-squared on this model is .36, which indicates a mild relationship, but is substantially smaller than the R-squared reported by our second regression. This model primarily contributes to this study by allowing us to run an F-test on the excluded variables. The hypothesis on this F-test was rejected, meaning that at least one of the excluded variables (sact2019, ventCap, or gdpgrowth) has an impact on job growth rates. This leads us to run a second F-test, this one excluding ventCap and gdpgrowth.

Model 4: Multiple Linear Regression #3

Model 3 omits variables with little significance from our second linear regression, namely GDP growth and venture capital. Based on the results from a P-stat test, it is likely that the reported coefficient on these two variables was incidental, meaning that they serve little value in the model. All other variables are at least significant at 90%.

Model 4: $jgr2019 = \beta_0 + \beta_1(ssr2019) + \beta_2(sact2019) + \beta_4(pcths) + \beta_6(taxClimate) + \beta_6(openness) + u$

This model features 51 variables. Our estimate is as follows.

Estimation 4: $jgr2019 = -3.19 + 4.21(ssr2019) + .98(sact2019) + .06(pcths) - .004(taxClimate) - 1.88(openness) + u$

Figure eight provides a summary of our results.

Figure 8 – Summary of Multiple Linear Regression #3

y	β_0	β_1	β_2	β_3	β_4	β_5
jgr2019	*-3.19*	*4.21*	*0.98*	*.06*	*-.004*	*-1.88*

**Appendix 1.6 displays the Stata command used and the results returned for Figure 8 and estimated equation 4.*

This model displays a R-squared value of 0.51, exactly like our second estimation. The model has a lower intercept coefficient again, which indicates that our second regression underestimated the relationship just like our first regression. This model predicts a 4.21% increase in a state's job growth rate with a perfect start-up survival rate, all else held constant. It is worth briefly interpreting the coefficient on taxClimate. Although the coefficient is negligible (-.004) the sign indicates that the higher a state's index, (which corresponds with a harsher tax climate for businesses) the lower their job growth rate will be. The sign on openness is also worth interpretation. A state's openness to competition and innovation might serve as a leading indicator for innovation success, but it might simply just make a crowded startup field more competitive.

Figure nine is a summary of our three models and provides insights into the individual statistical significance of the variables

Figure 9 – Summary of Models, Individual Significance

	SLR	MLR #2	MLR #3	MLR #4
intercept	1.11 (2.00)	-2.52 (2.50)	3.06 (1.97)	-3.19 (2.41)
ssr2019	4.60* (2.53)	3.52* (2.00)	3.90* (2.16)	4.21** (1.91)
sact2019		1.03 (1.22)		.98*** (0.26)
gdpgrowth		-.05 (0.08)		
pcths		.06*** (0.02)	-.008 (0.00)	.06*** (0.02)
ventCap		.02 (0.05)		
taxClimate		-.02*** (0.00)	-.021*** (0.00)	-.02*** (0.00)
openness		-2.10** (1.04)	-.87 (0.61)	-1.88*** (.60)
R-squared	0.06	0.51	0.36	0.51
Adjusted R-squared	0.04	0.43	0.31	0.46

*Significant at *10%, **5%, ***1%*

From this table, we can see that the ssr2019 coefficient was significant at the 10% level for our first two models, and significant at the 5% level in our third model. Pcths and taxClimate are both significant at the 1% level in both the second and third model, and their coefficient does not change with the exclusion of insignificant variables in the third regression. Openness demonstrated individual significance at the 5% and 1% level on our second and third model, respectively.

V. Extensions

This section will be organized into two sections. In section A, this paper conducts an F-Test to find the joint significance for gdpgrowth and ventCap. In the second section, we analyze different functional forms for our first model.

A. Robustness

An F-Test is conducted on the variables sact2019, gdpgrowth, and ventCap. These variables are chosen because they demonstrate a high degree of collinearity and are all individually statistically insignificant in our second model. The hypothesis is as follows:

$$H_0: \beta_3 = \beta_4 = \beta_6 = 0$$

The F-test runs as follows, testing for joint significance to the 5% level.

$$\frac{(0.51-0.0072)/3}{(1-0.51)/42} = 14.48$$

14.48 > 2.84, so in this case, we reject the hypothesis that sact2019, gdpgrowth, and ventCap have no impact on job growth rates. Rejecting this hypothesis does not allow us to say which variables have an effect on job growth rates, it just allows us to conclude that at least one affects the dependent variable.

Based on this conclusion, another F-test was conducted which excluded gdpgrowth and ventCap. This hypothesis is based off of regression #2 and is written:

$$H_0: \beta_4 = \beta_6 = 0$$

The F-test calculates as follows, testing for joint significance to the 5% level.

$$\frac{(0.51-0.51)/2}{(1-0.51)/45} = 0$$

0 < 3.23, so this F-test allows us to fail to reject the hypothesis. We can conclude that gdpgrowth and ventCap are not jointly significant and have no impact on job growth rates in a state. Based on these results, we exclude gdpgrowth and ventCap from our fourth regression.

B. Demographic Shifts

Demographics are an important proxy that indicate the potential for success for a given start-up. This is because shifts in demographics usually have implications for the area where the shift is occurring. An area with a positive growth in population might indicate a healthy housing and labor market, which are both contribute to the survival of a start-up.

Because of this, a fifth model was designed using a dummy variable representing population growth. The fourth model was estimated with the addition of a dummy variable representing whether a states population grew in 2019.

Model 5: $jgr2019 = \beta_0 + \beta_1(ssr2019) + \beta_2(sact2019) + \beta_4(pcths) + \beta_6(taxClimate) + \beta_6(openness) + \beta_7(poplation\_growth)u$

The results of our estimation are below.

Estimation 5: $jgr2019 = -3.21 + 4.33(ssr2019) + .98(sact2019) + .06(pcths) - .02(taxClimate) - 1.85(openness) - 0.9(population\_growth) +$

Figure 9 – Summary of Model #5, Individual Significance

Regression with dummy variable: MLR #6	
ssr2019	4.33** (1.93)
sact2019	0.98*** (0.26)
pctHS	0.06*** (0.02)
taxClimate	-0.02*** (0.004)
openness	-1.85*** (0.61)
population_growth	-0.09 (0.14)
intercept	-3.21

	(2.43)
Observations	51
R-squared	0.52
Adjusted R-squared	0.45

*Standard errors in paratheses. Significant at *10%, **5%, ***1%*

Summary located in Appendix 1.7

The results from this model indicate that population growth has an insignificant statistical impact on job growth rates. This might be because a growing population is primarily a response to job growth rates rather than a catalyst for it. It might indicate that start-up success is unrelated to growth in population and the factors that it indicates. A simple explanation to this phenomenon might be explained by the rise of the internet. Demographics have been historically helpful in predicting the ease of labor market-matching, but due to the rise of the internet and telework, the ability of companies to recruit talented workers and grow is less and less limited by geographic and spatial concerns. This is certainly an area for future research.

VI. Conclusion

Based on the research performed, these seems to be strong evidence for a relationship between start-up success and job growth rates. The results support our hypothesis, namely that a high level of start-up survival will exhibit a positive relationship with a state's job-openings rate when controlled for levels of startup activity and other factors influencing job growth. The relationship between start-up survival and job growth rates is the strongest relationship examined by any of our models and has important implications for policy and future researchers.

Policymakers should focus on cultivating an environment that is conducive to start-up survival. This includes passing initiatives that ease tax burdens, lower entry barriers, and encouraging high school completion. Our research indicates that start-ups contribute to job growth when tax burdens are low, and when their product or service receives less out-of-state competition. This observation confirms the fact that the tech revolution is continually altering the landscape for start-ups, changing the dynamics of competition and survival.

Additional research can be done to better estimate the marginal impact of raising tax burdens on start-ups. This study establishes a clear baseline that high taxes undermine start-up survival and job rates, but the use of an index to approximate tax burden does not inform policymakers about the actual impact of reform.

As we look to expedite our recovery from the labor-market conditions created by covid-19, it is important that we recognize the role that start-ups play. Fostering environments conducive to startup success is perhaps one of the more powerful tools we have to create new jobs and increase the rate of recovery.

References

Carlino, G., & Drautzburg, T. (2020). The role of startups for local labor markets. *Journal of Applied Econometrics*, *35*(6), 751–775. https://doi.org/10.1002/jae.2783

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives*, *28*(3), 3–24. https://doi.org/10.1257/jep.28.3.3

Educational Attainment by State 2021. (2019). World Population Review. https://worldpopulationreview.com/state-rankings/educational-attainment-by-state

Fonseca, R., Lopez-Garcia, P., & Pissarides, C. A. (2001). Entrepreneurship, start-up costs and employment. *European Economic Review*, *45*(4–6), 692–705. https://doi.org/10.1016/s0014-2921(01)00131-3

GDP by State | U.S. Bureau of Economic Analysis (BEA). (2020, December). US Bureau of Economic Analysis. https://www.bea.gov/data/gdp/gdp-state

Gräbner, C., Heimberger, P., Kapeller, J., & Springholz, F. (2020). Understanding economic openness: a review of existing measures. *Review of World Economics*, *157*(1), 87–120. https://doi.org/10.1007/s10290-020-00391-1

JOLTS Experimental State Estimates. (2019, December). Bureau of Labor Statistics. https://www.bls.gov/jlt/jlt_statedata.htm

KAUFFMAN INDICATORS OF ENTREPRENEURSHIP. (2019). Kauffman Indicators of Entrepreneurship. https://indicators.kauffman.org/

Mike Juang, special to CNBC.com. (2017, July 19). *A secret many small-business owners share with Mark Zuckerberg*. CNBC. https://www.cnbc.com/2017/07/19/survey-shows-majority-of-business-owners-lack-college-degree.html

Stebens, S. S., & Sauter, M. S. (2020, February 30). *Best and Worst States for Business*. Wall St. 24/7. https://247wallst.com/special-report/2020/02/13/best-and-worst-states-for-business-2/2/

US Census Bureau. (2021, April 20). *National Population Totals: 2010–2019*. The United States Census Bureau. https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-total.html

Walczak, J., & Cammenga, J. (2021, April 28). *2021 State Business Tax Climate Index*. Tax Foundation. https://taxfoundation.org/2021-state-business-tax-climate-index/

Appendix

Table 1.1

```
. sum jgr2019 ssr2019 sact2019 aggdemandg pcths ventcap taxclimate openess
```

Variable	Obs	Mean	Std. Dev.	Min	Max
jgr2019	51	4.749183	.4888827	3.741667	6.441667
ssr2019	51	.7899271	.0266707	.6917558	.907182
sact2019	51	.1102824	.7662353	.00171	5.475
aggdemandg	51	3.784314	12.79225	-.1	93.2
pcths	51	88.2178	9.080479	27.4	93.6
ventcap	50	1.878	1.854315	.1	9.6
taxclimate	51	25.90196	14.71361	1	50
openess	51	.119343	.1537886	0	.9

Table 1.2

```
. corr jgr2019 ssr2019 sact2019 aggdemandg pcths ventcap taxclimate openess
(obs=50)
```

	jgr2019	ssr2019	sact2019	aggdem~g	pcths	ventcap	taxcli~e	openess
jgr2019	1.0000							
ssr2019	0.2258	1.0000						
sact2019	0.2150	0.0302	1.0000					
aggdemandg	0.2061	0.0269	0.9983	1.0000				
pcths	-0.1064	-0.0463	-0.9567	-0.9562	1.0000			
ventcap	-0.3046	0.1222	-0.0762	-0.0619	0.0849	1.0000		
taxclimate	-0.5110	-0.0312	-0.2427	-0.2516	0.1809	0.2005	1.0000	
openess	0.1741	0.0070	0.7275	0.7406	-0.6442	0.3115	-0.5559	1.0000

Table 1.3

```
. regress jgr2019 ssr2019
```

Source	SS	df	MS			
				Number of obs	=	51
				F(1, 49)	=	3.29
Model	.752905297	1	.752905297	Prob > F	=	0.0756
Residual	11.1974075	49	.228518521	R-squared	=	0.0630
				Adj R-squared	=	0.0439
Total	11.9503128	50	.239006257	Root MSE	=	.47804

jgr2019	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ssr2019	4.600976	2.534781	1.82	0.076	-.4928577	9.69481
_cons	1.114747	2.003411	0.56	0.580	-2.911258	5.140753

Table 1.4

```
. regress jgr2019 ssr2019 sact2019 aggdemandg pcths ventcap taxclimate openess
```

Source	SS	df	MS
Model	5.67912135	7	.81130305
Residual	5.36828955	42	.127816418
Total	11.0474109	49	.225457365

Number of obs	=	50
F(7, 42)	=	6.35
Prob > F	=	0.0000
R-squared	=	0.5141
Adj R-squared	=	0.4331
Root MSE	=	.35751

jgr2019	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ssr2019	3.524759	1.998906	1.76	0.085	-.5091966	7.558714
sact2019	1.030964	.2978434	3.46	0.001	.4298911	1.632036
aggdemandg	-.048545	.0754262	-0.64	0.523	-.2007612	.1036712
pcths	.0601692	.0203267	2.96	0.005	.0191484	.1011901
ventcap	.0190717	.0488702	0.39	0.698	-.0795523	.1176957
taxclimate	-.0235261	.0064441	-3.65	0.001	-.0365307	-.0105214
openess	-2.102275	1.045041	-2.01	0.051	-4.211253	.0067029
_cons	-2.52134	2.496578	-1.01	0.318	-7.559639	2.516959

Table 1.5

```
. regress ssr2019 pcths taxclimate openess
```

Source	SS	df	MS
Model	.000257545	3	.000085848
Residual	.035308894	47	.000751253
Total	.035566439	50	.000711329

Number of obs	=	51
F(3, 47)	=	0.11
Prob > F	=	0.9513
R-squared	=	0.0072
Adj R-squared	=	-0.0561
Root MSE	=	.02741

ssr2019	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pcths	-.0002524	.0005811	-0.43	0.666	-.0014214	.0009166
taxclimate	-.0001574	.0003327	-0.47	0.638	-.0008267	.000512
openess	-.0143618	.0408441	-0.35	0.727	-.0965295	.0678059
_cons	.8179852	.0580932	14.08	0.000	.7011167	.9348537

Table 1.6

```
. regress jgr2019 ssr2019 sact2019 pcths taxclimate openess

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(5, 45)        =      9.55
       Model |  6.15266151         5  1.2305323   Prob > F        =    0.0000
    Residual |  5.79765133        45  .128836696   R-squared       =    0.5149
-------------+----------------------------------   Adj R-squared   =    0.4609
       Total |  11.9503128        50  .239006257   Root MSE        =    .35894

------------------------------------------------------------------------------
     jgr2019 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ssr2019 |   4.207189   1.911894     2.20   0.033     .3564375     8.05794
    sact2019 |   .9829085   .2629802     3.74   0.001     .4532392    1.512578
       pcths |   .0603896    .019914     3.03   0.004     .0202807    .1004985
  taxclimate |  -.0228164   .0043799    -5.21   0.000    -.0316379   -.0139948
     openess |  -1.880345   .5999716    -3.13   0.003     -3.08875   -.6719402
       _cons |  -3.194632   2.413482    -1.32   0.192    -8.055635    1.666371
------------------------------------------------------------------------------
```

Table 1.7

```
. regress jgr2019 ssr2019 sact2019 pcths taxclimate openess dum_popgrowth

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(6, 44)        =      7.92
       Model |  6.20501584         6  1.03416931   Prob > F        =    0.0000
    Residual |    5.745297        44  .130574932   R-squared       =    0.5192
-------------+----------------------------------   Adj R-squared   =    0.4537
       Total |  11.9503128        50  .239006257   Root MSE        =    .36135

-------------------------------------------------------------------------------
      jgr2019 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
      ssr2019 |   4.332165   1.934841     2.24   0.030     .4327499    8.231581
     sact2019 |   .9788102   .2648274     3.70   0.001     .4450855    1.512535
        pcths |   .0604443   .0200481     3.01   0.004       .02004    .1008486
   taxclimate |  -.0234923   .0045367    -5.18   0.000    -.0326355   -.0143491
      openess |    -1.8487   .6060693    -3.05   0.004    -3.070153   -.6272479
dum_popgrowth |   -.086744   .1369913    -0.63   0.530    -.3628318    .1893438
        _cons |  -3.214259   2.429907    -1.32   0.193    -8.111414    1.682896
-------------------------------------------------------------------------------
```