

# Corporate Venture Capital and Firm Scope

Yifei Zhang   
Peking University HSBC Business School  
tseyifeizhang@phbs.pku.edu.cn

## Abstract

This study examines whether and how corporate venture capital (CVC) spurs changes in firm scope. Using two text-based measures of firm scope, I provide evidence that CVC investments are strongly correlated with subsequent changes in firm scope among CVC parent firms, including seeding emerging businesses and creating new segments or divisions. Further evidence is consistent with an experimentation view, with more promising ventures having a stronger strategic impact on the scope changes of parent firms. Moreover, the study finds that post-CVC scope changes are primarily built internally and rarely involve killer acquisitions. These changes create value for CVC parents.

## I. Introduction

Understanding the scope and boundaries of firms is a fundamental topic in the fields of economics and finance.<sup>1</sup> However, the empirical research on the determinants of firm scope is limited (Hoberg and Phillips (2018)).<sup>2</sup> Surprisingly, there is also a lack of studies exploring the relationship between corporate innovation strategies and the dynamics of firm scope, despite the widely held belief that innovation drives the growth and evolution of firms and economies (Bena and Li (2014), Seru (2014), and Frésard, Hoberg, and Phillips (2020)). This study aims to fill this gap in the literature by examining the relationship between corporate venture capital (CVC) investments, an open corporate innovation strategy, and the dynamics of firm scope. CVC investments are popular among large industry leaders and provide a valuable opportunity to explore the impact of innovation strategies on firm scope. By investigating the effects of CVC investments on firm scope, this study contributes to the empirical literature on firm scope and expands

---

I thank the anonymous reviewer and Jarrad Harford (the editor) for their helpful suggestions. I am indebted to Ulrich Hege for his guidance and suggestions. I also thank Rustam Abuzov (discussant), Milo Bianchi, Patrick Coen, Florian Ederer, Caroline Gene (discussant), Di Li, Baixiao (Tony) Liu, Song Ma, Andrew Metrick, Sophie Moinas, Filip Mrowiec, Sebastien Pouget, Armin Schwienbacher (discussant), Merih Sevilir (discussant), Guillaume Vuillemy (discussant), Dong Yan (discussant), Xiang Zheng (discussant), and conference and seminar participants of 2021 Eastern Finance Association, 2021 European Finance Association – Doctoral Tutorial, 2021 WEFI Student Workshop, 2022 American Finance Association, 2022 MFA, and 2022 European Finance Association for their helpful comments. All errors are mine.

<sup>1</sup>Important milestones include transaction cost theory (Coase (1937), Williamson (1985)) and the property rights approach (Grossman and Hart (1986), Hart and Moore (1990)).

<sup>2</sup>The traditional literature on conglomerates often assumes firm scope as a given and rarely delves into the determinants of firm scope.

our understanding of the relationship between corporate innovation strategies and firm dynamics.

A CVC program is a venture capital (VC) arm affiliated with an established firm. In recent decades, CVC has emerged as a significant tool in the open innovation strategies of numerous prominent companies, including industry leaders such as Apple, Google, and Microsoft. Furthermore, owing to its distinctive characteristics, CVC presents a unique opportunity to explore a novel strategy for firm growth and scope change, known as the experimentation strategy.<sup>3</sup>

This study aims to examine the impact of CVC on firm scope changes. The hypothesis is based on anecdotal and survey evidence, which suggests that the establishment of a CVC program can assist the CVC's parent firm (e.g., Google) in identifying new business opportunities and directions. There are several ways in which a CVC program or CVC investments can facilitate the identification of these emerging business opportunities. First, during the pursuit of potential deals, a CVC parent can gain access to new business models and opportunities that may not be readily available within the firm's existing boundaries (Chesbrough (2002)). Second, through interactions with startup managers subsequent to investments, a CVC parent firm often acquires valuable first-hand information regarding emerging business opportunities and new markets (Keil, Autio, and George (2008)).<sup>4</sup>

I document in my data that, given newly identified business opportunities from CVC investment, a CVC parent firm frequently integrates those new businesses into its current business domain, thus reshaping its firm scope. Additional evidence supports the notion that the strategic impact of these changes is stronger for ventures that show greater promise. Furthermore, my findings indicate that these changes in firm scope, following CVC investments, predominantly occur through internal development rather than through "killer acquisitions," which account for only approximately 6% of the observed scope changes. In killer acquisitions, a CVC parent firm acquires a startup from its investment portfolio (Cunningham, Ederer, and Ma (2021)).

To measure changes in firm scope, I have developed two distinct measures. First, I employ textual analysis to identify emerging businesses and quantify the number of newly added emerging businesses in the annual 10-K business descriptions of publicly listed firms. I determine emerging businesses using "emerging phrases," which represent the top 5% most commonly used short phrases in the business vocabulary of VC-backed startups for a given year.<sup>5</sup> Second, I construct a text-based firm scope measure inspired by the work of Hoberg and Phillips (2021).

<sup>3</sup>Compared with other instruments that firms have at their disposal to foster innovation (e.g., in-house efforts to carry out R&D and create new intellectual property, acquisitions of research results or innovative startups, the recruitment of employees with new expertise), CVC offers the advantage that firms' initial investment decisions, as well as metrics of investment outcomes, can be observed hence offering an exciting view on the use of experimentation in firm strategy.

<sup>4</sup>In Keil et al.'s (2008) interview-based study, they conclude as follows: "Once the corporate venture capital managers had made investments in interesting ventures and formed insights concerning potential capability voids within the incumbent, they frequently connected these startups with [CVC parent firm's] senior managers to provide senior managers with a window on new technologies and business models. Influencing decision-making in this way is frequently an informal process in which it is important to expose the right decision-makers to the right external contexts."

<sup>5</sup>For a visual representation, please refer to Figure 2.

This measure allows me to identify industries where a firm establishes new segments.<sup>6</sup>

Using these metrics, I discover a strong correlation between CVC investments and subsequent changes in firm scope. Specifically, compared to industry-year peers that do not have CVC investments, a CVC corporate parent, on average, adds 1.5 (100%) more emerging phrases to its 10-K annual business description. This “treatment” effect of CVC deals is primarily observed within the 2 years following the investment, with no discernible effect before or 3 years later. Additionally, there is a strong connection between the emerging business of a startup involved in a CVC deal and the subsequent incorporation of emerging phrases by the CVC parent firm conducting the deal.<sup>7</sup> Regarding the second measure, I have also found similar results. Firms with CVC investments are 54% more likely to create new segments in a new VEIC industry within 2 or 3 years after the deal. This suggests that CVC investments have a notable influence on the expansion of firm scope by encouraging the establishment of new segments in previously unexplored industries.

Despite the significant correlation identified in the analysis, an alternative explanation for the findings is that CVC investments are merely correlated with the presence of novel technologies in the economy, and it is these technology changes that drive the observed changes in firm scope, rather than the CVC investments themselves. In a similar scenario, companies with outdated products and services are more inclined to seek out innovative ideas through CVC deals (Ma (2020)). Consequently, it is plausible that even in the absence of CVC deals, these companies would eventually change firm scope, thereby introducing an omitted variable bias concern.

To better identify the causal impact of CVC deals on changes in firm scope, I use the introduction of new (nonstop) airlines connecting the locations of the CVC firm headquarters and the startup involved in the CVC deal. This approach allows for two distinct identification exercises to be conducted. The first identification strategy closely aligns with the approach utilized by Bernstein, Giroud, and Townsend (2016), who employ the introduction of new (nonstop) airline routes in the United States as a proxy for capturing the effects of VC monitoring and involvement on the future success of portfolio companies.<sup>8</sup> After the establishment of a new

---

<sup>6</sup>Unlike Hoberg and Phillips (2021), who utilize FIC or NAICS industry vocabulary, my text-based segment measure is based on the VentureXpert industry (VEIC) classification and vocabulary. The overall procedures closely align with those outlined by Hoberg and Phillips (2021), with the key distinction lying in the vocabulary and classification used.

<sup>7</sup>In other words, CVC deals related to a specific emerging business, such as artificial intelligence, can predict the inclusion of corresponding emerging phrases, such as “artificial intelligence,” in the parent firm’s 10-K.

<sup>8</sup>One major difference between the identification strategy in Bernstein et al. (2016) and my exercise is the way in which we define and exploit the introduction of new (nonstop) airlines. While Bernstein et al. (2016) consider the first-time introduction of a nonstop flight between two locations, irrespective of the airline companies involved, I focus on the introduction of new (nonstop) airlines at the carrier level. This means that if Delta Airlines introduces a new nonstop flight between Dallas and Boston, I consider it as a new airline, even if American Airlines had already been operating a nonstop flight between the same destinations prior to Delta’s entry. In contrast, Bernstein et al. (2016) only consider the first airline company to establish a nonstop route, such as American Airlines in this example. Adding an extra airway is correlated with better on-time performance, greater service quality, more flexible time schedules, and

(nonstop) airline route, CVC firm managers are more inclined to visit startups within their portfolio located at the destination of the new airline (Bernstein et al. (2016)). This increased interaction provides the CVC parent firm with valuable opportunities to gain insights into new business models, market opportunities, and emerging business operations associated with the startup (Keil et al. (2008)). Finally, due to this frequent interaction, the CVC parent firm should be more likely to integrate emerging businesses from the startup. By measuring new airlines 1 year after the deal year (see Figure 5), I find strong evidence supporting the idea that the introduction of a nonstop flight is positively associated with a CVC parent firm's identification of emerging business opportunities and subsequent integration measured by emerging phrases.

In the second exercise, I follow the approach of Giroud (2013) and use new airlines as a shock to the potential deal flow. I analyze an empirical choice model that examines which startups a given CVC firm chooses to invest in. Once again, I document that CVC investments, which are plausibly driven by new airlines, lead to future changes in firm scope for CVC parent firms.

These findings help address potential endogeneity concerns by considering that the presence of new airlines is likely independent of disruptive technologies or any other unobserved firm characteristics. This is achieved by controlling for fixed effects related to locations and years. One specific concern is the technology agglomeration effect, where novel technologies tend to concentrate in regions like Silicon Valley. However, this concern is mitigated by including location  $\times$  year fixed effects in the analysis.

To further understand the causality, I study the channel through which CVC helps identify new business opportunities and ultimately spurs firm scope change. My empirical evidence is consistent with the experimentation view of CVC investment. According to the experimentation view, CVC enables corporations to explore and experiment with various business opportunities, with managers facing uncertainty regarding the outcomes of these opportunities. Each CVC deal can be seen as an experiment that creates a real option for potential new lines of products or activities (Keil et al. (2008)). Through interaction with startup managers and active participation in their operations, CVC parent firms gain valuable firsthand information, which I refer to as a "signal" in this study, about the future potential of the relevant business. These signals are crucial in identifying promising and new business opportunities while avoiding potential pitfalls. The experimentation strategy is a logical response to identifying good business opportunities in the VC industry, where uncertainty is pervasive (Kerr, Nanda, and Rhodes-Kropf (2014), Ewens, Nanda, and Rhodes-Kropf (2018)).

To test this CVC experimentation view, I estimate two discrete choice models (McFadden (1973)) to examine the industry choices made by CVC parent firms in terms of establishing new segments or seeding emerging businesses. For each CVC parent firm, I utilize the ex post performance of the startups in which they have invested (including IPO, acquisition, bankruptcy, and patenting information) as a proxy for the private signals received by the CVC firm. These private

---

ultimately, reduced ticket prices (Borenstein and Rose (1994), Mazzeo (2003)). All these factors may encourage CVC managers to visit the startup at the destination of a flight more frequently.

signals are incorporated into the empirical strategy based on the eventual observed investment outcomes. The findings reveal that, conditional on CVC investments, receiving a positive signal from an invested startup is associated with a higher likelihood of the CVC parent company choosing to establish a new segment or incorporate emerging phrases in the startup's industry. Conversely, receiving a negative signal is associated with a lower likelihood of making these choices. Furthermore, the study demonstrates that CVC parents only react to the private signals obtained from their own CVC deals and do not react to public information, such as industry IPO waves. This highlights the importance of the private information obtained through CVC investments in driving the decision-making process of CVC parent firms.

Overall, the study is related to three broad strands of literature. First, this study is related to the canonical literature on firm scope, dating back to Teece (1980) and Panzar and Willig (1981) in economics and Lang and Stulz (1994), Berger and Ofek (1995), and Lamont (1997) in finance. Recently, Hoberg and Phillips (2021) have provided evidence that firms in the 21st century tend to expand their businesses across related industries, resulting in a lack of diversification discount. This study builds upon this literature by uncovering a novel mechanism through which U.S.-listed firms change their scope. Specifically, it highlights the role of CVC experimentation as a means for firms to explore and expand into new industries.

Second, the study contributes to the literature on entrepreneurial finance such as Manso (2016) and Ewens et al. (2018). Specifically, Ewens et al. (2018) find that recent VC firms adopt a new experimentation strategy in their investments, the alleged "spray and pray" strategy, especially after the cost of starting software and Internet-related ventures drops significantly. There is one key difference between their experimentation and my experimentation – the goal. While VC firms engage in experimentation to search for "unicorns," CVC firms aim to figure out optimal growth directions for company's future.

Third, the study contributes to the VC literature and, more precisely, the CVC literature pioneered by Gompers and Lerner (2000) and Hellmann (2002). Previous studies have documented that firms in more competitive industries (Fulghieri and Sevilir (2009), Kim, Gopal and Hoberg (2016)), in industries with higher technology uncertainty (Basu, Phelps, and Kotha (2011)) and low intellectual property protection (Dushnitsky and Lenox (2005a)), with lower institutional ownership (Tian and Ye (2018)), and firms experiencing deterioration of internal innovation (Ma (2020)) are more likely to conduct CVC investments. I complement the aforementioned studies by relating CVC to firm scope changes of its parent corporation. Another closely related CVC study is Shan (2019). He studies how the ex ante product and technology distances between startups and established firms influence decisions between acquisitions and CVC investments by established firms. In contrast, my study addresses the broad questions whether CVC investments lead to changes in firm scope, and whether there is evidence in support of the experimentation view of CVC strategies.<sup>9</sup>

<sup>9</sup>Although he also uses textual measure based on the 10-K and startup's businesses, his measure significantly differs from mine. He uses the textual measure to measure the ex ante difference of

The rest of the paper proceeds as follows: [Section II](#) introduces the background of CVC and develops the hypotheses. [Section III](#) describes the data and summary statistics. [Section IV](#) provides the basic facts between CVC and firm scope changes. [Section V](#) provides the identification between CVC and firm scope changes. [Section VI](#) studies CVC as an experimentation process. [Section VII](#) provides the additional analyses. [Section VIII](#) concludes the study.

## II. Background and Hypothesis Development

This section starts with the institutional background of CVC. A CVC deal is formally defined as minority equity investment by an established corporation in a privately held entrepreneurial company (Dushnitsky (2012)). Alternatively, one can interpret a CVC program as a VC arm affiliated with an established corporation (e.g., Google Venture affiliated with Google).

CVC departs from traditional VC firms in three key dimensions. First, while traditional VC firms typically raise capital from external limited partners, CVC programs primarily rely on funding from their corporate parents. Second, around two-thirds of CVC programs do not operate with a dedicated fund structure that has a fixed fund lifetime. Instead, they often have a more flexible approach similar to “discretionary” or “evergreen” funds. They invest when venture opportunities arrive (MacMillan, Roberts, Livada, and Wang (2008)).

The third and crucial aspect of CVC programs is that, in addition to seeking financial returns, they also aim to achieve strategic goals for their corporate parents. These include identifying new technologies, exploring new growth opportunities, and fostering innovation within existing business units (Siegel, Siegel, and MacMillan (1988), MacMillan et al. (2008)). The literature on CVC has widely recognized the significance of these strategic goals. Chesbrough (2002) argues that if CVC investments were solely driven by the pursuit of financial gains and were disconnected from the corporation’s strategies and operational capabilities, shareholders of the CVC parent company would be better off investing in traditional independent VC funds instead.

Strategic management scholars have conducted interview-based surveys with CVC managers from around the world to gain insights into their strategic goals (Dushnitsky (2012)). These surveys have highlighted two recurring strategic objectives of CVC programs. The first strategic objective is to provide a window for new technology, which aligns with the concept of open innovation. The second strategic objective, which is often overlooked, is that CVC investments can help identify new business opportunities for CVC corporate parents.

For instance, a recent survey conducted by the National Venture Capital Association (NVCA) examines 48 large CVC programs and finds that more than half of the CVC managers identified the critical strategic aim of their program as being the identification of new markets and business directions (MacMillan et al. (2008)). Other survey evidence supporting the CVC objective of finding new business opportunities includes Winters and Murfin (1988), Sykes (1990), McNally

---

businesses and technologies between startups and established firms, while I construct the emerging phrases to proxy emerging businesses in the economy.

(1997), and Ernst and Young (2009). This study's hypothesis is developed on this second objective.

Accordingly, I argue that CVC programs and CVC investments can help firms to identify new business opportunities. There are several ways in which a CVC program or investment can facilitate the discovery of emerging business prospects. First, during the search for potential deals, a CVC parent gains access to new business models and opportunities that may not be readily available within the boundaries of the firm (Chesbrough (2002)). Second, through active engagement with startup managers post-investment, a CVC parent firm typically gains valuable first-hand information and insights into emerging business opportunities and new markets (Keil et al. (2008)).<sup>10</sup>

After identifying an emerging business, it is common for a CVC parent firm to integrate the new business into its existing operations, leading to a change in the firm's scope. To measure these changes, I employ two distinct but complementary approaches. On the one hand, I use textual analysis to identify emerging businesses in the U.S. economy and further measure the business integration of those emerging businesses by CVC parents through SEC annual 10-K filings. On the other hand, I develop a text-based measure of firm scope, building upon the methodology proposed by Hoberg and Phillips (2021).

A natural follow-up question concerns how CVC helps identify new business opportunities. In this study, I argue that it is a learning-through-experimentation story. Generally, this story is in line with the "long-shot bets" feature of VC investments documented in the literature with few "unicorn" startups reaching big successes (Bergemann and Hege (2005)).

The experimentation hypothesis postulates that CVC investments allow a corporate to experiment with various business opportunities before making large-scale investment decisions, reflecting that corporate managers are uncertain about their final results. Therefore, each CVC deal can be regarded as an experiment that creates a real option for a potential new line of products or activities (Keil et al. (2008)). Through interacting with startup managers and participating in the startup's operation, a CVC parent firm can receive valuable information (referred to as a signal in this study) about the future potential of this new business. Crucially, the signal can contain both soft and hard information, that is not available without investments or interactions with CVC-backed startups.<sup>11</sup> A CVC parent finally pins down the best business option according to the various positive and negative signals received from multiple experiments.

In the strategic management literature, this view is supported by Keil et al. (2008), who conducted several interviews with several senior managers of CVC

<sup>10</sup>“Once the venture capital managers had made investments in interesting ventures and formed insights concerning potential capability voids within the incumbent, they frequently connected these startups with [CVC program] senior managers to provide senior managers with a window on new technologies and business models. Influencing decision-making in this way is frequently an informal process in which it is important to expose the right decision-makers to the right external contexts.” (Keil et al. (2008))

<sup>11</sup>Keil et al. (2008) argues that the knowledge regarding emerging business from CVC-backed startups are usually non-codified or colloquial information. Accessing the knowledge is possible only if the firm accesses to the community, the VC industry.

programs and argued that CVC is a process of “disembodied experimentation” in learning knowledge from CVC-backed startups. In Keil et al.’s (2008) interviews, a CVC manager recalls that

if the [venture] turns out to be something important, you have to put in your own machines (page 1485). Sometimes we just speak up and say: ‘That will never work. I have seen it! Guys, that’s complete nonsense, I have seen the total opposite [failure] here in a start-up.’ (page 1490).

This view is also a good application of Ewens et al.’s (2018) experimentation theory. Ewens et al. document that many VCs adopt a “spray and pray” strategy in response to the reduced cost of initiating businesses in the software and Internet-related industry. More specifically, VCs spray their deals on more ventures in the early investment stage and also abandon more when they receive bad signals from startups. Interestingly, the software and Internet-related industry has the most intensive CVC deal activity.

This view is also supported by Lerner (2012) who argues that CVC has the function of leveraging limited resources to pursue or test a variety of technology options. Its cost-saving function is crucial when an established firm needs to test a large number of technology options. CVC also helps quickly pull the plug of unpromising initiatives in the experiments, while the inside project will never stop optimally, as the internal R&D manager has a strong incentive to hide unfavorable signals (Seru (2014)).

### III. Data and Sample Selection

#### A. CVC Sample

The raw data of my CVC sample are extracted from the Thomson Reuters SDC VentureXpert Database. Following Chemmanur, Loutskina, and Tian (2014) and Ma (2020), I start with a list of 1,248 U.S. corporate-affiliated VC firms as reported by VentureXpert.<sup>12</sup> I then manually link these CVC program names with the historical names of CRSP and Compustat firms (provided by WRDS) by checking various sources from Google, Factiva, LexisNexis, and PitchBook. This step helps me to identify the corporate parent(s) of each CVC program. As VentureXpert sometimes mislabels some CVC programs as IVCs or other types, I conduct an extensive search among all VC types following Hellmann, Lindsey, and Puri (2008) and supplement the above beginning CVC list with an additional set of 35 CVC firms. Taken together, I obtain 623 unique CVC firms (programs) affiliated with either CRSP firms or Compustat firms from 1980 to 2017.

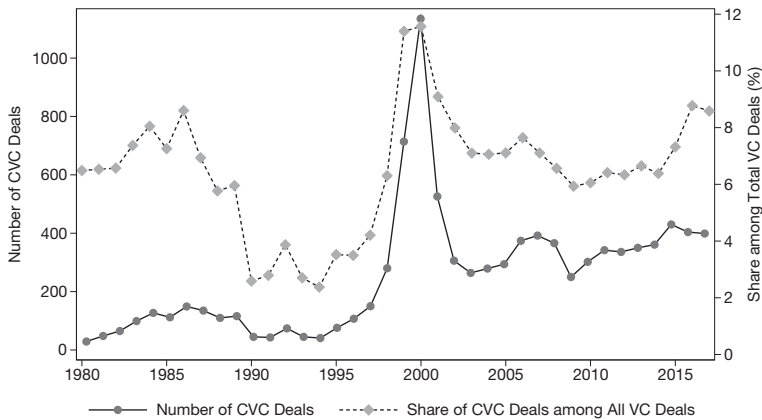
In the next step, I impose extra filters on these 623 CVC firms/programs by requiring the following: i) the corporate parent of CVC is incorporated in the United States and is not operating in financial industries (SIC code starting with 6), ii) only

<sup>12</sup>In detail, these VCs are classified as a Non-Financial Corporate Affiliate or Subsidiary Partnership, Venture/PE Subsidiary of Non-Financial Corporation, Venture/PE Subsidiary of Other Companies NEC, Venture/PE Subsidiary of Service Providers, Direct Investor/Non-Financial Corporation, Direct Investor/Service Provider, SBIC Affiliate with Non-Financial Corporation, and Non-Financial Corporate Affiliate or Subsidiary. In addition, I require that VC firms are located in the United States.



FIGURE 1  
Corporate Venture Capital Deals by Calendar Year

Figure 1 displays the annual corporate venture capital (CVC) investments initiated by U.S. public corporations (excluding financial firms) using data from the Compustat database. The left axis represents the number of CVC deals made each year, while the right axis represents the share of CVC deals as a percentage of all VC deals recorded in VentureXpert. The data used in this analysis primarily come from SDC VentureXpert, and the data set covers the period from 1980 to 2017.



a single corporate parent is matched to the CVC, and iii) the CVC program is not initiated by the financial division of a company, such as the GE Capital Equity Group or Exxon Pension Fund, as these CVCs most likely seek financial, rather than strategic, goals in their investments.

The final CVC sample contains 497 CVC programs launched by 448 unique public corporations, investing in CVC at least once in the sample period, with around 11,300 deals. Figure 1 plots the annual aggregate CVC investments by U.S. public (nonfinancial) corporations in the Compustat database. Investments are measured using i) the number of deals (left axis) and ii) the fraction of deals among all VC deals (right axis).

## B. Sample for Firm Scope Change

### 1. Textual Data on Emerging Business

To obtain a textual measure capturing time-varying emerging businesses in the U.S. economy, I combine two text sources from VentureXpert and the SEC digital filing system (EDGAR).

First, I download a detailed business description for each U.S.-based VC-backed startup from VentureXpert.<sup>13</sup> I group a set of startups' detailed business descriptions into a yearly single corpus, which contains all active VC-backed startups receiving VC funding in a given year. I drop common words and stop words and form short phrases (each containing two single words) for any two

<sup>13</sup>One caveat of this approach is that those startups' business descriptions are not historical but are updated to the date of data downloading.

adjacent words in each sentence.<sup>14</sup> Next, I define each year's "emerging phrases" set as those short phrases that are most widely used by the VC-backed startup community during that year. More precisely, I select the top 5% of the most frequently used short phrases from the yearly startup corpus. Approximately, 30 short phrases that represent businesses that are too general (e.g., "business service" or "product service") are manually excluded.

Ideally, each emerging phrase represents an emerging business that is popular among startup communities. I refer to these as "emerging phrases" under the implicit assumption that any popular business in the VC industry should be novel and emerging relative to any business of U.S.-listed firms. Figure 2 shows several examples of the emerging phrase sets in word clouds. As shown in the figure, the emerging phrases evolve significantly over time. For example, the emerging phrase set often relates to "Internet" and "e-commerce" during the 2000s (the Internet bubble period), while, in the most recent year of my sample, more "tech buzzwords" are included, such as "artificial intelligence," "virtual reality," "online platform," and "digital health."

I then search for these emerging phrases in business descriptions of the listed companies. I obtain the U.S. public-listed firms' business descriptions from the annual 10-K filings, following Hoberg and Phillips (2016). First, I download all 10-K filings from the SEC EDGAR system using Python automation scripts. Following Hoberg and Phillips (2016), I extract the Item 1 (Business Description) as my text source of firms' business in my regression sample, including those 450 U.S. firms with CVC investments and other firms without CVC deals.

In the main analysis, I search for emerging phrases in each 10-K filing in order to identify any emerging businesses that are newly integrated by CVC parents as well as by other public-listed firms with no CVC investments. The detailed procedure is illustrated in Figure 3.

## 2. Text-Based Segments Data

Using emerging phrases as a measure of firm scope has one potential limitation: Adding an emerging phrase in a company's 10-K business description does not guarantee significant changes in firm scope or business (i.e., the measure only captures granular business changes). Alternatively, one can measure firm scope using a more traditional method (i.e., counting the industrial segments operated by a firm in a certain year) by tracking Compustat business segment data. However, both Villalonga (2004) and Hoberg and Phillips (2021) document that the Compustat business segment data set fails to adequately reflect the actual business segments or industries in which a firm is engaged. To address this challenge, I follow Hoberg and Phillips (2021) to construct a text-based segment measure.

The text-based segment measure is based on the 8 VEIC industries defined by VentureXpert.<sup>15</sup> VEIC covers fewer industries than NAICS or SIC industry

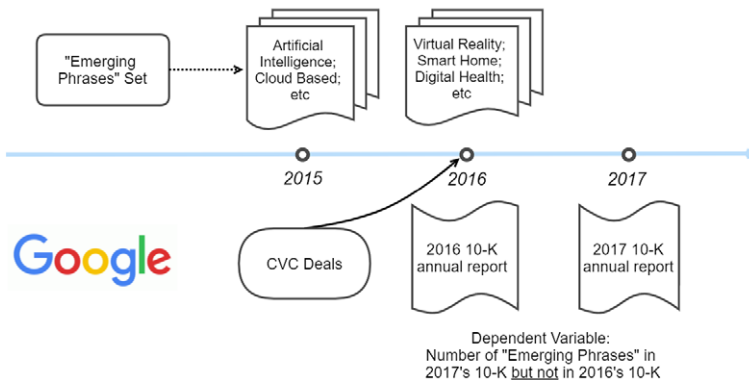
<sup>14</sup>The Supplementary Material lists those common words and stop words. Stop words are mainly from the NLTK.

<sup>15</sup>These industries include biotechnology, communications, computer hardware, computer software, internet specific, semiconductor, medical health, and non-high-tech.



FIGURE 3  
Regression Design in Table 2: An Example

Figure 3 provides an explanation of the regression design outlined in Table 2. To illustrate the process, let us consider the example of Google. Suppose Google's CVC program, Google Ventures, invested in startups in 2016. During that year, the set of emerging phrases for 2016 includes terms such as "virtual reality," "digital health," and "smart home." The next step is to search for these three emerging phrases in Google's 2017 10-K annual report, specifically in the Item 1 Business Description section. The dependent variable in the regression analysis is then defined as the count of the 2016 emerging phrases that are newly added to Google's 2017 10-K business description. The underlying intuition behind this regression design is that when Google invests in CVC in 2016, it helps the company identify new business opportunities, such as those related to digital health or smart home technologies. One year after the investment (in this case, 2017), Google is expected to be more likely to integrate these emerging areas into its own business operations.



bigram tokens (excluding common words and stop words) used by startups that operate in the corresponding VEIC industry.<sup>16</sup>

Next, for each firm  $i$  in year  $t$  and for each industry  $j$ , I calculate the following score:

$$(1) \quad \text{SCORE}_{i,t,\text{VEIC}-j} = \frac{\# \text{ of words in } (V_{i,t} \cap V_{\text{VEIC}-j})}{\# \text{ of words in } V_{\text{VEIC}-j}},$$

where  $V_{i,t}$  and  $V_{\text{VEIC}-j}$  denote the vocabulary in the 10-K business description of firm  $i$  in year  $t$  and the vocabulary of a specific VEIC industry  $j$ , respectively. As a result, the score evaluates how many fractions of the industry  $j$  vocabulary are used in each 10-K filing.

Following Hoberg and Phillips (2021), to compute how many VEIC industries a given firm might operate in, I identify a fixed threshold SCORE above which we deem a firm having  $\text{SCORE}_{i,t,\text{VEIC}-j} \geq \text{SCORE}$  to be operating in industry  $j$  in the given year  $t$  (Hoberg and Phillips (2021) call this an "operating pair"). Again, following Hoberg and Phillips (2021), I let SCORE be equal to the top 2% threshold of all  $\text{SCORE}_{i,t,\text{VEIC}-j}$  observations in the data set.

In the last step, I define a firm that creates a new VEIC-based segment if this firm starts to operate in the industry  $j$  in year  $t$  ( $\text{SCORE}_{i,t,\text{VEIC}-j} \geq \text{SCORE}$ ) and does not in the previous year ( $\text{SCORE}_{i,t-1,\text{VEIC}-j} < \text{SCORE}$ ).

<sup>16</sup>Following Hoberg and Phillips (2021), I lock in each industry's vocabulary and do not allow industry vocabulary to vary with time. This is in contrast to the time-varying property of emerging phrases.

### C. Other Data Sources and Summary Statistics

As mentioned in the Introduction, the study utilizes the introduction of new nonstop U.S. airlines as a means to show causality in estimating the impact of CVC investments on firm scope changes. To gather information on U.S. nonstop airline routes, the T-100 Domestic Segment Database from the United States Department of Transportation is utilized. Additionally, detailed information on U.S. Metropolitan Statistical Areas (MSA) is obtained from the U.S. Census Bureau. These MSA data provide essential geographical and demographic information that helps in understanding the locations and characteristics of the areas where the CVC firms and startups are based.

Table 1 provides the summary statistics of a firm–year sample that is created by merging the CVC sample, two firm scope measures, and the Compustat data set. In line with existing CVC literature, industries identified by their 3-digit Standard Industrial Classification (SIC) codes that do not have any CVC activities throughout the entire sample period are excluded from the analysis. To address extreme values and outliers, all variables (except for dummy variables) are winsorized at the 1st and 99th percentiles. The detailed construction of variables used in the analysis can be found in the Supplementary Material.

In Table 1, on average, a firm conducts six CVC deals in a single year conditional on having CVC investments. Furthermore, firms with CVC investments are more inclined to incorporate emerging phrases into their 10-K filings in the following 2 years. Additionally, they are more likely to establish new businesses, as indicated by text-based segment measures, within 3 years after making these investments. Lastly, companies with CVC investments tend to be larger, more profitable (as measured by ROA), and more likely to be conglomerates.

## IV. CVC and Change of Firm Scope

I begin by presenting evidence of the correlation between CVC and changes in firm scope. One of the main challenges in conducting these tests is the lack of a suitable metric for measuring changes in firm scope that involve the integration of emerging businesses. While it is reasonable to expect CVC firms to integrate emerging businesses following CVC investments, it is not straightforward to measure this empirically. To address this issue, I offer two approaches. The first approach, discussed in Section IV.A, utilizes text-based analysis of “emerging phrases” to measure changes in firm scope. The second approach, outlined in Section IV.B, involves the creation of text-based segments as a means of measuring firm scope.

### A. Evidence on Emerging Phrases

In this section, I estimate the following regression:

$$(2) \quad \text{EMERGING\_PHRASES}_{i,t+1} = \beta \text{D\_CVC}_{i,t} + \gamma \mathbf{X}_{i,t} + v_t \times l_j + (\tau_i) + \varepsilon_{i,t},$$

where  $\text{EMERGING\_PHRASES}_{i,t+1}$  denotes the number of year  $t$ 's “emerging phrases” (those top 5% word pairs most popular among startups (see Figure 2)) that are newly added into firm  $i$ 's 10-K Item 1 (business description) in year  $t + 1$ . (By mentioning “newly added,” I imply that the phrases appear in year  $t + 1$ 's

TABLE 1  
Summary Statistics

Table 1 provides summary statistics for a firm-year sample used in most of the regressions. The sample includes all observations recorded in Compustat from 1995 to 2017, with the exclusion of foreign firms (firms incorporated outside of the United States) and firms in financial industries (SIC industry codes starting with 6). Additionally, industries with no CVC activity throughout the entire sample period are completely excluded. To be included in the sample, each firm must have 10-K filings reported to the SEC in both the present and previous year. The sample is then divided into two parts based on whether a firm has engaged in any CVC deals during that year. The variable NUM\_CVC\_DEALS represents the number of CVC deals initiated by a firm in a given year. The variable NUM\_EMERGE\_PHRASES\_ADDED ( $t + k$ ) represents the number of new emerging phrases, as defined in Figure 2, that are newly added to the firm's 10-K filing in year  $t + k$ . The variable D\_NEW\_TEXT\_BASED\_SEGMENT ( $t + k$ ) is a dummy variable that equals 1 if the firm establishes new segments in year  $t + k$ . Segment establishments are identified using the text-based analysis method outlined in Hoberg and Phillips (2021), with segments defined at the VEIC industry level. All variables, except for the dummy variables, are winsorized at the 1% and 99% levels.

	D(CVC) = 1				D(CVC) = 0			
	Mean	Median	Std. Dev.	No. of Obs.	Mean	Median	Std. Dev.	No. of Obs.
NUM_CVC_DEALS	5.578	2	13.832	1,294	0.000	0	0.000	69,667
NUM_EMERGE_PHRASES_ADDED ( $t + 1$ )	2.221	1	3.253	1,136	0.699	0	1.598	58,439
NUM_EMERGE_PHRASES_ADDED ( $t + 2$ )	2.007	1	3.085	1,015	0.676	0	1.541	48,889
D_NEW_TEXT_BASED_SEGMENT ( $t + 1$ )	0.143	0	0.417	1,136	0.039	0	0.216	58,439
D_NEW_TEXT_BASED_SEGMENT ( $t + 2$ )	0.140	0	0.407	1,015	0.039	0	0.214	48,889
D_NEW_TEXT_BASED_SEGMENT ( $t + 3$ )	0.149	0	0.423	897	0.037	0	0.207	40,891
NUM_VEIC_SEGMENTS	0.842	0	1.146	1,294	0.219	0	0.543	69,667
FIRM_SIZE	9.210	9.361	2.058	1,289	4.749	4.758	2.541	65,510
TOBINS_Q	2.772	2.002	3.134	1,127	5.906	1.704	30.576	61,966
ROA	0.128	0.142	0.219	1,290	-0.274	0.083	1.897	66,887
BOOK_LEVERAGE	0.329	0.284	0.325	1,280	0.285	0.208	0.655	67,992
CASH	0.229	0.184	0.199	1,294	0.226	0.119	0.254	68,835
CAPITAL_EXPENDITURE	0.054	0.038	0.063	1,285	0.061	0.032	0.094	67,178
D_CONGLOMERATE	0.427	0	0.495	1,294	0.228	0	0.420	69,667

10-K but not in year  $t$ 's 10-K.) This measure aims to capture the integration of new businesses that are both new compared to the existing businesses of the CVC parent and new to the overall U.S. economy.<sup>17</sup> D\_CVC is a dummy equal to 1 if a firm conducts at least one CVC deal in year  $t$ . Firm-level controls ( $X$ ) include FIRM\_SIZE, TOBINS\_Q, ROA, R&D, LEVERAGE, CAPX, CASH, SALES\_GROWTH, HHI, and D\_CONGLOMERATE (lagged), as well as two mechanical textual measures: the number of any new short phrases appearing in the 10-K Item 1 and the total length of 10-K Item 1.

Before delving into the regression results, I illustrate the regression design in Figure 3. To illustrate the process, let us consider the example of Google. Suppose Google's CVC program, Google Ventures, invested in startups in 2016. During that year, the set of emerging phrases for 2016 includes terms such as "virtual reality," "digital health," and "smart home." The next step is to search for these three emerging phrases in Google's 2017 10-K annual report, specifically in the Item 1 Business Description section. The dependent variable in the regression analysis is then defined as the count of the 2016 emerging phrases that are newly added to Google's 2017 10-K business description. The underlying intuition behind this regression design is

<sup>17</sup>The popularity of businesses among VC-backed startups suggests that they are new and emerging. Venture capitalists typically finance businesses that have innovative characteristics such as new technologies, business models, industries, or products. These emerging characteristics make these businesses attractive investment opportunities for venture capitalists.

that when Google invests in CVC in 2016, it helps the company identify new business opportunities, such as those related to digital health or smart home technologies. One year after the investment (in this case, 2017), Google is expected to be more likely to integrate these emerging areas into its own business operations.

Table 2 corroborates that CVC investments are strongly correlated with adding new emerging phrases. As shown in column 1 (column 5), on average, a CVC parent will add 0.78 (0.68) more emerging phrases compared to its industry-year peers in the first year (second year) after investment.<sup>18</sup> This amount of increase translates into 100% of the sample average (0.75). When I switch to within-firm variations by adding firm fixed effects (FEs), the coefficients in columns 2 and 6 are comparable and positively significant. More specifically, the firm FE helps rule out the case that these results are driven by firms that constantly add emerging phrases and enjoy a high product fluidity (Hoberg, Phillips, and Prabhala (2014)).

Further, I present a continuous measure capturing the relatedness between the startup's business in a CVC deal and the CVC parent's own business in columns 3–4 and 7–8 of Table 2.<sup>19</sup> Theoretically, there are arguments both in favor of and against the idea that closely related CVC deals should help promote emerging business integration. On the one hand, investing in a tightly linked startup can indeed facilitate the entry of a CVC parent firm into the corresponding industry. This ease of entry can be attributed to the high asset redeployability of similar businesses or industries (Kim and Kung (2016)).

On the other hand, Keil et al. (2008) document that CVC firms seldom invest in ventures that are too close to their existing technologies or businesses. Instead, they focus on semi-distant domains that are neither too close, nor too far. Chesbrough (2002) argues that a CVC parent benefits in investing in (distant) strategic whitespace of its businesses, therefore offering an option-like strategic upside for future business shifts or expansions. In other words, a CVC parent company may not get much fresh knowledge about potential business opportunities from investing in a startup that is too closely tied. In fact, the coefficients of the interaction term are all insignificant, showing that deal relatedness does not matter in the process of emerging business integration.

Figure 4 examines the usage of new “emerging phrases” in the 10 years around each CVC deal. The point estimates (from OLS) and confidence intervals are taken from the following regression specification:

$$(3) \quad \text{EMERGING\_PHRASES}_{i,t} = \sum_{k=1}^{+5} \gamma_k \text{D\_CVC\_BEFORE}(k)_{i,t} + \sum_{k=1}^{+5} \alpha_k \text{D\_CVC\_POST}(k)_{i,t} + \beta \mathbf{X} + \tau_i + \nu_t + \varepsilon_{i,t},$$

<sup>18</sup>Industry by year fixed effects help rule out common industry shocks that drive changes in firm scope, as studied in Harford (2005) and Maksimovic and Phillips (2008).

<sup>19</sup>I thank the JFQA referee's suggestion to establish this continuous metric. This metric is calculated as the cosine similarity between the business vocabularies of the startup and CVC parent firm. The vocabulary of the startup and CVC parent company is represented by two vectors, each of which has an entry equal to 1 when the word appears in the business description. In text-based analysis, it is referred to as the “bag of words approach.” A similar method is also used in Shan (2019).

TABLE 2  
CVC Investments and Firm Scope Change: Adding Emerging Phrases

Table 2 presents regressions about CVC and changes in firm scope measured by emerging phrases. The regression sample consists of all Compustat firms incorporated in the United States, with 10-K filings of year  $t$  and  $t - 1$  searchable in SEC, and are not in financial industries. Industries (defined as 3-digit SIC) with no CVC activity during the whole sample period are excluded entirely. The dependent variable is the number of "emerging phrases" newly added in the next year (or in the second year)'s 10-K business description. The emerging phrases are those top 5% most frequently used word pairs (excluding stop words and common words) in the detailed business descriptions of all VC-backed startups receiving VC funding in a given year. Columns 1–4 (5–8) count those "emerging phrases" appearing in year  $t + 1$  ( $t + 2$ )'s 10-K Item 1 but not in the year  $t$  ( $t + 1$ ). The  $t$ -statistics are shown in parentheses, and standard errors are clustered by firm. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

OLS	No. of "Emerging Phrases" Newly Added in 10-K Item 1							
	Added in Year $t + 1$ 's Business				Added in Year $t + 2$ 's Business			
	(Not Already in Year $t$ 's Business)				(Not Already in Year $t + 1$ 's Business)			
	1	2	3	4	5	6	7	8
D_CVC	0.776*** (7.36)	0.655*** (5.65)	0.833*** (7.59)	0.688*** (5.72)	0.676*** (6.53)	0.433*** (3.81)	0.708*** (6.45)	0.428*** (3.62)
D_CVC × RELATED_DEAL (continuous measure)			5.343 (1.43)	5.665 (1.41)			-0.571 (-0.16)	0.00973 (0.00)
NUM_NEW_TOKENS (in $t + 1$ 10-K Item 1)	0.000770*** (33.06)	0.000829*** (29.65)	0.000769*** (33.06)	0.000828*** (29.65)				
NUM_NEW_TOKENS (in $t + 2$ 10-K Item 1)					0.000769*** (32.20)	0.000822*** (29.87)	0.000769*** (32.21)	0.000822*** (29.88)
Firm-level controls Year × industry FE	Yes		Yes		Yes		Yes	
Year FE		Yes		Yes		Yes		Yes
Firm FE		Yes		Yes		Yes		Yes
No. of obs.	53,074	52,060	53,074	52,060	48,958	48,087	48,958	48,087
Adj. $R^2$	0.379	0.410	0.379	0.410	0.363	0.400	0.363	0.400

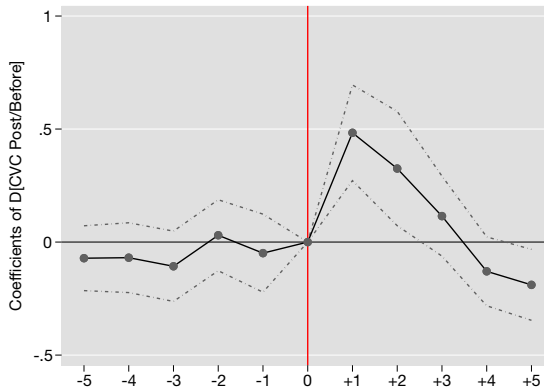


FIGURE 4  
Firm Scope Change (Measured by “Emerging Phrases”) Around CVC Deals

Figure 4 examines the emerging phrases usage in the years around each CVC deal. The estimates (OLS) and confidence intervals are taken from the following regression specification:

$$\text{EMERGING\_PHRASES}_{i,t} = \sum_{k=1}^{+5} \gamma_k \text{D\_CVC\_BEFORE}(k)_{i,t} + \sum_{k=1}^{+5} \alpha_k \text{D\_CVC\_POST}(k)_{i,t} + \beta \mathbf{X} + \tau_i + \nu_t + \varepsilon_{i,t},$$

where the dependent variable on the left-hand side represents the count of new emerging phrases that are added to a firm’s 10-K Item 1 section. These emerging phrases appear in the current year ( $t$ ) but not in the previous year ( $t - 1$ ). The emerging phrases, as depicted in Figure 2, are derived from the top 5% most popular short phrases (excluding common words and stop words) found in the business descriptions of all VC-backed startups in a given year. The variable set  $\{\text{D\_CVC\_BEFORE}(k)\}_{k=1}^{+5}$  consists of dummy variables that take a value of 1 if the year is  $k$  years before a CVC deal was invested by the firm. Similarly,  $\{\text{D\_CVC\_POST}(k)\}_{k=1}^{+5}$  represents dummy variables for the 5 years after a CVC deal was invested. These dummy variables capture the impact of CVC deals on firm scope changes over time. To account for firm-specific and year-specific effects, the regressions include firm and year fixed effects. Standard errors are clustered at the firm level to address potential correlation within firms. The confidence intervals are calculated at the 90% confidence level. The covariates represented by  $\mathbf{X}$  include various firm-level characteristics such as FIRM\_SIZE, TOBINS\_Q, ROA, R&D, LEVERAGE, CAPX, CASH, SALES\_GROWTH, HHI, FIRM\_AGE, and D\_CONGLOMERATE (lagged).



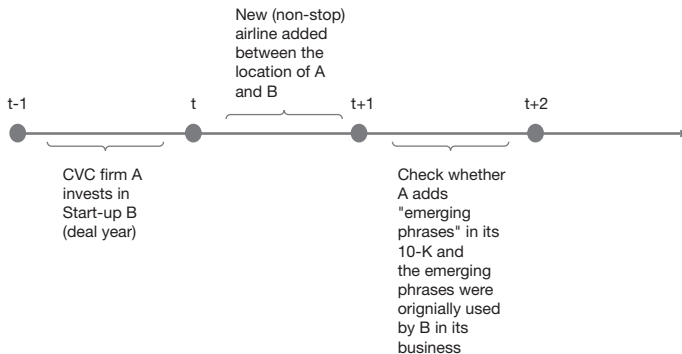
where  $\text{EMERGING\_PHRASES}_{i,t}$  simply counts the number of newly added year  $t - 1$ 's emerging phrases in year  $t$ 's 10-K.  $\{\text{D\_CVC\_BEFORE}(k)\}_{k=1}^{+5}$  and  $\{\text{D\_CVC\_POST}(k)\}_{k=1}^{+5}$  denote a set of 10 dummies in the  $[-5, +5]$  year window around each CVC deal. For example,  $\text{D\_CVC\_POST}(k = +3)$  equals 1 if a firm-year observation is the third year after a CVC deal.  $\tau_i$  and  $\nu_t$  are the firm and year fixed effects, respectively.

Figure 4 illustrates that the “treatment” effect of CVC deals is most prominent within the 2 years following the investment, with no significant effect observed before or 3 years after the investment. This finding aligns with expectations since older CVC investments, such as those made in 2010, may not be able to identify current emerging business opportunities in 2022. Similarly, future deals, such as those in 2030, may not have an impact on identifying contemporary business opportunities. In essence, CVC deals are most effective in identifying and capitalizing on contemporaneous business opportunities.

Next, do CVC parents incorporate new emerging phrases that directly correspond to the specific CVC deals? For example, CVC deals related to artificial intelligence should predict the adding of emerging phrases such as “artificial intelligence” into the 10-K. Table 3 answers this question. Specifically, the table

FIGURE 5  
Regression Setup in the First Airline Natural Experiment

Figure 5 presents a regression design related to the introduction of a new nonstop airline route, and the results are reported in Table 6. The analysis investigates the relationship between a CVC deal, the establishment of a new nonstop airline route, and the integration of emerging businesses. The process begins by observing a CVC deal initiated by firm A, which invests in startup B during year  $t$  (deal year). The analysis then examines whether there is a newly established nonstop airline route between the locations of firm A's headquarters and startup B. In year  $t+2$ , the analysis checks whether firm A adds new emerging phrases to its 10-K report and whether these newly added phrases originated from startup B's business description in year  $t$ . The aim is to assess whether CVC firm A integrates emerging businesses from startup B following the investment. The underlying intuition is as follows: if a new nonstop airline route is established, it increases the likelihood of firm A visiting startup B after the investment. This increased interaction and proximity may enable firm A to identify new business opportunities associated with the investment. Consequently, the expectation is that firm A will integrate emerging businesses from startup B, as reflected in the addition of new emerging phrases in its 10-K report.



categorizes each emerging phrase into one of the 8 VentureXpert VEIC industries based on the industry of the startups that use the emerging phrase to describe their businesses.<sup>20</sup> The main diagonal of the table indicates that CVC deals in industry  $j$  tend to predict the inclusion of industry  $j$  specific emerging phrases in a firm's 10-K annual report. This suggests that CVC deals within a particular industry are associated with the addition of emerging phrases that are relevant to that industry into the firm's 10-K.

One important consideration is how long emerging phrases added by a CVC parent are retained in subsequent annual 10-K reports. To address this concern, we examine Panels C and D of Figure A1 in the Supplementary Material. Panel D illustrates the distribution of years of survival for each emerging phrase after it is incorporated into the business of a CVC parent. On average, these phrases remain in the annual 10-K reports for approximately 2.5 years. Moreover, more than 75% of the phrases continue to be present in the subsequent annual 10-K reports.

## B. Evidence on Text-Based Segments

The analysis in the previous section focuses on capturing granular changes in firm scope through emerging phrase usage in 10-Ks. This section complements the previous analysis by considering more significant changes in firm scope, specifically the creation of text-based segments. The construction of text-based segments, as outlined in Section IV.B, follows Hoberg and Phillips (2021).

<sup>20</sup>This sorting is non-exclusive. Take an example: artificial intelligence is sorted into both computer software and internet-specific industries.

TABLE 3

## Industry-Specific CVC Investments and Emerging Phrases Adding

Table 3 presents the regressions analyzing the relationship between industry-specific corporate venture capital (CVC) investments and the usage of industry-specific emerging phrases by U.S. public firms. The dependent variable in the regressions is the number of industry-specific emerging phrases newly added to a firm's annual 10-K Item 1 in year  $t$ . Each emerging phrase is categorized into one of 8 VEIC industries, based on the industry of the startup that utilizes the phrase. The main independent variables are dummies representing industry-specific CVC investments made in the past 3 years. These CVC investments are also categorized into the 8 VEIC industries. The regression model includes several firm controls, such as FIRM\_SIZE, TOBINS\_Q, return on assets (ROA), research and development (R&D) expenditure, LEVERAGE, cash holdings (CASH), SALES\_GROWTH, capital expenditures (CAPX), Herfindahl-Hirschman index (HHI), dummy variable for conglomerate status (D\_CONGLOMERATE), and FIRM\_AGE. The  $t$ -statistics are reported in parentheses, and standard errors are clustered at the firm level. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

VEIC Industry	Number of Newly Added Emerging Phrases (VEIC-Industry-Specific Phrases)							
	Biotechnology	Communication	Computer Hardware	Computer Software	Internet Specific	Medical Health	Non-High-Tech	Semiconductor
	1	2	3	4	5	6	7	8
D[CVC deal – biotechnology] (past 3-year)	0.092** (2.390)	0.021 (0.474)	0.036 (1.573)	0.007 (0.096)	-0.075 (-0.734)	0.022 (0.745)	0.017 (0.852)	0.029 (0.748)
D[CVC deal – communication] (past 3-year)	-0.008 (-1.290)	0.135*** (2.748)	-0.008 (-0.426)	0.027 (0.373)	0.154* (1.863)	-0.008 (-0.985)	-0.000 (-0.019)	0.046* (1.761)
D[CVC deal – computer hardware] (past 3-year)	0.014 (1.307)	0.022 (0.445)	-0.004 (-0.148)	-0.040 (-0.459)	0.044 (0.395)	-0.003 (-0.207)	-0.005 (-0.258)	-0.031 (-0.948)
D[CVC deal – computer software] (past 3-year)	-0.010 (-1.086)	0.040 (1.087)	0.005 (0.360)	0.141** (2.378)	0.073 (0.985)	0.013 (1.279)	0.001 (0.083)	-0.002 (-0.122)
D[CVC deal – Internet specific] (past 3-year)	0.012** (2.076)	0.013 (0.406)	0.014 (1.230)	0.221*** (2.980)	0.366*** (3.811)	-0.003 (-0.563)	0.004 (0.277)	0.019 (1.152)
D[CVC deal – medical health] (past 3-year)	0.028 (1.062)	-0.034 (-0.683)	-0.029 (-1.384)	0.087 (1.324)	0.122 (1.532)	0.013 (0.594)	0.012 (0.508)	0.028 (1.027)
D[CVC deal – non-high tech] (past 3-year)	-0.016 (-1.497)	-0.001 (-0.040)	0.011 (0.729)	0.066 (1.033)	0.038 (0.571)	-0.017 (-1.388)	-0.039*** (-2.640)	0.005 (0.252)
D[CVC deal – semiconductor] (past 3-year)	-0.017** (-2.496)	0.001 (0.012)	0.006 (0.347)	-0.021 (-0.340)	-0.084 (-1.119)	-0.001 (-0.163)	0.004 (0.287)	0.084*** (2.885)
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	73,188	73,188	73,188	73,188	73,188	73,188	73,188	73,188
Adj. $R^2$	0.138	0.211	0.048	0.269	0.302	0.055	0.048	0.082

Table 4 reports the correlation between CVC investments and segment-level firm scope. In Panel A, the dependent variable measures the number of VEIC industries in which a firm operates. Panel B, on the other hand, uses a dummy variable that equals 1 if a firm creates a new segment, indicating its entry into a new VEIC industry.<sup>21</sup> To capture significant changes in firm scope, both measures are evaluated in year  $t + 2$  or  $t + 3$ , considering that substantial business changes require time.<sup>22</sup>

Panel A of Table 4 illustrates that CVC deals have a positive impact on firm scope in both year  $t + 2$  and  $t + 3$ . The effect of CVC deals on firm scope increases steadily from  $t + 2$  to  $t + 3$ . It is important to note that all regressions control for firm scope in year  $t$ , thereby capturing the impact on changes in firm scope. Moving on to Panel B, when examining the creation of text-based segments, a similar pattern emerges. CVC investments are found to predict a greater number of future segment creations in VEIC industries. Notably, the effect is significantly stronger in  $t + 3$ .

In conclusion, Table 4 suggests that CVC deals play a crucial role in helping CVC parent firms identify emerging business prospects. This, in turn, facilitates the expansion of their operations and allows them to enter new VEIC industries. These findings align with the general knowledge acquisition narrative put forth by Dushnitsky and Lenox (2005b) and Ma (2020). In their narratives, CVC parent firms primarily acquire knowledge about cutting-edge technology, as evidenced by the quantity and quality of their patents. On the other hand, in the context of our study, CVC firms acquire knowledge about new business opportunities, which subsequently leads to changes in their business scope.

In the Supplementary Material, I conduct an additional robustness check by running similar regressions using Compustat Industrial Segment data, where the data and variable constructions follow Seru (2014) and Matvos, Seru, and Silva (2018). Table A.1 in the Supplementary Material shows that CVC deals lead to greater chances of adding new SIC-3 segments. A minor distinction is that the result is driven exclusively by unrelated CVC deals.

Overall, the analysis indicates that the implementation of a CVC program is associated with changes in firm scope, as observed through both measures. However, it is important to acknowledge the potential presence of endogeneity issues that could influence the baseline results. These endogeneity concerns originate from two distinct sources. Firstly, firms that possess outdated business models or technologies may be more inclined to utilize CVC as a means to acquire new ideas. This, in turn, leads to changes in firm scope by introducing new businesses and discarding obsolete ones (Ma (2020)). Secondly, a specific new business model or technology might incentivize managers to engage in CVC investments for the purpose of learning and understanding it. This contrasts with the narrative that suggests CVC facilitates the identification of new business opportunities. In this

<sup>21</sup>This choice of using VEIC industries, as opposed to NAICS or SIC, allows the analysis to focus on changes in firm scope that are most relevant to CVC, while excluding those that are unrelated. For example, CVC investments and experiments may contribute to the establishment of a company's new artificial intelligence division, whereas a new segment in the agriculture sector would have little to do with CVC investments.

<sup>22</sup>It is worth noting that positive and significant coefficients are observed when firm scope is measured in year  $t + 2$  or  $t + 3$ , but not in year  $t + 1$ .

TABLE 4  
CVC Investments and Text-Based Firm Scope Changes

Table 4 studies CVC investments and subsequent changes in firm scope of those CVC parent firms. In Panel A, the dependent variable is Hoberg and Phillips (2021) text-based firm scope measure, counting how many VEIC industries a firm currently operates. The construction of the measure follows Hoberg and Phillips (2021) (detailed description in the text), with the only exception that I use the 8 VEIC industries to define segments and use the VEIC industry vocabularies in the text-based analysis. In Panel B, the dependent variable is a dummy of whether a firm establishes a new VEIC-based segment. Firm controls include FIRM\_SIZE, TOBINS\_Q, ROA, R&D, LEVERAGE, CASH, SALES\_GROWTH, CAPX, HHI, D\_CONGLOMERATE, and FIRM\_AGE. The *t*-statistics are shown in parentheses, and standard errors are clustered by firm level. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

*Panel A. Number of Text-Based Segments*

Measured in	Hoberg and Phillips Text-Based Firm Scope							
	Year <i>t</i> + 2				Year <i>t</i> + 3			
	1	2	3	4	5	6	7	8
D(CVC)	0.141*** (4.01)	0.0660 (1.60)	0.148*** (4.05)	0.0734* (1.80)	0.195*** (4.03)	0.148*** (3.14)	0.201*** (3.98)	0.147*** (3.10)
D(CVC) × RELATED_DEAL_MEASURE (continuous measure)			2.079* (1.82)	1.400 (1.20)			0.881 (0.56)	0.389 (0.30)
TEXT_BASED_FIRM_SCOPE (year <i>t</i> )	0.743*** (83.02)	0.265*** (15.94)	0.743*** (83.00)	0.265*** (15.94)	0.687*** (59.89)	0.134*** (7.25)	0.687*** (59.89)	0.134*** (7.25)
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × industry FE	Yes		Yes		Yes		Yes	
Year FE		Yes		Yes		Yes		Yes
Firm FE		Yes		Yes		Yes		Yes
No. of obs.	39,555	38,753	39,555	38,753	33,397	32,734	33,397	32,734
Adj. <i>R</i> <sup>2</sup>	0.666	0.766	0.666	0.766	0.607	0.764	0.607	0.764

*Panel B. Dummy of Adding New Text Segments*

Measured in	D[Add New Text Segments]							
	Year <i>t</i> + 2				Year <i>t</i> + 3			
	1	2	3	4	5	6	7	8
D(CVC)	0.0710*** (4.61)	0.0243 (1.43)	0.0718*** (4.43)	0.0179 (1.02)	0.0833*** (4.71)	0.0556*** (2.95)	0.0898*** (4.90)	0.0608*** (3.10)
D(CVC) × RELATED_DEAL_MEASURE (continuous measure)			-1.114** (-1.98)	-1.225** (-2.21)			-0.158 (-0.22)	-0.0843 (-0.12)
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × industry FE	Yes		Yes		Yes		Yes	
Year FE		Yes		Yes		Yes		Yes
Firm FE		Yes		Yes		Yes		Yes
No. of obs.	39,555	38,753	39,555	38,753	33,397	32,734	33,397	32,734
Adj. <i>R</i> <sup>2</sup>	0.012	0.093	0.012	0.093	0.007	0.080	0.007	0.080

scenario, a CVC parent company may decide to initiate a new business venture within a certain industry before making crucial CVC investments. Regrettably, the existing CVC literature does not extensively address these complex endogeneity concerns. In the subsequent section, I will outline my approach to addressing these endogeneity issues.

## V. Exogenous Variation on CVC and Firm Scope Change

This section establishes a plausible causal relationship between CVC and subsequent changes in firm scope by examining the introduction of new

U.S. airlines that operate nonstop flights. I employ two different identification methods to analyze this relationship. In the first method, I adopt the approach used by Bernstein et al. (2016) and introduce a shock to the potential information flow from startups to CVC firms following their deals. In the second method, I follow the approach outlined by Giroud (2013) and directly introduce a shock to the CVC deal flow by using airlines as a proxy.

### A. New Airlines and Knowledge Acquisitions

The first identification strategy closely follows the methodology employed by Bernstein et al. (2016). In their study, they utilize the introduction of new (nonstop) U.S. airline routes as a proxy for the monitoring and involvement effects of venture capitalists (VCs) on the future success of portfolio companies. The key assumption underlying their empirical strategy is that VC partners are more likely to visit a portfolio startup frequently following the introduction of a new (nonstop) airline route between the startup and the VC headquarters locations. To test this hypothesis, Bernstein et al. (2016) conducted a survey involving 306 senior VC partners. The survey results indicated that 83% of the respondents planned to increase their visiting frequency after the launch of a new nonstop airway to the startup.

By employing a similar rationale, the establishment of a new airline can also influence the inclination of a manager from the parent firm of a CVC to visit a startup within its portfolio that is situated at the destination served by the new airline. Consequently, the CVC parent firm becomes more likely to acquire knowledge regarding new business models, market opportunities, and emerging business prospects associated with the startup. Ultimately, this heightened visit increases the likelihood of the CVC parent firm integrating emerging businesses that are pertinent to the startup's operations.

In an interview with several CVC senior managers, Keil et al. (2008) summarize that visiting and interacting with startup managers is essential in the process of accessing knowledge about new business opportunities or models, as the knowledge is often in the form of non-articulated narratives or behavioral nature.<sup>23</sup> Additionally, being an active part of the startup community is crucial for accessing soft knowledge within the VC industry (MacMillan et al. (2008), Keil, Maula, and Wilson (2010)).

The regression strategy illustrated in Figure 5 focuses on CVC deals exclusively within the United States. The analysis begins by considering a CVC deal initiated by firm A, which invests in startup B in a given year (referred to as the deal year, denoted as  $t$ ). Subsequently, in the following year ( $t + 1$ ), the study examines whether a new nonstop airline route is established between the locations of firm A's headquarters and startup B. Finally, in year  $t + 2$ , the analysis investigates whether firm A includes new emerging phrases in its 10-K report and determines if these phrases originate from startup B's business description in year  $t$ . The underlying

<sup>23</sup>Keil et al. (2008) describe an example: “[Our President] is currently sitting in [locations]. He is meeting many of these companies, which is giving him a [very realistic] view of these companies. It is very difficult for a senior manager to get this first-hand information ... this kind of [tacit] knowledge transfer is only possible by working intensely together. By only sitting together once in a while you don't get that ....”

rationale is as follows: if a new nonstop airline route is established, it is more likely that the CVC firm will visit startup B after the investment. Consequently, the CVC firm has a greater chance of identifying new business opportunities associated with the investment. Ultimately, this analysis aims to observe a higher likelihood of firm A integrating emerging businesses derived from startup B.

One major difference between the identification strategy in Bernstein et al. (2016) and my exercise is the way in which we define and exploit the introduction of new (nonstop) airlines. While Bernstein et al. (2016) consider the first-time introduction of a nonstop flight between two locations, irrespective of the airline companies involved, I focus on the introduction of new (nonstop) airlines at the carrier level. This means that if Delta Airlines introduces a new nonstop flight between Dallas and Boston, I consider it as a new airline, even if American Airlines had already been operating a nonstop flight between the same destinations prior to Delta's entry. In contrast, Bernstein et al. (2016) only consider the first airline company to establish a nonstop route, such as American Airlines in this example. This broader approach to identification allows me to capture the effects of adding extra airways, which are correlated with better on-time performance, greater service quality, more flexible time schedules, and ultimately reduced ticket prices (Borenstein and Rose (1994), Mazzeo (2003)). All these factors may encourage CVC managers to visit the startup at the destination of a flight more frequently. Additionally, my sample begins in 1995 with data from SEC 10-K filings, as opposed to Bernstein's study which primarily focuses on first-time nonstop airlines in the 1970s and 1980s.

Table 5 tabulates the summary statistics of my CVC deal sample, where I split the sample into treatment and control groups. The treatment group comprises 421 deals in which there is at least one new (nonstop) airline established in year  $t + 1$  (1 year after the deal year) between the locations of the CVC firm's headquarters and the startup. In the treatment group, the average and median numbers of new airlines added are 1.5 and 1, respectively. This addition is significant and influential to airline competition, given that the median number of existing airlines (carriers) operating between the two locations is 5. The annual median enplaned passenger number is about 18,148 for those newly added airlines, equivalent to 28% of the annual enplaned passengers aggregated by existing airline companies. Furthermore, the treatment group contains deals with longer distances between the CVC firm and startup as well as fewer existing airline companies already offering nonstop flights.

Panel A of Table 6 shows the main results of the empirical exercise. All regressions are at the CVC deal level. The dependent variable is a dummy equal to 1 if a CVC parent adds emerging phrases to the 10 K from its portfolio startup in that deal, as shown in Figure 5. The dependent variable is measured in year  $t + 2$ , immediately after the possible new airline introduction. The design of this timing also follows the results in Figure 4 showing that a CVC firm is most likely to add emerging phrases within 2 years of CVC investment. Besides, I control for the number of existing airline companies offering nonstop airlines, the startup's age, and the number of co-investors in a given deal, and so forth. Finally, for robustness checks, I add various fixed effects, with the most high-dimensional one having locations-by-year fixed effects for both the startup and the CVC firm in a deal.

TABLE 5  
Summary Statistics in the Natural Experiment of Airline

Table 5 presents basic summary statistics for a natural experiment involving the introduction of new nonstop airlines. The sample is based on CVC deal-level data between 1995 and 2017, including deals where location information for both the CVC firm and the investing startup's headquarters is available. The sample is divided into a treatment group and a control group. The treatment group consists of deals where a new nonstop airline is introduced between the locations of the startup and the CVC firm in the year following the deal (refer to Figure 5 for timing details). New nonstop airlines are identified at the airline company level. The remaining deals form the control group. No. of existing nonstop airlines: The number of current existing airline companies offering nonstop airways between the startup and CVC firm locations. No. of passengers before: The count of enplaned passengers in existing airlines. No. of newly added nonstop airlines: The number of new nonstop airlines offered by airline companies between the two locations in year  $t + 1$ .

	Treated Group New (Nonstop) Airline Introduced				Control Group No New (Nonstop) Airline Introduced				Test of Mean Difference
	Mean	Median	Std. Dev.	No. of Obs.	Mean	Median	Std. Dev.	No. of Obs.	p-Value
<i>Deal-Level Sample</i>									
No. of passengers before	244,906	64,023	407,804.8	421	268,427	5,280	544,378.2	5,959	0.000
No. of existing nonstop airlines	5.670	5	4.479	421	3.461	2	5.612	5,959	0.000
No. of newly added nonstop airlines	1.486	1	0.929	421					
No. of passengers in new airlines	27,979.46	18,148	33,708.99	421					
Distance (miles)	1,167.201	855.568	1,013.190	421	975.510	597.829	928.038	5,959	0.001
Startup in CA?	0.460	0	0.499	421	0.526	1	0.499	5,959	0.000
Startup in NY?	0.097	0	0.297	421	0.074	0	0.262	5,959	0.001
Startup in MA?	0.121	0	0.327	421	0.101	0	0.302	5,959	0.017
CVC firm in CA?	0.452	0	0.498	421	0.570	1	0.495	5,959	0.000
CVC firm in NY?	0.094	0	0.293	421	0.061	0	0.239	5,959	0.000
CVC firm in MA?	0.043	0	0.203	421	0.052	0	0.221	5,959	0.150
Startup age (deal year)	4.386	4	3.169	409	4.224	3	3.124	5,237	0.078
No. of co-investors	5.357	5	3.392	409	5.620	5	3.528	5,959	0.006
<i>CVC Parent Firm Characteristics</i>									
ln(SALE)	9.439	10.186	1.959	402	9.394	10.176	1.984	5,932	0.415
ROA	0.159	0.180	0.243	401	0.161	0.172	0.221	5,903	0.843
R&D	0.088	0.090	0.234	409	0.087	0.089	0.151	5,958	0.912
CASH	0.235	0.216	0.177	407	0.240	0.215	0.183	5,947	0.357
CAPX	0.065	0.052	0.061	405	0.070	0.055	0.069	5,890	0.143

In Panel A of Table 6, the coefficients of I(Intro. New Airline) [ $t + 1$ ] are positive and significant across all specifications with different fixed effect combinations, consistent with the intuition that the introduction of a nonstop flight is positively associated with a CVC parent firm's identification of emerging business opportunities and its subsequent emerging business integration. Next, regarding to Panel B, I interact the treatment dummy with a distance dummy (median split). The results show that deals with longer distances mainly drive the effects in Panel A, again consistent with Bernstein et al.'s (2016) findings: VC managers who face longer distance barriers are more likely to react to the introduction of new nonstop flights.

## B. New Airlines and CVC Deal Flow

In the second exercise, I follow Giroud's (2013) study on investment of subsidiaries and local plants. Specifically, I use the U.S. new airlines to shock the activities of CVC investments and then examine whether these CVC deals plausibly



TABLE 6  
Introduction of New (Nonstop) Airlines and Emerging Phrases Adding

Table 6 presents analyses of the introductions of new nonstop airlines and changes in firm scope for CVC parent firms. The sample is based on CVC deal-level data between 1995 and 2017, including deals where location information for both the CVC firm and the invested startup's headquarters is available. Panel A focuses on the dependent variable, which is the number of emerging phrases newly added to the CVC firm's 10-K in year  $t + 2$  (2 years after the CVC deal year). These emerging phrases must have been previously used by the startup in the deal year ( $t$ ). The variable "I(INTRO\_NEW\_AIRLINE) [ $t + 1$ ]" equals 1 if a new nonstop airline is added between the locations of the startup and CVC firm in the deal. New airlines are identified at the airline company level. "NUM\_EXISTING\_AIRLINE [ $t$ ] (deal year)" and "NUM\_PASSENGERS [ $t$ ] (deal year)" are measured in the deal year ( $t$ ). Refer to Figure 5 for a detailed illustration of the timing. In Panel B, "I(INTRO\_NEW\_AIRLINE) [ $t + 1$ ]" is further interacted with a distance measure between the locations of the CVC firm and startup. Panels C and D provide similar results using text-based segment measures as dependent variables. The  $t$ -statistics are shown in parentheses, and standard errors are double clustered by firm and startup levels. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Emerging Phrases

	Dependent Variable: Dummy of Adding Emerging Phrases in 10-K				
	Measured in $t + 2$				
	1	2	3	4	5
I(INTRO_NEW_AIRLINE) [ $t + 1$ ]	0.00883** (2.57)	0.00903** (2.58)	0.00989*** (2.65)	0.0101*** (2.66)	0.00931** (2.30)
NUM_EXISTING_AIRLINE [ $t$ ] (deal year)	-0.000307 (-0.45)	-0.000237 (-0.03)	-0.000283 (-0.41)	0.0000776 (0.01)	-0.00117* (-1.66)
NUM_PASSENGERS [ $t$ ] (deal year)	0.0104 (1.30)	0.00982 (1.19)	0.00980 (1.16)	0.00914 (1.05)	0.00840* (1.77)
STARTUPS_AGE	-0.000997 (-0.50)	-0.00165 (-0.81)	-0.00108 (-0.52)	-0.00174 (-0.83)	-0.000966* (-1.82)
I(SEED_OR_EARLY_STAGE)	-0.00315 (-0.47)	-0.00347 (-0.50)	-0.00331 (-0.46)	-0.00366 (-0.49)	-0.000455 (-0.09)
NUM_CO-INVESTORS	0.000786 (0.99)	0.000812 (0.98)	0.000807 (0.98)	0.000831 (0.97)	0.000795 (1.36)
ROUND_AMOUNT (1,000 USD)	-1.28e-08 (-0.61)	-1.07e-08 (-0.55)	-1.27e-08 (-0.60)	-1.07e-08 (-0.53)	-2.09e-08 (-1.34)
CVC patent firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	
Startup FE	Yes	Yes	Yes	Yes	
CVC location FE		Yes		Yes	
Startup location FE		Yes		Yes	
Startup industry FE			Yes	Yes	Yes
CVC location $\times$ year FE					Yes
Startup location $\times$ year FE					Yes
No. of obs.	4,263	4,263	4,091	4,091	5,182

Panel B. Emerging Phrases (Interact with the Distance)

	Dependent Variable: Dummy of Adding Emerging Phrases in 10-K				
	Measured in $t + 2$				
	1	2	3	4	5
I(INTRO_NEW_AIRLINE) [ $t + 1$ ] $\times$ LONG_DISTANCE	0.0151* (1.70)	0.0154* (1.72)	0.0159* (1.76)	0.0172** (1.98)	0.0159** (2.22)
I(INTRO_NEW_AIRLINE) [ $t + 1$ ] $\times$ SHORT_DISTANCE	0.00381 (1.07)	0.00384 (1.06)	0.00438 (1.08)	0.00441 (1.06)	0.00533 (1.20)
CVC patent firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	
Startup FE	Yes	Yes	Yes	Yes	
CVC location FE		Yes		Yes	
Startup location FE		Yes		Yes	
Startup industry FE			Yes	Yes	Yes
CVC location $\times$ year FE					Yes
Startup location $\times$ year FE					Yes
No. of obs.	4,263	4,263	4,091	4,091	5,182

(continued on next page)

TABLE 6 (continued)  
Introduction of New (Nonstop) Airlines and Emerging Phrases Adding

Panel C. Create New Text-Based Segments

	Dependent Variable: Dummy of Creating New Segments				
	Measured $int + 2$				
	1	2	3	4	5
I(INTRO_NEW_AIRLINE) [ $t + 1$ ]	0.0226** (2.16)	0.0227** (2.14)	0.0272** (2.25)	0.0273** (2.23)	0.0241** (2.03)
CVC patent firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	
Startup FE	Yes	Yes	Yes	Yes	
CVC location FE		Yes		Yes	
Startup location FE		Yes		Yes	
Startup industry FE			Yes	Yes	Yes
CVC location $\times$ year FE					Yes
Startup location $\times$ year FE					Yes
No. of obs.	4,263	4,263	4,091	4,091	5,182

Panel D. New Text-Based Segments (Interact with the Distance)

	Dependent Variable: Dummy of Creating New Segments				
	Measured $int + 2$				
	1	2	3	4	5
I(INTRO_NEW_AIRLINE) [ $t + 1$ ]	0.0235*	0.0240*	0.0268*	0.0273*	0.0226**
$\times$ LONG_DISTANCE	(1.77)	(1.79)	(1.95)	(1.96)	(2.37)
I(INTRO_NEW_AIRLINE) [ $t + 1$ ]	0.0214**	0.0213*	0.0265**	0.0263*	0.0271*
$\times$ SHORT_DISTANCE	(1.99)	(1.95)	(1.98)	(1.94)	(1.68)
CVC patent firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	
Startup FE	Yes	Yes	Yes	Yes	
CVC location FE		Yes		Yes	
Startup location FE		Yes		Yes	
Startup industry FE			Yes	Yes	Yes
CVC location $\times$ year FE					Yes
Startup location $\times$ year FE					Yes
No. of obs.	4,263	4,263	4,091	4,091	5,182

driven by new airlines lead to future changes in firm scope, such as adding new emerging phrases or creating new VEIC segments.

The intuition is that following a new (nonstop) airline, a CVC firm will be more willing to invest in a startup located at the new airline's destination since the new airline provides more convenient ways to learn relevant business opportunities after investment. To implement this idea, I develop and estimate a simple choice model about which startups a given CVC firm chooses to invest in a given year. The set of alternative startups includes all VC-backed startups that have received funding from any CVC programs in my sample and are currently active in receiving new funding. Table 7 provides the estimate of this model, where the sample is at the CVC firm  $\times$  year  $\times$  startup level. In columns 1 and 2, the dependent variable is a dummy equal to 1 if the CVC firm chooses a given startup to invest in the given year. The coefficients show that the probability of choosing the startup to invest in increases following a new (nonstop) airline in the previous year. In columns 3–6, I focus on the second stage in the identification, that is, whether these CVC deals plausibly driven by new airlines lead to future changes in business scope, such as adding new emerging phrases or creating new VEIC segments. Since the results in columns 1 and 2 are not strong enough to implement a formal 2SLS estimate, I use a reduced

TABLE 7  
Introduction of New Airlines, CVC Deal Flow, and Emerging Phrases Adding

Table 7 estimates an empirical choice model about which startup a given CVC parent firm will choose to invest and the CVC parent firm's subsequent changes in firm scope. The sample is at the CVC firm  $\times$  year  $\times$  startup level. For each CVC firm  $i$  in year  $t$ , the firm chooses which startups to invest. The set of alternative startups includes all VC-backed startups actively receiving funding in the given year. The dependent variable in columns 1 and 2 is a dummy equal to 1 if the CVC firm chooses the given startup to invest in year  $t + 1$ . The dependent variable in columns 3 and 4 is a dummy equal to 1 if the CVC firm adds emerging phrases into its 10-K in year  $t + 2$ , and the emerging phrases are from the given startup's business descriptions.  $I(\text{INTRO\_NEW\_AIRLINE}) [i]$  is equal to 1 if there is a new nonstop airline added between the location of the startup and CVC firm in the choice model. New airlines are identified at the airline company (carrier) level. Airline controls follow Table 6. Firm-level controls follow Table 4. The  $t$ -statistics are shown in parentheses, and standard errors are double clustered by firm and startup levels. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	D(CVC)		EMERGING_PHRASES		NEW_VEIC_SEGMENTS	
	Measured in					
	$t + 1$		$t + 2$		$t + 2$	
	1	2	3	4	5	6
$I(\text{INTRO\_NEW\_AIRLINE}) [i]$	0.0121** (2.23)	0.0107** (2.08)	0.00436** (2.03)	0.00407* (1.77)	0.00828* (1.68)	0.00637 (1.27)
Other airline controls	Yes	Yes	Yes	Yes	Yes	Yes
CVC firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes
CVC firm FE	Yes	Yes	Yes	Yes	Yes	Yes
CVC location $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
Startup location $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
Startup industry FE		Yes		Yes		Yes
No. of obs.	1,379,515	1,376,254	1,280,656	1,280,656	1,281,344	1,281,344
Adj. $R^2$	0.232	0.233	0.485	0.500	0.222	0.234

form model and regress changes in firm scope directly using the new airlines. In columns 3–6, I corroborate that new airlines positively predict that the CVC firm will add emerging phrases from the invested startups or create a new segment in the same industry as the startup.

## VI. CVC Signal and Firm Scope Change

In this section, I delve into the mechanism through which CVC plays a crucial role in identifying new business opportunities and catalyzing changes in firm scope. I propose that CVC operates as a learning-through-experimentation process. Here, a CVC parent firm initially engages in a “spraying” strategy, akin to the approach outlined in Ewens et al. (2018). This involves pursuing deals across various technology or business options. Subsequently, the firm patiently awaits signals that reveal the potential of these options. Finally, armed with these signals, the CVC parent firm responds by making informed decisions on where to launch new business activities.

### A. CVC Signal and Emerging Phrase

This section aims to test the experimentation hypothesis by examining the presence of emerging phrases in 10-K filings. The underlying idea is that a CVC parent firm does not establish a new business in every industry it explores through CVC deals. Instead, only industries that show positive post-investment signals, as received from startups, are considered. This feedback loop, where information from CVC investments informs decision-making, is a crucial aspect of the

TABLE 8  
CVC Investment Signal and Emerging Phrases Usage

Table 8 studies signals received by CVC parent firms and subsequent changes in firm scope as measured by emerging phrase usage. I estimate a McFadden (1973)-style discrete choice model regarding which VEIC industry a firm chooses to expand its businesses. The regression sample is at the firm  $\times$  year  $\times$  VEIC industry level. Each observation represents a choice for a firm to expand its businesses. The expansion of businesses is proxied by industry-specific emerging phrases added into 10-K. The dependent variable is a dummy equal to 1 if a firm adds at least one emerging phrase belonging to the VEIC industry  $j$  in year  $t$ .  $D\_CVC\_VEIC\_J$  is a dummy of whether this firm has conducted any CVC investments in the VEIC industry  $j$  in the past 3 years. The construction of signal variables is illustrated in the text. The  $t$ -statistics are shown in parentheses, and standard errors are clustered by firm level. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	Dummy of Adding Emerging Phrases in VEIC $j$					
	1	2	3	4	5	6
$D\_CVC\_VEIC\_J$	0.237*** (5.279)	0.247*** (5.851)	0.247*** (5.804)	0.241*** (5.691)	0.247*** (5.866)	0.251*** (5.858)
$D\_CVC\_VEIC\_J \times IPO\_SIGNAL$	0.066*** (2.649)					
$D\_CVC\_VEIC\_J \times ACQUISITION\_SIGNAL$		0.000 (0.011)				
$D\_CVC\_VEIC\_J \times HIGH\_VALUE\_SIGNAL$			0.151** (2.269)			
$D\_CVC\_VEIC\_J \times ACQUIRED\_BY\_CVC\_FIRM$				0.201*** (3.040)		
$D\_CVC\_VEIC\_J \times BANKRUPTCY\_SIGNAL$					-0.084 (-1.106)	
$D\_CVC\_VEIC\_J \times PATENT\_SIGNAL$						0.564*** (3.514)
$D\_CVC\_VEIC\_J \times NUM(STARTUPS\_INVESTED)$	0.018*** (2.380)	0.024*** (2.930)	0.017*** (4.120)	0.023*** (4.054)	0.025*** (4.305)	0.025*** (4.409)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
VEIC industry $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	616,544	616,544	616,544	616,544	616,544	616,544
Adj. $R^2$	0.177	0.177	0.177	0.177	0.177	0.177

experimentation process. In other words, if changes in firm scope are driven by experimentation, the data should reveal that firms react to both positive and negative information updates (signals) they receive from their CVC investments when making decisions to launch new activities and divisions.

Table 8 presents the development and estimation of a McFadden discrete choice model (McFadden (1973)), which examines industry choices for emerging business integration. The measure of emerging business integration used in this analysis is the inclusion and utilization of emerging phrases in 10-K filings. By employing this model, we aim to gain insights into the CVC experimentation hypothesis.

The unit of observation in this model is at the firm–year–industry level. Each observation represents an alternative (industry  $j$ ) in which a firm  $i$  can choose to create an emerging business in a particular year  $t$ . The choice set consists of 8 VEIC industries, namely biotechnology, communications, computer hardware, computer software, Internet-specific, semiconductor, medical health, and non-high-tech. The dependent variable is a binary variable. It takes a value of 1 if firm  $i$  in year  $t$  adds at least one emerging phrase in its 10-K filing, and this emerging phrase belongs to industry  $j$ . The main control variable in the model is denoted as  $D\_CVC\_VEIC\_J$ . It is equal to 1 if a firm has conducted CVC deals in industry  $j$  within the past 3 years.

To test the CVC experimentation hypothesis, I interact `D_CVC_VEIC_J` with each CVC signal variable iteratively. However, it is challenging to measure the signal variables since researchers are never able to observe the real information transmission from startups to CVC firms. A classic type of soft information that only insiders can access is how well a possible new business will fit with the parent's existing business.

Consequently, in [Table 8](#), I use the startup's ex post measurable performances as proxies of signals. Importantly, no signal variable is observable to a CVC parent before the investment. In column 1, the signal is measured by the number of startups that are invested by a CVC firm and finally exit through IPO. In other words, the signal variable is based on information from a firm's past 3-year CVC investments in industry  $j$ .<sup>24</sup>

To provide an example, let us consider Google Venture's investments in industry  $j$  in the past 3 years leading up to year  $t$  (Figure A.2 in the Supplementary Material provides this example). Suppose Google Venture has invested in a total of five startups during this period. Among these five startups, three of them have successfully gone public (IPO). In this case, the signal variable, which represents the number of startups that have exited through IPO, would be equal to three. It is important to note that the number of deals, which in this case is five, will also be included as a control variable in the regression analysis.

In this case, it is assumed that Google cannot directly observe the value of the signal (equal to 3) at the time of the decision for emerging business integration. This is because the IPO dates of the three startups will naturally occur after year  $t$ , which is the decision-making year of the model. However, Google has access to important information about the potential of its investments in these five startups through various means, such as sitting on boards, participating in operational management, and communicating with syndicated venture capitalists. This assumption is supported by Bergemann and Hege (2005), MacMillan et al. (2008), and Dushnitsky (2012).

Crucially, the private signal obtained by Google is proxied in the empirical strategy by the eventual IPO outcome. In [Table 8](#), specifically in column 1, the results are consistent with the hypothesis that firms are more likely to create a new business in an industry when they receive a positive signal from their past 3-year investments in that industry. The coefficient of the interaction term indicates that there is a 56% increase in the unconditional probability of establishing new businesses for a 1-standard-deviation increase in the IPO signal.

Moreover, I control the number of startups invested in the past 3 years in industry  $j$  and denote it as `NUM_STARTUPS_INVESTED`. It is essential because the more you invest, the more IPO startups you will have.

In addition to the IPO signal, I also construct other signal variables based on startup acquisition, bankruptcy, and patent information. These signal variables are constructed using a similar method as the IPO signal.

<sup>24</sup>The 3-year window is also consistent with the window used in the `D_CVC_VEIC_J` dummy. This is also why I interact `D_CVC_VEIC_J` dummy with signal variables.

In column 2 of [Table 8](#), the acquisition signal is measured by the number of portfolio startups in industry  $j$  that eventually exit through acquisitions, mostly acquired by third parties. It is noted that the acquisition signal does not influence a CVC firm's decisions on firm scope. This finding is consistent with the view presented by Kerr et al. (2014), which suggests that most startup acquisitions, with a missing transaction value, are fire sales. In other words, startups eventually fail and sell themselves under the pressure of VC firms.

In contrast, in column 3 of [Table 8](#), a positive and significant coefficient is found when the CVC signal is measured by the number of acquisitions with a high acquisition value (top-tercile and non-missing) in industry  $j$ . Moving to column 4, it is observed that if a CVC parent firm acquires a startup from its portfolio of investing companies in industry  $j$ , it strongly predicts that the CVC firm will ultimately establish a new business in the relevant industry. However, it is noted that such occurrences are rare, as shown in [Table 10](#). Finally, in column 5, the bankruptcy signal is measured by the number of bankruptcy startups in the portfolio. It is found to be negatively associated with emerging business integration. However, the coefficient is insignificant.

## B. CVC Signal and New Segment Creation

[Table 9](#) presents and estimates a similar discrete choice model for firms' decisions to create new VEIC-based segments. The empirical model design follows [Table 8](#), with the only difference being that the dependent variable is a dummy variable equal to 1 if a firm establishes a new segment in the VEIC industry  $j$  in year  $t$ . The construction of segments is detailed in [Section VI.B](#), following the approach outlined by Hoberg and Phillips (2021).

The results in [Table 9](#) indicate that CVC firms react similarly to a series of signals when considering significant and segment-level changes in firm scope. Specifically, the acquisition signal positively predicts the decision to establish a new segment. On the other hand, the bankruptcy signal negatively correlates with the decision to establish a new segment, implying that CVC firms are less likely to create new segments when they have a higher number of bankruptcy startups in their portfolio.

A few more comments and tests are necessary.

*1. Placebo Test with Public VC Signals:* In order to test the experimental hypothesis, it is important to use private signals that are perceived by firms with CVC deals. To conduct a placebo test in the Supplementary Material, I have created public VC signals based on IPO, acquisition, bankruptcy, and patent information. However, the outcomes of these tests do not align with the results shown in [Tables 8](#) and [9](#). This highlights the importance of private information in the decision-making process, which is consistent with the existing CVC literature on information acquisition (Dushnitsky and Lenox (2005b), Basu et al. (2011), and Ma (2020)). The use of private signals by CVC firms gives them an advantage in identifying and capitalizing on new business opportunities.

*2. Alternative Explanation:* It is acknowledged that this section only tests one possible mechanism and does not rule out other underlying mechanisms that may explain how CVC spurs the discovery of new business opportunities and

TABLE 9  
CVC Investment Signal and Establishing New Segment

Table 9 studies signals received by CVC parent firms and subsequent changes in firm scope as measured by establishing new segments. I estimate a McFadden (1973)-style discrete choice model regarding which VEIC industry a firm chooses to expand its businesses. The regression sample is at the firm  $\times$  year  $\times$  VEIC industry level. Each observation represents a choice for a firm to expand its businesses. The expansion of businesses is proxied by establishing a new segment which is identified by the text-based analysis following Hoberg and Phillips (2021). The dependent variable is a dummy equal to 1 if a firm newly establishes a segment in VEIC industry  $j$  in year  $t$ .  $D\_CVC\_VEIC\_J$  is a dummy of whether this firm has conducted any CVC investments in the VEIC industry  $j$  in the past 3 years. The construction of signal variables is illustrated in the text. The  $t$ -statistics are shown in parentheses. All coefficients are multiplied by 100 for readability. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: Dummy of Adding Text-Based New Segment in VEIC $j$					
	1	2	3	4	5	6
D_CVC_VEIC_J	1.916*** (15.989)	1.997*** (16.761)	1.996*** (16.750)	1.943*** (16.254)	2.000*** (16.786)	1.999*** (16.742)
D_CVC_VEIC_J $\times$ IPO_SIGNAL	0.550*** (6.339)					
D_CVC_VEIC_J $\times$ ACQUISITION_SIGNAL		0.173*** (2.684)				
D_CVC_VEIC_J $\times$ HIGH_VALUE_SIGNAL			1.456*** (6.675)			
D_CVC_VEIC_J $\times$ ACQUIRED_BY_CVC_FIRM				1.883*** (5.699)		
D_CVC_VEIC_J $\times$ BANKRUPTCY_SIGNAL					-1.452*** (-2.653)	
D_CVC_VEIC_J $\times$ PATENT_SIGNAL						0.142 (0.138)
D_CVC_VEIC_J $\times$ NUM(STARTUPS_INVESTED)	-0.075*** (-4.979)	-0.097*** (-3.186)	-0.086*** (-5.441)	-0.034*** (-2.668)	-0.013 (-1.001)	-0.022* (-1.742)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
VEIC industry $\times$ year FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	686,488	686,488	686,488	686,488	686,488	686,488
Adj. $R^2$	0.014	0.014	0.014	0.014	0.014	0.014

TABLE 10  
Post CVC Firm Scope Change and Acquisition

Table 10 provides summary statistics about post-CVC changes in firm scope and acquisitions. All events are at the VEIC industry level. Panel A studies changes in firm scope that are measured by emerging phrase usage. In Panel A, adding emerging business is identified as an event that a CVC parent firm adds at least one industry  $j$  related emerging phrase in its 10-K and, within the past 3 years, the firm has conducted at least one CVC deal in the industry  $j$ . Column 1 counts the total number of events; column 2 counts how many events are not accompanied with acquisitions of VC-backed startups. In other words, in the data, the CVC parent firm does not acquire any VC-backed startups that operate in industry  $j$  in the past 3 years or in year  $t$ . Columns 3 and 4 count the events involved with acquisitions of startups. Panel C considers acquisitions of a third-party startup, while column 4 counts acquisitions of startups that are from the CVC firm's portfolio. Panel B studies the events of establishing new VEIC segments. Each event is identified if a CVC parent firm establishes a new segment in industry  $j$  and the CVC firm has conducted CVC deals within the past 3 years in industry  $j$ .

Total Events	Events Without Acquisition Involved	Events With Acquiring Third-Party Startup	Events With Acquiring CVC Program's Portfolio Companies (Startups)
1	2	3	4
<i>Panel A. Adding Emerging Businesses</i>			
1,176	794	313	69
100%	67.52%	26.62%	5.87%
<i>Panel B. Establishing New Text-Based Segments</i>			
146	97	40	9
100%	66.44%	27.40%	6.16%

subsequent changes in firm scope. One alternative explanation is network effects. CVC firms, by accessing VC communities and building networks with independent VCs, can gather first-hand information on emerging business opportunities. This idea is tested and supported in the Supplementary Material, where an alternative instrument is developed using each CVC program's network with independent VC firms and the Granular IV approach (Gabaix and Koijen (2020)). Although this approach may not be truly exogenous, it provides a new method for constructing instruments in the VC literature.

3. *First-Mover Advantage*: One may question why CVC firms cannot passively learn from successful startups through public signals and subsequently change their firm scope. It is argued that CVC firms benefit from being the “first mover” and can prevent startups from significantly disrupting their business (Kim et al. (2016)). By actively investing in and integrating with startups, CVC firms can gain insights and control over emerging technologies and markets, giving them a strategic advantage over competitors.

## VII. Additional Analyses

### A. Post-CVC Firm Scope Change and Acquisition

In Section VI, an important question that remains unanswered is how a CVC parent firm implements changes in its firm scope, specifically whether it does so through internal development or external acquisitions. This section aims to address this question by examining post-CVC changes in firm scope and acquisitions of startups.

To analyze this, the same sample used in Section VI is utilized, where the unit of observation is at the level of firm  $\times$  year  $\times$  VEIC industry. Changes in firm scope are measured either by the usage of emerging phrases or by text-based segment creation. For the former, the usage of emerging phrases is aggregated at the VEIC industry level, similar to the approach in Table 3.

It is worth noting that in the CVC literature, Benson and Ziedonis (2010) investigate a sample of acquisitions of VC-backed startups by CVC corporate parents and find interesting results. They observe that when a CVC parent acquires a startup from its portfolio companies, the event-study daily returns are generally negative. On the other hand, acquiring a third-party startup tends to result in a positive return.

Table 10 presents the total number of events (firm scope changes) that occur within 3 years following a CVC deal. To focus on firm scope changes that are presumably driven by CVC investments, it is further required that the industry of the CVC deal and the industry with changes in firm scope are the same VEIC industry.

Panel A of Table 10 reports the emerging business integration measured by emerging phrases, while Panel B focuses on new VEIC segments established by CVC firms. There are 1,176 events related to new emerging phrase additions and 146 events related to new segment creation following CVC deals. The frequency of segment creation is lower, as it captures more significant changes in firm scope.



To understand how a CVC parent firm executes changes in firm scope, three different options are considered in columns 2–4 of Table 10. In column 2, the number of events without any acquisitions of VC-backed startups is counted. Approximately, 66% of post-CVC changes in firm scope are not associated with acquisitions, indicating that they are executed through internal development. Approximately, 26% of events are accompanied by acquisitions of third-party startups, which are startups in which the focal CVC parent firm does not invest. In contrast, only about 6% of the events are associated with acquisitions of startups from the focal CVC firm's investing portfolio. This suggests that post-CVC firm scope changes are not primarily driven by killer acquisitions (Cunningham et al. (2021)).

This finding aligns with the existing CVC literature, which highlights the cautious relationship between a CVC parent and a CVC-backed startup (Dushnitsky and Shaver (2009)). Startups are concerned that their breakthrough technology or new business model may be appropriated by the CVC parent company sponsoring them. Additionally, many co-investors, particularly independent VC firms, prefer that the investing startup is not acquired by a CVC investor, especially through a killer acquisition. Moreover, a CVC parent corporation must consider its long-term reputation before making a killer acquisition (Chesbrough (2002)).

## B. Post-CVC Firm Scope Change and Value Creation

The existing literature highlights that CVC creates value for its corporate parents (Dushnitsky and Lenox (2006), Ma (2020)). However, the specific channel through which this value creation occurs remains an empirical question. In the last analysis, a connection is established between post-CVC changes in firm scope and post-CVC value creation, and the regression results are presented in Tables 11 and 12.

To measure the value creation, TOBINS\_Q is used, following the approach of Dushnitsky and Lenox (2006) and Ma (2020). The dependent variable is TOBINS\_Q measured in year  $t + 1$ ,  $t + 2$ , or  $t + 3$ , while CVC activity is measured in year  $t$ . Firm scope changes following CVC investments are measured between year  $t$  and  $t + k$  (the year in which TOBINS\_Q is measured), ensuring that TOBINS\_Q is measured after or in the year of the changes in firm scope. Table 11 measures firm scope changes using emerging phrases, while Table 12 analyzes it through text-based segment creation.

In Table 11, the interaction term,  $D(CVC) \times \text{NUM\_EMERGING\_PHRASES\_ADDED}$ , is found to be positively significant. This indicates that CVC investments, when accompanied by emerging business integration, create value for the CVC corporate parent compared to CVC firms without any changes in firm scope. Interestingly, the coefficients of the standalone variable  $D(CVC)$  are not significant, suggesting that changes in business scope are a critical factor in explaining how CVC investments add value. The insignificant and negative coefficients of the standalone variable  $\text{NUM\_EMERGING\_PHRASES\_ADDED}$  suggest that adding emerging phrases alone, without CVC investments, slightly destroys value. However, this effect is economically weaker than the interaction term. This finding is consistent

TABLE 11  
CVC Investments, Emerging Business Adding and Value Creation

Table 11 presents regressions about CVC investments, emerging phrases additions, and value creation of CVC corporate parents. All regressions are at the firm  $\times$  year sample. The dependent variable is TOBINS\_Q measured in year  $t + 1$ ,  $t + 2$ , or year  $t + 3$ . D(CVC) ( $t$ ) is a dummy equal to 1 if a firm conducts CVC investments in year  $t$ . In all regressions, firm FE and industry  $\times$  year FE are added. Firm controls include FIRM\_SIZE, TOBINS\_Q, ROA, R&D, LEVERAGE, CASH, SALES\_GROWTH, CAPX, HHI, D\_CONGLOMERATE, and FIRM\_AGE. The  $t$ -statistics are shown in parentheses, and standard errors are clustered by firm level. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: TOBINS_Q				
	Measured in				
	$t+1$	$t+2$	$t+3$	$t+2$	$t+3$
	1	2	3	4	5
D(CVC) ( $t$ )	0.221 (0.845)	-0.074 (-0.377)	-0.207 (-1.044)	0.117 (0.565)	-0.147 (-0.739)
NUM_EMERGING_PHRASES_ADDED in $t+1$	-0.224 (-1.458)	-0.236 (-1.466)	-0.067 (-0.381)		
D(CVC) ( $t$ ) $\times$ NUM_EMERGING_PHRASES_ADDED in $t+1$	0.747*** (2.997)	0.604*** (2.823)	0.363* (1.776)		
NUM_EMERGING_PHRASES_ADDED in $t+2$				-0.216 (-1.310)	-0.081 (-0.571)
D(CVC) ( $t$ ) $\times$ NUM_EMERGING_PHRASES_ADDED in $t+2$				0.360* (1.883)	0.341** (1.980)
Firm controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
No. of obs.	47,624	39,991	33,608	39,991	33,608
Adj. $R^2$	0.377	0.344	0.352	0.344	0.352

TABLE 12  
CVC Investments, New Text-Based Segments, and Value Creation

Table 12 presents regressions about CVC investments, new segment establishments, and value creation of CVC corporate parents. All regressions are at the firm  $\times$  year sample. The dependent variable is TOBINS\_Q measured in year  $t + 1$ ,  $t + 2$ , or year  $t + 3$ . D(CVC) ( $t$ ) is a dummy equal to 1 if a firm conducts CVC investments in year  $t$ . In all regressions, Firm FE and industry  $\times$  year FE are added. Firm controls include FIRM\_SIZE, TOBINS\_Q, ROA, R&D, LEVERAGE, CASH, SALES\_GROWTH, CAPX, HHI, D\_CONGLOMERATE, and FIRM\_AGE. The  $t$ -statistics are shown in parentheses, and standard errors are clustered by firm level. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: TOBINS_Q				
	Measured in				
	$t+1$	$t+2$	$t+3$	$t+2$	$t+3$
	1	2	3	4	5
D(CVC) ( $t$ )	-0.334 (-1.417)	-0.160 (-0.704)	-0.275 (-1.159)	-0.183 (-0.783)	-0.185 (-0.785)
D(NEW_TEXT_SEGMENT) in $t+1$	-0.144 (-0.585)	-0.090 (-0.361)	-0.253 (-1.013)		
D(CVC) ( $t$ ) $\times$ D(NEW_TEXT_SEGMENT) in $t+1$	0.836** (2.056)	0.524 (0.978)	0.777* (1.694)		
D(NEW_TEXT_SEGMENT) in $t+2$				-0.406 (-1.374)	-0.225 (-1.021)
D(CVC) ( $t$ ) $\times$ D(NEW_TEXT_SEGMENT) in $t+2$				0.802* (1.874)	0.910** (2.261)
Firm controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry $\times$ year FE	Yes	Yes	Yes	Yes	Yes
No. of obs.	53,706	45,067	37,744	45,067	37,744
Adj. $R^2$	0.360	0.331	0.339	0.331	0.339

with the conglomerate literature, which argues that the organization of multiple segments or businesses can destroy value (Lang and Stulz (1994)). A very similar result is documented in Table 12.

## VIII. Conclusion

Understanding the scope and boundaries of firms is a crucial topic in economics and finance. However, there is limited empirical research on the determinants of firm scope and the relationship between corporate innovation strategies and firm scope dynamics. This study aims to fill this gap by examining the impact of CVC investments on firm scope changes.

CVC programs, which are VC arms affiliated with established firms, have become popular in corporate innovation strategies. They provide opportunities for firms to identify new business models and opportunities that may not be available within their existing boundaries. Usually, through interactions with startup managers, CVC parent firms gain valuable information about emerging business opportunities and new markets. This study documents that CVC investments frequently lead to the integration of new businesses into the parent firm's current business domain, thereby reshaping firm scope.

The study introduces two measures to assess changes in firm scope. The first measure uses textual analysis to identify emerging businesses in annual business descriptions. The second measure identifies industries where a firm establishes new segments. The findings show a strong correlation between CVC investments and subsequent changes in firm scope. Firms with CVC investments add more emerging phrases to their business descriptions and are more likely to create new segments in previously unexplored industries.

To establish causality, the study utilizes the introduction of new airlines connecting the locations of CVC firms and startups involved in CVC deals. The increased interaction between CVC parent firms and startups due to new airlines provides opportunities for gaining insights into new business models and market opportunities. The study finds that the introduction of new airlines is positively associated with the identification of emerging business opportunities and subsequent integration by the CVC parent firm.

The study addresses potential endogeneity concerns by controlling for fixed effects related to locations and years. It also investigates the channel through which CVC investments facilitate the identification of new business opportunities. The findings support the experimentation view of CVC investment, where each CVC deal is seen as an experiment creating real options for potential new lines of products or activities. Through interactions with startups, CVC parent firms gain valuable firsthand information, or signals, about the future potential of relevant businesses.

In conclusion, this study provides empirical evidence of the impact of CVC investments on firm scope changes. It highlights the role of CVC in identifying emerging business opportunities and expanding firm scope. The findings contribute to the understanding of the relationship between corporate innovation strategies and firm dynamics.

## Supplementary Material

To view supplementary material for this article, please visit <http://doi.org/10.1017/S0022109023001370>.

## References

- Basu, S.; C. Phelps; and S. Kotha. "Towards Understanding Who Makes Corporate Venture Capital Investments and Why." *Journal of Business Venturing*, 26 (2011), 153–171.
- Bena, J., and K. Li. "Corporate Innovations and Mergers and Acquisitions." *Journal of Finance*, 69 (2014), 1923–1960.
- Benson, D., and R. H. Ziedonis. "Corporate Venture Capital and the Returns to Acquiring Portfolio Companies." *Journal of Financial Economics*, 98 (2010), 478–499.
- Bergemann, D., and U. Hege. "The Financing of Innovation: Learning and Stopping." *RAND Journal of Economics*, 36 (2005), 719–752.
- Berger, P. G., and E. Ofek. "Diversification's Effect on Firm Value." *Journal of Financial Economics*, 37 (1995), 39–65.
- Bernstein, S.; X. Giroud; and R. R. Townsend. "The Impact of Venture Capital Monitoring." *Journal of Finance*, 71 (2016), 1591–1622.
- Borenstein, S., and N. L. Rose. "Competition and Price Dispersion in the U.S. Airline Industry." *Journal of Political Economy*, 102 (1994), 653–683.
- Chemmanur, T. J.; E. Loutskina; and X. Tian. "Corporate Venture Capital, Value Creation, and Innovation." *Review of Financial Studies*, 27 (2014), 2434–2473.
- Chesbrough, H. W. "Making Sense of Corporate Venture Capital." *Harvard Business Review*, 80 (2002), 90–99.
- Coase, R. H. "The Nature of the Firm." In *The Nature of the Firm. Origins, Evolution, and Development*, O. E. Williamson and S. G. Winter, eds. New York, NY: Oxford University Press (1937), 18–33.
- Cunningham, C.; F. Ederer; and S. Ma. "Killer Acquisitions." *Journal of Political Economy*, 129 (2021), 649–702.
- Dushnitsky, G. "Corporate Venture Capital in the Twenty-First Century: An Integral Part of Firms' Innovation Toolkit." In *The Oxford Handbook of Venture Capital*, D. Cumming, ed. New York, NY: Oxford University Press (2012), 156–210.
- Dushnitsky, G., and M. J. Lenox. "When Do Firms Undertake R&D by Investing in New Ventures?" *Strategic Management Journal*, 26 (2005a), 947–965.
- Dushnitsky, G., and M. J. Lenox. "When Do Incumbents Learn from Entrepreneurial Ventures? Corporate Venture Capital and Investing Firm Innovation Rates." *Research Policy*, 34 (2005b), 615–639.
- Dushnitsky, G., and M. J. Lenox. "When Does Corporate Venture Capital Investment Create Firm Value?" *Journal of Business Venturing*, 21 (2006), 753–772.
- Dushnitsky, G., and J. M. Shaver. "Limitations to Interorganizational Knowledge Acquisition: The Paradox of Corporate Venture Capital." *Strategic Management Journal*, 30 (2009), 1045–1064.
- Ernst and Young. "Global Corporate Venture Capital Report." (2009).
- Ewens, M.; R. Nanda; and M. Rhodes-Kropf. "Cost of Experimentation and the Evolution of Venture Capital." *Journal of Financial Economics*, 128 (2018), 422–442.
- Frésard, L.; G. Hoberg; and G. M. Phillips. "Innovation Activities and Integration Through Vertical Acquisitions." *Review of Financial Studies*, 33 (2020), 2937–2976.
- Fulghieri, P., and M. Sevilir. "Organization and Financing of Innovation, and the Choice Between Corporate and Independent Venture Capital." *Journal of Financial and Quantitative Analysis*, 44 (2009), 1291–1321.
- Gabaix, X., and R. S. Koijen. "Granular Instrumental Variables." NBER Working Paper No. 28204 (2020).
- Giroud, X. "Proximity and Investment: Evidence from Plant-Level Data." *Quarterly Journal of Economics*, 128 (2013), 861–915.
- Gompers, P., and J. Lerner. "Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuation." *Journal of Financial Economics*, 55 (2000), 281–325.
- Grossman, S. J., and O. D. Hart. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94 (1986), 691–719.
- Harford, J. "What Drives Merger Waves?" *Journal of Financial Economics*, 77 (2005), 529–560.
- Hart, O., and J. Moore. "Property Rights and the Nature of the Firm." *Journal of Political Economy*, 98 (1990), 1119–1158.

- Hellmann, T. "A Theory of Strategic Venture Investing." *Journal of Financial Economics*, 64 (2002), 285–314.
- Hellmann, T.; L. Lindsey; and M. Puri. "Building Relationships Early: Banks in Venture Capital." *Review of Financial Studies*, 21 (2008), 513–541.
- Hoberg, G., and G. Phillips. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy*, 124 (2016), 1423–1465.
- Hoberg, G., and G. Phillips. "Conglomerate Industry Choice and Product Language." *Management Science*, 64 (2018), 3735–3755.
- Hoberg, G.; G. Phillips; and N. Prabhala. "Product Market Threats, Payouts, and Financial Flexibility." *Journal of Finance*, 69 (2014), 293–324.
- Hoberg, G., and G. M. Phillips. "Scope, Scale and Competition: The 21st Century Firm." Working Paper, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3746660](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3746660), (2021).
- Keil, T.; E. Autio; and G. George. "Corporate Venture Capital, Disembodied Experimentation and Capability Development." *Journal of Management Studies*, 45 (2008), 1475–1505.
- Keil, T.; M. V. Maula; and C. Wilson. "Unique Resources of Corporate Venture Capitalists as a Key to Entry into Rigid Venture Capital Syndication Networks." *Entrepreneurship Theory and Practice*, 34 (2010), 83–103.
- Kerr, W. R.; R. Nanda; and M. Rhodes-Kropf. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives*, 28 (2014), 25–48.
- Kim, H., and H. Kung. "The Asset Redeployability Channel: How Uncertainty Affects Corporate Investment." *Review of Financial Studies*, 30 (2016), 245–280.
- Kim, K.; A. Gopal; and G. Hoberg. "Does Product Market Competition Drive CVC Investment? Evidence from the U.S. IT Industry." *Information Systems Research*, 27 (2016), 259–281.
- Lamont, O. "Cash Flow and Investment: Evidence from Internal Capital Markets." *Journal of Finance*, 52 (1997), 83–109.
- Lang, L. H., and R. M. Stulz. "Tobin's Q, Corporate Diversification, and Firm Performance." *Journal of Political Economy*, 102 (1994), 1248–1280.
- Lerner, J. *The Architecture of Innovation: The Economics of Creative Organizations*. Boston, MA: Harvard Business Press (2012).
- Lerner, J., and R. Nanda. "Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn." *Journal of Economic Perspectives*, 34 (2020), 237–261.
- Ma, S. "The Life Cycle of Corporate Venture Capital." *Review of Financial Studies*, 33 (2020), 358–394.
- MacMillan, I.; E. Roberts; V. Livada; and A. Wang. *Corporate Venture Capital (CVC) Seeking Innovation and Strategic Growth*. National Institute of Standards and Technology (2008).
- Maksimovic, V., and G. Phillips. "The Industry Life Cycle, Acquisitions and Investment: Does Firm Organization Matter?" *Journal of Finance*, 63 (2008), 673–708.
- Manso, G. "Experimentation and the Returns to Entrepreneurship." *Review of Financial Studies*, 29 (2016), 2319–2340.
- Matvos, G.; A. Seru; and R. C. Silva. "Financial Market Frictions and Diversification." *Journal of Financial Economics*, 127 (2018), 21–50.
- Mazzeo, M. J. "Competition and Service Quality in the U.S. Airline Industry." *Review of Industrial Organization*, 22 (2003), 275–296.
- McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, P. Zarembka, ed. Cambridge, MA: Academic Press (1973), 105–142.
- McNally, K. *Corporate Venture Capital: Bridging the Equity Gap in the Small Business Sector*. London: Routledge (1997).
- Panzar, J. C., and R. D. Willig. "Economies of Scope." *American Economic Review*, 71 (1981), 268–272.
- Seru, A. "Firm Boundaries Matter: Evidence from Conglomerates and R&D Activity." *Journal of Financial Economics*, 111 (2014), 381–405.
- Shan, H. "Corporate Venture Capital and the Boundaries of the Firm." 2nd Emerging Trends in Entrepreneurial Finance Conference, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3145341](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3145341) (2019).
- Siegel, R.; E. Siegel; and I. C. MacMillan. "Corporate Venture Capitalists: Autonomy, Obstacles, and Performance." *Journal of Business Venturing*, 3 (1988), 233–247.
- Sykes, H. B. "Corporate Venture Capital: Strategies for Success." *Journal of Business Venturing*, 5 (1990), 37–47.
- Teece, D. J. "Economies of Scope and the Scope of the Enterprise." *Journal of Economic Behavior & Organization*, 1 (1980), 223–247.
- Tian, X., and K. Ye. "Less is More: Institutional Investors and Corporate Venture Capital." PBCSF-NIFR Working Paper, available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3241866](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3241866) (2018).

- Villalonga, B. "Diversification Discount or Premium? New Evidence from the Business Information Tracking Series." *Journal of Finance*, 59 (2004), 479–506.
- Williamson, O. E. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York, NY: The Free Press (1985).
- Winters, T. E., and D. L. Murfin. "Venture Capital Investing for Corporate Development Objectives." *Journal of Business Venturing*, 3 (1988), 207–222.