Journal of Business & Entrepreneurship

Volume 33 | Number 2 Article 3

July 2024

Cyclical Interest in Entrepreneurship and the Virginia Economy

Daniel Farhat
Radford University, d.farhat.radford@gmail.com

1. Introduction

Entrepreneurship fuels prosperity. Innovation and new firm creation lead to higher-value products, efficient processes, and more employment opportunities. Predicting various entrepreneurial trends and analyzing factors correlated with entrepreneurial activity helps policymakers and non-profit organizations cultivate a thriving entrepreneurial ecosystem. Big data offers new opportunities to study entrepreneurship from a different perspective: general interest. Tracking shifts in public fascination with entrepreneurship may offer fresh insights about entrepreneurial environments.

This study looks at how often people in Virginia search for entrepreneurship-related keywords on Google. Data measuring how frequently select search terms are used over time is collected and combined to build an entrepreneurial interest index. Compared to other types of data, keyword search data is arguably richer and can unlock a deeper understanding of the entrepreneurial spirit. Data on business licenses or first-year revenues, for example, may miss earlier-stage activities. Search data may be able to capture a portion of this activity since pro-active entrepreneurs tend to rely on internet search as a low-cost way to solve business problems from the start. Unlike business-level data, search data reveals how captivated the general population is with entrepreneurship and paints a more complete picture of entrepreneurial culture. Further, the data is very convenient to access. Google provides keyword search frequency data on its Google Trends platform for free, along with easy-to-use tools to filter and download the data for various analyses.

The dynamic correlation between the entrepreneurial interest index and other data is assessed to identify potential factors influencing, or influenced by, entrepreneurial mindsets. Economic factors are explored since data on the economy is readily available. This analysis helps to determine if the connections between the state of the economy and entrepreneurial outcomes discussed in past research extend to general entrepreneurial interest. The economic data in this study relates to employment and earnings, leading economic indicators, home prices, and applications for business licenses.

The investigation below is limited to a specific geographic area. This is done for two reasons. First, many policymakers and organizations interested in improving entrepreneurship work locally. Understanding the regional entrepreneurial spirit is more useful compared to a national or international study as a result. Second, combining data across structurally different regions may hide true relationships. A significant factor driving entrepreneurial interest in one location may not matter in another. This study focuses on Virginia as a case study, however the methods used here can be applied to look at other states in future research.

This study examines short-term changes rather than long term trends. A flexible trend is independently identified for each variable used. Next, the deviation from the trend is calculated (the cycle component). How two cycles are dynamically correlated is studied: is the relationship positive or negative, does one cycle lead or lag the other, etc. To account for shifting dynamics in the data, we construct correlations for 60-mo (5-year) windows across time. This deviations-from-trend approach may be especially useful when it comes to creating short-term strategies that improve the entrepreneurial ecosystem.

The analysis shows both expected and unexpected results. The correlation between interest in entrepreneurship and the labor market has shifted over the past 18 years. Broadly, a good labor market (above-trend earnings and employment) positively correlates to entrepreneurial interest. This supports the idea of opportunity driving entrepreneurial interest as opposed to necessity. The future state of the economy has a roughly negative, lagging relationship with the index. This result highlights the forecast potential of entrepreneurial mindsets. The correlation between home prices and interest in entrepreneurship was negative in the period after the Great Recession, which differs from other research findings when taken together with the results from the labor market. This result comes from how home prices fluctuate in Virginia: above-trend home prices sometimes occur when the economy performs poorly. The correlation between interest in entrepreneurship and business applications is unusually mild. This finding implies that measures of interest in entrepreneurship contain different information from data on actual firm launches and produces new insights.

The rest of this article proceeds as follows. Relevant literature of interest is described. The method used to produce an index representing broad interest in entrepreneurship is explained. Dynamic correlations with the economic data are analyzed. Finally, concluding remarks and future work are discussed.

2. Motivating Literature

Two strands of inquiry motivate this project: culturomics and entrepreneurship studies.

2.A. Culturomics

Culturomics is the computational analysis of large-scale textual data to study public interest, mindsets and cultural trends. This area of research assumes that word choices reflect thought patterns. Therefore, measuring how frequently we use different terms and phrases serves as an indicator for collective mentalities. Michel et al. (2011) illustrates this concept using the frequency of word use in books to study public interest in people (e.g. Marc Chagall), historical events (e.g. The Civil War), social ideas (e.g. feminism and religion), and scientific knowledge (e.g. evolution). Their data reveals information about fame, suppression, and cultural memory.

Large-scale datasets are publicly available for culturomics research. Google plays a key role in collecting culturomic data and making it accessible through user-friendly tools. Google Ngram Viewer focuses on words used in books while Google Trends tracks search queries on the Google search engine. Both tools present data visually for quick inspection, and Google Trends allows convenient .csv downloading for deeper statistical analysis.

Existing studies show that culturomic data has forecasting potential. The most popular example of this uses Google Trends to predict influenza outbreaks. People attempt to self-diagnose flu symptoms online before seeking medical attention. As a result, the volume of flu-related keyword searches on Google spikes before government agencies can see the outbreak in clinic data. See Ginsberg et al. (2008), Polgreen et al. (2008), and Carneiro and Mylonakis (2009) for examples. These studies show that online flu symptom searches can detect epidemics 7-21 days ahead of public health agencies, helping them respond more efficiently.

Culturomics analysis lets us explore sensitive, stigmatizing, or abstract topics hidden from traditional data. For instance, Farhat (2017) and Jena et al. (2013) study sexually transmitted infections. Farhat and Viitanen (2017) looks at mental health. The main argument in these studies is that people turn to online searches for information about these stigmatizing, sensitive topics before seeking professional help. Clinical data may underestimate the true prevalence of these health issues as a result, and internet search data may reveal new trends to consider. Farhat et al. (2019) uses keyword search data to track interest in the general notion of hunger. The argument used in this study is that people rely on online searches to help resolve their food insecurity or to find hunger-related charities to donate to. The objective is to produce a broad measure representing interest in hunger that policymakers and non-profit organizations can use to assess program needs.

This project applies culturomics methods to uncover and analyze hidden interest in entrepreneurship in Virginia. One source of interest is the entrepreneurs themselves. Entrepreneurs are known to be self-starters with limited resources. An entrepreneur might turn to the internet for information to address business challenges as an inexpensive alternative to hiring professional assistance. This behavior occurs over time as the start-up develops, and will have happened even if the venture ultimately fails to launch. Online search data may therefore capture early-stage entrepreneurial activity with more accuracy. An additional source of interest is the general population. Anyone can perform an online search using an entrepreneurship-related keyword. Search data contains information about the greater community's overall entrepreneurial mindset and generates broad insights about the health of the entrepreneurial ecosystem as a result. Better strategies for policymakers and non-profit organizations can be built if new correlating relationships between entrepreneurial interest and various types of data are found.

The analysis in this study relates to existing studies that use Google Trends data to look at entrepreneurship. Semerci et al. (2022) draws a connection between internet search and entrepreneurial mentality by conducting a technical analysis on the popularity of keyword searches related to female entrepreneurship. Gómez Martínez et al. (2014) connects "entrepreneurship" searches to business ownership rates for various countries using panel data methods. Similarly, Huynh (2019) connects keyword searches describing entrepreneurship to new business creation in Vietnam using a variety of time series techniques. The approach here uses simpler tools compared to these studies, yet builds on them by looking at a wider variety of entrepreneurial keywords and extending the types of economic data considered.

2.B. Entrepreneurial Activity

Entrepreneurs are stimulated by forces in the entrepreneurial system, economic incentives and personal characteristics. Isenberg (2011) and Shwetzer et al. (2019) describe the components of the entrepreneurial ecosystem and how it supports start-ups. This ecosystem includes government institutions, socio-cultural norms, supporting industries, education and training, and market attributes. The method below incorporates select aspects of the entrepreneurial ecosystem by choosing keywords related to the financial services and legal frameworks used by new firms.

Uhlaner and Thurik (2007) outline the economic incentives driving people to start-ups. They describe "pull" factors (opportunities/benefits that attract people into entrepreneurship from their current state) and "push" factors (conflicts/problems with a person's current state that drives them to entrepreneurship as an alternative). Pull factors include new technologies, improved access to venture finance, and good business conditions. Push factors include poverty and unemployment. Macroeconomic data captures trends in these factors, producing patterns in entrepreneurial activity that connect to business cycles. Koellinger and Thurik (2012) and Thurik (2014) explore these cyclical patterns further. The analysis done below directly relates to this strand of literature using a measure of entrepreneurial interest instead of metrics for actualized businesses. The macroeconomic data used is obtained from the Federal Reserve Bank of St. Louis FRED Database and relates to employment, future business conditions, housing, and new business creation.

It is important to note that keywords associated with entrepreneurship and features of the entrepreneurial ecosystem can also connect to other types of businesses. "Entrepreneurship", "enterprise", and "small business" are concepts often used in the discussion of new ventures. Carland et al. (1984) concisely outlines key differences. "Entrepreneurship" involves identifying opportunities, innovating to create entirely new products, and opening new markets to deliver those products to consumers. Entrepreneurs face a high degree of uncertainty and risk since they have limited information about production and market response for new products. Some business activities are distinctly associated with entrepreneurship (for example, filing patents, seeking incubator support, or crowdfunding on platforms like Kickstarter).

"Enterprise" is a broad term for business activity that includes both entrepreneurship and the management of businesses selling established products. Several start-up tasks are done both by entrepreneurs and non-entrepreneurial new enterprises (such as filing business licenses or applying for loans at banks). "Small businesses" are simply firms with few employees and low annual revenues. The exact definition of a small business varies by country.[1] Many entrepreneurs start out as small businesses until they pin down production strategies and see profits. Policies, support industries, cultural norms, and the availability of finance may support small business of all kinds, both entrepreneurial and enterprising. Most keywords used in the study here are exclusive to entrepreneurial ventures, but some keywords do overlap with enterprise and small business. As a result, the entrepreneurial interest index created will contain some information about new firm interest in general. Blundell et al. (2021) provides a comprehensive background on entrepreneurship for researchers interested in refining the list of keywords in future studies.

When it comes to personal characteristics, Thompson (1999) and Omerzel Gomezelj and Kušce (2013) describe traits held by entrepreneurs and the different dimensions of business they pursue. These traits include risk tolerance, creativity, flexibility, independence, self-confidence, competitiveness, and locus of control. As Thompson (1999) notes, these attributes are common in entrepreneurs but not exclusive to them. To keep this study focused and direct, keywords related to personal characteristics are not included in the construction of the entrepreneurial interest index. Creating an index representing entrepreneur attributes and analyzing its relationship to entrepreneurial and macroeconomic data is left as a task for a different study.

[1] For example, the US Small Business Administration (2022) has a table defining small business size by North American Industry Classification System (NAICS) code.

3. Creating an Index for Cyclical Interest in Entrepreneurship

The first step to constructing an index representing interest in entrepreneurship is to create an appropriate list of keywords/phrases that may be used in online searches. Google Gemini, a large-language AI model similar to ChatGPT, helps with this task by generating a preliminary list of terms to consider.[2] The Google Gemini recommendations are edited to improve the likelihood of viable results from Google Trends; terms that are too broad (e.g. "marketing" and "human resources") are cut and additional terms directly related to entrepreneurship are added (e.g. "entrepreneur" and "crowdfunding").

Next, terms from the preliminary list are fed, one by one, into Google Trends to obtain a set of monthly time series data representing search frequency. The search date is limited to 1/2006 to 12/2023. Location is restricted to Virginia only. All possible categories of terms are considered to account for entrepreneurial influences from a variety of sources. Appendix 1 provides details on how these choices impact the data extracted from Google Trends. Terms yielding little or no data are eliminated. Figure 1 below shows the final list of terms used in the index.

Figure 1: Keywords/phrases used to compile the entrepreneurship interest index.

General	Finance	Innovation
• entrepreneur • entrepreneurial • entrepreneurship	• angel investor • crowdfunding • venture capital	• business ideas • innovation

Legal	Planning and Startup
• business license • business permit • copyright • intellectual property • patent • trademark	• business plan • how to start a business • starting a business • startup

This study focuses on short-term patterns only; long-term trends are left for future analysis. The Hodrick-Prescott (HP) Filter is used to independently extract a flexible trend for each of the 18 series listed in Figure 1. The deviation from the flexible trend is calculated to represent the cycle component of each series. The cycle components are combined in each time period using a simple

---

[2] The free version of Google Gemini is used. The following prompt was used: "Generate a list of keywords and phrases directly related to entrepreneurship. Categorize them into theme groups."

weighted sum to form an aggregate cycle series. The weights in the sum are determined by maximizing the sum of the squared correlations between the final index and each cycle component.[3] The combined cycle is re-scaled to take on a value between 0 and 100. The end result is the broad cyclical index of entrepreneurship interest (or "BEI") used in the rest of the analysis.[4] The correlations between this index and its components are reported in Figure 2 below. The index itself is shown in Figure 3. Note that more complex methods of extracting short-term cycles and constructing an index are available and worthy of further study. Focusing on a simple, easy-to-implement method serves as a useful first step.

Figure 2: Broad interest index correlations with its components.

Keyword/phrase	**BEI**
entrepreneur	0.625
entrepreneurial	0.361
entrepreneurship	0.483
angel investor	0.156
crowdfunding	0.277
venture capital	0.599
business ideas	0.587
innovation	0.778
business license	0.779
business permit	0.130
copyright	0.806
intellectual property	0.675
patent	0.847
trademark	0.718
business plan	0.753
how to start a business	0.506
starting a business	0.613
startup	0.712

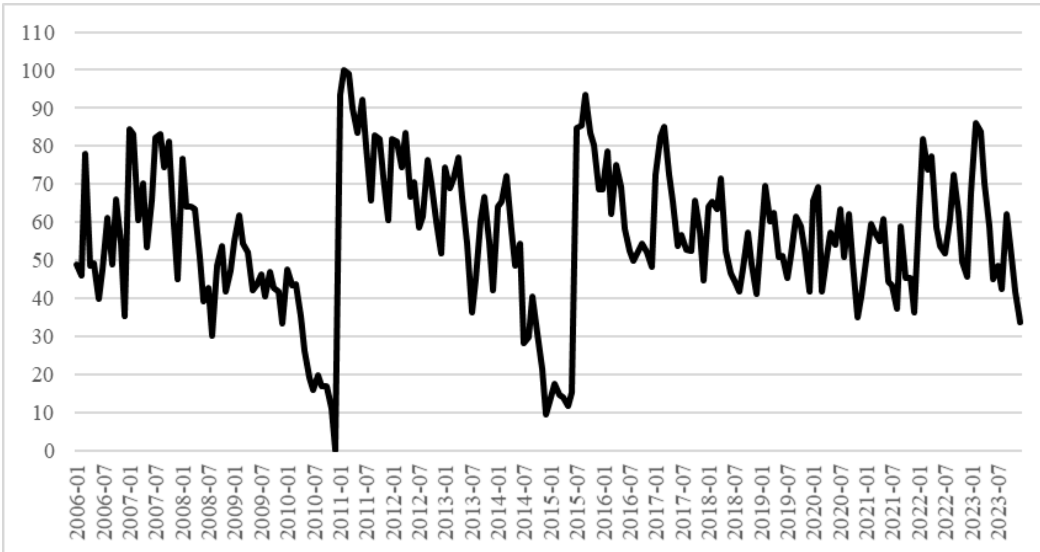
[3] If we denote the weighs as α, the keyword cycles as *k*, the number of terms as *I*, the time period at *t*, and the combined cycle as C, the numerical optimization problem we solve is:

$$Choose\ \alpha_i\ to\ Max\ \sum_{i=1}^{I} Corr(C, k_i)^2\ \ s.t.\ \ C_t = \sum_{i=1}^{I} \alpha_i \times k_{i,t}$$

The numerical optimization tool in Excel (Solver) is used for this purpose.

[4] Note that a "Direct Index of Entrepreneurship Index" was also created using only the terms from the "General" and "Planning and Startup" categories. The intention was to see if incorporating terms related to the functions of an entrepreneur (like getting a business permit or finding an angel investor) produce substantially different patterns compared to more obvious searches for "entrepreneur" or "startup". The Direct Index was nearly identical to the Broad Index, with a correlation of 0.93. The Broad Index alone sufficiently captures the patterns we are interested in.

Figure 3: Broad cyclical interest in entrepreneurship index (BEI).



4. Correlations Between the BEI and Economic Data

Dynamic correlations between the BEI and short-term movements in macroeconomic data for Virginia are analyzed to identify how the state of the regional economy relates to entrepreneurial mindsets in the Commonwealth. Four categories of data are explored: employment and earnings, leading economic indicators, housing, and planned business. Employment and earnings relate to labor market outcomes which influences interest in necessity and opportunity entrepreneurship. The expected state of the economy connects to incentives to innovate and the planning process for launching new businesses, prompting interest in certain start-up activities. Housing prices impact an entrepreneur's access to credit since real estate can be used as collateral for business loans, thereby affecting interest in entrepreneurship via finance options. Examining business applications helps identify how interest in entrepreneurship may turn into action. These 4 categories connect to themes used in studies looking at manifested entrepreneurship, allowing us to add information about mindsets and early-stage activities to their findings. Further, we require state-level data at the monthly frequency to match the structure of the Google Trends data. Data for these four categories has the correct structure and is readily available for download from the Federal Reserve Bank of St. Louis FRED Database.

The HP Filter is applied to each economic time series to extract the short-term cycle component. This is the same method of detrending used to craft the BEI. A rolling window approach is then used to construct a series of correlation statistics between the BEI and the economic cycles. This approach allows us to identify shifts in relationships and contain the impact of extreme data points. For each month, the correlation coefficient between the BEI and the cycle component of the relevant economic data is computed using the previous 60 months.[5]

Temporal correlations using up to 6-month leads/lags are calculated to pin down dynamic relationships. In other words, we compute 13 correlation measures for each month using the previous 60-months of data: Corr(BEI_t, $Data_{t+j}$) for j = -6 to 6.[6] The 5% level of statistical significance for a calculation using 60 periods is associated with a correlation value of approximately ±0.25. The lead/lag with the most extreme significant correlation value is identified.

4.A. Employment and Earnings

The relationship between entrepreneurial interest and employment may be either positive or negative. On one hand, low employment and low earnings might serve as a push factor driving people to launch new ventures as a way to cope. On the other hand, high employment and high earnings might make people more confident about earning profits, pulling them into launching a new business. To identify the correlation between entrepreneurial interest and employment/earnings, the constructed BEI is compared to the cycles of the following FRED data for Virginia:

Variable	**FRED Description**	**Period Analyzed**
Unemployment rate	Unemployment Rate in Virginia, Percent, Monthly, Seasonally Adjusted.	1/2006 to 11/2023
Labor force participation rate	Labor Force Participation Rate for Virginia, Percent, Monthly, Seasonally Adjusted.	1/2006 to 11/2023
Average hourly earnings	Average Hourly Earnings of All Employees: Total Private in Virginia, Dollars per Hour, Monthly, Not Seasonally Adjusted.	1/2007 to 11/2023

Figure 4 below reports findings from analyzing the computed correlations. Each graph shows the value of the most extreme correlation (left axis) and the lead/lag in which that correlation was observed (right axis). For example, Figure 4A shows the most extreme correlation value between the BEI and the unemployment rate in 6/2016 is -0.65 and is associated with a lead-lag period of -2. In other words, data from the 6/2011 to 6/2016 window suggests interest in

[5] For example, the data point for 12/2010 is calculated using data from 1/2006 to 12/2010. The data point for 1/2011 is constructing using data from 2/2006 to 1/2011.

[6] Care must be taken when interpreting the correlations. When the lead-lag period is positive, we are calculating the correlation between the BEI today and the economic cycle in the future (i.e. the economic cycle lags behind the BEI). When the lead-lag period is negative, we are calculating the correlation between the BEI today and the economic cycle in the past (i.e. the economic cycle leads the BEI).

entrepreneurship will likely be above trend if the unemployment rate was below trend 2 months in the past.

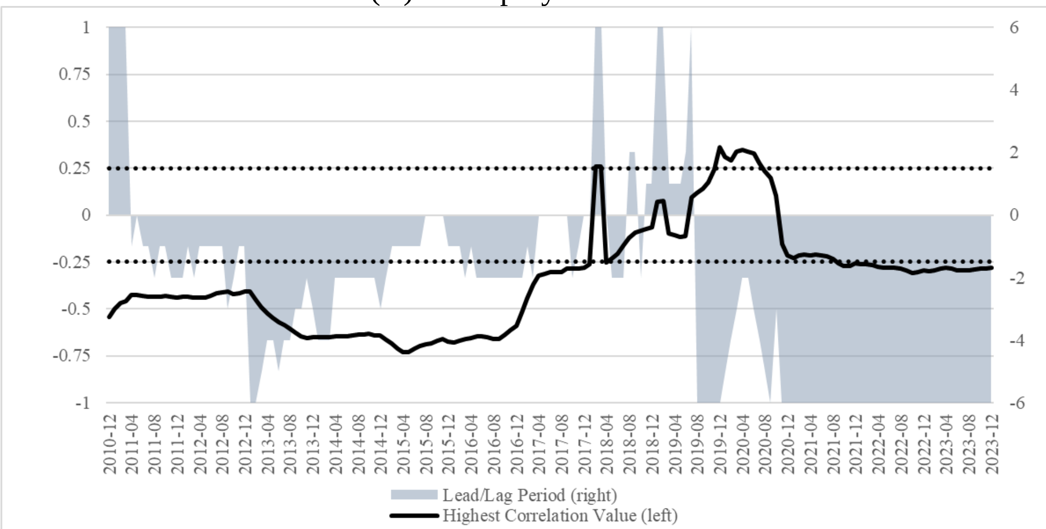
The results shown in Figure 4A-B suggest a shifting relationship exists between interest in entrepreneurship and labor market outcomes. Good labor market outcomes (below-trend unemployment, above trend hourly earnings) seem to be positively related with entrepreneurial interest on average, and shifts in these economic variables generally occurred *before* shifts in the BEI (i.e. have a leading relationship to the BEI). These results imply that Virginian interest in entrepreneurship is driven by pull factors rather than push factors. Good labor market performance may give potential entrepreneurs confidence that they will find customers with disposable income interested in buying their new products. Higher earnings might result in more personal funds to use in start-up financing. Lower unemployment rates might soothe an entrepreneur's fears about finding a job should their business fail. However, the correlating relationships are not stable. The correlations shift as the rolling window begins to capture data from 2018 and 2019. This is likely due to sharp changes in employment data occurring during the Covid Pandemic.

The correlation between the BEI and the labor force participation rate exhibits sharp switches. A spike in interest in start-ups followed by a rise in labor market participation may be a sign of the job-creating impact of entrepreneurship (a positive relationship, with labor market participation lagging). We see this effect in the years prior to 2015[7] and after 2020 in Figure 4C. However, a fall in the labor market participation rate can create a pool of people potentially looking for entrepreneurial opportunities to support their household income or enhance their retirement experience (a negative relationship, with labor market participation leading). We see this impact between 2015 and 2020. Why these switches are occurring and if these patterns are a quirk of the data may be of interest to those studying the cyclical features of labor markets.

[7] The initial correlation window in Figure 4C is less than 60-months since data for average hourly earnings begins in 2007. Data is used as it becomes available. A consistent 60-month window proceeds from the 1/2012 data point onwards. This allows us to make the most of the data and construct figures with consistent time axes.

Figure 4: Most extreme correlations between the BEI and employment/earnings data for Virginia.

(A) Unemployment Rate



B) Average Hourly Earnings

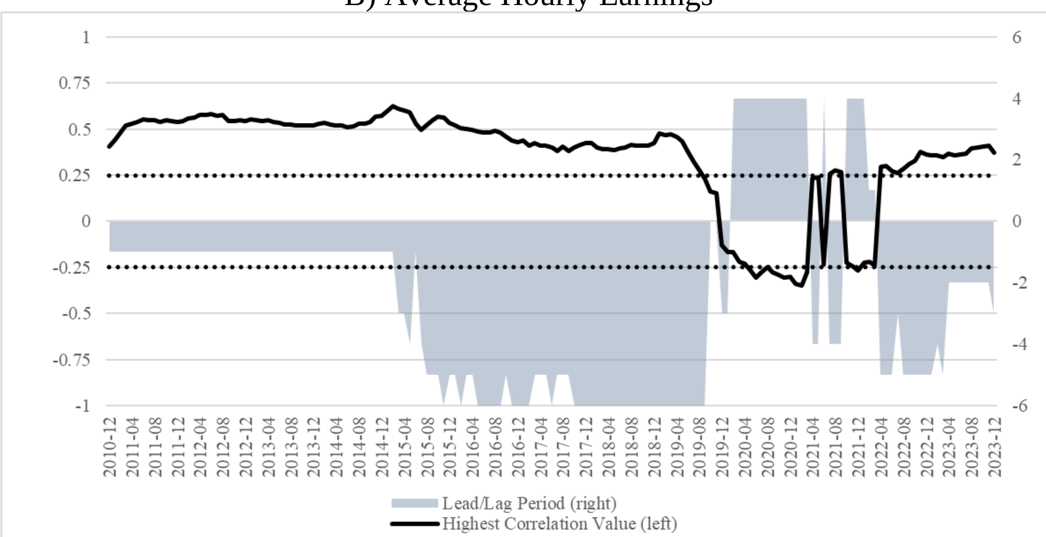
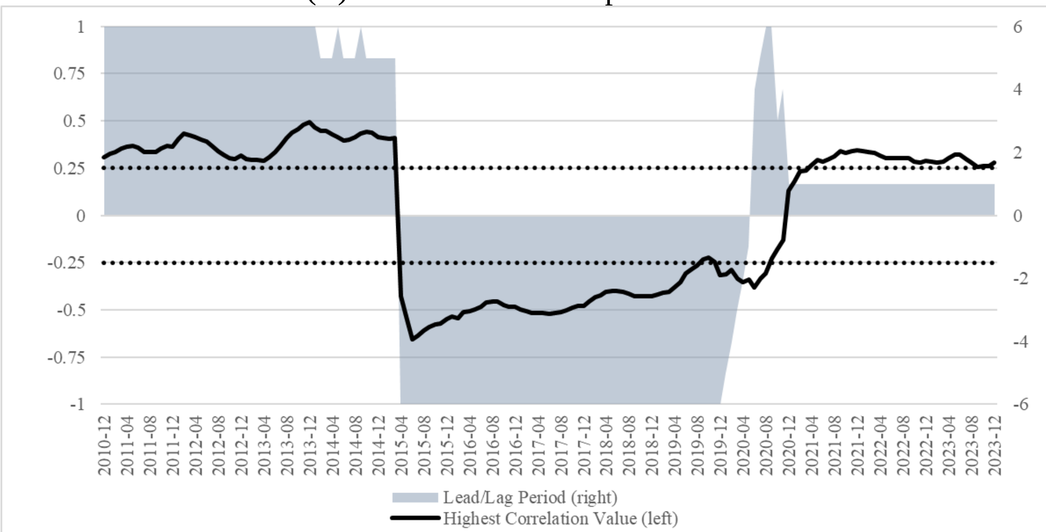


Figure 4: Continued.

(C) Labor Force Participation Rate



4.B. Leading Indicators

Studies relating entrepreneurship to the business cycle tend to focus on outcomes: business launches and the actual state of the economy. Thurik (2014), for example, tests the hypothesis that entrepreneurship is pre-cyclical (positively correlated with the business cycle, and leading). The BEI may capture entrepreneurial activities leading up to business launch, and the FRED database provides a leading economic indicator for Virginia. An opportunity is available for us to study the forecasting potential of entrepreneurial interest.

The correlation between the BEI and the leading indicator may be either positive or negative. Thurik (2014) notes entrepreneurs invest substantial efforts in their start-ups. If they believe a recession is coming, would-be entrepreneurs may delay their plans to avoid a failure. Waiting until the economy improves and higher profits are more likely appears to be a sensible strategy. However, Thurik (2014) also acknowledges the role of necessity entrepreneurship: an expected economic crisis might encourage people to launch their own business to supplement their income. Devece et al. (2016) and Peris-Ortiz et al. (2014) agree that the state of the macroeconomy impacts entrepreneurial motivation, but opportunity recognition and innovation activity are also important. Both studies show a reduction in the number of entrepreneurial firms during recessions in Spain, however firms that focus on innovation activities and opportunity recognition are more resilient during slow-downs.

We use the following data to explore how recession forecasts might impact entrepreneurial interest in Virginia:

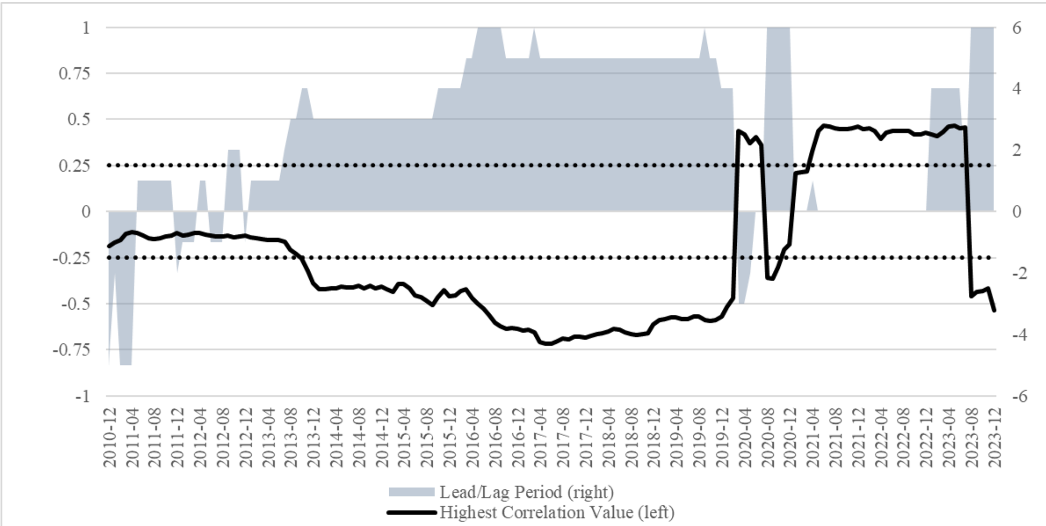
Variable	FRED Description	Period Analyzed
Leading index	Leading Index for Virginia, Percent, Monthly, Seasonally Adjusted. The leading index for each state predicts the six-month growth rate of the state's coincident index. In addition to the coincident index, the models include other variables that lead the economy: state-level housing permits (1 to 4 units), state initial unemployment insurance claims, delivery times from the Institute for Supply Management (ISM) manufacturing survey, and the interest rate spread between the 10-year Treasury bond and the 3-month Treasury bill.	1/2006 to 2/2020

Figure 5 illustrates the dynamic relationship between the BEI and the leading economic indicator for Virginia. The results show that the cyclical components of these two series are negatively correlated in general, with the leading indicator *lagging* behind the BEI for much of the time span. In other words, an increase in the BEI above trend today will be followed by a decrease in the leading indicator 3 to 5 periods in the future. This finding is consistent with Thurik (2014). The correlating relationship shifts after 2020, but this change requires more data and further analysis to verify accurately.[8]

The leading indicator gives us information about the expected state of the economy in the future and the labor market gives us information about the state of the economy in the past and present. Combining the results from Figure 5 with those from Figure 4A-B point to a complex dynamic relationship between entrepreneurial interest and overall economic performance. Prior to 2018, above-trend entrepreneurial interest occurs when the economy is projected to worsen in the future. At the same time, entrepreneurial interest tends to be high when the labor market was strong in the recent past (below-trend unemployment above-trend wages). Entrepreneurial interest therefore peaks if the economy was recently strong but is forecast to weaken, and falls if the economy was recently poor but is forecast to improve. It is as if current opportunities and future necessity together shape entrepreneurial mindsets. Explaining this dynamic relationship further is a source of future research.

[8] The correlation window in later periods of Figure 5 is less than 60-months since data for the leading economic indicator ends in 2/2020. Data is phased out as it becomes unavailable, shrinking the window (e.g. the correlations in 2022 are constructed using data from 2018 to 2020 whereas the correlations in 2023 are constructed using data from 2019 to 2020). This allows us to utilize as much data as possible and construct figures with consistent time axes, but means we need more data on the leading economic indicator to analyze the post-pandemic patterns fully.

Figure 5: Most extreme correlations between the BEI and leading economic indicator data for Virginia.



4.C. Home Value

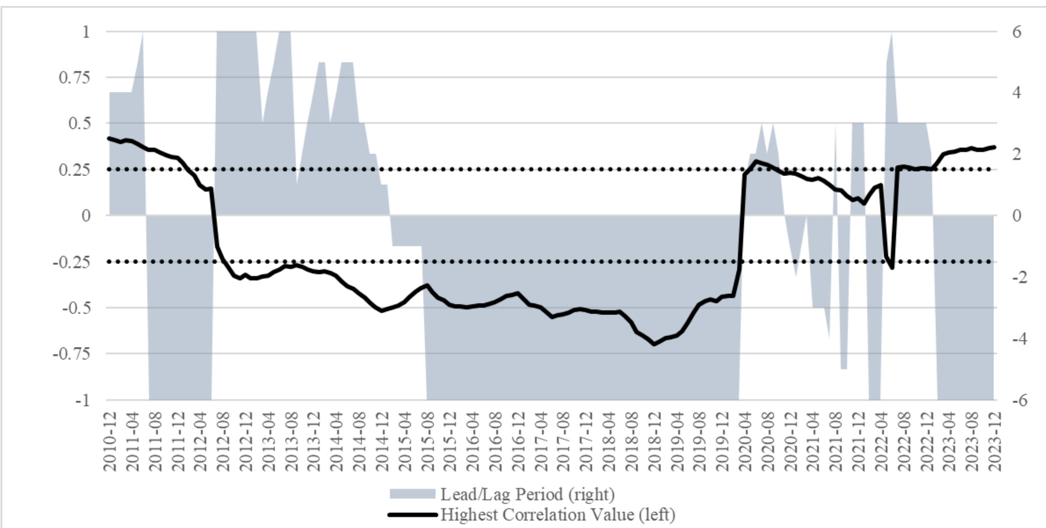
Recent studies looking at the connection between entrepreneurship and home prices note that real estate is used as collateral by many entrepreneurs. An increase in home prices therefore helps some entrepreneurs secure larger business loans to fund their start-up. The size of the effect depends on how many prospective entrepreneurs face collateral constraints, how many entrepreneurs acquire funds from other sources, and the alternative opportunities associated with rising home prices. For example, Schmalz et al. (2017), Berggren and Wilhelmsson (2019), and Kerr et al. (2022) identify positive correlations between home equity growth and entrepreneurship, though the effect is mild in some cases. Hu et al. (2019) finds a negative impact of house price on entrepreneurship for China, suggesting labor market effects (home prices rising imply good job market opportunities in existing business) and real estate investment (rising home prices make the return to investing in real estate better compared to investing in new ventures) keep people from pursuing entrepreneurship despite the increase in the value of collateral. As before, studies in this area use actual launches and not measures of general interest.

To explore the correlating relationship between the BEI and home prices for Virginia, we use the following data:

Variable	FRED Description	Period Analyzed
Home value index	Zillow Home Value Index (ZHVI) for All Homes Including Single-Family Residences, Condos, and CO-OPs in Virginia, Dollars, Monthly, Smoothed Seasonally Adjusted. The Zillow Home Value Index or ZHVI is a smoothed, seasonally adjusted measure of the typical home value and market changes across a given region and housing type. It reflects the typical value for homes in the 35th to 65th percentile range.	1/2006 to 12/2023

Figure 6A shows a broadly negative correlation between entrepreneurial interest and home prices, with home prices leading when data from 2015 to 2019 is a part of the rolling window. When home prices rise above their trend, the BEI will tend to fall below its trend six or more months later. The relationship shifts in 2015 (likely due to the dissipating effects of the Great Recession that greatly impacted the real estate sector) and after 2020 (likely due to data limitations and the impact of the pandemic). This finding differs from earlier studies that suggest a positive relationship should be present. Further, this result is only partially consistent with the negative correlations found by Hu et al. (2019), with the labor market results found in section 4A above being a point of departure.

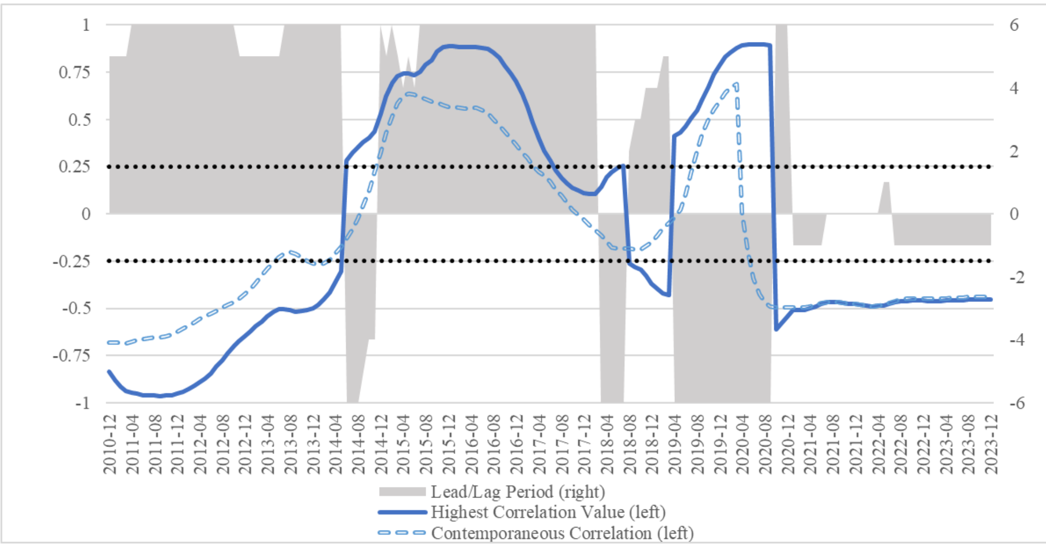
Figure 6A: Most extreme correlations between the BEI and the home value index for Virginia.



The logic in Hu et al. (2019) is built on the notion that a strong labor market (low unemployment and high wages) appears alongside rising real estate prices as both can describe an economic boom. The Virginia data used in this study does not consistently exhibit this pattern. Figure 6B shows the dynamic correlation pattern between the home price index and the unemployment rate.

Above-trend real estate prices did occur along with below-trend unemployment prior to 2014. However, this relationship reversed for much of the 2015-2020 period. Multiple shifts in the lead/lag structure occur, with home prices generally leading unemployment before 2019. Taken together, we see a cooling real estate market preceding a strengthening labor market, followed by rising interest in entrepreneurship for a large part of the 2015-2020 period.

Figure 6B: Contemporaneous and most extreme correlations between the home value index and the unemployment rate for Virginia.[9]



Although a hot housing market creates credit opportunities for entrepreneurs, rising housing prices may not drive Virginian's thinking towards start-ups because they are sometimes associated with poor economic conditions. Alternatively, growth in home prices may pose more challenges than opportunities for would-be entrepreneurs (such as higher mortgage obligations). A deeper look into the Virginia real estate market, its contribution to economic performance, and its role in entrepreneurial finance may help identify these effects.

[9] Figure 6B shows the most extreme value for Corr(Housing Prices$_t$, Unemployment Rate$_{t+j}$) for j = -6 to 6, the period in which the extreme value occurs, and the contemporaneous correlation (j=0).

4.D. Business Applications

The FRED database contains data on business applications for Virginia, allowing us to see the dynamic relationship between the BEI and actual new firm creation. Although all the firms in the applications data are not entrepreneurial per se, we can draw broad conclusions about new firm mindsets and new firm actualization using the BEI. This is possible since some of the terms used to construct the BEI apply to tasks done by all new businesses, hence the BEI partially captures general interest in new firm creation.

The following variables are used to explore the relationship between the BEI and business launches:

Variable	**FRED Description**	**Period Analyzed**
Business applications with planned wages	Business Applications with Planned Wages: Total for All NAICS in Virginia, Number, Monthly, Seasonally Adjusted. High-Propensity Business Applications (HBA) that indicate a first wages-paid date on the IRS Form SS-4. The indication of a wages-paid date is associated with a high likelihood of transitioning into a business with payroll.	1/2006 to 12/2023
High propensity business applications.	High-Propensity Business Applications: Total for All NAICS in Virginia, Number, Monthly, Seasonally Adjusted. High-Propensity Business Applications (HBA): A subset of Business Applications (BA) that contains all applications with a high-propensity of turning into a business with a payroll, based on various factors.	1/2006 to 12/2023
Total business applications	Business Applications: Total for All NAICS in Virginia, Number, Monthly, Seasonally Adjusted. The core business applications series that correspond to a subset of all applications for an Employer Identification Number (EIN). Includes all applications for an EIN, except for applications for tax liens, estates, trusts, certain financial filings, applications outside of the 50 states and DC or with no state-county geocodes, applications with certain NAICS codes in sector 11 (agriculture, forestry, fishing and hunting) or 92 (public administration) that have low transition rates, and applications in certain industries (e.g. private households, civic and social organizations).	1/2006 to 12/2023
Projected business formation	Projected Business Formations Within Four Quarters: Total for All NAICS in Virginia, Number, Monthly, Not Seasonally Adjusted. The projected number of employer businesses that originate from Business Applications (BA) within four quarters from the quarter of application.	1/2006 to 12/2023

A positive, lagging correlation is expected since we hope increased interest in entrepreneurship produces more new firms later on. However, a negative relationship is possible. More firms entering the market means stronger competition and fewer sales, ceteris paribus. This might dissuade interest in entrepreneurship. Figure 7 shows mixed results. Positive correlations between the BEI and business applications (7A-C) are present when data before 2015 is considered. During this period, applications *lead* the BEI. An exception to this is with total business applications (7C), which lag the BEI when the rolling window covers data before 2013 and between 2014 and 2015. Although we expect greater interest in entrepreneurship to result in more business applications down the line, seeing applications lead interest is not impossible. Perhaps noticing more new firms on the scene inspires people to become entrepreneurs.

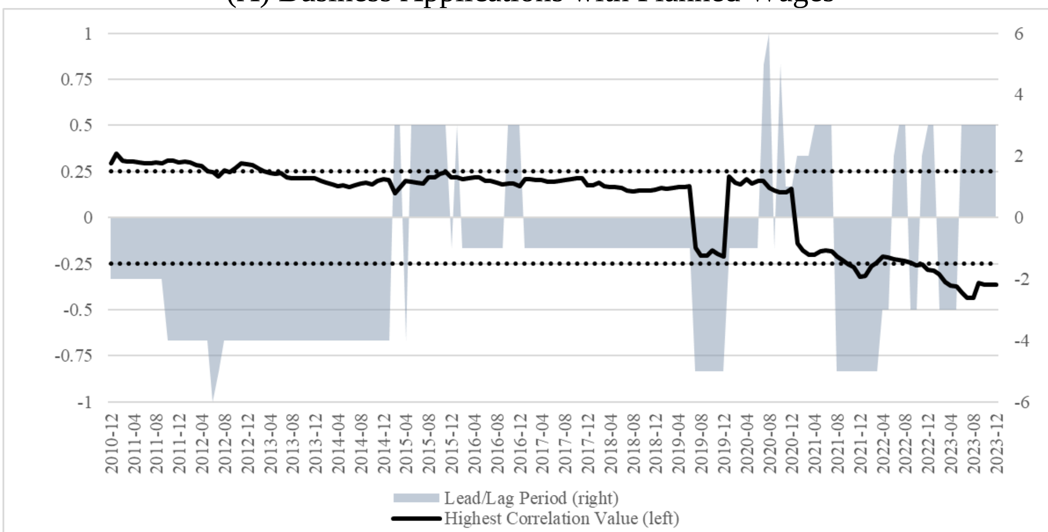
When the rolling window covers data from 2015-2020, the correlations become less significant. The correlations turn negative when data after 2020 is considered, with less stable lead/lag results. Since business applications are still leading the BEI for many periods, it is possible that new business applications rising above trend are providing new employment opportunities, increasing market competition, and driving interest away from entrepreneurship. Why this shift occurs is worthy of further study.

The correlations between the BEI and business applications is overall quite mild (close to the ±0.25 cutoff for significance in most periods). This implies the BEI is providing different information compared to data generated from new business licenses. We expect this since the BEI is capturing entrepreneurial activity at various stages in the launch process, not just near the launch date. Further, the BEI captures interest from everyone, not just those engaged in a start-up. The gap between interest and action demands further exploration, presenting a new opportunity for research.

Projected business formation, the number of businesses formed within 1 year of the business application, seems to be positively related to the BEI (pre-2015 and after 2020) and lag entrepreneurial interest by 1 period (Figure 7D). This supports the logic that entrepreneurs are engaging in keyword search for information as they work on their venture, culminating in a launch shortly thereafter. The strong correlations between the BEI and terms like “copyright”, “patent”, “business plan”, and “business license” (Figure 2) imply entrepreneurs are actively working on their idea and are perhaps close to formally starting their businesses.

Figure 7: Most extreme correlations between the BEI and new business applications for Virginia.

(A) Business Applications with Planned Wages



(B) High Propensity Business Applications

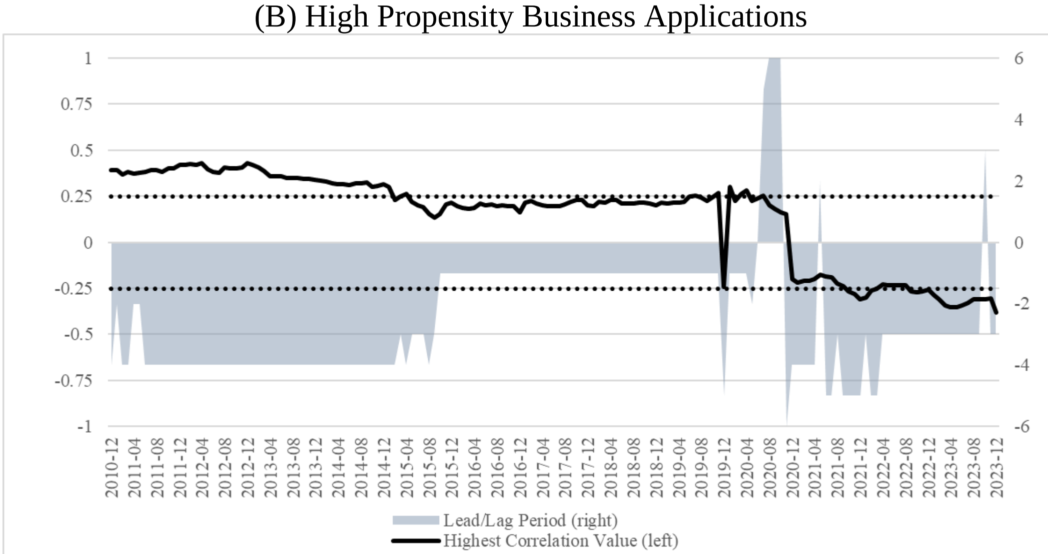
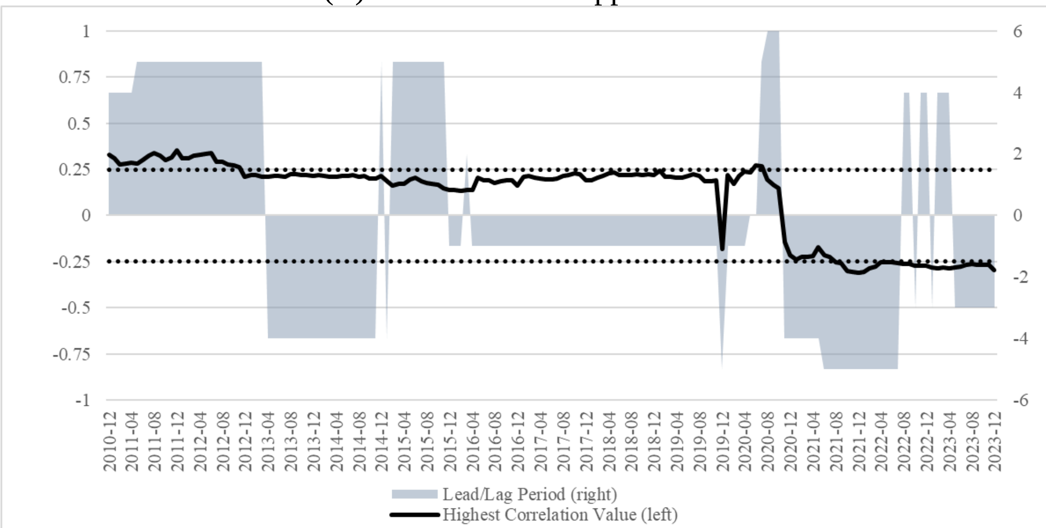
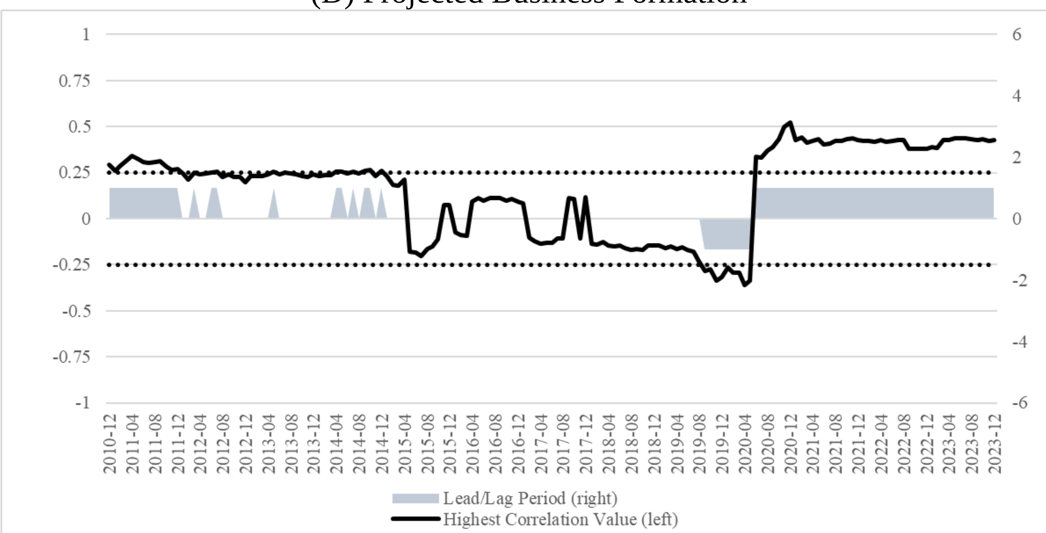


Figure 7: Continued.

(C) Total Business Applications



(D) Projected Business Formation



5. Concluding Remarks

5.A. Summary of Findings

Crafting an index using internet keyword search frequency gives us a convenient way to broadly capture general interest in entrepreneurship. Many new insights are discovered from performing a cyclical analysis for Virginia. First, the dynamic correlations between the BEI and the various types of economic data we consider exhibit shifts, with 2015 and 2019-2020 represent turning points in most cases. Second, results from the labor market support the idea of opportunity-driven entrepreneurship: when the labor market is good, there is more interest in entrepreneurship. Correlations to the labor force participation rate illustrate how entrepreneurial interest might be drawing people into employment. Next, results from the leading economic indicator imply that entrepreneurial interest has some forecasting potential. Also, we do observe some connections between home prices and entrepreneurial interest that are novel yet also consistent with previous studies. Finally, results from looking at business applications show that the BEI is providing new information compared to metrics connected to real business activity.

Timing is a key theme in this study. Analyzing leading/lagging correlations produces new outcomes to investigate. Interest is shown to rise if the labor market was strong in the recent past. Interest is also shown to rise when the economy is projected to be weak. Further investigation shows that rising real estate prices do not always appear when the Virginia economy performs well, and tends to result in lower entrepreneurial interest. The weak correlation between business applications and entrepreneurial interest suggest additional factors are at play as early-stage entrepreneurial activity leads to actual business launches. The use of a 60-month rolling window shows that these relationships are not static.

5.B. Policy Implications

Using keyword search data to construct an entrepreneurship interest index and conducting a correlation analysis can be useful to policymakers and non-profit organizations. Smallbone (2020) provides a general discussion about entrepreneurship policy. Key policy issues from Smallbone (2020) that connect to the approach used in this study include (1) paying attention to the context of entrepreneurship to identify trends and needs, (2) relying on evidence in policy development, and (3) improving how policy is evaluated. Keyword search data provides hidden or costly-to-discover information about entrepreneurial culture and activities. Further, the data is high-frequency (monthly), regional, and easy-to-handle. This helps identifying trends and forecasting patterns to use in policy development. For example, we can look into specific connections, such as entrepreneurial interest by industry type, to create policies. Also, watching for sharp shifts in particular keywords in response to policy changes may be a useful policy evaluation tool.

Two exact types of policy the BEI analysis can support are (1) the promotion of an entrepreneurial culture and (2) the development of start-up support programs. Programs seeking to improve entrepreneurial culture can monitor changes in the entrepreneurial interest as measured by the BEI before and after some treatment is applied. For example, an upswing in the BEI after the launch of a positive entrepreneurship marketing campaign provides early evidence that the campaign is reaching people even though we cannot measure the actual impact until after new businesses develop.

Monitoring individual keyword trends helps start-up support programs tailor their services. The types of keywords searched for tell us what information people need about business activities (like "how to get a business license" or "how to file for a patent"). Highly-searched keywords point to topics that incubators, educators, and policymakers should focus their sights on. However, current developments in AI are starting to take the place of traditional search engines. Obschonka and Audretsch (2020) and Mu and Zhao (2024) note that AI can play a role in entrepreneurial education and training, making for a possible policy overlap between AI, business operations, and higher education.

5.C. New Research Paths

This study brings to light new research opportunities.

- **Geographic Comparison.** The analysis in this study can be done for other US states to identify differences between the Virginia experience and those elsewhere. Common threads and points of divergence will improve our understanding of entrepreneurial interest.
- **Alternative Indices.** Using different keywords to construct the interest index can help explore other aspects of the entrepreneurial process. Constructing an index representing entrepreneurial traits (like risk tolerance and creativity) may be useful in studying the prevalence of these personal characteristics in the entrepreneurial ecosystem.
- **Alternative Data and Trend Modeling.** Applying different methods for statistically modeling cycles is a worthy econometric exercise to conduct to evaluate the robustness of the findings. An additional analysis focusing on long-term trends will yield information about the overall growth of entrepreneurial interest. Connecting the BEI to other types of regional,

monthly data will help develop new topics in the study of entrepreneurial interest.

- **Labor Force Participation.** The relationship between the labor force participation rate and the BEI exhibited stark shifts in 2015 and 2020. Further investigation into this shifting relationship may help us understand how people outside the formal labor force might view entrepreneurship. Policymakers may find strategies that help generate opportunities for discouraged workers.
- **Real Estate and Finance.** Additional investigation into the connection between entrepreneurial interest and regional access to finance might be of value, particularly when looking at the spatial distribution of entrepreneurs across a region. This study focused on the relationship between entrepreneurial interest and real estate prices at the state-level. A more local analysis (city/county level) may generate further insights. Other types of region-based financial assets and credit options that entrepreneurs might take advantage of may yield useful results, such as interest rates at local banks.
- **Interest Versus Action.** The weak correlation between the BEI and actual business applications suggests there are still new things to discover about how entrepreneurial interest leads to the creation of new firms. For example, it is possible that internet search is filtering ideas to improve efficiency which reduces the correlation (i.e., would-be entrepreneurs using online search to resolve business challenges may halt their efforts after discovering too many complications). A theme-group approach looking at which groups in Figure 1 correlate most to business applications can help identify which types of entrepreneurial interests lead to action.
- **Economic Performance.** This study showed the past state of the economy and the projected future state of the economy impact entrepreneurial interest differently across time. It was also shown that some variables conventionally associated with a strong economy do not always manifest simultaneously. Investigating other features of a strong economy (such as increases in consumer spending and inventory orders) would give us a more comprehensive perspective on the economic factors that encouraging entrepreneurial interest.
- **The Covid-19 Pandemic and Post-2020.** Further analysis of the breaks in the correlating relationships occurring after 2020 can be done as more data becomes available. In-depth exploration of the shifts brought on during the Covid-19 pandemic is a unique research opportunity. The internet played a key role during the crisis, and treating the pandemic as a natural

experiment may yield new patterns about online interest in entrepreneurship.

5.D. Final Remarks

Keyword search analysis has potential when it comes to monitoring entrepreneurial activity. Our understanding of the entrepreneurial process and the nature of the entrepreneurial ecosystem can be improved by looking at culturomics data. The policy landscape can benefit from the added information that an entrepreneurship interest index can provide. Leveraging big data tools opens many new lines of research on regional entrepreneurial activity.

6. AI Disclaimer

This project utilized Google Gemini, a large language AI model. Gemini (formally known as Google Bard) was used to produce an initial list of potential keywords for the analysis and to recommend language edits to written content. At no time was AI used to create original content, analyze data or evaluate results.

7. Appendix

Appendix 1: Notes on Google Trends Data Collection

There are four notes to be aware of regarding data collection from Google Trends. The first is that Google Trends data starts in 2004. However, the early era of internet search (2004 and 2005) can produce extreme or erratic patterns that are difficult to fit. A starting date of 2006 is chosen to avoid these patterns as they may affect the data in other periods.

Second, Google Trends data is a scaled index of internet search frequency; the date with the highest search volume is set to 100 and all other periods' frequencies are set to a percentage of that date's volume. For example, the data in Figure A1 below shows that September 2022 had the highest volume of searches for the term "entrepreneur", and December 2022's volume of searches was 51% of September's. Note that the actual search volume is not given (i.e. we do not know how big a volume an index number of 100 really is), but the pattern we are interested is still present.

Figure A1: Google Trends data for "entrepreneur" searches.

Period	"entrepreneur" index
2022-07	54
2022-08	71
2022-09	100
2022-10	79
2022-11	71
2022-12	51

The third note is that data is obtained from Google Trends one keyword at a time to avoid a scaling issue. Although it is possible to feed multiple terms into Google Trends with the hopes of acquiring a lot of data simultaneously, doing this results in all data being scaled relative to the term with the highest volume. This is because Google Trends is designed to help with frequency comparisons. For example, the data in Figure A2 below is obtained by asking Google Trends to report trends for "entrepreneur" and "innovation" simultaneously. The results show that searches for "innovation" in March 2017 had the highest volume. The index number for January 2017 for "entrepreneurship" suggests that the volume searches for entrepreneurship during that month was 25% of the volume of searches *for "innovation" in May 2017.* Because "entrepreneurship" is a less popular search term relative to "innovation", all the entrepreneurship index numbers are diminished compared to the numbers in Figure A1. If we ask for terms one at a time, the index numbers will be scaled to relative to the highest volume *of that search term* (i.e. each time series will have a max of 100 somewhere in the 2006-2023 period). This produces the clearest patterns, however we are not able to compare volumes across terms.

Figure A2: Google Trends data for "entrepreneur" searches and "innovation" searches, simultaneous.

Period	"entrepreneur" index	"innovation" index
2017-01	25	86
2017-02	23	94
2017-03	24	100
2017-04	24	85
2017-05	21	86

The final note is that terms can represent different concepts. For example, suppose you obtain Google Trends data for the search term "shark". People may be looking for aquariums, books with a specific title, types of toys, documentaries or movies about sharks, or for information about the popular entrepreneurship show *SharkTank*. People may have also misspelled their search and were really looking for "stark" or "spark". Multiple intentions can be included in the results produced by Google Trends. It is possible to use a categorical sort feature to narrow down meaning, however this analysis allows for all categories. We want to capture general interest and overall mindset from searches with a wide variety of intentions.

8. References

Berggren, B., A. Fili, and M. Wilhelmsson. 2019. "Homeownership and Entrepreneurship: A Regional and Industrial Analysis of House Prices and Startups." International Journal of Housing Markets and Analysis 12(3): 456-473.

Blundel, R., N. Lockett, C. Wang, and S. Mawson. 2021. Exploring Entrepreneurship: Practices and Perspectives, 3rd ed. London: SAGE Publications.

Carland, J. W., Hoy, F., Boulton, W. R., and Carland, J. A. C. 1984. "Differentiating Entrepreneurs From Small Business Owners: A Conceptualization." Academy of Management Review 9(2): 354-359.

Carneiro, H. A., and E. Mylonakis. 2009. "Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks." Clinical Infectious Diseases 49(10): 1557-1564.

Devece, C., Peris-Ortiz, M., and Rueda-Armengot, C. 2016. "Entrepreneurship During Economic Crisis: Success Factors and Paths to Failure." Journal of Business Research 69(11): 5366-5370.

Farhat, D. and T. Viitanen. 2017. "Business 'Psych'cles: A Close Look at Mental Health and State-Level Economic Performance Using Google Search Data." Review of Economic Analysis 9(2): 107-125.

Farhat, D. 2017. "Awareness of Sexually Transmitted Disease and Economic Malady Using Search Engine Query Data." International Journal of Business and Economics 16(1): 101-108.

Farhat, D., D. Kunkel, and J. Quesenberry. 2019. "Peaked Interest: Public Interest in Hunger and the Economic Cycle." Journal of Applied Business and Economics 21(7): 25-38.

Federal Reserve Bank of St. Louis. 2024. FRED Economic Data. Available at https://fred.stlouisfed.org/.

Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M.S. Smolinski, and L. Brilliant. 2008. "Detecting Influenza Epidemics Using Search Engine Query Data." Nature 457(7232): 1012-1014.

Gómez Martínez, R., M. Prado Román, and C. Mercado Idoeta. 2014. "Google Search Activity as Entrepreneurship Thermometer." In New Challenges in Entrepreneurship and Finance: Examining the Prospects for Sustainable Business Development, Performance, Innovation, and Economic Growth, edited by N. Peris-Ortiz and J. M. Sahut, 225-233. Cham: Springer International Publishing.

Google Books Ngram Viewer. 2024. Available at https://books.google.com/ngrams/.

Google Gemini. 2024. Available at https://gemini.google.com/.

Google Trends. 2024. Available at https://trends.google.com/trends/.

Hu, M., Y. Su, and W. Ye. 2019. "Promoting or Inhibiting: The Role of Housing Price in Entrepreneurship." Technological Forecasting and Social Change 148: 119732.

Huynh, T. L. D. 2019. "Which Google Keywords Influence Entrepreneurs? Empirical Evidence from Vietnam." Asia Pacific Journal of Innovation and Entrepreneurship 13(2): 214-230.

Isenberg, D. 2011. "The Entrepreneurship Ecosystem Strategy as a New Paradigm for Economic Policy: Principles for Cultivating Entrepreneurship." Presentation at the Institute of International and European Affairs.

Jena, A. B., P. Karaca-Mandic, L. Weaver, and S. A. Seabury. 2013. "Predicting New Diagnoses of HIV Infection Using Internet Search Engine Data." Clinical Infectious Diseases 56(9): 1352-1353.

Kerr, S. P., W. R. Kerr, and R. Nanda. 2022. "House Prices, Home Equity and Entrepreneurship: Evidence from US Census Micro Data." Journal of Monetary Economics 130: 103-119.

Koellinger, P. D., and R. Thurik. 2012. "Entrepreneurship and the Business Cycle." Review of Economics and Statistics 94(4): 1143-1156.

Michel, J. B., Y. K. Shen, A. P. Aiden, A. Veres, M.K. Gray, Google Books Team, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, and J. Orwant. 2011. "Quantitative Analysis of Culture Using Millions of Digitized Books." Science 331(6014): 176-182.

Mu, Q., and Y. Zhao. 2024. "Transforming Entrepreneurship Education in the Age of Artificial Intelligence." Resources Data Journal 3: 2-20.

Obschonka, M., and D. B. Audretsch. 2020. "Artificial Intelligence and Big Data in Entrepreneurship: A New Era Has Begun." Small Business Economics 55: 529-539.

Omerzel Gomezelj, D., and Kušce, I. 2013. “The Influence of Personal and Environmental Factors on Entrepreneurs' Performance.” Kybernetes 42(6), 906-927.

Peris-Ortiz, M., Fuster-Estruch, V., and Devece-Carañana, C. 2014. “Entrepreneurship and Innovation in a Context of Crisis”. In Entrepreneurship, Innovation and Economic Crisis: lessons for Research, Policy and Practice, edited by K. Rüdiger, M. Peris-Ortiz and A. Blanco-González, 1-10. Cham: Springer International Publishing.

Polgreen, P. M., Y. Chen, D. M. Pennock, F. D. Nelson, and R. A. Weinstein. 2008. “Using Internet Searches for Influenza Surveillance.” Clinical Infectious Diseases 47(11): 1443-1448.

Schmalz, M. C., D. A. Sraer, and D. Thesmar. 2017. “Housing Collateral and Entrepreneurship.” The Journal of Finance 72(1): 99-132.

Semerci, A. B., A. A. Özgören, and D. İçen. 2022. “Thoughts on Women Entrepreneurship: An Application of Market Basket Analysis with Google Trends Data.” Soft Computing 26(19): 10035-10047.

Shwetzer, C., A. Maritz, and Q. Nguyen. 2019. “Entrepreneurial Ecosystems: A Holistic and Dynamic Approach.” Journal of Industry-University Collaboration 1(2): 79-95.

Smallbone, D. 2020. “Entrepreneurship Policy: Issues and Challenges.” Understanding the Development of Small Business Policy, edited by T. M. Cooney, 6-23. London: Routledge.

Thompson, J. L. 1999. “The World of the Entrepreneur: A New Perspective.” Journal of Workplace Learning 11(6): 209-224.

Thurik, R. 2014. “Entrepreneurship and the Business Cycle.” *IZA World of Labor*.

Uhlaner, L., and R. Thurik, R. 2007. “Postmaterialism Influencing Total Entrepreneurial Activity Across Nations.” Journal of Evolutionary Economics 17(2): 161-185.

US Small Business Administration. 2022. “Table of Small Business Size Standards”. Available at https://www.sba.gov/document/support-table-size-standards.