J Technol Transf (2018) 43:1466–1492
https://doi.org/10.1007/s10961-016-9540-1

Venture capital as a catalyst for commercialization and high growth

Ryan Kelly[1] · **Hankook Kim**[1]

Published online: 8 December 2016

Abstract We use Canadian data linking information on venture capital (VC) financing with firm-level administrative data to compare performance of VC-backed and non-VC-backed firms. The richness of the data allows us to incorporate a wide range of firm-level information into creating a control group based on propensity-score matching. In particular, we use typical covariates reflecting firm performance and characteristics (e.g., size, age, industry, location) as well as measures of firm-level innovation such as research and development (R&D) expenditures that are often thought to be associated with the potential for high growth and the probability of receiving VC financing. Our results show R&D expenditures not only attract VC, but are also increased more intensely for VC-backed firms than non-VC-backed counterparts over the short-run. Further, we show VC-backed firms enjoy greater growth in wages and scale over the 5-year period. Overall, our results provide empirical evidence that VC financing is associated with the acceleration of the innovation and commercialization process accompanied by greater growth in wages and scale.

Keywords Venture capital · Innovation · Firm performance · Matching estimation

JEL Classification L25 · L26 · O32

The views and opinions expressed in this paper are those of the authors alone and do not represent, in any way, the views or opinions of Innovation, Science and Economic Development Canada or of the Government of Canada.

✉ Ryan Kelly
ryan.kelly@canada.ca

Hankook Kim
hankook.kim@canada.ca

1 Innovation, Science and Economic Development Canada, 235 Queen St, Ottawa, ON, Canada

1 Introduction

The significance of venture capital (VC) as a financing instrument for high-growth firms has been well documented. VC investors rely on specialized knowledge to identify early-stage firms with the potential for high growth, and proceed to buy large equity stakes in these firms with the objective of helping them to mature into profitable enterprises. As part of its growth strategy, VC provides not only financing but also close monitoring, expertise, and industry contacts to the investee firm. These additional supports could greatly improve a firm's chances for success given that VC-backed firms tend to be young and possess technologies, ideas or assets with strong potential, but often lack the experience or the proven sales base to secure a sufficient level of financing to expand.

Given the managerial support and monitoring provided through VC investments, it is not surprising to observe the impact of VC manifests itself through a number of different aspects of a firm's operations. In particular, the literature has identified VC's active role at the firm level in product market strategies and commercialization (Hellmann and Puri 2000), human resources practices (Hellmann and Puri 2002), innovation strategies (Da Rin and Penas 2007; Colombo et al. 2016) and patenting (Arqué-Castells 2012). VC's influences lead to positive firm outcomes such as employment, sales growth and a reduced likelihood of failure (Engel and Keilbach 2007; Bertoni et al. 2011; Puri and Zarutskie 2012). The positive effects seem to last even after firms go public in terms of superior operating returns (Jain and Kini 1995). Studies also show the positive outcomes are measurable at the aggregate level. In particular, Kortum and Lerner (2000) show industries receiving more VC have greater patenting rates while Samila and Sorensen (2010, 2011) show firm formation, employment and patenting in a region are positively affected by supply of VC. See Manigart and Wright (2013) for a comprehensive summary of research on how VC affects its portfolio firms.

While the literature has identified certain positive effects of VC on its portfolio firms, there are remaining gaps. In particular, the relationship between innovation and VC remains unclear. On one hand, VC seems to have positive effects on innovation as captured in changes of a firm's strategies to do research and development (R&D) (Da Rin and Penas 2007; Colombo et al. 2016). On the other hand, the studies (Engel and Keilbach 2007; Audretsch et al. 2012; Arvanitis and Stucki 2014; Hoenen et al. 2014; Lahr and Mina 2016) suggest VC has little effects on innovation as measured by the number of patents. The economic rationale is that firms obtain patents as a way to signal the marketability of their innovations and attract VC investment, and as such, post-VC, there is little increase in the number of patents the firm produces. However, one study (Arqué-Castells 2012) does find the positive effects of VC on the number of patents, underscoring the not-so-clear relationship between VC and innovation.

The studies based on patents may not provide a complete picture on the relationship between VC and innovation. While patents can be considered an output of R&D, firms have a choice to patent or not, independent of underlying firm-level innovative activities, and as such, patents may not always be closely related to firm-level innovation. Indeed, several studies in the literature highlight the factors leading to a higher patent intensity without changing the innovation performance at the firm (Kortum and Lerner 1999; Hall and Ziedonis 2001; Mowery et al. 2001). In relation to VC, patenting could be used to signal high quality R&D as part of the efforts to attract VC regardless of the underlying R&D intensities.

A related measure of innovation that is often used in the literature is R&D expenditures. As an input to innovation, R&D expenditures are more likely to be readily influenced by shifts in the firm strategy. In fact, a patent is cumulative of several years' investment, and as such, the number of patents may not vary as much as R&D expenditures per change in efforts to innovate. In the literature, one study (Arvanitis and Stucki 2014) examines the effects of VC on the probability of performing R&D and fails to find any effects. However, since the firms looking for VC are likely in high-tech sectors, we expect most of them to be conducting R&D regardless of whether they are successful in getting VC or not. Accordingly, R&D intensity or expenditures could be a better measure in benchmarking a difference in innovation between VC-backed and non-VC-backed firms.

In this paper, we use a newly constructed VC dataset from Canada that combines Thomson Reuters' VC data and Statistics Canada's administrative data to examine the effects of VC on firm performance. Specifically, we construct a control group that mirrors our sample of VC-backed firms across a number of firm characteristics at the time of the first VC investment, and then examine growth differences in select performance metrics between the two groups of firms, including R&D expenditures. The richness of the data allows us to utilize a wide array of covariates that capture the essence of VC financing more than most studies. In particular, the covariates used in this paper not only include typical measures of firm characteristics and performance such as total assets, sales, employment, wages, profits and age, but also measures of innovation performance by a firm such as whether the firm performs any R&D, R&D expenditures and participation in a government R&D support program that includes non-monetary support (Industrial Research Assistance Program, IRAP). Using these covariates, we estimate the probability of receiving VC investment, which helps us answer the question whether an innovative firm attracts VC. With the estimated probabilities of receiving VC, we construct a control group that shares similarities with the firms in the treatment group, and then compare both groups of firms to examine whether VC affects a firm's growth performance not only in scale and profitability, but also in innovation as measured by R&D expenditures and the quality of job creation as proxied by wages.

We find that a firm's previous record on innovation has positive effects on the probability of receiving VC. In particular, R&D expenditures, incidence of R&D and participation in the R&D support program all have strong positive effects on the probability of receiving VC. This confirms findings in the literature, particularly on how having patents help firms obtain VC financing (Engel and Keilbach 2007; Audretsch et al. 2012; Hoenen et al. 2014). Further, our results on the positive effects from IRAP are consistent with the findings in Colombo et al. (2016), that there is a positive link between an external government-sponsored R&D program and VC investment. The covariates on scale (e.g., sales, total assets, employment) also have positive effects on the probability while VC is not necessarily attracted to profitable firms.

Our results on the growth performance comparison show VC-backed firms increase their R&D expenditures more rapidly than comparable non-VC-backed firms. However, the effects are evident only in the short-run (i.e., one year after the investment). These results are consistent with Arqué-Castells (2012) in a sense that positive effects of VC on innovation are more pronounced in the short-run. Further, we show the lack of VC's effects on patenting as suggested by some studies does not provide a full picture on how VC affects firm-level innovation. Indeed, our results show VC ramps up R&D expenditures in the early years after the VC investment, which could be associated with VC's efforts to quickly commercialize from existing research (i.e., product development). This is consistent with a previous study by Hellmann and Puri (2000), which shows VC shortens a product's time to market. Ultimately, this paper underscores the importance of examining a broader set of innovation measures, including R&D expenditures, in evaluating the effects of VC on innovation.

A key missing piece in the literature is whether innovation as a result of VC investment translates into the creation of sustainable high-paying jobs. Innovation is considered an important contributor to the overall economic growth, and as such, is expected to enhance firm-level productivity and contribute to wages growth. Our results show that VC-backed firms experience higher wages growth than non-VC-backed counterparts, and this difference is evident not only in the first year after the VC investment, but also after five years. VC's positive effects seem to be manifested not only though innovation itself, but also through more concrete economic benefits such as wages growth that could be driven by innovation. These results provide important insight into how VC could shape firm performance through innovation and add an important dimension to the discussions in the existing literature.

We also examine the scale growth of the VC-backed firms and find results that are consistent with the existing literature. In particular, we find that VC-backed firms enjoy a growth premium relative to non-VC-backed firms following VC investment in terms of total assets, numbers of employees, revenues and sales. However, VC's more rapid growth in scale does not translate into superior performance in profitability as the growth in profits per employee between the control and treatment groups are similar. This does not mean VC-backed firms' profitability is deteriorating after the VC investment. Instead, both VC-backed and non-VC-backed firms in the sample experience positive growth in profits per employee that is statistically indistinguishable. This is consistent with the finding in the literature that VC's main focus is on growth rather than profitability (Puri and Zarutskie 2012).

Our results have strong implications for public policy supporting VC as a way to promote innovation. In particular, we show VC could be an important financial intermediary in helping a firm go from existing basic research (e.g., patents) to commercial products with increased R&D expenditures in the early years after the investment. As such, if the objective of public policy were to boost the innovation ecosystem by ensuring firms get on a financially sustainable path of continuing innovation, VC could be an important element as it helps firms financially gain from their innovation. While this paper highlights the importance of VC in public policy consideration, it does not provide evidence on whether a certain government support has been successful. This is particularly relevant in Canada—the source country of our data—since much of the literature on VC in Canada shows the shortcomings of VC support programs (Cumming and MacIntosh 2004, 2006, 2007; Brander, et al. 2010). Regardless of how important VC is to the innovation ecosystem, care is still required in designing an effective support to VC. Research on how different types of VC affect firm performance in innovation and growth could help in this regard, as identified by Manigart and Wright (2013) as a key knowledge gap in the literature.

The paper proceeds as follows: Sect. 2 describes the data; Sect. 3 presents our analytical framework and methodology; Sect. 4 discusses the results; and Sect. 5 concludes.

2 Data

We utilize a dataset developed through a linkage between Thomson Reuters' VC database (Thomson Data)—covering information on Canadian VC investments—and a number of administrative enterprise-level databases at Statistics Canada (STC Data). The administrative datasets include the Corporate Income Tax Returns (T2), the Statement of

Remuneration Paid (T4) and the Business Registry (BR). These data were further linked to external data from the Industrial Research Assistance Program database (IRAP Data). The dataset was originally developed for a 2013 industry report on economic contributions of firms backed by VC in Canada by the Canadian Venture Capital Association (CVCA) and contains 1545 firms. The details on how the linkage was built are summarized in "Appendix 1", but also can be found in CVCA (2013).

2.1 Data linkage

The sample of the dataset is drawn from the population of incorporated enterprises employing at least one individual.[1] We refer to these units as firms in our analysis. Each database in the STC Data contains common firm-specific identifiers, which can be used to establish a deterministic linkage of all records across different databases in the STC Data. These data provide detailed firm-level information covering basic firm characteristics, typical items on financial statements and employee information. Specifically, total assets, total revenue, sales, net income, retained earnings, gross profits and R&D expenditures come from the T2 while total payroll and the number of Individual Labor Units[2] (ILUs) come from the T4. The BR provides the date of incorporation, the province of operation and the primary North American Industry Classification System (NAICS) code.

The STC Data are linked to two other datasets, Thomson Data and IRAP Data. The former provides information detailing the amount and date of all VC investments received by a given Canadian firm going back to the early 1990s. This information is used to identify the treatment and control groups. The latter provides information on the amount and date of any IRAP funding a firm in Canada has received. Alongside the STC Data, this information helps us identify covariates for matching estimation. In particular, we anticipate IRAP-receiving firms are more likely to have a similar focus on technology and growth opportunities as those seeking VC investment. Further, given the similar support offered through IRAP funding, controlling for IRAP will reduce the possibility of attributing any difference in performance resulting from IRAP to VC.

The linkage of the STC Data to these datasets was done based on names and addresses including postal codes and cities.[3] Not all firms in the BR had exactly the same names and addresses as their suspected counterparts in the Thomson Data and IRAP Data. Accordingly, some judgements are made to improve the linkage rate. For example, firms are considered the same if they have small variations in their names but identical addresses. Further, some of the linkages are established manually by experts familiar with the particular companies, i.e.,

[1] The definition of an enterprise can be found at http://www.statcan.gc.ca/concepts/definitions/ent-eng.htm: *The enterprise, as a statistical unit, is defined as the organisational unit of a business that directs and controls the allocation of resources relating to its domestic operations, and for which consolidated financial and balance sheet accounts are maintained from which international transactions, an international investment position and a consolidated financial position for the unit can be derived.*

[2] ILU is a continuous measure of the annual number of employees for a firm. Each individual receiving a T4 represents 1.0 ILU that is allocated to one or more firms according to the fraction of the individual's total annual payroll contributed by the particular firm. This measure will partially account for firms that hire employees later in the year as well as individuals working more than one job, but does not account for differences in the number of hours worked across employees.

[3] The IRAP Data does not contain information on the operating street address or postal code, and as a result, is linked based on the enterprise name, city and province. While the linkage is made using less information, the IRAP linkage achieves a higher initial rate of concordance due to higher quality matches among the firm names. This is not surprising as the recipients of IRAP funding are required to use their legal names, and as a result, are more likely to be exactly identified in the BR.

analysts at the Small Business Branch of Innovation, Science and Economic Development Canada. Ultimately, this dataset has firm-level information on 1545 VC-backed firms. See "Appendix 1" and CVCA (2013) for more information on the linkage between the STC Data and the external databases.

2.2 Longitudinalization

The linked dataset is essentially a series of annual cross-sectional data over the 1999–2009 period, which is not suitable for a robust assessment on growth performance of a firm. In particular, the identifications assigned to firms in the STC Data, called BRIDs as they are assigned through the BR, can change over time for reasons that are unrelated to the destruction or creation of a firm (e.g., legal name change). Further, mergers and acquisitions (M&As) are not identified or treated in the STC Data. To address these shortcomings and ensure growth performance is appropriately measured and attributed to the correct firm, the STC Data were longitudinalized through a labor tracking methodology. In particular, movements of employees were tracked based on T4 filings and corresponding employees' Social Insurance Numbers (SINs) to identify any structural change in a firm as an entity. For example, consider a case where a particular BRID associated with a firm that employs 50 people ceases to report any economic activity in one year. If a new BRID contemporaneously enters the sample, which employs more than 25 of the individuals under the previous BRID, and in the absence of significant relationships with other BRIDs, it would be concluded that the two BRIDs belong to the same firm. We apply this procedure to a matched sample of VC-backed and non-VC-backed firms to create a single longitudinal record for each firm in the paring. Essentially, we maintain the continuity of a firm that is acquiring others while attempting to end the record of a firm that is merged with or acquired by others. See "Appendix 2" for further information on the longitudinalization.

3 Matching VC-backed and non-VC-backed firms

3.1 Firms available for matching

Our approach involves matching VC-backed firms to a control group at the time of the first round of VC financing and examining their subsequent divergence in growth performance. As such, it is imperative to exclude firms without financial information at the time of the first VC financing from our analysis. This exclusion involves firms receiving the first VC financing prior to 1999 as the STC Data does not go back beyond that year. This leads to a loss of 501 firms, reducing our sample from 1545 to 1044 VC-backed firms. Further, even among the firms receiving the first VC financing in the post-1999 period, some had no information from the T2 or T4 in the year of their first VC financing or belonged to an industry where it was there by itself or with only one other firm. Given that the VC financing information spans over a period of 20 years (1990–2009) and VC tends to focus on a narrow selection of industries, a NAICS code with an extremely small number of VC-backed firms likely suggests misclassification.[4] We exclude these observations to arrive at

[4] Although ubiquitous for their ease of use, NAICS codes are not necessarily ideal proxies for the market in which a firm competes as NAICS codes categorize firms based on production rather than on demand. VC investors focus on markets where they have technical knowledge or some form of expertise. Accordingly, a lack of investment activity among VC funds in a VC-financed firm's NAICS code could indicate the assigned NAICS code does not accurately reflect the firm's main market, or the firm may not have suitable analogues within that NAICS code.

Table 1 Benchmarking VC-backed firms to general population during first year of VC financing

Variable	Times larger than general population[a]	# of Observations	*P* value
Total assets	2.359	662	0.000
Sales	1.383	662	0.002
Employment	2.428	662	0.000
Wages	1.155	662	0.000
Age	0.610	662	0.000
Gross profits	1.595	662	0.000
Gross margin[b]	0.765	482	0.000
Gross profits per employment[b]	0.588	485	0.000
R&D expenditures[c]	1.818	432	0.000

The general population refers to all enterprises operating in Canada reporting the necessary financial information and operating in the relevant 4-digit NAICS codes during the 1999–2009 sample (2,573,663 observations)

[a] $\frac{1}{n}\sum_{i=1}^{n}\left(X_i/\bar{X}_{Non\_VC\_ind\_year,i}\right)$, where i indexes VC-backed firms during the first year of financing, and $\bar{X}_{Non\_VC\_ind\_year,i}$ refers to the average X of non-VC-backed firms in the general population in the same 4-digit NAICS industry of firm i during the year of firm i's first round of VC financing

[b] Based on firms with positive profits

[c] Based on R&D performers only

662 VC-backed firms, a potential treatment group that can be matched to similar firms among the general population.

We test to determine how our sample of VC-backed firms at the time of their first VC financing compares to the general population of firms in Canada over the 1999–2009 sample period. A straight comparison between averages of these respective groups will not be very informative since the majority of the firms in the VC-backed sample are concentrated near the beginning of the sample period in contrast to the general population. Further, the VC-backed sample exhibits a different industry distribution, which is disproportionately concentrated within professional services. To address these concerns, we calculate the average ratio of the values for the VC-backed firms relative to the average value for their corresponding 4-digit NAICS category and year of financing. In essence, we control for industry and cohort effects.

Table 1 shows these ratios, which are all statistically significant from one at the 1% level or lower. As expected, the results suggest our sample of 662 VC-backed firms is significantly different from the general population of firms in Canada across several covariates that could be important in determining firm growth outcomes. The fact that VC financing is concentrated among a very distinct sub-set of the general population presents methodological challenges as how these covariates influence growth could involve complicated non-linear relationships. Given the difficulty and potential biases that could result from specifying an incorrect functional form for firm growth, we opt for a matching methodology, which allows us to deal with this complexity by avoiding having to decide on a functional form while controlling for a large set of covariates that could affect growth. As highlighted in Table 1, a matching methodology that produces closely aligned matches across multiple covariates is critical to construct an appropriate control group for our analysis.

3.2 Matching estimator

Our intention is to take advantage of the richness of the data and match firms over a large set of continuous covariates. To achieve this, we use propensity score matching as first outlined in Rosenbaum and Rubin (1983, 1985). This strategy involves estimating the probability of receiving a treatment, or propensity score, and matching treatment observations to potential controls with a similar propensity score. In effect, the propensity score summarizes the covariate distribution for each observation while accounting for each covariate's significance in explaining the treatment outcome. Matching on the propensity score allows us to include more firm characteristics into the matching without requiring exact (or near exact) matching over each covariate.

Our propensity score matching estimator involves three steps. First, we fit a logit model to calculate the predicted probability of receiving VC for the entire population as in Eq. (1).

$$\Pr(VC_{it} = 1) = \alpha_c + \alpha_t + \alpha_{ind} + \beta X_{it} + \varepsilon_{it}, \tag{1}$$

where VC_{it} is a dummy variable taking the value of one when a firm receives VC for the first time.[5] X_{it} is the full set of the desired covariates, α_c is the constant, α_t and α_{ind} are time- and industry-specific intercepts, and ε_{it} is the standard econometric error term.

Second, we define a tolerance threshold or caliper based on the propensity score calculated in the first step to create a set of potential matches for each VC-backed firm that falls within their respective cohort based on the province, industry and financing year of the VC-backed firm.[6] VC financing is a highly idiosyncratic event and is concentrated among firms that differ significantly from the general population. Accordingly, some VC-backed firms may not have a suitable match within their cohort, and a caliper is a common method to ensure that the 'nearest neighbour' for a given VC-backed firm is included only if it is reasonably similar. We select a caliper of 0.25 standard deviation of the linear propensity score to balance between potential selection issues from omitting observations and biases from including poor matches.[7] If there is no match for an observation from the group of VC-backed firms that satisfies the pre-defined threshold, it is dropped from the sample as it has no suitable analogue among non-VC-backed firms.

Third, we construct a control group by pairing each VC-backed firm in the treatment group to a corresponding non-VC-backed firm (i.e., closest in terms of the linear propensity score) within the set of potential matches described in the second step. More explicitly, let P denote the linear propensity score[8] and I_0 the set of non-VC-backed firms within the

[5] A number of firms receive multiple rounds of VC financing, which is a standard practice in the industry. However, for the matching purpose, our focus is on how firm characteristics affect a firm's probability of becoming VC-backed (i.e., receiving any kind of VC investment). Accordingly, for the VC-backed firms, only the observations up to and including the first round of VC financing are included in the regression.

[6] By limiting the potential set of matches to the cohort based on the VC-backed firm's province, industry and financing year, we are performing an exact matching over these covariates. This strategy makes sense since these covariates are categorical, and as such, it is possible to obtain a reasonably high match rate even through exact matching. Further, location and industry are not just important variables determining the likelihood of receiving VC financing. An identical match on these covariates allows us to fully control for industry-location-time-specific shocks.

[7] See Rosenbaum and Rubin (1985) for a discussion on selecting a caliper size.

[8] Defined as $p = \ln(\Pr(VC = 1)/(1 + \Pr(VC = 1)))$.

same 4-digit NAICS code and province during the first year of VC financing. The match for firm i, M_i, is defined as:

$$M_i = \left\{ \min_j \parallel abs(p_i - p_j) \parallel, j \in I_0 | abs(p_i - p_j) < (0.25)\sigma_p \right\}. \quad (2)$$

We follow a nearest-neighbour matching strategy on the linear propensity score without a replacement and remove each pair of VC-backed and non-VC-backed firms from the pool once they are matched. Given that some of the non-VC-backed firms could be the closest match for multiple VC-backed firms, the order in which the matches are made could impact which firms are selected for the control group.[9] Sensitivity tests using different orderings as well as matching with a replacement yield qualitatively and quantitatively the same results.[10] Following these steps, a difference in means test between the treatment and control groups yields a consistent estimate of the Average Treatment effect on the Treated (ATT) under standard assumptions,[11] or more colloquially the impact of VC financing on VC-backed firms.

As we match firms over a large set of covariates, which would impose greater information reporting requirements, we leave out some VC-backed firms, and as a result, the matched sample may not be fully reflective of the whole Canadian VC market. However, most of the sample attrition comes from simply including basic covariates that are standard in the literature (e.g., sales, employment, age and industry),[12] which requires firms to be captured in all three of the administrative data sources (i.e., T2, T4 and BR). Further, the focus of this paper is on innovation and firm performance, and as such, if firms do not report on measures of innovation, they cannot be used to answer our research question. To test our results, we show, in Sect. 4.4, an alternative methodology (*exact matching*) ends up producing a less satisfactory control group, particularly in R&D performance.

Ultimately, to estimate a robust measure of VC's impact, one would have to account for large selection effects that are present in how VC selects a firm to invest. As noted in Lerner (2009), a VC investor could take up to 160 h to screen a potential investment. Our intention is to minimize these selection effects by including as many covariates as possible and examine whether the performance premium for VC-backed firms remains after controlling for a more expansive set of covariates including measures of innovation. As there is no a priori reason to suspect our matched samples significantly lack the representativeness of the Canadian VC market, the increased risk of issues related to restrictive matching seems minor in comparison to the increased accuracy of the control group.

9 We order the treatment firms based on their date of first financing to maximize the length of the longitudinal records of the treatment and control groups.

10 The sensitivity tests are not reported due to residual disclosure concerns that would violate Statistics Canada's confidentiality requirements.

11 Our matching estimator requires two basic assumptions. First, there is a set of covariates, X, such that firm outcomes are independent of the treatment (i.e., VC financing) after controlling for these characteristics. As we are estimating the ATT, only mean independence (i.e., $E[Y(0), Y(1)] \perp (VC|X)$) is necessary. See Heckman et al. (1997, 1998) for more information. Second, for all X, there is a positive probability of either receiving or not receiving a treatment (i.e., there is a region of common support).

12 See Engel and Keilbach (2007), Bertoni et al (2011), Puri and Zarutskie (2012) and Arvanitis and Stucki (2014).

Table 2 Logit—Pr(VC first financing = 1)

Independent variables	Coefficient	SE	*P* value
Ln total assets	3.527***	0.384	0.000
Asinh sales	0.021	0.038	0.583
Ln employment	0.538***	0.107	0.000
Ln wages	6.270***	1.736	0.000
Asinh retained earnings	−0.043***	0.005	0.000
Asinh revenue	0.018	0.014	0.200
Asinh net income	−0.029***	0.005	0.000
Age	−0.076***	0.012	0.000
Asinh R&D expenditures	0.270***	0.045	0.000
Ln total assets squared	−0.107***	0.013	0.000
Asinh sales squared	−0.002	0.003	0.415
Ln employment squared	−0.056***	0.019	0.003
Ln wages squared	−0.282***	0.082	0.001
Asinh retained earnings squared	−0.008***	0.001	0.000
Asinh revenue squared	−0.009***	0.002	0.000
Asinh net income squared	0.014***	0.002	0.000
Age squared	0.001**	0.000	0.021
Asinh R&D expenditures squared	−0.009***	0.003	0.004
IRAP dummy	1.154***	0.097	0.000
R&D 2000 dummy	1.973***	0.253	0.000
Industry effects	Yes		
Year effects	Yes		
N	2,573,663	Pr > ChiSq	0.000
LR test	5059.62	LL	−3603.92
Pseudo R^2	0.412		

*** 1% significance level; ** 5% significance level; * 10% significance level

The sample consists of all firms filing both T2 and T4 spanning 1999–2009 within NAICS 4-digit industries where at least three firms have received VC financing at some point during the 1990–2009 period. All values are in either natural logs (Ln) or the inverse hyperbolic sine transformation (Asinh), which is similar to a log transformation but is defined at zero and for negative values. $\text{Asinh}(y_i) = \log\left(y_i + \left(y_i^2 + 1\right)^{1/2}\right)$

4 Empirical results

4.1 Matching results: propensity score matching

Using propensity score matching requires us to estimate a model to produce predicted probabilities of receiving a treatment. We report the results from the logit estimation of the propensity score in Table 2. We include all available variables that could affect either firm outcomes or the probability of receiving VC financing. We also include the corresponding squared values for all the continuous variables to allow for non-linear relationships. The variables used in the logit estimation include total assets, sales, the number of ILUs, wages, retained earnings, revenues, net income, age, R&D expenditures, a dummy variable for firms financed in 1999 that conducted R&D in 2000,[13] industry fixed-effects based on 4-digit NAICS codes and year fixed-effects.

In addition to these standard variables capturing the firms' activities related to their operations and innovation, we include a dummy variable indicating whether the firm receives government support through the IRAP funding. The dummy variable for IRAP is not directly related to operations or innovation, rather it represents whether a firm is successful in obtaining financial and mentoring support from a government program designed to promote innovation at the firm level. This variable likely captures some of the residual management capabilities that are not captured by the standard financial variables but affect the firm's growth outcome or the probability of receiving VC investment. To receive IRAP funding, not only do firms need to pass a vetting process to demonstrate the proposed project or the business plan is commercially viable and sufficiently innovative, but the due diligence process also examines the management capability of the firm. IRAP funding is arguably one of the most robust characteristics to establish a potential control group as IRAP targets similar firms as VC and provides similar supports (e.g., funding, mentoring, networking and linkage). As such, it is included in our logit estimation.

The coefficient estimates largely conform to general knowledge of the VC industry. In particular, the coefficients suggest firm size as measured by total assets, employment, sales and revenue is positively correlated with the probability of a firm receiving VC financing. However, the positive correlation eventually fades away and becomes negative as the coefficient estimates for the squared terms of the size variables are all negative. This is consistent with VC investing in relatively small firms, but generally avoiding very small or micro sized firms—64% of VC-backed firms in the treatment group have more than 10 employees. The coefficient estimate for age is negative, suggesting VC tends to invest in younger firms. Profitability measures, net income and retained earnings, are negatively correlated with the probability of receiving VC financing. This likely stems from VC investing in early-stage firms that tend to be subject to negative profitability as they grow. In addition, firms with large profits can self-finance or secure financing through more traditional channels, thus less likely to seek VC financing. Our results show VC looks to invest in early-stage young firms with growth potential, which is consistent with VC's focus on growth as suggested in the literature (Engel and Keilbach 2007; Bertoni et al. 2011; Puri and Zarutskie 2012).

With respect to innovation measures, we find significant positive correlations of previous records of innovation with the probability of receiving VC. The coefficient estimates for IRAP, R&D dummy variables and R&D expenditures are all positive and statistically significant. Further, the magnitudes of these estimates are large. For example, the coefficient estimate for receiving IRAP suggests the propensity score increases by over 300% from receiving IRAP. As R&D expenditures are a continuous variable, we need to consider the sample distribution to assess the impact of R&D expenditures on the probability of receiving VC financing. An important distinction between R&D expenditures and other financial variables is that it is common for firms to report $0 of R&D expenditures. Accordingly, a firm could experience a much higher impact to their propensity score from becoming an R&D performer than from a range of percentage changes in assets or wages, which have coefficient estimates larger than other variables. To state this in numbers, among the sample of non-R&D performers in the treatment and control groups, a discrete

[13] R&D expenditures data is not available for the year 1999. Note that we do not include firm observations from 1999 when calculating growth rates for R&D in the subsequent sections. The dummy variable is for the sole purpose of achieving better matching among innovative firms financed in 1999.

change in R&D expenditures from \$0 to \$15,000 (a relatively small amount, but a large implied percentage gain) would lead to an average 460% increase in their estimated probabilities of receiving VC. The significant positive effects of firm-level innovation on receiving VC are consistent with the findings in the literature that show patents and government-sponsored R&D programs are positively linked to VC involvement (Engel and Keilbach 2007; Audretsch et al. 2012; Hoenen et al. 2014; Colombo et al. 2016).

Our matching strategy based on propensity scores leads to an 82% match rate, resulting in 544 matched pairs. Table 3 shows differences between the matched samples of VC-backed and non-VC-backed firms. They are very similar in all dimensions and the *P* values of the differences are high enough to suggest the two groups of firms are similar in a statistical sense. In addition to standard hypothesis testing, we present the results from an alternative balance measure—the standardized difference in means (SDM). This measure is often considered more reliable in the context of assessing changes in balance between two samples as standard tests are subject to a higher probability of encountering type II errors when the sample size is reduced. Rubin (2001) suggests that SDM values exceeding 0.5 can be problematic. As shown in Table 3, the SDM values for all of our covariates of interest and that of the linear propensity score are at or below 0.05, suggesting the treatment and control groups are well balanced.

The comparison based on the number of R&D performers and IRAP recipients also highlights the strong similarity between the control and treatment groups. As shown in Table 4, both groups have nearly identical counts of R&D performers and IRAP recipients, suggesting a similar degree of focus on innovation across the two groups. These results reinforce our confidence in the extent to which the control group resembles the treatment group at the time of the first VC investment.

Moving beyond the mean comparisons and simple firm counts, we compare the actual distribution of the linear propensity score for the treatment and control groups in Fig. 1.[14] The kernel density estimates show that the treatment and control groups have a nearly identical distribution for the linear propensity score. This visual inspection is confirmed by the Kolmogorov–Smirnov equality of distributions test, which has a *P* value very close to one, suggesting that the two distributions are statistically indistinguishable. The treatment and control groups are not only similar in mean but also in their distributions of the linear propensity score.

The results in Tables 2, 3 and Fig. 1 highlight the robustness of our matching estimator. In particular, both groups share a similarity not only over the overall propensity score, but also over important individual dimensions such as size, age, wages and profitability. There is also a similarity over firm-level innovation as measured by R&D expenditures and IRAP participation, which are particularly relevant to the key research question of this paper to examine the effects of VC on innovative activities of firms.

4.2 Growth comparisons

With a matched sample of VC-backed and non-VC-backed firms, we can compare their respective growth rates over various timeframes. For each group of the firms, the growth rates are computed as follows:

[14] We cannot replicate this exercise for other variables while maintaining Statistics Canada's confidentiality requirements.

Table 3 Comparison between matched pairs—propensity score matching

Covariate	Mean VC	Mean control group	Difference	P value of difference	Standardized difference in means[a]
Ln total assets	14.485	14.453	0.032	0.735	0.020
Asinh[b] sales	11.618	11.835	−0.217	0.539	−0.037
Ln employment	2.717	2.765	−0.049	0.544	−0.036
Ln wages	10.683	10.674	0.009	0.787	0.016
Asinh[b] retained earnings	−8.089	−7.835	−0.254	0.727	−0.021
Asinh[b] revenue	13.507	13.454	0.053	0.850	0.012
Asinh[b] net income	−7.572	−7.305	−0.267	0.707	−0.023
Age	4.952	5.042	−0.090	0.796	−0.016
Asinh[b] R&D expenditures	8.228	8.456	−0.227	0.566	−0.035
Asinh[b] gross profits	9.617	9.201	0.416	0.410	0.052
Linear propensity score	−4.729	−4.754	0.025	0.860	0.011

VC obs = 544; control group obs = 544

[a] Standardized difference in means = $(\bar{X}_{VC} - \bar{X}_{Control})/s_{VC}$

[b] Asinh refers to the inverse hyperbolic sine transformation—$\text{Asinh}(y_i) = \log\left(y_i + \left(y_i^2 + 1\right)^{1/2}\right)$. Asinh is similar to a log transformation but is defined at zero and for negative values

Table 4 Comparison between matched pairs in R&D and IRAP—propensity score matching

	Count		Percentage	
	VC	Control group	VC (%)	Control group (%)
R&D performer in the year of match	336	347	61.8	63.8
R&D performer during the sample period	412	396	75.7	72.8
IRAP recipient in the year of match	91	88	16.7	16.2
IRAP recipient during the sample period	152	156	27.9	28.7

VC obs = 544; control group obs = 544

$$y\ year\ growth = \frac{\sum_{i=1}^{n}\left(\ln X_{i,t+y} - \ln X_{i,t}\right)}{n}, \tag{3}$$

where y indicates the number of years in the time period used to compute the growth rate. We compute growth rates over one-, three-, and five-year intervals. Commensurate with calculating growth, we only use firm observations that have positive values in both periods under consideration. Our approach to calculate growth rates based on log differences is different from a more standard approach in the literature, where the growth rate from Davis et al. (1998) (DHS henceforth) is widely used (See Haltiwanger et al. 2013). We argue that the properties of growth calculated through log differences better reflect VC's investment strategies. VC often invests in multiple high-risk ventures under the expectation that many would fail, but the losses would be more than off-set by a small number of spectacular successes. The DHS growth rate is bounded between −200% and +200%, which limits the

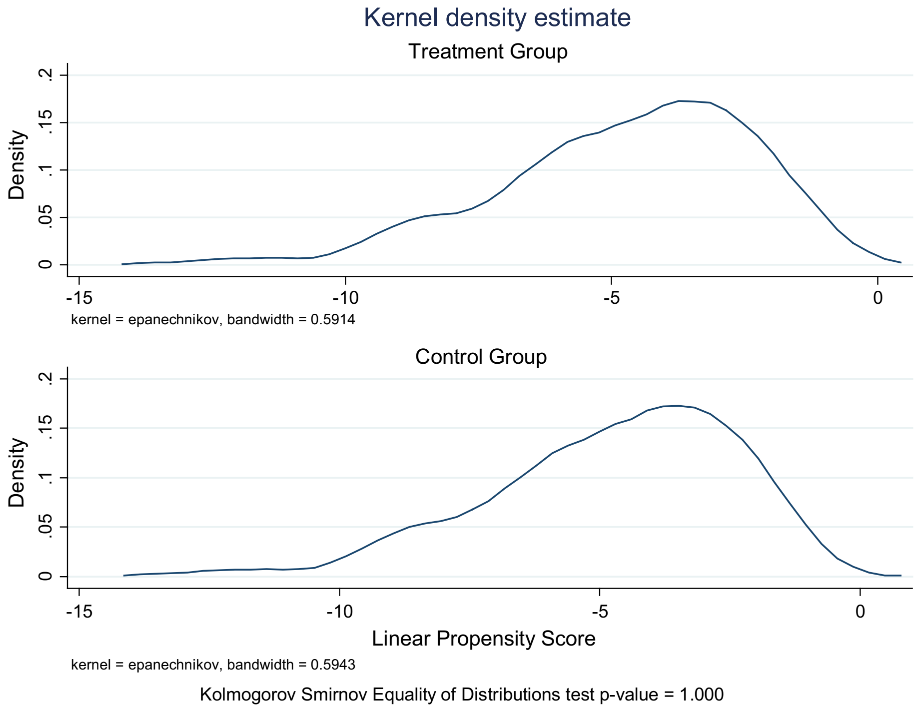

Fig. 1 Kernel density plot of the linear propensity score

impact of these successes. Further, the primary advantage of the DHS growth rate is that it partially mitigates regression-to-the-mean effects. While this is generally an important consideration in calculating growth rates, our focus is on the difference between two growth rates and there is no reason to believe that VC-backed firms could be more subject to mean reversion than their counterparts. If anything, given VC's focus on continuing growth, the growth rates of VC-backed firms are more likely to avoid mean reversion in later periods.

The fact that our data is longitudinalized as described in Sect. 2.2 means, in our computations, we omit firm values if the firm experience an exit (i.e., merged with or acquired by another firm prior to the end of the period). This ensures we measure only the growth achieved during the time when the firm is part of the VC's portfolio. While we omit observations after the firm is acquired or merged, the records after the firm acquire another firm (i.e., growth through acquisitions) are maintained since such a transaction is an integral part of VC's growth strategy. In the end, our computations include more than *organic growth*,[15] but exclude exits based on mergers and acquisitions. Since our methodology is not perfect (i.e., we will never know exactly what happens to the firms unless they are interviewed), there could be bias in our results. However, the bias is

[15] *Organic growth* in reference to employment generally refers to net job creation or destruction independent of employment reallocation. However, among our data, employment changes within the firm include non-*organic* changes (e.g., mergers, acquisitions or divestitures). While non-*organic growth* may amount to nothing more than job shuffling when viewed at the industry level, such growth could reflect important changes in capacity at the firm level.

downward on the growth rates of VC-backed firms, which would strengthen our results that VC-backed firms enjoy growth premium over their counterparts. See "Appendix 2" for more details on the identification of these exits and the longitudinal record keeping of the firms.

All the performance measures referring to dollar amounts are in nominal terms as industry- and time-specific price deflators are unavailable, and as a result, all growth rates reflect nominal growth except those relating to employment or gross margin. While this may seem to limit the interpretation of our results, the growth comparisons between VC-backed and non-VC-backed firms remain robust as, by construction, the treatment and control groups have an identical year and industry profile at the time of their match (i.e., both groups facing similar price inflation over the sample period).

4.2.1 R&D expenditures and wages

Firms looking to develop and commercialize new goods and services are a prime target of VC investors. These innovative activities are likely a focal point of VC financing, monitoring and mentorship support. Accordingly, we investigate growth in two measures related to innovation: R&D expenditures and wages.

R&D expenditures are likely to provide more detailed information on how VC affects firm-level innovation than the measures used in previous studies such as the number of patents or propensity to conduct R&D. While the latter measures can show whether a firm's focus rests on innovation, they do not fully capture, or not at all in case of the propensity, changes in intensity of innovation, particularly subsequent to a structural change in the firm. R&D expenditures can change instantly in response to an ownership change, and are arguably the best measures to gauge the increased intensity of innovative activities at the firm level after VC makes investment.

Table 5 shows the results on R&D expenditures. VC-backed firms experience a higher growth rate in R&D expenditures than their non-VC-backed counterparts by about 16 percentage-points. The difference is also statistically significant at the 5% level. However, the growth premium for VC-backed firms seems to be short-lived as the difference is not statistically significant over the three- and five-year periods. This could reflect VC's ramped up efforts to develop and commercialize the technologies developed prior to the VC's investment.

Our results on R&D expenditures offer an important context to the findings in the literature on how VC affects firm-level innovation. While our results are consistent with Arqué-Castells (2012), which shows VC's positive effects on a firm's patenting in the first two years after the VC investment, they contradict a number of other studies showing VC's lack of impact on patenting or R&D propensity (e.g., Engel and Keilbach 2007; Arvanitis and Stucki 2014; Lahr and Mina 2016). This discrepancy could be driven by a couple of factors. First, patents or R&D propensity may not be as responsive to VC investment as R&D expenditures. After all, a patent represents years of R&D, which suggests VC may have little instant effects on a firm's capacity to patent more, and a decision to do R&D is more likely to be made before any VC investment. Second and more interestingly, VC's focus may have little to do with nudging a firm to create more innovations over the long-run. Rather, VC's main interest could be in ensuring successful product development and acceleration of the commercialization process of the existing innovation, which would be more critical to the sales growth of the firm (and VC does help firms increase sales as shown later in this section). These increased efforts to commercialize are likely captured in increased R&D expenditures. Indeed, Hellmann and Puri (2000) show VC investment is

Table 5 Growth rates for treatment and control groups—R&D expenditures and wages

	Mean VC (%)	Mean control group (%)	Difference (%)	*P* value of difference	# VC	# Non-VC
R&D expenditures						
1 Year growth	25.3	9.2	16.1**	0.012	292	294
3 Year growth	24.9	8.9	16.0	0.230	178	172
5 Year growth	48.9	29.4	19.5	0.230	99	103
Wages						
1 Year growth	8.2	4.6	3.6*	0.061	496	485
3 Year growth	16.7	11.8	4.9	0.106	322	337
5 Year growth	29.4	19.3	10.1**	0.029	193	225

P values are based on a two-sided test. *** 1% significance level; ** 5% significance level; * 10% significance level

associated with improved commercialization performance of a firm as demonstrated through a faster time to market for new products.

Table 5 also shows the results on wages. The VC-backed firms have a 3.6 percentage-points growth premium in wages over the non-VC-backed firms during the first year after the financing. The cumulative growth premium expands to 10.1 percentage-points over the 5-year period. The differences are statistically significant at the 10 and 5% levels for the 1- and 5-year period, respectively. The wages growth is sustained over the long-run, which suggests it is not purely driven by the injection of new VC funds, nor is it the result of downsizing as we show with the results on scale growth that VC-backed firms actually grow in employment. The real value creation through the utilization of the high-skill labor seems to take place after the VC investment.

The results on wages shed light on how VC affects not only firm-level innovation, but also the firm's performance. The development of new products and commercialization often require highly-skilled labor, and as such, intensified efforts in these areas could result in an increase in wages as firms try to meet the new requirements for high value-added employment. The fact that the increase in wages is sustained over the long-run suggests the firms are successful in transforming the increase in quality of the labor input into value creation on the output market. This evidence provides an important insight into how enhanced innovation as a result of VC investment could shape firm performance, adding a new dimension to the existing discussions in the literature.

4.2.2 Scale growth

We evaluate VC's impact on the growth of firm scale. In the literature, a number of studies show VC-backed firms experience higher growth in size than non-VC-backed firms, which our results corroborate using more extensive covariates, including those related to innovation, to control for selection effect. In Table 6, we compare the growth rates between the treatment and control groups for four measures of firm scale: employment, total assets,

Table 6 Growth rates for treatment and control groups—firm scale measures

	Mean VC (%)	Mean control group (%)	Difference (%)	P value of difference	# VC	# Non-VC
Total assets						
1 Year growth	19.6	0.1	19.5***	0.001	487	493
3 Year growth	38.7	−10.2	48.8***	0.000	334	360
5 Year growth	53.5	1.1	52.4***	0.001	215	254
Employment						
1 Year growth	34.4	6.6	27.7***	0.000	496	485
3 Year growth	41.5	−3.5	45.0***	0.000	322	337
5 Year growth	50.6	3.6	47.0***	0.000	193	225
Revenues						
1 Year growth	52.9	28.6	24.2***	0.002	470	468
3 Year growth	93.7	36.7	57.0***	0.000	323	333
5 Year growth	107.3	40.3	67.0***	0.001	203	232
Sales						
1 Year growth	48.7	27.2	21.6**	0.022	383	403
3 Year growth	100.1	47.2	53.0***	0.001	262	273
5 Year growth	137.4	56.0	81.4***	0.001	156	181

P values are based on a two-sided test. *** 1% significance level; ** 5% significance level; * 10% significance level

revenues and sales. The estimates of the differences between the treatment and control groups are all positive and statistically significant at least at the 5% level. The magnitudes of the differences are substantial. For the differences over the 5-year period, the gap ranges from 47 percentage-points for employment to 81.4 percentage-points for sales. Some of these differences come from lagging growth for firms in the control group. For example, with respect to total assets and employment, firms in the control group experience little change while treated firms experience substantial growth. However, for measures related to the output market (i.e., revenues and sales), even the firms in the control group do experience significant growth over the 5-year period (40.3 and 56% in revenues and sales, respectively), and the difference is driven by much more substantial growth experienced by VC-backed firms (107.3 and 137.4% in revenues and sales, respectively).

Most of the growth over the 5-year period occurs by the third year. For all four measures, the share of growth occurring up to the third year is upwards of 70%. For revenues, the share is greater than 87%. These results highlight the VC's focus on immediately improving the performance of the firm, which likely includes providing financing to help planned expansions. This is particularly noticeable in hiring of new labor as more than 67% of cumulative growth for employment occurs in the first year alone.

The consistency of superior performance for VC-backed firms across all four measures of scale growth provides strong evidence that VC is associated with high growth. Further, the results for revenues and sales suggest VC-backed firms do not simply increase their assets or hire new employees with the new infusion of VC funds. Instead, VC-backed firms successfully grow their businesses through increased sales, whether through expansion to meet demand in their current market or by successfully entering new markets to tap additional sources of sales.

4.2.3 Profitability

We compare the profitability between VC-backed and non-VC-backed firms using two measures: gross margin and gross profits per employee. The results are in Table 7. We find no statistically significant divergence in profitability between the treatment and control groups over any period of time. However, that does not mean VC-backed firms suffer a significant decline in profitability as gross profits per employment for VC-backed firms go up by 70.2% over the 5-year period. This suggests, at least for surviving VC-backed firms, the rapid increase in size has little negative impact on the capacity to generate profits over the long-run. Our findings reinforce the previous results (Puri and Zarutskie 2012) that VC's impact is more pronounced on growth than profitability.

4.3 Robustness check: balanced sample

As explained in the previous sections, the performance comparison between VC-backed and non-VC-backed firms is based on all the firms that are successfully matched. However, the sample's initial year and industry balance could be disrupted if there are asymmetrical exits among the matched pairs of treatment and control firms. In particular, the growth rate comparisons could be influenced by survivor bias if poorly performing firms are more

Table 7 Growth rates for treatment and control groups—profitability measures

	Mean VC (%)	Mean control group (%)	Difference (%)	*P* value of difference	# VC	# Non-VC
Gross margin						
1 Year growth	−1.2	−0.5	−0.7	0.717	341	338
3 Year growth	0.5	6.4	−5.9	0.211	240	243
5 Year growth	0.8	5.7	−5.0	0.255	138	159
Gross profits/employment						
1 Year growth	14.3	26.7	−12.5	0.183	341	333
3 Year growth	54.6	48.1	6.4	0.668	235	233
5 Year growth	70.2	71.3	−1.1	0.958	135	152

P values are based on a two-sided test likely to remain in the sample among the control group relative to the treatment group or vice versa. In such cases, the estimates on the differences in growth rates between VC-backed and non-VC-backed firms could be biased downwards if VC financing keeps poorly performing firms afloat, or upwards if VC takes action to accelerate the exit of poorly performing firms. Further, the year-industry balance may be needed to ensure both groups of firms face the same price inflation over the sample period of any duration.

To assess how these concerns could affect our results, we estimate the growth rates based on a restricted sample of the treatment and control groups that only includes matched pairs where both components of the matched pair report the necessary information needed to calculate the growth rates. This mitigates potential bias on the growth rate differences as a result of the differences in survival by equalizing the number of firms showing, ex-ante, good or bad prospects among the pool of surviving firms in the treatment and control groups.

The growth rates based on the restricted balanced sample are presented in Table 8. The results based on the balanced sample are qualitatively and quantitatively the same as our main results. In particular, we confirm the key observations that VC-backed firms experience (1) greater R&D expenditures in the short-run, (2) an increase in wages over the long-run, (3) superior growth performance in scale and (4) little difference in profitability.

4.4 Robustness check: *exact matching*

We choose propensity score matching based on a large set of covariates including those related to innovation, which could lead to a set of matches that are robust based on the whole of covariates but highly restrictive. Our choice of the approach is largely based on the availability of different covariates in our data. However in the literature, some utilize *exact matching* over a small set of covariates. Puri and Zarutskie (2012) is one such study, in which the authors perform *exact matching* over age, industry classification, geographical region and employment size. While *exact matching* is the simplest and the most robust way to match two observations for a given covariate, it can be highly data-intensive as finding the exactly the same numerical figures among observations simultaneously over various continuous covariates can be very difficult. Indeed, it would be near impossible to perform *exact matching* over the large set of covariates used in this paper unless we choose to reduce the set of covariates or categorize them into a smaller set. This comes with a cost as the likelihood of violating the assumption of mean independence would increase (i.e., firm outcomes may not be independent of the treatment after controlling for the characteristics), resulting in inconsistent estimates.

To verify the quality of the matching based on propensity score, we perform *exact matching* using our data over a small number of mostly discrete covariates: 4-digit NAICS codes (industry), province of operation (location), years since incorporation (age) and the number of ILUs from the T4 (size). This is similar to Puri and Zarutskie (2012) except that our employment figure is more refined than a simple headcount and includes part-time workers or employees joining the firm mid-year. As a result, our employment figure or ILU is not only continuous but also includes decimal points, which makes it less likely to have two exact observations as compared to a headcount. To facilitate *exact matching* over ILU, we use a 10% rule, where firms are flagged as a potential match if their ILUs are either within 10% of the ILU of the treatment firm or within one ILU if the treatment firm has less than 10 ILUs. Our *exact matching* results in 377 matched pairs or a 57% match rate.

Table 9 compares the treatment and control groups based on *exact matching* over several covariates. Since *exact matching*, by definition, will result in exactly the same

Table 8 Growth rates for treatment and control groups—balanced sample

	Mean VC (%)	Mean control group (%)	Difference (%)	*P* value of difference	# VC	# Non-VC
R&D expenditures						
1 Year growth	24.5	6.5	17.9***	0.009	234	234
5 Year growth	36.7	7.9	28.8	0.348	37	37
Wages						
1 Year growth	8.4	5.3	3.1	0.120	458	458
5 Year growth	33.9	14.3	19.6***	0.006	106	106
Sales						
1 Year growth	47.8	28.5	19.2*	0.062	304	304
5 Year growth	141.2	58.9	82.3**	0.015	72	72
Employment						
1 Year growth	34.6	6.7	27.8***	0.001	458	458
5 Year growth	47.4	5.2	42.2***	0.008	106	106
Gross margin						
1 Year growth	−0.5	−0.9	0.4	0.864	236	236
5 Year growth	−0.1	4.9	−5.0	0.437	58	58
Gross profits/employment						
1 Year growth	18.7	27.6	−8.8	0.409	231	231
5 Year growth	65.7	79.4	−13.6	0.660	55	55

P values are based on a two-sided test. *** 1% significance level; ** 5% significance level; * 10% significance level

location and industry, there is no need to compare the sample over those covariates. The upper section of Table 9 compares the pairs over other covariates used in the matching, and as expected, the matched pairs share very similar employment levels and an identical age profile. The bottom section of the table shows the comparison over financial variables (e.g., total assets, sales, profitability), wages and R&D expenditures. While *exact matching* produces very close matches over a small number of covariates used for matching, the results in the bottom section show the matches are significantly different in all the other measures important for VC financing and firm growth outcomes.

In Table 10, we delve deeper and examine how measures related to innovation (i.e., propensity to perform R&D, IRAP participation) differ between firms based on *exact matching*, and the results show a significant difference between the treatment and control groups in these measures. While 242 VC-backed firms perform R&D at the time of the matching, only 87 non-VC-backed firms do the same. In percentage, this amounts to a gap

Table 9 Comparison between matched pairs—*exact matching*

Covariate	Mean VC	Mean control group	Difference	*P* value of difference	Standardized difference in means[a]
Covariates used					
Ln employment	2.193	2.185	0.008	0.931	0.007
Age	3.694	3.694	0.000	1.000	0.000
Other pertinent measures					
Ln total assets	14.081	12.949	1.132	0.000	0.749
Asinh[b] sales	9.877	13.141	−3.264	0.000	−0.531
Ln wages	10.754	10.426	0.328	0.000	0.620
Asinh[b] retained earnings	−10.825	0.704	−11.529	0.000	−1.247
Asinh[b] revenue	12.389	13.953	−1.563	0.000	−0.338
Asinh[b] net income	−9.926	1.156	−11.082	0.000	−1.159
Asinh[b] R&D expenditures	9.435	3.212	6.223	0.000	1.034
Asinh[b] gross profits	8.346	11.851	−3.506	0.000	−0.457

VC obs = 377; control group obs = 377

[a] Standardized difference in means = $(\bar{X}_{VC} - \bar{X}_{Control})/s_{VC}$

[b] Asinh refers to the inverse hyperbolic sine transformation—$\text{Asinh}(y_i) = \log\left(y_i + \left(y_i^2 + 1\right)^{1/2}\right)$. Asinh is similar to a log transformation but is defined at zero and for negative values

Table 10 Comparison between matched pairs in R&D and IRAP—*exact matching*

	Count		Percentage	
	VC	Control group	VC (%)	Control group (%)
R&D performer in the year of match	242	87	64.2	23.1
R&D performer during the sample period	277	123	73.5	32.6
IRAP recipient in the year of match	76	18	20.2	4.8
IRAP recipient during the sample period	116	23	30.8	6.1

VC obs = 377; control group obs = 377

of more than 40 percentage-points in shares of firms performing R&D between the control and treatment groups. A similar difference exists in receiving IRAP. There are 76 and 18 firms receiving IRAP at the time of the matching in the groups of VC-backed and non-VC-backed firms, respectively. This amounts to 20.2% of firms in the treatment group receiving IRAP as compared to 4.8% of non-VC-backed firms. The significant gap in these metrics remains regardless of whether the comparison is at the time of the matching or over the whole sample period. These results suggest that, for our sample, *exact matching* over industry, location, age and size does not produce control and treatment groups that are similar with respect to the number of R&D performers or IRAP recipients.

5 Conclusion

In this study, we compare VC-backed and non-VC-backed firms across several firm performance measures controlling for a wide range of firm characteristics, including those related to innovation that are less extensively used in the previous papers. In particular, we control for R&D expenditures and participation in a government support program for R&D in matching estimation, which helps us examine the impacts of VC on innovation based on growth in R&D expenditures after the VC investment. We find that the growth in R&D expenditures by VC-backed firms is greater than that of non-VC-backed counterparts. The growth premium is the most pronounced immediately after the VC investment, which is consistent with VC involvement accelerating the commercialization process through quicker product development based on existing research and technological know-hows.

Our results also show VC-backed firms enjoy higher wages growth, suggesting they are successful in adding more high-value-added employment than non-VC-backed firms. This adds a new dimension to discussions on how VC shapes firm performance through innovation as high wages can be considered a tangible economic benefit of enhanced innovation and commercialization performance spurred by VC investment. With respect to growth in scale and profitability, we find results reflective of the existing literature that VC helps firms scale up without improving profitability relative to non-VC-backed firms. This confirms the prevailing view that VC's focus rests on growth rather than profitability.

From a policy perspective, our results highlight VC's potential role in helping firms be more innovative. In particular, VC could be an important financial intermediary targeting firms that seek to make financial gains from their research and scientific know-hows while approaching later stages in the innovation and commercialization cycle. Accordingly, VC could be critical to maintaining a large swath of innovative firms in the economy by placing firms in a financially sustainable path of continuing innovation. This could be a complementary element to broad public support for innovation that focuses on supporting firms at the early stage of research and development. That said, there are knowledge gaps concerning the effectiveness of publicly supported VC in enhancing a firm's innovation and commercialization cycle. Accordingly, future work into the performance of different types of VC, including various government programs that are designed in different ways (e.g., tax credits vs. direct support), on how they shape firm performance is warranted.

Acknowledgements We would like to thank Statistics Canada, the Canadian Venture Capitalist Association and the Small Business Branch of Innovation, Science and Economic Development Canada for assistance and comments throughout this project. In particular, we would like to thank Jim Valerio, Shane Dolan, Anne-Marie Rollin and Younes Errounda for their tireless efforts in the data development and descriptive analysis that made this research possible. We would also like to thank Thomas Hellmann, Javier Miranda, Mariagrazia Squicciarini, Robin Prager, Leonard Sabetti and participants at the 2014 OECD Conference on Entrepreneurship, Innovation and Enterprise Dynamics, the 12th Annual International Industrial Organization Conference and the 47th Annual Canadian Economics Association Meetings for their helpful comments and suggestions. Finally, we would like to thank two anonymous referees whose suggestions tremendously improved this paper.

Appendix 1: Linking VC and IRAP data to STC data

In this appendix, we summarize the linkages among the VC, IRAP and STC Data. Both the VC (from Thomson Reuters) and IRAP Data are external to the firm-level databases at Statistics Canada. Accordingly, the linkages are done through names and addresses given that there is no common firm identifier among these data. CVCA (2013) provides full

Table 11 Linkage of Thomson VC data to BR

Name	Address	Postal code	City	Quality flag	Counts	Percentage (%)	Cumulative percentage (%)
Same	Same	Same	Same	D1	208	7.5	7.5
Same	Same			D2	60	2.2	9.7
Same		Same		D3	812	29.4	39.1
Same			Same	D4	227	8.2	47.3
Similar	Same	Same	Same	M1	32	1.2	48.5
Similar	Same			M2	69	2.5	51.0
Similar		Same		M3	96	3.5	54.5
Similar			Same	M4	86	3.1	57.6
Manual linkage					463	16.8	74.3
Unlinked					709	25.7	100.0
Total					2762	100.0	

details on how the linkages are done. Table 11 provides the initial linkage results on the Thomson Data.

A number of VC-backed firms have suspected linkages in the BR, but these potential counterparts are not identical in either their names or addresses. This is not surprising as the data collection method used by Thomson Reuters do not necessarily produce the legal names of the firms. For these records, analysts from Statistics Canada and Small Business Branch of Innovation, Science and Economic Development Canada made the linkages manually, leveraging their expertise and public sources of information on these firms. There are 463 such firms, accounting for 16.8% of the total available for linkage. Including these firms, the initial linkages resulted in 74.3% of 2762 or 2053 firms linked to the BR.

Following the initial linkages, information from other STC Data such as T2, T4 and PD7 remittance slips are used to further check the validity of the linkages. In particular, we are concerned with firms that are identified as financed by VC but do not have any economic activities around the time of their first financing. The linkages of these firms are potentially erroneous as VC-backed firms are expected to be economically active. We remove the linkage if the firm is not identified in the BR within two years of the first VC financing (372 firms)[16]; if the firm reports no economic activity (T2, T4 or PD7 reporting) in the year of or the year following the VC investment (81 firms)[17]; if the firm reports revenues exceeding $50 million at the time of the first VC financing (11 firms); and if the firm suffers from inconsistencies between exit outcomes from the STC and Thomson Data (44 firms). These additional filters are similar to dropping outliers in a statistical analysis as we are dropping linkages that do not make economic sense. After applying these additional checks, we have 1545 VC-backed firms linked to the STC Data.

[16] Firms receiving their first round of VC financing prior to 1999 are removed if they are not present in the BR in 1999 or 2000.

[17] Economic activity is defined as filing a T2, T4 or PD7 payroll remittance form. Firms receiving their first round of VC financing prior to 1999 are removed if they show no signs of economic activity in 1999 and 2000.

Table 12 Linkage of IRAP data to BR

Name	City	Province	Quality flag	Counts	Percentage (%)	Cumulative percentage (%)
Same	Same	Same	D1	5187	67.1	67.1
Same	Same		D2	961	12.4	79.5
Same		Same	D3	6	0.1	79.6
Similar	Same	Same	M1	192	2.5	82.0
Similar		Same	M3	122	1.6	83.6
Unlinked				1268	16.4	100.0
Total				7736	100.0	

The linkage results for the IRAP Data are in Table 12. While it contains less information on the operating addresses of the firms, the IRAP Data have a much higher rate of quality linkage with 80% of the sample reporting an identical name in both the IRAP Data and the BR. This is likely the result of firms having to provide their legal names in applying for IRAP while the Thomson Data is indirectly collected through investing VC funds and limited partners. Overall, 83.6% of 7736 firms or 6468 firms in the IRAP Data are linked to the STC Data.

Appendix 2: Longitudinalization through labor tracking

An identification assigned to an enterprise in the BR, or BRID, is not designed to track a given enterprise over time. In particular, an enterprise's BRID can change for reasons unrelated to any structural change (e.g., a change in the legal name). Conversely, an enterprise could maintain the same BRID, even after a substantial structural change including M&As.

Since we are examining growth metrics, it is important to correctly follow a firm over time (i.e., using longitudinalized data). Accordingly, the labor tracking methodology developed by Statistics Canada is implemented to identify firms over time. This procedure involves following masses of employment (i.e., individuals identified by their SINs in the T4 tax files) moving from one BRID to another. Depending on the nature of these relationships, we amended the respective BRID entries to arrive at a single longitudinal record for each firm within the treatment and control groups using the year they are matched as the base year.

Under the labor tracking, relationships among BRIDs are only identified when one BRID either starts or stops filing the T4 tax information, which we call a T4 birth or T4 death, respectively. Once a T4 death or birth occurs, Statistics Canada determines whether there is sufficient evidence of a relationship with another BRID by examining the size of the firms and the proportion of shared employees among the predecessor and successor BRIDs. The thresholds to determine whether a relationship exists are summarized in Table 13. Approximately, one-third of the BRIDs associated with firms in the treatment and control groups are involved in one or more labor tracking relationships meeting these criteria.

There are three basic types of labor tracking relationships: (1) death-to-birth, (2) death-to-continuer and (3) continuer-to-birth. While specific events in the firm's life cycle cannot be identified with absolute certainty, these relationships as identified by the methodology roughly translate into a false death (i.e., the same firm reporting under a new BRID), an acquisition and a spin-off, respectively. Further, a BRID may be involved in multiple

Table 13 Thresholds for shared employment to identify labor tracking relationships

	Size of T4 birth/death					
	>250 Employees	>50 Employees	>15 Employees	>7 Employees	>5 Employees	5 Employees
Proportion of shared employees	25%	30%	50%	50% if target is a birth/death 60% if target is a continuer	70%	100%

relationships (e.g., two deaths to a birth roughly corresponding to a merger). These relationships can become quite complex when a BRID is involved with many different types of relationships spanning several other BRIDs—a situation not uncommon among BRIDs corresponding to large firms.

For our purpose, we do not want to focus solely on *organic growth* within firms. Our objective is to measure the performance of VC-backed firms, and growth through acquisitions could be part of their strategy to expand the operations. At the same time, if the VC-backed firm is the target of a merger or acquisition, this is likely an exit for the VC investors and ideally the end of the longitudinal record. Regarding spin-offs, it remains unclear whether the new BRID reflects a separate entity or a change in the firm's reporting practices. Accordingly, Statistics Canada connects records when the labor tracking suggests there has been a false death, and ends records when the labor tracking suggests the firm in the treatment or control group is acquired, or there was a substantial spin-off.

Statistics Canada applies a threshold based on 50% of shared employees for evaluating all the labor tracking relationships.[18] In particular, for a potential false death, the predecessor and successor BRIDs are connected into one longitudinal record if the successor BRIDs share 50% or more of the employees of the predecessor. For cases suggesting the firm in the treatment or control group acquiring another firm, the record is ended if the former represents less than 50% of the combined employment in the merged firm. For potential spin-offs, the record is ended if the spin-off exceeds 50% of the employment of the firm in the treatment or control group. Finally, for complex cases where a BRID is involved in many relationships, the 50% threshold is applied cumulatively, ending the record if the total employment associated with all the acquisitions and spin-offs exceeds 50% of the firm's employment, whether in the treatment or control group.

The Thomson Data contains information on VC exits including initial public offerings, M&As and business failures. These data are used to supplement the exits identified through the labor tracking among firms in the treatment group.

References

Arqué-Castells, P. (2012). How venture capitalists spur invention in Spain: Evidence from patent trajectories. *Research Policy, 41*(5), 897–912.

[18] The threshold of 50% to determine whether a firm is the dominant component in the labour tracking relationship is not arbitrary and based on discussions between analysts from Statistics Canada and Innovation, Science and Economic Development Canada. In particular, thresholds based on figures greater than 50% are considered (e.g., 80%), and these are deemed to be too restrictive as firms routinely experience changes in employees through regular employee churns that account for more than 20% of their employees.

Arvanitis, S., & Stucki, T. (2014). The impact of venture capital on the persistence of innovation activities of start-ups. *Small Business Economics, 42*(4), 849–870.

Audretsch, D. B., Bönte, W., & Mahagaonkar, P. (2012). Financial signaling by innovative nascent ventures: The relevance of patents and prototypes. *Research Policy, 41*(8), 1407–1421.

Bertoni, F., Colombo, M. G., & Grilli, L. (2011). Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. *Research Policy, 40*(7), 1028–1043.

Brander, J. A., Egan, E., & Hellmann, T. F. (2010). Government Sponsored versus Private Venture Capital: Canadian Evidence. In J. Lerner & A. Schoar (Eds.), *International differences in entrepreneurship* (pp. 275–320). Chicago, IL: University of Chicago Press.

Canadian Venture Capital Association. (2013). *The performance of Canadian firms that received venture capital financing*. Retrieved from http://www.cvca.ca/wp-content/uploads/2014/07/VC_Study_Final_English_September_4_2013.pdf.

Colombo, M. G., D'Adda, D., & Pirelli, L. H. (2016). The participation of new technology-based firms in EU-funded R&D partnerships: The role of venture capital. *Research Policy, 45*(2), 361–375.

Cumming, D. J., & MacIntosh, J. G. (2004). Canadian labour-sponsored venture capital corporations: Bane or boon? In A. Ginsberg & I. Hasan (Eds.), *New venture investment: Choices and consequences* (pp. 169–200). Amsterdam: Elsevier Science Academic Press.

Cumming, D. J., & MacIntosh, J. G. (2006). Crowding out private equity: Canadian evidence. *Journal of Business Venturing, 21*(5), 569–609.

Cumming, D. J., & MacIntosh, J. G. (2007). Mutual funds that invest in private equity? An analysis of labour-sponsored investment funds. *Cambridge Journal of Economics, 31*(3), 445–487.

Da Rin, M., & Penas, M. F. (2007). *The effect of venture capital on innovation strategies*. NBER Working Paper No. 13636.

Davis, S. J., Haltiwanger, J. C., & Schuh, S. (1998). *Job creation and destruction*. Cambridge, MA: MIT Press.

Engel, D., & Keilbach, M. (2007). Firm-level implications of early stage venture capital investment—An empirical investigation. *Journal of Empirical Finance, 14*(2), 150–167.

Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979–1995. *RAND Journal of Economics, 32*(1), 101–128.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics, 95*(2), 347–361.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies, 64*(4), 605–654.

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies, 65*(2), 261–294.

Hellmann, T., & Puri, M. (2000). The interaction between product market and financing strategy: The role of venture capital. *Review of Financial Studies, 13*(4), 959–984.

Hellmann, T., & Puri, M. (2002). Venture capital and the professionalization of start-up firms: Empirical evidence. *Journal of Finance, 57*(1), 169–197.

Hoenen, S., Kolympiris, C., Schoenmakers, W., & Kalaitzandonakes, N. (2014). The diminishing signaling value of patents between early rounds of venture capital financing. *Research Policy, 43*(6), 956–989.

Jain, B. A., & Kini, O. (1995). Venture capitalist participation and the post-issue operating performance of IPO firms. *Managerial and Decision Economics, 16*(6), 593–606.

Kortum, S., & Lerner, J. (1999). What is behind the recent surge in patenting? *Research Policy, 28*(1), 1–22.

Kortum, S., & Lerner, J. (2000). Assessing the contribution of venture capital to innovation. *RAND Journal of Economics, 31*(4), 674–692.

Lahr, H., & Mina, A. (2016). Venture capital investments and the technological performance of portfolio firms. *Research Policy, 45*(1), 303–318.

Lerner, J. (2009). *Boulevard of broken dreams: Why public efforts to boost entrepreneurship and venture capital have failed—and What to do about it*. Princeton, NJ: Princeton University Press.

Manigart, S., & Wright, M. (2013). Venture capital investors and portfolio firms. *Foundations and Trends in Entrepreneurship, 9*(4–5), 365–570.

Mowery, D. C., Nelson, R. R., Sampat, B. N., & Ziedonis, A. A. (2001). The growth of patenting and licensing by U.S. universities: An assessment of the effects of the Bayh-Dole act of 1980. *Research Policy, 30*(1), 99–119.

Puri, M., & Zarutskie, R. (2012). On the life cycle dynamics of venture-capital- and non-venture-capital-financed firms. *Journal of Finance, 67*(6), 2247–2293.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*(1), 33–38.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 2*(3), 169–188.

Samila, S., & Sorenson, O. (2010). Venture capital as a catalyst to commercialization. *Research Policy, 39*(10), 1348–1360.

Samila, S., & Sorenson, O. (2011). Venture capital, entrepreneurship, and economic growth. *Review of Economics and Statistics, 93*(1), 338–349.