Sant'Anna School of Advanced Studies, Institute of Management Paper Series_ 2024 Winter Issue

Artificial Intelligence in Venture Capital operations: an empirical analysis

Umberto Ronco
Sant'Anna School of Advanced Studies

Roberto Barontini*[1]
Sant'Anna School of Advanced Studies

Abstract

This paper explores the adoption and perceived impact of artificial intelligence in venture capital (VC) firms. Although AI adoption in financial services is increasing, its integration within VC firms remains largely underexplored, in particular as regards its application in business operations.

Based on data obtained from a questionnaire distributed among European venture capitalists, it is observed that the adoption of artificial intelligence has increased markedly since 2022, with screening emerging as the most prevalent application. Statistical models suggest that firms with employees with strong ICT backgrounds are more likely to adopt AI. Although AI reduces due diligence time, its overall effect on long-term benefits remains inconclusive, perhaps due to the limited available data.

Keywords: Venture capital, Artificial intelligence, Technological change, Start-ups, Entrepreneurial finance.

[1] **Corresponding author**

1. Introduction

The diffusion of artificial intelligence (AI) is radically reshaping business operations and creates the opportunity to increase the efficiency of processes in industrial and service activities (Ransbotham, 2017; Rao and Greenstein, 2022).

This study focusses on the application of AI in venture capital (VC) operations, an issue that is largely neglected in the literature. In fact, although VC investments toward AI-related start-ups surge (Ernst & Young, 2024; Struta, 2024), it is not clear whether VCs use AI internally and whether this has a positive impact on their activity. This is particularly interesting because VC operations imply qualitative assessment of teams and business models, and investments are based on negotiations, so they represent a very challenging area of application for artificial intelligence.

This gap in the literature may be related to different reasons. First of all, and this is also a great limitation of this work, it is very difficult to collect relevant data on the issue. Indeed, the current gold standard for collecting data is through interviews or questionnaires, and this methodology is problematic due to the scarcity of venture capitalists (especially in Europe) and to the low response rates, since venture capitalists do not disclose light-heartedly private information. There are also limitations regarding the use of the questionnaire itself, since responses could be biased due to self-selection issues, and also qualitative data based on self-assessment can be biased as well.

Despite these considerations, the main source of data is a questionnaire, distributed to 458 European venture capitalists, containing the most relevant information on the adoption of AI and the benefits or the difficulties related to its use in VCs business activity.

The first aim is to try to understand whether VCs are adopting AI tools in their internal operations. The second is to shed some light on the use of these technologies, extending previous literature, such as Arroyo et al. (2019), Gao (2023), Maurer et al. (2024), Setty et al. (2024) and Wang and Xue (2022), which tried to support the use of AI algorithm for screening start-ups. Another aim of this work is to look at the data in a quantitative way, trying to identify patterns that lead to the adoption of AI and to evaluate whether AI affects the efficiency of VC funds.

This study is structured as follows. Section 1 analyzes the relevant literature and develops the research questions. Then, in Section 2 the methodologies are explained, from the design of the questionnaire to the data integration process and the econometric methodologies. Subsequently, Section 3 includes qualitative and quantitative analyses. After that, Section 4 discusses research limitations and future research opportunities.

2. Literature Review and Hypotheses Development

Venture capital is a subset of the private equity universe, and it consists of "independently managed, dedicated capital focusing on equity or equity-linked investments in privately held, high-growth companies" (Hall and Lerner, 2010). The targets of venture capital are start-ups, young and innovative companies with relevant financial needs, necessary to support firm growth. Besides that, founders of start-ups often lack managerial background, and therefore they also look for advice, mentorship, and network (Amit et al., 1998). Venture capitalists are able to provide founders with these different resources and, therefore, have been addressed as a driver of innovation that could benefit the entire economy (Mollica and Zingales, 2007; Sun et al. 2020).

The main issue regarding venture capital funds is the high mortality of the start-ups they invest in. That happens, of course, because the firms backed by venture capital are supposed to be highly innovative, regarding products with little or nonexisting markets.

The selection of the most promising projects is considered to be one of the most important parts of the VC activity (Andreoli and Rouwelaar, 2024; Nadeau, 2010; Streletzki and Schulte, 2013), and therefore it is not surprising that artificial intelligence has been proposed as a useful tool for this task.

It is not easy to define Artificial Intelligence (Sheikh et al., 2023; Mueller, 2023; Schmidt, 2019), however the most widespread ones fundamentally concern AI as developing computers that can "understand, think, and learn" in a human-like manner, highlighting different skills, such as learning from experience, adaptability to new data, and the development of autonomous algorithms (Duan et al., 2019). AI in finance has contributed, since its first diffusion, to the rise of productivity, supporting the decision-making process, and increasing efficiency (Rao and Greenstein, 2022).

Concentrating on the literature considering the use of artificial intelligence in venture capital operations, we selected relevant works from Scopus and then integrated results with the search on Google Scholar.

Searching the string *"venture AND capital AND AI AND operations"* among all fields, we found a pool of 1,662 documents. Subsequently, a time filter was selected, as the technology we are referring to and its diffusion is quite recent, so papers published after 2014 were selected, with 1,581 remaining documents. The next step of the selection of the relevant papers was based on the subject areas (*"Business, Management and Accounting"* and *"Economics, Econometrics and Finance"*), that led to 1,004 papers. The last screening regarded 47 keywords, which were considered inherent, related to innovation, venture capital, artificial intelligence, technology transfer, and entrepreneurship. At the end of this process, 374 relevant papers emerged. All remaining works have been briefly evaluated and only 15 that were related to the purpose of this paper have been analyzed in detail (the results are included in appendix, Table A1).

First of all, Mueller (2023) conducted a series of interviews with private equity (PE) and AI experts. This work focuses on the deployment of AI in portfolio companies of the PE fund; in particular, findings suggest that PE can effectively accompany portfolio companies to adopt AI. Another interesting finding, which is also more important for the scope of this work, concerns the internal use of AI by private equity funds. In fact, the interviews conducted with PE managers showed that, while the majority did not use AI (around 85%), most of them (around 83%) showed interest in its adoption and recognize this technology as a key driver of efficiency improvement.

Sanchez (2020) interviewed different senior managers of PE and VC funds, focusing on the importance of AI in decision-making and the likelihood of its short-term adoption. Results show that, even though at the time of the interview (2020 or earlier) none of the respondents had already adopted AI internally, most of them thought that in a 3-years horizon it would have been possible.

Following these works, the first research question is:

RQ 1: What is the current rate of AI-adoption among VCs and has there been a significant increase in adoption of AI tools by VC firms?

Jain (2018) analyzes the decision-making process of VC firms in relation with the feasibility of AI integration within each of the six steps of the decision-making process, finding that deal selection is the phase most suited for AI support, followed by valuations, post-investment value-added, deal sourcing, exits, and, last, deal structure.

Chalmers et al. (2021) focus on the efficiency boost of AI for VC investments, and, partially, include directly VC operations. The authors highlight that the proliferation of start-ups makes it more useful for VCs the adoption of artificial intelligence, supporting the screening and the decision-making phase.

Schmidt (2019) interviewed 12 VC managers and AI experts who believed that – although artificial intelligence was used primarily for the selection of deals – it could be implemented at every step of venture capital activity. The study predicts that the use of AI can improve decision making by, on the one hand, lowering uncertainty and bias and, on the other, boosting productivity and efficiency through a change in their business processes. Consistently, Sanchez (2020) finds that experts think that the adoption would be mostly beneficial for the lengthy due diligence phase and data gathering and analysis.

Thus, in this paper we aim to understand whether it is confirmed that the screening phase is actually the most fit for a proper AI adoption and implementation. So, it follows:

RQ 2: Is the Screening phase the one in which AI is used the most and the one in which AI is more useful?

Then, our objective is to empirically identify some predictive factors of the adoption of AI tools, since expectations are not clear. For example, it can be argued that seed or pre-seed investors could be more inclined to adopt AI since they are usually investing in the most advanced technologies. On the other hand, it can be argued that pre-seed and seed investing, precisely because it regards the most advanced technologies, may be unsuited for tools that rely on past data such as machine learning (Bonelli, 2022).

The main factor which could be expected to have a positive impact on the adoption of AI relates to the background of the employee (as will be explained, this is referred to the time before the adoption of AI) for what regards ICT skills and knowledge. In fact, it is straightforward to think that the knowledge of these technologies makes them more likely to be adopted.

RQ 3: Is the engineering background of the employees a positive predictor of AI adoption?

Expectations on the effects of the adoption of AI tools on VC firms' operations are also mixed. Bonelli (2022) emphasises some problematic aspects especially for seed or pre-seed investments, pointing out that technologies such as machine learning rely on past data to discover patterns and make predictions, so they may not be appropriate for VC firms, that must invest in the frontier of technology to be competitive on the market. The author through a quantitative analysis demonstrates that a machine learning approach, trying to avoid failures, misses also the best opportunities (and it is well known that without these deals VC is not a profitable asset class).

On the other hand, Fernández Tamayo et al. (2023) argue that machine learning can predict the success of private equity funds because of its ability to avoid biases. In particular, when deciding which fund to invest in, humans often rely on quantitative data, such as past performance, which are shown to be completely unrelated to the future performance. This prediction can be improved, instead, by looking at qualitative data that may be able to capture a significant spread in returns. This approach of looking at qualitative features, both of people and companies, could overcome the issues presented by Bonelli (2022), because the ability to capture qualitative information through AI applications may be the driver of success in the VC selection process.

In this strand of the literature lies Brandt and Stefansson (2018). The authors, by comparing the findings obtained through interviews with venture capitalists, assert that the personality of founders can be labeled and analyzed by specific AI tools in an effective way. Similarly, Lyonnet and Stern (2022) used ML algorithms to detect qualitative features in a sample of start-ups and of their founders. They find out that VCs overestimate some representative features of success, do not invest in predictable successes and, on the contrary, invest in predictable failures. Most of the biases exposed in this work regard the gender of the founders, geographic location of the start-up and background of the founders. They infer that algorithmic decision support may be used to broaden the scope of VC investments and the diversity of the founders, overcoming interiorized biases. Other papers are consistent with this conclusion (Gao, 2023; Arroyo et al., 2019; Maurer et al., 2024; Wang and Xue, 2022; Setty et al., 2024), supporting the possibility of conducting an effective screening using AI.

Some studies lie in the middle of these views. Hu (2024) proposes an AI model to screen start-ups and, finding out that it can be an improvement compared to the current rules of thumb used by venture capitalists, but there are also clear doubts concerning the current applicability of AI for these kind of tasks. In particular, the author raises two different issues: the first concerns the availability and quality of data, the second the technology itself, as anyone who handles it should be fully aware of its weaknesses, limitations and scope of application.

From the previous discussion it emerges that most of the existing literature believe that the adoption of AI tools could be beneficial for VCs, boosting efficiency and effectiveness of VC operations, but the evidence is still inconclusive. This brings us to the last research question of this work.

RQ 4: Does the adoption of AI tools positively affect VC firm operations?

3. Methodology and data

3.1 The dataset

Data used for this work are based on a questionnaire submitted to European venture capitalists and then integrated with additional information.

A questionnaire has been designed and sent to 458 venture capitalists, receiving 31 responses, with a response rate of 6.8%. The questionnaire is reported in the Appendix, Table A2.

The questionnaire has been developed following some principles: firstly, it must contain relevant information not publicly available; then, it must be as fast and simple as possible, in order to incentivize responses; lastly, it should be clear and flexible (questions have been constructed to have mostly multiple answers or predefined options).

The questionnaire consists of eight brief sections. The first section contains information for the identification of the respondent. The second section asks for general information about the fund, such as year of establishment and fund asset under management (AUM). The third section regards investment segment, location, and the stage of invested firms. Every question in this section requires answers in percentage terms, recalling the distribution of the parameters (location, stage, or sector) within VC activity.

The fourth section is composed of only one question that asks if the respondent employs AI in fund operations. From this question, the path of AI users and not AI users is divided. The fifth section is for AI adopters only and regards the use of AI in the fund. So, the questions are about the year of adoptions, the kind of AI employed, and the phases in which AI is used. This section is fundamental for every analysis that follows because it details the use of AI beyond the mere adoption rate.

The sixth section is, instead, for non-AI-users. In fact, in this section the only question is on the reason why respondents do not use AI.

The seventh section of the questionnaire is for AI users only and qualitatively asks if the adoption of AI helps improve operational efficiency. For efficiency, two measures are considered: the duration of due diligence time and the number of investments per year. Also, in which phase AI has proved to be most helpful is asked.

Lastly, the eighth section includes questions addressed to both AI-users and non-AI-users. These questions concern the quantitative relation between artificial intelligence and the two previous proxies for efficiency analyzed: due diligence time duration and investments per year. These data are going to represent the dependent variable in the econometric analysis regarding the effects of AI adoption.

After collecting the questionnaires, the data integration process began. This process is necessary because the questionnaire asks only inside information, so publicly available data that may be relevant for the research had to be integrated. Some variables may be a good predictive factor for AI adoption; other capture characteristics that are interesting to consider as control variables.

A variable that may represent a good predictor of AI adoption is the academic or professional background of VC employees and managers, since human capital is needed to effectively adopt such new technologies, and, furthermore, employees more interested in AI make it more likely the adoption of the technology. This variable will be considered the independent variable in the analysis of Section 3.2, therefore, information on the background of VC employees and managers was the first to be added to the dataset.

Of course, to check for a possible causal relationship, this information should be referred to a period prior to the adoption of AI. Since in the questionnaire the year of AI adoption is requested, in our dataset the variable named "ICT Background" has been added, which measures the percentage of VC employees who hold a background that assumes a deep knowledge of ICT technology, such as engineering, in the year before the adoption. The selection was made manually, considering the LinkedIn profiles of the VC employees.

Another important factor that is not present in the questionnaire but may contain relevant information is the location of the fund.

3.2. Methodology

The empirical analysis is divided into two parts and applies different approaches. The first one (Section 3.1) focuses on the data obtained through the questionnaire in a descriptive and qualitative way. Instead, the second one (Section 3.2 and 3.3) develops econometric models based on the available data.

In particular, Section 3.2 builds some models with a binary dependent variable, AI adoption, which captures whether the respondent uses AI tools (AI adoption=1) or not (AI adoption=0). As mentioned above, a causal relation could be expected between the educational and work background of VC employees and the adoption of AI tools, as a higher percentage of employees with deep knowledge of ICT could increase the probability of the adoption of AI tools.

Control variables, due to constraints regarding the actual size of the dataset, have been included in the models just one at a time. This approach is used throughout the quantitative analysis, with the aim of reducing multicollinearity and overfitting issues that could emerge including all the control variables contemporarily.

The standard regression model used for binary dependent variables is logistic regression (Hosmer et al., 2013). As robustness tests, we also estimated probit models (Hsiao, 1996), Naive Bayes methods (Friedman et al., 1997), and the K-Nearest Neighbors (Cover and Hart, 1967). The last two methodologies are selected because they work properly with few observations (Naive Bayes methods) and perform nonparametric estimations that can capture to more easily complex relationships (K-Nearest Neighbors).

Section 3.3 has the aim of answering RQ 4 - that is, whether the adoption of AI tools positively affects VC firm operations – from a quantitative perspective. In this case, the dependent variables, Due Diligence Duration and Number of Investments, are expressed, respectively, in the number of weeks and the number of investments per year, while the independent variable is represented by AI adoption, a dummy variable that takes value = 1 if the respondent adopts AI tools, 0 otherwise. The models involved in the analysis are linear models, due to their simplicity and the small number of observations in the data-set, and also because we do not find in the literature any theoretical reason supporting the use of non-linear models.

So, the equations are:

$$Due\ Diligence\ Duration = \beta_1 * AI\ adoption + \beta_2 * Controls + \epsilon$$

and

$$Number\ of\ Investments = \beta_1 * AI\ adoption + \beta_2 * Controls + \epsilon$$

where Controls are variables built from the questionnaire, respectively:

- Asset under management (AUM), a proxy of the size of the VC fund.
- Investment Headquarter, a dummy variable which takes value 1 if the fund comes

from Italy and 0 otherwise.

- Investment Location, a dummy variable, which takes value 1 if the fund makes at least 70% of its investments in Italy and 0 otherwise.
- Investment Stage, a categorical variable that divides respondents among 4 categories, *Early Stage investors* (more than half of investments made in Pre-Seed or Seed phases), *Series A investors* (more than half of investments made in Series A phase), *Series B or later investors* (more than half of investments made in Series B or later phases) and *Stage Agnostic investors* (none of the above criteria is met).
- Investment Sector: it is a categorical variable capturing the sectoral specialization of the fund.

4. Results and analysis

4.1. Qualitative analysis

First of all, it is interesting to look at the distribution of the stage of investments highlighted by the respondents. Figure 1 shows that the most common stages are Seed and, especially, Series A, therefore perfectly matching what is considered to be the *core* of venture capital activity. On the other hand, there are a few respondents that invest consistently at a pre-seed stage or at a Series B or later stage.

Next, Figure 2 displays the distribution of the sectors in which the respondents invest. The first thing to observe is that AI technology is particularly targeted by VCs, since it is the sector with the lowest number of non-investors in the sample. It is also interesting to note that four respondents have 100% of their investments in one sector only, highlighting that specialization within a unique sector is quite rare.

Figure 3 shows the geographical distribution of investments, that highlights a substantial concentration in Italy: 10 (or 23) respondents out of 31 locate 100% (at least 50%) of their investments in Italy. Furthermore, it can be observed that there is a strong centrality of Europe, as could be expected considering the location of venture capitalists.

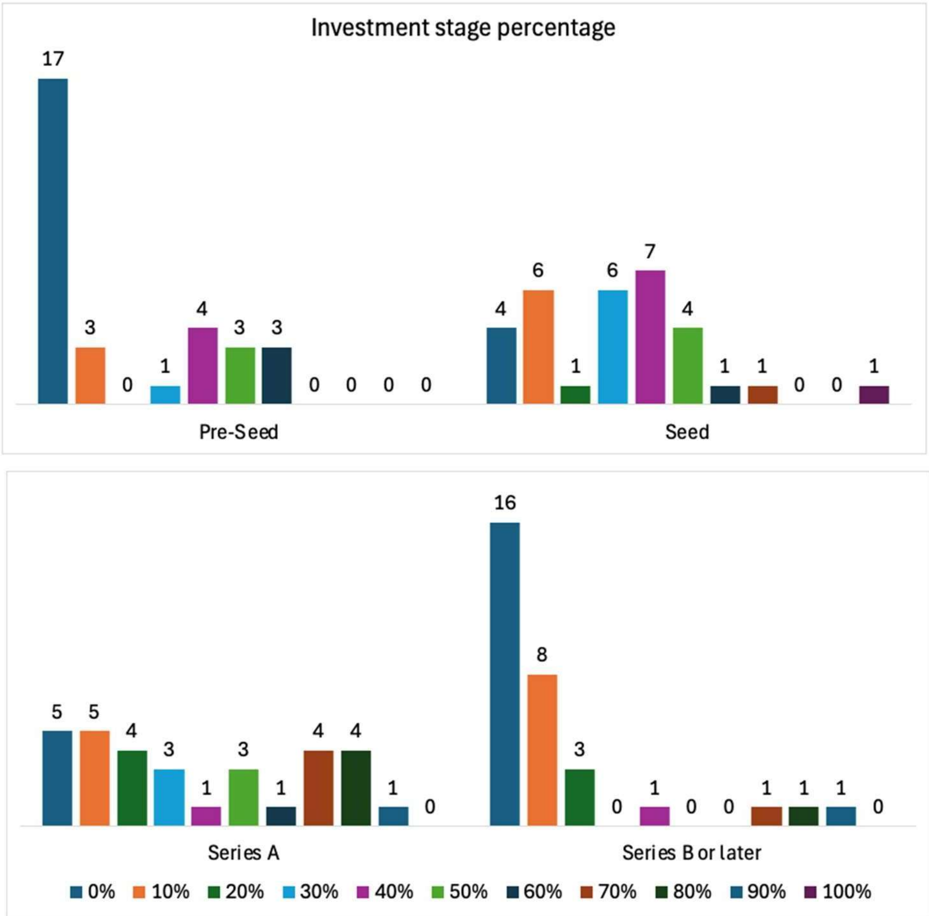


Figure 1: Investment stage percentage

Now, let us look at the data regarding the adoption of AI tools. Figure 4a shows that 71% of the respondents currently use AI tools, confirming the predictions of some authors. In fact, while older empirical studies showed little or no internal use of AI by PE and VC funds (Sanchez, 2020; Schmidt, 2019), the professionals interviewed predicted an increase in future adoption, which has been confirmed by our data.

In addition, Figure 4b refers to the different types of AI tools that have been adopted. The greater part of AI adopters uses Natural Language Processing and Generative AI which, of course, have become very common in the last couple of years, thanks to the large diffusion of tools such as ChatGPT and its API.

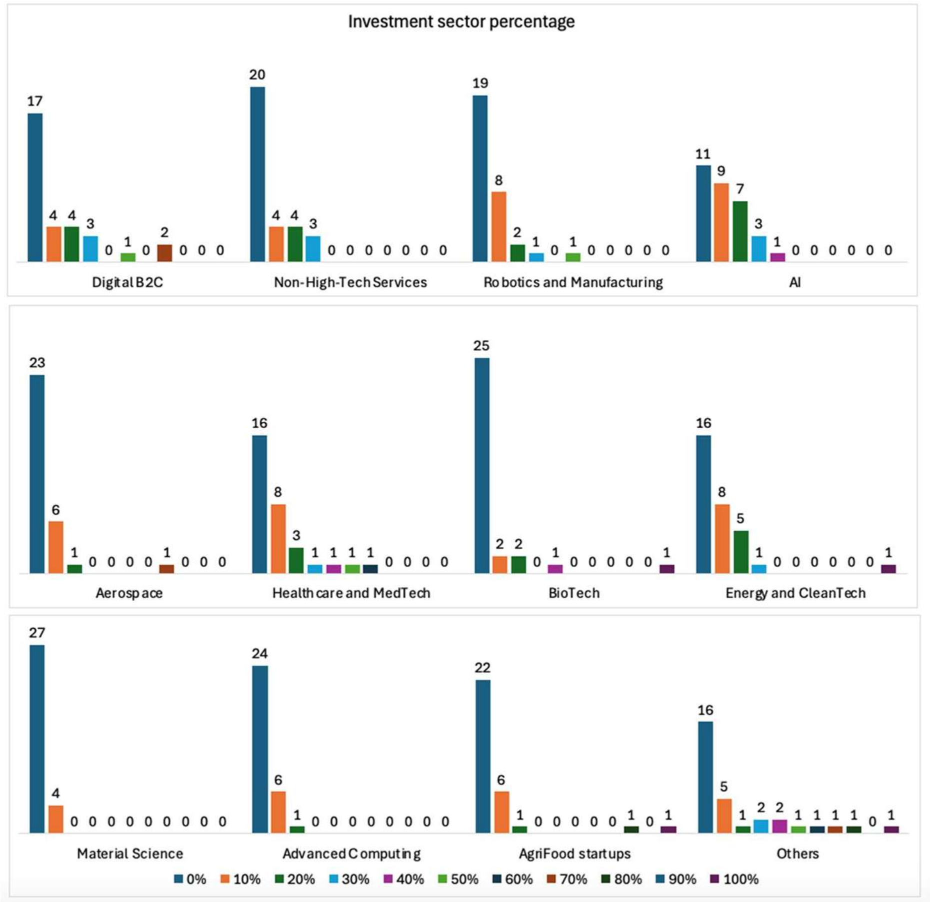


Figure 2: Investment sector percentage

Another confirmation of the prediction of the literature (Sanchez, 2020; Schmidt, 2019) can be found in Figure 5a, which shows that 21 (or 19) of the 22 respondents adopted AI in their internal processes in 2022 or later (or in 2023 or later). The results show huge progress in the adoption of artificial intelligence in recent years and clearly confirm RQ 1.

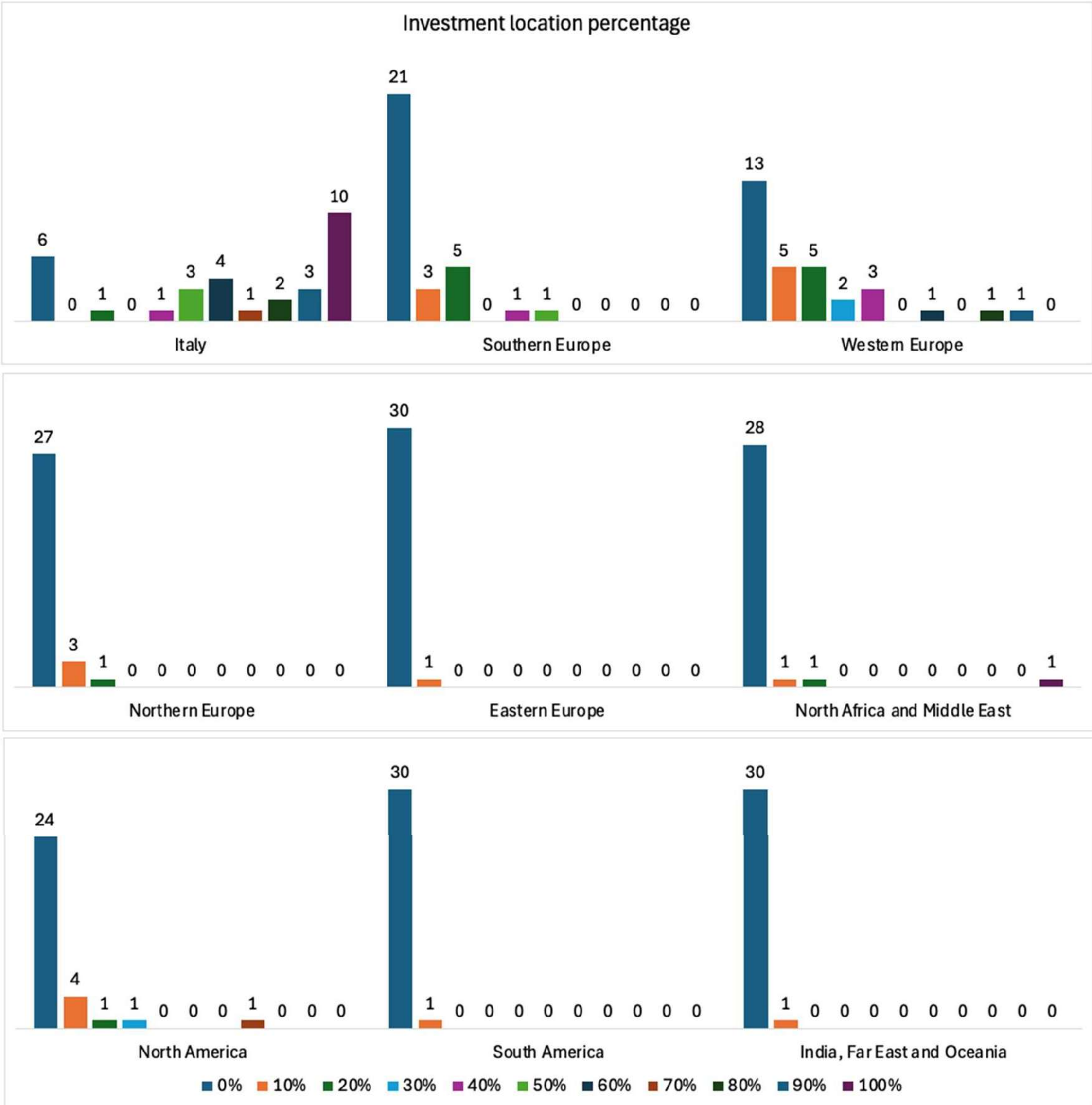
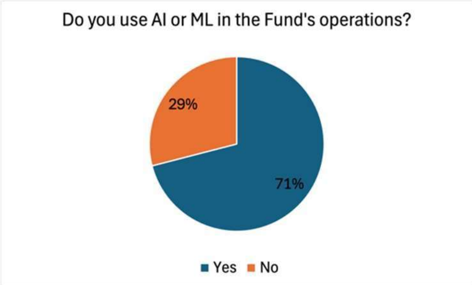
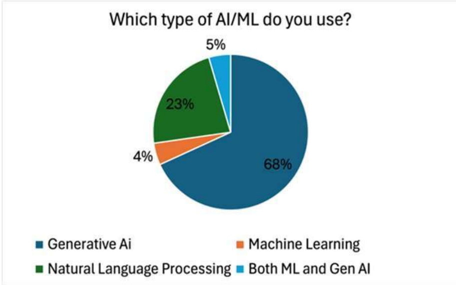


Figure 3: Investment location, percentage on total investments



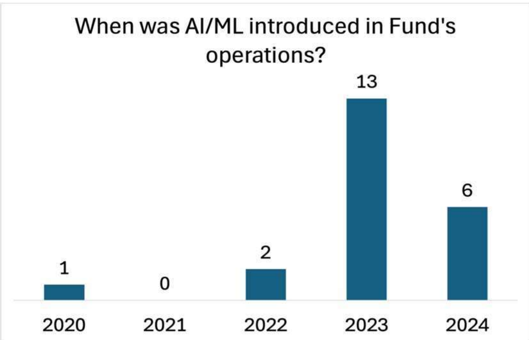
(a) AI adoption rate



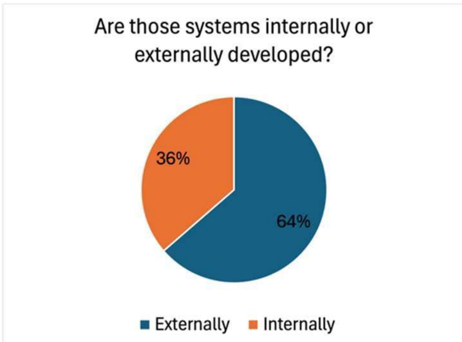
(b) Types of AI

Figure 4: Adoption rate and types of AI used among respondents

Figure 5b shows an interesting result, concerning the internal or external development of the AI tools used by venture capitalists. As expected, the majority of respondents relies on externally developed AI tools, however a relevant 36% of the sample affirm that they use an internally developed AI technology, showing a significant effort.



(a) Year of introduction of AI



(b) AI development

Figure 5: Internal or external development of adopted AI tools and year of AI adoption

Figure 6 tries to answer to the first part of RQ 2, regarding the phases in which AI is used more frequently by VCs. As predicted by the literature, the largest number of respondents use AI to screen startups (19 AI-adopters out of 22, i.e. 86%), followed by due diligence (with a remarkable frequency of 73%), internal organization of the fund (68%) and data analysis (64%). The other phases follow behind, with deal structuring receiving zero positive responses.

Data provided therefore confirm the RQ 2, though they show a notable frequency of adoption of artificial intelligence in many phases involving complex administrative activities. The most "high-level" investment phases, deal structuring startup valuation and the final investment decision are the ones in which artificial intelligence is actually less easily applicable.

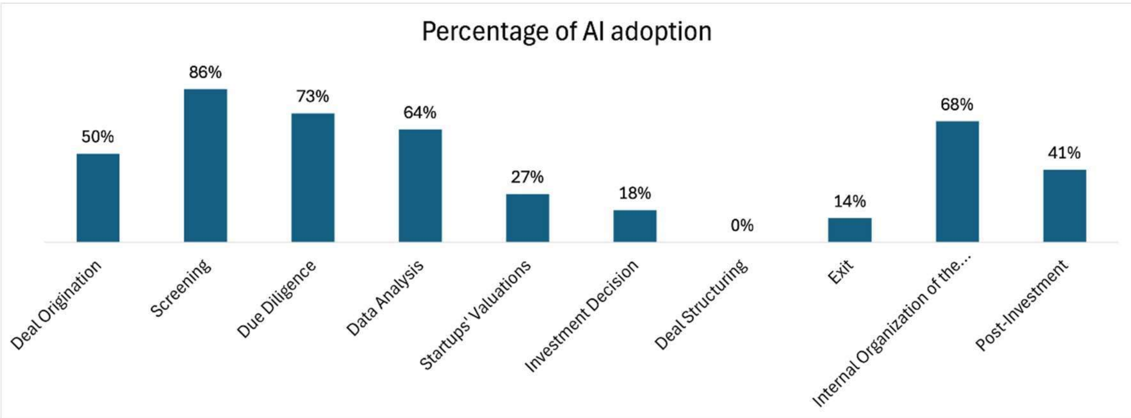
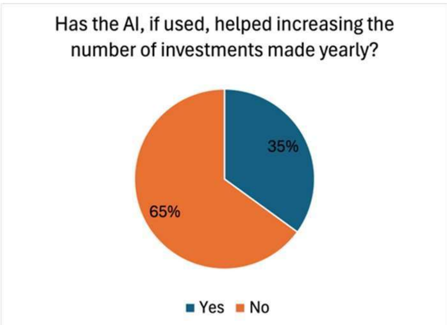


Figure 6: Phases in which AI tools are used

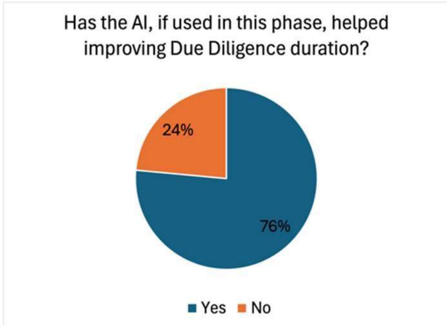
Next figures show the perceived impact of AI tools on the number of investments made yearly (Figure 7a) and the due diligence duration (Figure 7b).

The number of investments can be considered as a proxy of internal efficiency because, *ceteris paribus*, more efficient processes should lead to an increase in deals screened, analyzed, and then also in investments. Figure 7a shows that 35% of respondents state that AI helped to increase the number of investments made yearly, while the other 65% does not. May be that this increase does not entirely catch the benefits of the AI, due to its recent adoption by almost all respondents, however even this result could be considered a notable outcome, considering the high selectivity nature of venture capital activity.

Figure 7b exhibits the answers about whether AI helped lowering the due diligence duration. Data show that a considerable 76% of AI-adopters affirm that AI helps in lowering the time needed for the due diligence phase, supporting the hypothesis that AI improves operational efficiency of VC firms, in particular in most complex and bureaucratic administrative activities.



(a) AI help on number of investments



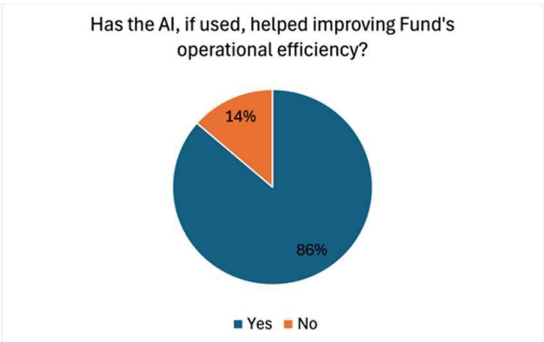
(b) AI help on due diligence duration

Figure 7: AI correlation with the increase in the number of investments and improvement in due diligence duration

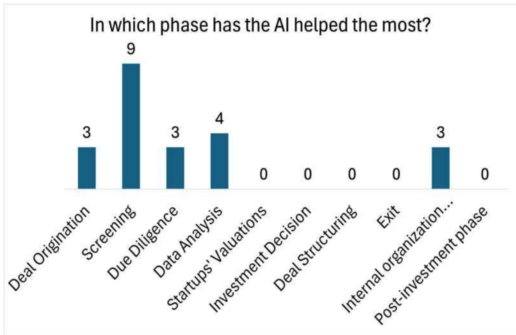
Concluding the qualitative analysis on the perceived effects of AI, Figure 8a refers to the question of whether AI improves the operational efficiency of the fund, while Figure 8b highlights the phase in which AI is more useful for VC firms.

Looking at Figure 8a, it can be seen that 86% of the respondents support the positive effect of AI on the operational efficiency of the VC firm. This result clearly confirms RQ 4, even if in a very general and qualitative way, based on the mere perception of the respondents.

On the other hand, Figure 8b completes the answer to RQ 2, showing the phases of VC operations in which AI tools are considered more useful by adopters. As argued in Section 1, according to the literature the screening phase is expected to be the one in which AI is more useful. Results confirm that the screening phase is the one in which AI is useful the most, followed by data analysis, deal origination, due diligence and internal organization of the fund.



(a) AI help on operational efficiency



(b) Phases where AI helped the most

Figure 8: AI help on the improvement of number of investments and due diligence duration

Lastly, a question is asked to non-AI-adopters which regards the reason why they do not adopt AI tools. The responses, given as a brief text, generally highlight that no useful VC-specific tool has been found, while one respondent affirms that it is due to lack of resources, and one affirms that they are working on adopting AI tools. These results confirm, clearly, the intention of VCs to explore in the future the opportunity to adopt this technology.

4.2. Determinants of the adoption of AI

The results of the logistic model are shown in Table 1, where the coefficients of the independent variable are positive and significant at a level of 5%. Therefore, our findings support the hypothesis RQ 3 that a higher percentage of employees with deep knowledge of ICT increases the likelihood of adoption of AI.

These results are confirmed in every specification. It is important to note that a model with all control variables could not be estimated, due to the low number of observations; however, model (1), with no control variables, according to the Akaike Information Criterion is the second most informative model (besides model 2, with Assets Under Management control). We should not expect a strong impact of omitted control variables on the results obtained for the main independent variable.

Table 1: Logistic Regressions Summary for AI Adoption

Dependent variable: AI Adoption						
	(1)	(2)	(3)	(4)	(5)	(6)
ICT Background	8.841**	10.209**	9.050**	9.159**	8.262**	13.678**
	(3.674)	(4.276)	(3.740)	(3.749)	(3.809)	(6.272)
AUM Control	No	Yes	No	No	No	No
Investor Headquarter Control	No	No	Yes	No	No	No
Investment Location Control	No	No	No	Yes	No	No
Investment Stage Control	No	No	No	No	Yes	No
Investment Sector Control	No	No	No	No	No	Yes

Observations	31	31	31	31	31	31
Akaike Inf. Crit.	31.954	30.941	33.714	33.454	33.773	36.673
Note:	*p<0.1; **p<0.05; ***p<0.01					

Next, Table 2 shows the results of the probit models. Also in this case, the coefficient of the independent variable is positive and significant at least on a 5% level for every different model (for the models with Investor Headquarter control, the Assets Under Management control and the Investment Location control this coefficient is significant at a 1% level). Also in this case, the base model, with no control, has the second lowest AIC, while the model with AUM control has the lowest one, confirming results of in the logistic regressions.

Table 2: Probit Regressions Summary for AI Adoption

Dependent variable: AI Adoption						
	(1)	(2)	(3)	(4)	(5)	(6)
ICT Background	5.250**	5.801***	5.407***	5.415***	5.109**	7.652**
	(2.048)	(2.225)	(2.081)	(2.060)	(2.172)	(3.323)
AUM Control	No	Yes	No	No	No	No
Investor Headquarter Control	No	No	Yes	No	No	No
Investment Location Control	No	No	No	Yes	No	No
Investment Stage Control	No	No	No	No	Yes	No
Investment Sector Control	No	No	No	No	No	Yes
Observations	31	31	31	31	31	31
Akaike Inf. Crit.	31.757	30.890	33.488	33.342	33.443	36.578
Note:	*p<0.1; **p<0.05; ***p<0.01					

The models built with the Naive Bayes method are shown in Table 3. This methodology, as explained in Section 2.2, provides a prediction of the dependent variable, *AI adoption* in this case, which is then evaluated for its accuracy. The structure of the controls is the same as in the previous models, so in the first one, there are no controls, while in the other there is one control per model. It is to be pointed out that the first model, without controls, has an accuracy of 77.42%, with just 2 models having a higher accuracy. These models are the ones with the Investment Stage control, with an accuracy of 80.65%, and with the Investment Sector control, with an accuracy of 83.87%. It is noticeable, though, that *ICT Background* captures the greatest part of the accuracy.

Table 3: Naive Bayes accuracy for AI Adoption

Predicted variable: AI Adoption	(1)	(2)	(3)	(4)	(5)	(6)
ICT Background	Yes	Yes	Yes	Yes	Yes	Yes
AUM Control	No	Yes	No	No	No	No
Investor Headquarter Control	No	No	Yes	No	No	No
Investment Location Control	No	No	No	Yes	No	No
Investment Stage Control	No	No	No	No	Yes	No
Investment Sector Control	No	No	No	No	No	Yes
Observations	31	31	31	31	31	31
Accuracy	77.42%	77.42%	77.42%	77.42%	80.65%	83.87%

In table 4 are shown the results regarding the K-Nearest Neighbors models. As in the previous case, this model predicts the values of a set of test observations, and then the models' performance is measured through the accuracy of these predictions. The dataset has been divided into train set and test set in the proportion of 70% and 30%. The results presented in Table 4 show that all the accuracy is captured by the independent variable, as the accuracy is the same in the model with no control as in the other models with the controls.

In summary, empirical evidence supports the hypothesis according to which a higher percentage of employees with a deep knowledge of ICT raises the probability to observe an adoption of AI tools by the VC firms. This happens both in regression models such as the logistic and the probit and in prediction models such as Naive Bayes and K-Nearest Neighbors.

Table 4: K-Nearest Neighbors accuracy for AI Adoption, K=5

Predicted variable: AI Adoption						
	(1)	(2)	(3)	(4)	(5)	(6)
ICT Background	Yes	Yes	Yes	Yes	Yes	Yes
AUM Control	No	Yes	No	No	No	No
Investor Headquarter Control	No	No	Yes	No	No	No
Investment Location Control	No	No	No	Yes	No	No
Investment Stage Control	No	No	No	No	Yes	No
Investment Sector Control	No	No	No	No	No	Yes
Observations	31	31	31	31	31	31
Accuracy	90.0%	90.0%	90.0%	90.0%	90.0%	90.0%

4.3. Effects of AI adoptions on firms' operations

In Section 3.1 the research question **RQ 4**, i.e. whether the adoption of AI tools positively affects VC firm operations, has been analyzed in a qualitative way, taking into account the results obtained with the questionnaire. The aim of this section is to look at the same problem from a quantitative point of view, estimating regressions where the dependent variables are *Due Diligence Duration* and *Number of Investments*. It must be noted that, while Due Diligence Duration captures the improvement in efficiency just for what regards one phase of the internal process, the Number of Investment variable is able to catch all the improvement in efficiency, at the end of the process.

The first analysis regards the t-test conducted on the dependent variables, (Table 5) used to check whether the difference of the means of the dependent variables is significant, when grouping the observations according to their value of the independent variable. So, in other words, for both *Due Diligence Duration* and *Number of investments*, the observations have been divided into two groups: one group of the variable *Due Diligence Duration* corresponds to the observations that have 0 in *AI adoption* and the other group corresponds to the observations that have 1 in *AI adoption*. The same thing happens also for *Number of Investments*. After this grouping, the average of the dependent variable of each group is calculated and the test calculates the statistical significance of this difference, starting from the null hypothesis that the difference is equal to 0.

The result of this test is shown in Table 5. It can be seen that AI-adopters have a statistically significant higher Number of Investments than non-AI-adopters, with a p-value of 0.09.

On the other hand, for what concerns the variable *Due Diligence Duration*, AI-adopters have on average, a shorter Due Diligence, but this difference is not statistically significant. This first result allows us to underline what expressed above about *Number of Investments* being possibly a more significant variable for the econometric purpose of this Section.

Table 5: T-test Results

Variable	Mean_Difference	Std_Error	t_value	p_value
Due Diligence Duration	−0.348	1.304	0.267	0.791
Number of Investments	2.308	1.316	−1.754	0.090

In Table 6 are presented the results for the regressions concerning the variable *Due Diligence Duration*. These results follow the t-test analysis, and no coefficient of the independent variable *AI adoption* is significant. It needs to be pointed out that the sign of the coefficient is always negative, which is what the theoretical background suggests, apart from the regression 5 with the In- vestment Stage control. However, this regression is the only one with a positive *AdjustedR*^{2}, so the model with the best fit. In this case, also a regression with all the controls together has been fitted, but as can be observed, it is the worst fit among all regressions for what concerns *AdjustedR*^{2}. Anyway, the conclusion that can be taken from Table 6 is that no relation is observable for the variable *Due Diligence Duration*.

On the other hand, in Table 7 are shown the results of the regressions for the variable *Number of Investments*. These results show positive coefficients for *AI adoption*, as expected from both the theoretical background and from the T-test in Table 5, and in some models the coefficient of the independent variable is also statistically significant. In particular, the model with the Investment Stage control presents a coefficient of *AI adoption* that is significant on a 0.05 level. It must be noticed that this model is also the best fit when looking at the *Adjusted* R^2. So, in this situation, the results are more concordant, as the coefficients of *AI adoption* have always the same, positive, sign, with also a more significant p-value. *Adjusted* R^2 are much better, as well, as they are all positive apart from the one in the model with the Investment Sector control.

Table 6: Linear Models Summary for Due Diligence Duration

Dependent variable: Due Diligence Duration							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AI adoption	−0.348	−0.635	−0.353	−0.363	0.555	−0.439	0.044
(Std. Err.)	(1.304)	(1.328)	(1.317)	(1.324)	(1.346)	(1.609)	(2.032)
AUM Control	No	Yes	No	No	No	No	Yes
Investor Headquarter Control	No	No	Yes	No	No	No	Yes
Investment Location Control	No	No	No	Yes	No	No	Yes
Investment Stage Control	No	No	No	No	Yes	No	Yes
Investment Sector Control	No	No	No	No	No	Yes	Yes
Observations	31	31	31	31	31	31	31
R^2	0.002	0.041	0.018	0.008	0.234	0.242	0.485
Adjusted R^2	−0.032	−0.027	−0.053	−0.063	0.116	−0.083	−0.104

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: Linear Models Summary for Number of Investments

Dependent variable: Number of Investments							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AI adoption	2.308*	2.352*	2.319*	2.323*	3.110**	2.182	2.726
(Std. Err.)	(1.316)	(1.367)	(1.288)	(1.336)	(1.352)	(1.640)	(1.938)
AUM Control	No	Yes	No	No	No	No	Yes
Investor Headquarter Control	No	No	Yes	No	No	No	Yes
Investment Location Control	No	No	No	Yes	No	No	Yes
Investment Stage Control	No	No	No	No	Yes	No	Yes
Investment Sector Control	No	No	No	No	No	Yes	Yes
Observations	31	31	31	31	31	31	31
R^2	0.096	0.097	0.164	0.101	0.313	0.299	0.583
Adjusted R^2	0.065	0.032	0.104	0.037	0.208	−0.001	0.107

Note: *p<0.1; **p<0.05; ***p<0.01

5. Limitations and future research

The main limitation of this work is related to the small number of responses to the questionnaire. In fact, this issue causes a strict limitation on the quantitative analysis, as models cannot be properly estimated. Indeed, it is not clear, when analyzing an econometric model, whether there is no relation between the variables or there is just a lack of data. Besides that, as explained in Section 2.2, models cannot include every control variable, so the models are not able to capture the true marginal effect of the independent variable.

The cause of this issue may also be nourished by a selection bias, according to which among all the people who received the questionnaire, only venture capitalists who have something to say about the theme of the research responded. This bias, if present, affects not only the availability of data but also its representativeness. In fact, it is straightforward that if, due to this bias, most of the venture capitalists not interested in AI (and therefore, non-adopters) have been ruled out of the analysis, the analysis itself is biased. Of course, the presence of this bias is not verifiable in this context.

Besides that, there is also an issue concerning the fact that, even if AI causes a boost in operational efficiency, it may not be visible now, due to the very recent adoption of AI technologies. In fact, even though AI could provide quick results, it may also not be the case, as in order to be fully capable of exploiting AI may take its time.

Another limitation, which may be considered as an element of originality of this work relates to the literature available on the subject. In fact, comparable analysis are not present in the literature, especially for what concerns the quantitative analysis. This fact may represent a problem as there is no benchmark not only for the results, but also for the methodologies and the hypotheses development.

While this situation can be thought as a limitation for the current work, it surely represents a stimulus for future research, also considered the relevance of the themes in question.

Future research on the topic should focus, first of all, on developing a common methodology to gather, structure and analyze data, also focusing on how to access more data in a more efficient and effective way, maybe working on commonly accepted proxy data. For what regards the aims to explore the reasons that lead to adopt AI technologies, the analytics of sources of big data, such as LinkedIn, may provide good results.

On the other hand, the research concerning the effects of the adoption of AI tools is extremely relevant not only for the VC sector. Of course, the passing of time will make the effects of this technology clearer, but a critical mass of scientific research is needed in order to build structured results. The first aim of this research should be to develop a way to gather large data efficiently, in order to be able to develop effective econometric models.

6. Conclusions

This work extends the literature on the integration of AI tools in the venture capital sector, analyzing data gathered through a questionnaire sent to European VCs.

Results are broadly consistent with research hypotheses, confirming the surge in AI adoption in the last years (**RQ 1**) and the use of artificial intelligence mostly in the screening phase (**RQ 2).** We also find that the ICT background of employees of a VC firm may is a good predictor of the adoption of AI, confirming hypothesis **RQ 3.**

With regard to the assumption that the adoption of AI raises the efficiency of VC operations (**RQ 4**) in this phase only qualitative data, based on direct questions included in the survey, have been considered, due to the lack of data on performance following the AI adoption. Linear regression models showed basically no significant results for the dependent variable *Due Diligence Duration*, while for the other dependent variable *Number of Investments* results show a positive and statistically significant coefficient.

This fact can be explained as *Number of Investments* encloses all the internal process and thus can benefit from the hypothetical increase in efficiency of the whole process, while *Due Diligence Duration* represents one single phase, where it could be more difficult to measure an effect of the independent variable *AI adoption.*

Of course, the small number of responses (n=31) affected all empirical analysis, that we hope to extend on a larger sample. However, despite the small number of observations, this paper finds some new evidence that future research should further develop, filling a gap in the existing literature.

Appendix

Table A1: Selected literature on the use of artificial intelligence by venture capital firms.

Source	Subject of analysis	Key Results	Empirical Analysis	Data used
Arroyo et al. (2019)	AI tool for VC	Demonstrates the possibility to develop an effective AI algorithm to improve VC decision making	Yes	Crunchbase
Bonelli (2022)	AI selection of VC projects	By using AI to select projects, VCs may find themselves funding less innovative firms	Yes	Crunchbase
Brandt et al. (2018)	VC founders assessment	Labels the personal characteristics of good founders and suggests an AI approach to select them	No	
Chalmers et al. (2021)	Implications of AI integrations for VCs.	AI, as a GPT, will be able to positively affect every aspect of entrepreneurship	No	
Fernàndez Tamayo et al. (2023)	PE funds	ML algorithms can be used to effectively capture PE funds characteristics and predict future success	Yes	Private Placement Memorandums of PE funds
Gao (2023)	AI tool for VC	Demonstrates the possibility to develop an effective AI algorithm to improve VC decision making	No	
Hu (2024)	AI integration in VC phases	Feasibility analysis of AI integration for VC activities, underlying the potential of AI tools in VC	No	
Jain (2018)	AI integration in VC phases	Feasibility analysis of AI integration for different phases of VC activities, with the screening phase being found as the fittest	No	
Lyonnet et al. (2022)	AI tool for VC	Demonstrates the possibility to develop an effective AI algorithm to improve VC decision making	Yes	French Administrative data
Maurer et al. (2024)	ML tool for Screening process in VC	A ML algorithm for the screening phase in that performs better than VC's current methodologies	Yes	LinkedIn

Mueller (2023)	PE funds and PE portfolio companies	Improvement of efficiency in portfolio companies and might lead to an improvement in efficiency in PE funds	Yes	Interviews with PE experts
Sanchez (2020)	AI role in PE and VC	AI could be an important part of internal operations of PE and VC funds but it is currently not	No	
Schmidt (2019)	AI use in VC decision making	Currently, AI is used at the beginning of the decision-making process, but its role could eventually become more important in different phases	No	
Setty et al. (2024)	AI tool for VC	Demonstrates the possibility to develop an effective AI algorithm to improve VC decision making	Yes	Crunchbase
Wang et al. (2022)	AI tool for VC	Demonstrates the possibility to develop an effective AI algorithm to improve VC decision making	Yes	Xinniu Data

Table A2: Structure of the questionnaire sent to VC firms

Question	Type of answer
General info of the fund	
When was the fund established (year)?	Numerical
What is fund's AUM (M EUR)?	Numerical
Investment stage, location and sector of the Fund	
Investment stage percentage, based on number of investments.	
Pre-Seed investments	Percentage
Seed investments	Percentage
Series A investments	Percentage
Series B or later investments	Percentage
Investment sector percentage, based on number of investments.	
Digital B2C startups	Percentage
Non-High-Tech Services startups	Percentage
Robotics and Manufacturing startups	Percentage
AI startups	Percentage
Aerospace startups	Percentage
Healthcare and MedTech startups	Percentage
BioTech startups	Percentage

Energy and CleanTech startups	Percentage
Material Science startups	Percentage
Advanced Computing (Quantum, Optical, Semiconductors) startups	Percentage
AgriFood startups	Percentage
Others	Percentage
Investment location percentage, based on number of investments.	
Italy	Percentage
Southern Europe (P, ES, GR, ML, SLO, HR, BUL, TK)	Percentage
Western Europe (FR, DACH, Benelux, UK, IRL)	Percentage
Northern Europe (DK, SW, FL, NO, IS)	Percentage
Eastern Europe (PL, CZ, SK, Baltics, UA)	Percentage
North Africa and Middle East	Percentage
North America	Percentage
South America	Percentage
India, Far East and Oceania	Percentage
AI use in VC operations	
Do you use AI or ML in the Fund's operations?	Yes/No
AI use details	
Which type of AI/ML do you use?	Multiple Choice
Are those systems internally or externally developed?	Yes/No
When was AI/ML introduced in Fund's operations (year)?	Numerical
In which phase is AI/ML used?	Multiple Choice
Reasons for not using AI	
Why do you not use AI?	Text
AI effects on Fund's efficiency	
Has the AI, if used in this phase, improved Due Diligence duration?	Yes/No
Has the AI, if used, helped increasing the number of investments made yearly?	Yes/No
Has the AI, if used, helped improving Fund's operational efficiency?	Yes/No
In which phase has the AI helped the most?	Multiple Choice
AI impact on Fund's operations	
How many weeks does the Due Diligence takes on average?	Numerical
What is the number of investments made by the Fund in the last year?	Numerical

References

Amit, R., Brander, J., & Zott, C. (1998). Why do venture capital firms exist? Theory and Canadian evidence. *Journal of business Venturing*, 13(6), 441-466.

Andreoli, J. J., & ten Rouwelaar, J. A. (2024). Venture capital investment selection: an exploratory assessment of the role of entrepreneur personality traits. *Venture Capital*, 26(2), 163-189.

Arroyo, J., Corea, F., Jimenez-Diaz, G., & Recio-Garcia, J. A. (2019). Assessment of machine learning performance for decision support in venture capital investments. *Ieee Access*, 7, 124233-124243.

Bao, Y., Du, X., & Ishii, N. (2002). Improving performance of the k-nearest neighbor classifier by combining feature selection with feature weighting. *Transactions of the Japanese Society for Artificial Intelligence*, 17(3), 209-216.

Bonelli, M. (2022). Data-driven investors, working paper available at SSRN: https://ssrn.com/abstract=4362173.

Brandt, M., & Stefansson, S. (2018). The personality venture capitalists look for in an entrepreneur: An artificial intelligence approach to personality analysis.

Chalmers, D., MacKenzie, N. G., & Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice*, *45*(5), 1028-1053.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, *13*(1), 21-27.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International journal of information management*, *48*, 63-71.

Ernst & Young (2024, Febbraio 28). Large AI deals boost investment in 2024 VC market. EY. https://www.ey.com/en_us/insights/growth/venture-capital-investment-trends

Fernández Tamayo, B., Braun, R., Lopez-de-Silanes, F., Phalippou, L., & Sigrist, N. (2023). Limited Partners versus Unlimited Machines; Artificial Intelligence and the Performance of Private Equity Funds. *Artificial Intelligence and the Performance of Private Equity Funds (December 15, 2023)*.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, *29*, 131-163.

Gao, H. (2023, February). Algorithm Based Venture Capital Analysis and Decision System for Large Health Industry Start-UPS. In *2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1-5). IEEE.

Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the Economics of Innovation* (Vol. 1, pp. 609-639). North-Holland.

Hosmer Jr, David W, Stanley Lemeshow, and Rodney X Sturdivant (2013). Applied logistic regression. John Wiley & Sons.

Hsiao, C. (1996). Logit and probit models. In *The econometrics of panel data: A handbook of the theory with applications* (pp. 410-428). Dordrecht: Springer Netherlands.

Hu, L. (2024). Current Study and Method on Artificial Intelligent-based on Venture Capital Decision. *Proceedings of the 4th International Conference on Economic Management and Big Data Applications*, ICEMBDA 2023.

Jain, C. (2018). Artificial intelligence in venture capital industry: opportunities and risks (Doctoral dissertation, Massachusetts Institute of Technology).

Lyonnet, V., & Stern, L. H. (2022, June). Venture capital (mis) allocation in the age of AI. In *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

Maurer, M., Buz, T., Dremel, C., & de Melo, G. (2024). Design and Evaluation of an AI-Augmented Screening System for Venture Capitalists, *Working paper* presented at the Thirty-Second European Conference on Information Systems (ECIS 2024), Paphos, Cyprus.

Mollica, M. & Zingales, L. (2007). The impact of venture capital on innovation and the creation of new businesses. *Unpublished working paper, University of Chicago*.

Mueller, M. M. (2023). Portfolio company value creation: When private equity deploys AI (Master's thesis, Universidade Catolica Portuguesa (Portugal)).

Nadeau, P. (2010). Venture capital investment selection: do patents attract investors?. *Strategic Change: Briefings in Entrepreneurial Finance*, *19*(7-8), 325-342.

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT sloan management review*, *59*(1).

Rao, A. & Greenstein, B. (2022). PwC 2022 AI business survey. *URL: https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-business-survey*.

Sanchez, O. (2020). The Role of Artificial Intelligence in Investment Decision Making: A Study of Senior Management Perceptions within Private Equity and Venture Capital Firms (Doctoral dissertation, Dublin, National College of Ireland).

Schmid, T., Hildesheim, W., Holoyad, T., & Schumacher, K. (2021). The AI methods, capabilities and criticality grid: a three-dimensional classification scheme for artificial intelligence applications. *KI-Künstliche Intelligenz*, *35*(3), 425-440.

Schmidt, C. M. (2019). The impact of artificial intelligence on decision-making in venture capital firms (Master's thesis, Universidade Catolica Portuguesa).

Setty, R., Elovici, Y., & Schwartz, D. (2024). Cost-sensitive machine learning to support startup investment decisions. *Intelligent Systems in Accounting, Finance and Management*, *31*(1), e1548, 1-17.

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Mission AI: The new system technology (p. 410). *Springer Nature*.

Streletzki, J. G., & Schulte, R. (2013). Which venture capital selection criteria distinguish high-flyer investments? *Venture Capital*, *15*(1), 29-52.

Struta, I. (2024). How top VCs are placing their growing AI bets. *S&P Global Market Intelligence.*

Sun, W., Zhao, Y., & Sun, L. (2020). Big data analytics for venture capital application: Towards innovation performance improvement. *International Journal of Information Management*, *50*, 557-565.

Wang, Y. M., & Xue, L. (2022). Application of decision tree algorithm in early entrepreneurial project screening. *Scientific Programming*, *2022*(1), 3584196.