

Picking Profitable Investments: The Success of Equal Weighting in Simulated Venture Capitalist
Decision Making

Jan K. Woike

Max Planck Institute for Human Development, Center for Adaptive Rationality (ARC)

Ulrich Hoffrage and Jeffrey S. Petty

University of Lausanne, Faculty of Business and Economics (HEC)

Author Note

Correspondence concerning this article should be addressed to Jan K. Woike (woike@mpib-berlin.mpg.de, Max Planck Institute for Human Development, Center for Adaptive Rationality (ARC), Lentzeallee 94, 14195 Berlin, Germany). We thank Suzanne de Treville, Marc Gruber, Sebastian Hafenbrädl, Chris M. White, the special issue editors, and three anonymous reviewers for helpful comments on previous versions of this manuscript, and the Swiss National Science Foundation for their financial support (SNF 100014-124512/1 and 100014_140503).

Abstract

Using computer simulation, we investigate the impact of different strategies on the financial performance of VCs. We compare simple heuristics such as equal weighting and fast and frugal trees with more complex machine learning and regression models and analyze the impact of three factors: VC learning, the statistical properties of the investment environment, and the amount of information available in a business plan. We demonstrate that the performance of decision strategies and the relative quality of decision outcomes change critically between environments in which different statistical relationships hold between information contained in business plans and the likelihood of financial success. The Equal Weighting strategy is competitive with more complex investment decision strategies and its performance is robust across environments. Learning only from those plans that the simulated VC invested in, drastically reduces the VC's potential to learn from experience. Lastly, the results confirm that decision strategies differ in respect to the impact of added information on the outcomes of decisions. Finally, we discuss real-world implications for the practice of VCs and research on VC decision making.

Keywords: decision making, VC, investment, simple heuristics, simulation

1. Introduction

Each year venture capital (VC) funds invest billions of dollars in entrepreneurs and companies that seek to commercialize their innovative products and services. VC-backed companies receive much needed capital as well as privileged access to financial and commercial networks, so being selected by a VC as a portfolio investment is considered by many to be one of the keys to success for innovative high growth ventures (Gorman & Sahlman, 1989; Megginson & Weiss, 1991). Yet, only one to three per cent of the hundreds of deals received each year by any given VC firm are ultimately selected as portfolio investments (Bruno & Tyebjee, 1985; Maier & Walker, 1987; Petty & Gruber, 2011). Given the perceived benefits of receiving VC investment combined with the apparent selectivity on the part of the VCs, it is little wonder that VC decision making has prompted multiple streams of research.

The existing research on VC decision making has focused on two main issues, the overarching selection process (Tyebjee & Bruno, 1984; Riquelme & Rickards, 1992; Fried & Hisrich, 1994) and the criteria used by VCs when evaluating deals (MacMillan et al., 1985; Dixon, 1991; Hisrich & Jankowicz, 1990; Hall & Hofer, 1993; Muzyka et al., 1996; Franke et al., 2008). Multiple studies have demonstrated that VCs generally rely upon a multi-stage screening and evaluation process that is designed to limit information asymmetry while reducing the probability of missing a potentially profitable opportunity, thus ultimately maximizing their profits (Tyebjee & Bruno, 1984; Hall & Hofer, 1993; Fried & Hisrich, 1994; Gompers, 1998). In terms of the actual investment decision, research has repeatedly shown that when forecasting the viability of a deal, its relative attractiveness is determined by the VC's assessment of the many associated informational cues or characteristics (Tyebjee & Bruno, 1984; MacMillan et al., 1985; Riquelme & Rickards, 1992; Zacharakis & Meyer, 2000). However, although different VCs

appear to engage in the same selection process and focus on the same general categories of cues when evaluating deals, all VCs do not necessarily employ the same forecasting and decision making strategies when assessing the potential profitability of deals and ultimately selecting portfolio investments (Gupta & Sapienza, 1992; Muzyka et al., 1996; Shepherd, 1999a; Shepherd & Zacharakis, 2002). Hoffman, one of the pioneering researchers in the field, concluded that “VCs are likely to differ substantially in their investment propensity and practices to the same set of investment situations” (1972, p.182).

However, two research questions central to VC decision making have received only limited attention in the literature: (1) how do different decision making strategies impact the VC’s overall performance, and (2) what are the factors that influence how well VCs can learn from outcome feedback. These issues are of particular importance in the VC context given the failure rates of new ventures and the varying rates of returns of VC investment portfolios (Sahlman, 1990; Mason & Harrison, 2002; Dimov & De Clercq, 2006). The objective of this paper is to estimate the relative effectiveness of different decision making strategies which may be employed by a VC while at the same time accounting for the impact of learning when screening and selecting deals over the life of a fund. This goal is achieved by means of computer simulations in which the relation between the variables describing a business plan and the success probability of this plan is systematically manipulated.

2. Related Literature

A recurring theme in the VC decision making literature over the past forty years has been the attempt to identify the informational cues that VCs deem to be relevant when reviewing an investment proposal or a business plan. Specifically, researchers have endeavored to develop lists

or categories of the key cues and then determine how these cues are used by the VC to evaluate and select, or more commonly reject, an investment proposal (Hoffman, 1972; MacMillan et al., 1985; Robinson, 1987; Muzyka et al., 1996; Zacharakis & Meyer, 1998; Kaplan & Strömberg, 2004). As a result, a multitude of cues have been recorded and there is general agreement between researchers that these cues can be grouped within the following broad categories: (a) product characteristics, (b) market characteristics, (c) a company's financial position and outlook, (d) the traits of the entrepreneur or management team, and (e) other cues such as the interest of another VC in a business plan under consideration or the ability of a VC to add value to a deal (Tyebjee & Bruno, 1984; Muzyka et al., 1996; Zacharakis & Meyer, 2000; Petty & Gruber, 2011).

Regardless of how these cues may be listed or arbitrarily grouped, it is important to note that there is no corresponding agreement amongst VCs as to which of the cues is considered to be the most important when evaluating and selecting a deal (Riquelme & Rickards, 1992; Hall & Hofer, 1993; Gompers & Lerner, 1999; Shepherd, 1999b). In reality, the perceived relative significance of any given cue is often context specific and may be influenced by the investment climate, characteristics of the VC firm, or the preferences of the individual VC.

Several experimental studies involving VCs have established that the evaluation of a deal is often influenced by idiosyncratic factors such as an individual's intuition, previous industry experience, educational background, or investing experience (Hisrich & Jankowicz, 1990; Riquelme & Rickards, 1992; Zacharakis & Meyer, 2000; Franke et al., 2008). However, despite being considered experts at selecting high potential deals and spending on average one-third of their time screening and evaluating deals (Robinson, 1987), there is little evidence that VCs are aware of their underlying decision making strategies. To the contrary, as a result of their

performance in a variety of study settings, the ability of VCs to accurately identify successful companies has often been called into question by researchers (Khan, 1987; Zacharakis & Meyer, 1998; Shepherd et al., 2003; Zacharakis & Shepherd, 2005).

Another area that has received little attention is the role of learning and its impact on VC investment decision making and firm performance. The few studies that have examined learning within VC firms (Chan et al., 1990; Bergemann & Hege, 1998; Sorenson & Stuart, 2001) have focused on contracting and portfolio management activities, which take place after the VC has made their investment decision. Unlike existing research, our study focuses on VC learning during the pre-investment phase of the process. Unlike the post-investment activities that are only performed with the few deals that are selected as portfolio companies, VCs repeatedly perform the pre-investment screening routine on each of the thousands of proposals they receive and conduct evaluations on hundreds of deals over the life of any given fund (Tyebjee & Bruno, 1984; Fried & Hisrich, 1994). Existing views of the VC's selection and evaluation of the cues that are used when making investment decisions have been built primarily upon survey and experimental design-based studies. As a result, the majority of the research in this stream remains cross sectional in design and only indirectly linked to the VC's combined objective to both enhance process efficiency and maximize the firm's return on investment (ROI) (Gifford, 1997; De Clercq & Sapienza, 2005; Shepherd et al., 2005).

Our simulations provide a comparison of the effectiveness of various forecasting and decision making strategies with prescriptive implications for VC decision making. As Herbert Simon (e.g. 1955) repeatedly asserted, strategies and environments have to be considered jointly. Building on this insight, the notion of ecological rationality (Hogarth & Karelaia, 2007; Todd et al., 2012) implies that the effectiveness of strategies can only be evaluated in a given context. We

report how well strategies for investment decisions perform in environments that vary along the following dimensions: (i) selectivity of outcome information (information about the success or failure of a business plan is available for all plans versus only for those plans a VC has invested in), (ii) the relative importance of cues, and (iii) the number of cues available to the VC. Note that our investigation on the influence of the selectivity of feedback is one of the first that addresses this important aspect of learning in VC decision making.

3. Setup

One way to determine the performance of strategies would be to implement what Brunswik (1955) called a representative design (see also Dhimi et al., 2004). In the present case this would require a randomly drawn set of real business plans, all described on the same set of features and with information available about their performance in case these plans had found an investor and had been realized. There are two reasons why we did not evaluate the decision strategies this way. The first is pragmatic: To the best of our knowledge, such a data set does not exist, nor is it likely to be collected and made accessible to academic inquiry in the near future. The second reason is of a methodological nature: Even if it existed and were at our disposal, simply observing how the strategies fare in a given set of business plans for which we would not manipulate anything, would make any causal claims and a systematic study of the dependency of strategies' success on particular information structures impossible.

In an attempt to overcome the difficulties associated with representative design, Hammond (1966) differentiated between the concepts of substantive situational sampling and formal situational sampling. The former implements the original idea of representative design by focusing on the content of the task and using real stimuli that have been representatively sampled

from the environment. The latter, in contrast, permits the construction and presentation of stimuli that are representative in terms of the formal informational properties of the environment (i.e., number of cues, their values, distribution, inter-correlations, and ecological validities). We followed Hammond's approach and created fictitious business plans that our simulated VCs used as input when making their success forecasts and investment decisions. This procedure gave us perfect control over the statistical structure within "our" world. It is important to add that the statistical properties in this artificial world were informed by the literature and our knowledge of the real world, that is, the generation of business plans was based on a set of plausible assumptions that had an empirical basis.

Each of the plans that we generated is associated with a set of k binary cues $\mathbf{C}_i \in \{0000, \dots, c_1c_2c_3c_k, \dots, 1111\}$, and a dichotomous (future) outcome variable which can take the values +1 (success or positive return on investment) and -1 (failure or negative return on investment). Each cue is probabilistically related to this outcome so that the vector of cue values (henceforth, the cue profile) is useful information when it comes to predicting whether a business plan will result in success or not. Note that the cues in the present simulations are devoid of any semantics, that is, they could include variables specified in the business plan (e.g., experience of the management team) or variables describing the broader context (e.g., market characteristics).

3.1. The Environments

The term "environment" refers to the set of probabilities used when creating a set of business plans. Specifically, an environment is characterized by the following probabilities. Each of the possible cue profiles is associated with a probability $p_o(\mathbf{C}_i)$ that it occurs, and with a probability $p_s(\mathbf{C}_i)$ that it will be successful, so that knowledge of this set of probabilities is

sufficient to determine the probability of success (π_s) for a randomly sampled business plan from this environment.

3.1.1. Cue Weighting Structures

We were interested in the question whether the performance of deal selection strategies depends on how cues differ from each other with respect to their statistical relationships with the outcome variable. Therefore, we generated environments in which the cues are equally important when determining outcomes, and other environments in which some cues are more important than others. Specifically, the outcome is determined by comparing a weighted sum of the cues and a random error component with a threshold value. In the Equal Weights environments, all cues are weighed equally. In the Arithmetic Weights environments, a constant value is added to cue weights for each subsequent cue in a fixed sequence, whereas in the Geometric Weights environments, cue weights in the sequence are multiplied by a constant factor. Parameter values and details for generating these environments for different numbers of cues are described in Appendix A.

3.1.2. Overall Success Probability

In our simulations, we created environments with a predetermined overall success probability. Because the individual success probabilities of the various cue profiles are known (see Appendix A), it was possible to manipulate the occurrence probabilities of the individual cue profiles such that the resulting overall success probability matched the predetermined probability. The details of this procedure are described in Appendix B.

3.2. The Decision Strategies

Even experts have been shown to have major difficulties when engaging in unaided judgmental forecasting (Dawes, 1979; Tetlock, 2005; Burgman et al., 2011). In the present paper, we use simulated VCs who execute four well-defined strategies to make these selections. Such a simulation approach

removes any noise that humans produce when using strategies unreliably, and hence allows us to determine the upper bound of aided decision making.

Each of the four deal selection strategies we consider in this paper is responsive to experience and integrates new information in systematic ways. However, the strategies differ with respect to how information about cue profiles is processed and how their parameters are computed from a set of business plans for which the cue profiles and outcome information are given. *Equal Weighting* is an algorithm that simply bases the decisions on the number of cue values in a cue profile that point to success. *Logistic Regression* is a procedure to estimate weights in a linear combination of cues, which, in turn, is used to make decisions. *CART* (Breiman et al., 1984) is a tree-building algorithm that is flexible enough to produce the same decisions as any other possible strategy for binary cues. The *Fast and Frugal Tree* heuristic (Martignon et al., 2003, 2008) is a tree-building algorithm as well, but it is more restricted in the sense that the set of trees it can possibly generate is reduced. Note that each of these strategies needs to estimate parameters, which is achieved on the basis of a sample of cue profiles and their associated outcome values, henceforth referred to as the learning sample L .

Equal Weighting combines the principle of unit weighting of cues, as instantiated by Dawes' Rule (see Dawes, 1979; Czerlinski et al., 1999; Hastie & Kameda, 2005) and index methods (see Burgess, 1939; Armstrong & Graefe, 2011) and the principle of making identical predictions for sets of objects with the same number of positive cue values, as, for example, implemented in the Mapping Model (von Helversen & Rieskamp, 2008). Fast and Frugal Trees are simple binary decision trees that look up cue values sequentially and that are potentially, depending on the cue value observed, able to make decisions without consulting further cues. They embody the principle of one-reason decision making and have been analyzed and tested in the context of the simple heuristics program Todd et al. (2012). Logistic regression and CART are procedures often used in machine learning and classification tasks across disciplines.

Ordered with respect to complexity when estimating their parameters, Equal Weighting is the simplest strategy, followed by the Fast and Frugal Tree, with Logistic Regression and CART being the most complex strategies of the set. The details of their implementation in the present study are specified in Appendix C.

3.3. The Learning Task: The Standard Condition and Its Variations

3.3.1. The Standard Condition

In this paper, we simulated a number of VCs over the period of one fund. During the life of a fund, four VCs were confronted sequentially with 2^{10} randomly generated business plans. The four VCs differed with respect to which of the four forecasting and deal selection strategies described above they used to make their 2^{10} investment decisions. A given condition, which was defined by a particular set of parameter combinations (described below), was realized 100 times, corresponding to 100 different sets of 2^{10} randomly generated business plans. Each of these different sets was shown to new VCs, so that none of the simulated VCs could use knowledge from a previous fund.

The following properties define what we refer to as the standard condition that we use as a reference point when presenting the results of the simulation: Each VC started with an experience base of 2^6 business plans with known cue profiles and outcome information, generated by the same process as the 2^{10} plans for which the VCs had to make their investment decisions. When learning its parameter values from the experience base before starting a new fund, each deal selection strategy used the same 2^6 plans. In the standard condition, each business plan was characterized by eight cues and their values, and all this information was given, without any cue reading errors, to the VCs. The plans were randomly generated using the arithmetic cue weighting structure displayed in Table 2 and an overall success probability $\pi_s = .2$.

For each of the 2^{10} business plans in a given fund, the VC had to decide whether or not to invest an arbitrarily chosen investment unit (the same for each investment). After a feedback lag of 2^6 plans (reflecting the fact that feedback on whether or not a business succeeds is usually delayed, Shepherd & Zacharakis, 2002), the business plan's cue profile and its outcome was added to the VC's experience base regardless of the investment decision. Each time a profile was added to this experience base, all strategies estimated their parameters again as described above and without any bias. An investment decision for any of these plans led to a return of 150% of the investment unit if the business was successful, and to its complete loss if it failed.¹

3.3.2. Variations

Taking the standard condition as a starting point, we manipulated a number of parameters. In this paper we focus on three parameters that we considered to be most interesting and for which we report the effects below: First, the *selectivity of the feedback* relates to the question of whether feedback on the outcome variable was provided for all plans or only for those in which a VC actually invested, resulting in a biased sample (compare Denrell, 2003). Second, the *cue weighting structure* (as described in Table 2) determined how predictive the cues were relative to each other (Zacharakis & Meyer, 2000). Third, the *number of cues* specified how many pieces of information the strategies had at their disposal when making decisions (Zacharakis & Meyer, 1998).

In addition, we manipulated a number of other factors: Overall success probability (Chrisman et al., 1989; Shepherd, 1999b), financial consequences of success and failure (Bruno & Tyebjee, 1986; Gompers & Lerner, 2001), size of the expertise base (Shepherd et al., 2003; Franke et al., 2008), delay of outcome

¹This translates into $o_{is} = .5$, $o_{if} = -1$, $o_{ns} = 0$, and $o_{nf} = 0$, and $\tau = \frac{2}{3}$, see Appendix C.

feedback (Einhorn & Hogarth, 1978; Harvey & Fischer, 2005), amount of error in determining success (Khan, 1987; Zacharakis & Meyer, 1998). These manipulations served as a sensitivity analysis and their effects were either minor or intuitive. More importantly, these robustness checks revealed that our major results seem not to hinge on specific parameter constellations, therefore specific results are not reported here.

4. Results

4.1. The Standard Condition

The main measure we use to evaluate the behavior and the monetary performance of the strategies is the accumulated profit, that is, the total ROI after the first k investments. Note that the number of investments should not be confused with the number of encountered plans. This difference is important as the decision strategies differ with respect to how many investments were made for a given number of presented plans, or, conversely, how many plans were presented before making k investments. If the fund closed before k investments were made, then the profit for k investments was defined as the profit that had accumulated at the time the fund closed.

Figure 1 displays the profit of the strategies after k investments, averaged across the 100 simulated funds, for the parameter configuration that we defined as the standard condition.

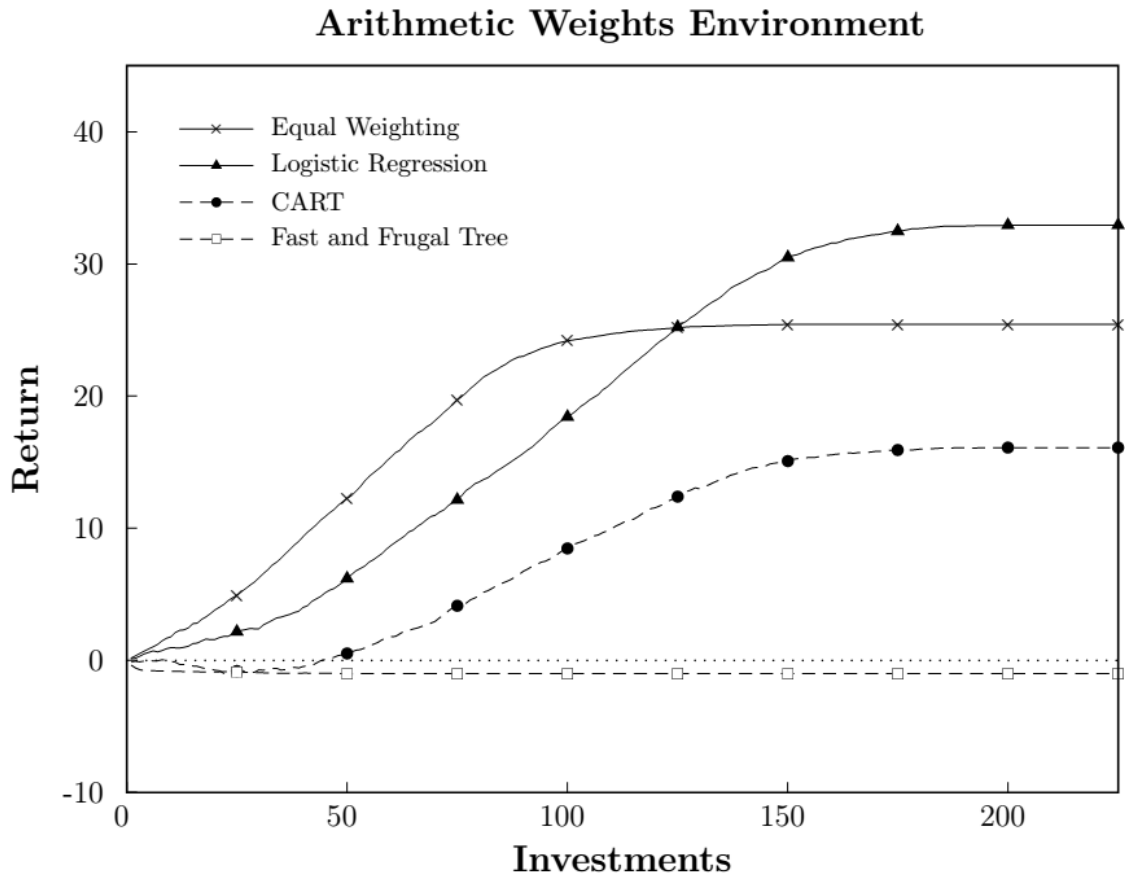


Figure 1: Return after k investments in the standard condition

The curves, except for the Fast and Frugal Tree, have similar shapes and can be divided into four distinct phases. In the first phase, the learning phase, profit gradually increased with the number of investments, reflecting the fact that VCs refined their parameters and continued to benefit from a growing experience base. In the second phase, the informed phase, the slope (i.e., the average difference with respect to accumulated return) remains relatively stable, reflecting the fact that the parameter estimates no longer changed substantially as new feedback was obtained (which amounts to saying that VCs were relatively well informed about their environment). In the fourth phase, the slope finally approaches zero, reflecting the fact that – by the time the fund was closed (i.e., after 2^{10} plans) – each VC had made a certain number of

investments (m_i) and the VC's profit is defined to be constant for any $k > m_i$. In the intermediate third phase, the slope smoothly decreases, reflecting the fact that the curve for a given strategy is aggregated over 100 funds, each managed by a different VC with a distinct m_i . Had we used twice as many plans, the learning phase would have looked exactly the same, the informed phase would have lasted more than twice as long (because there would be no necessity to start learning from scratch), and the final profit would have been more than twice as high (because the informed phase, during which profit is higher compared to the learning phase, would have lasted more than twice as long).

The Equal Weighting strategy had the shortest learning phase, ending at approximately 5–20 investments, and the highest slope during the informed phase. CART started negative while learning but was able to recover and to make positive profits. Still, it was less profitable than Equal Weighting and Logistic Regression, and it never reached their final level of profits. The Fast and Frugal Tree did not make any profit at any time.

Table 1 provides some additional information about the behavior and performance of the strategies in the standard condition. The measures introduced in this table correspond to the phases that we identified in Figure 1. The first data column reports the average return per investment computed over the first 25 investments (ROI_{1-25}). If a VC made less than 25 investments, this VC's investments were still included in the analysis. As Figure 1 shows, some strategies had a learning phase longer than 25 investments, but it also seems to be the case that none of them had completed learning before this point. Equal Weighting, which had the shortest learning phase, also made the most profit during this phase.

The second data column reports the return averaged across the 25 investments between the 51st and 75th (ROI_{51-75}). Again, not every VC contributed 25 data points to the analysis. As

Figure 1 shows, this period falls into the informed phase of Equal Weighting and the Regression strategy, while it seems to be at the very beginning of CART's informed phase. And again, Equal Weighting performed the best.

The next data column in Table 1 reports the average return across all investment decisions made by the VCs using a given strategy (ROI_{all}). The next four data columns report how many investments the 100 simulated VCs who used a given strategy made across the 210 plans. First, the variance and the range of this number (across the VCs who used the same strategy) were quite large. Second, the difference between the strategies with respect to the mean number of investments that VCs made was also considerable. The entries for the minimal and maximal number of investments roughly mark when the slope during the informed phase starts to decrease and when it approaches zero, respectively. Even though Equal Weighting had a higher slope during its informed phase, VCs who used this strategy invested, on average, less than those using Regression. As a consequence, the curves cross and Equal Weighting ended up with a lower final return (see Table 1, last two columns).

Table 1: Results for the Standard Condition.

	Return on Investment			Total Number of Investments				Final Return	
	ROI_{1-25}	ROI_{51-75}	ROI_{all}	Minimum	Maximum	Mean	SD	Mean	SD
Equal Weighting	.20	.31	.27	62	149	92.9	17.3	25.4	7.2
Logistic Regression	.09	.24	.21	113	203	157.7	17.9	32.9	8.4
CART	-.03	.14	.11	108	197	147.1	18.1	16.1	10.1
Fast and Frugal Tree	-.52	-.25	-.52	0	52	2.0	5.2	-1.0	2.4

The performance of the strategies, as summarized in Figure 1 and Table 1, can be related to the cue weighting structure. In the present environment with its linear and compensatory set of

weights (arithmetic cue weighting structure, see Table 2), neither of the two tree algorithms was competitive. The Fast and Frugal Tree did not find the single cue that is valid enough for one-reason decision making, simply because this cue did not exist. CART, in contrast, has the potential to deal with an arithmetic cue weighting structure, but this requires the construction of a complex tree with a high number of nodes, which, in turn, requires an extended experience base. Note that at a given stage during the construction of the tree, the business plans are distributed over its end nodes and a small sample of plans prevents these end nodes from being split up further. Even though Equal Weighting does not match the weighting structure in the present environment, its ability to make robust inferences Dawes (1979) was able to compensate for this mismatch – at least to such an extent that it was able to outperform the Logistic Regression.

4.2. Selectivity of Feedback

In the standard condition, the VCs could include all business plans in their experience base, no matter whether they had decided to invest or not to invest. While it is of vital importance to monitor the success or failure of business plans a VC invested in, such information about plans that a VC rejected is often not available – be it because outcome information did not reach the VC after a different investor funded the unmodified proposal, or because no other investor has been found and the business hence failed to materialize (Bruno & Tyebjee, 1986). A high confidence in one's ability to make good investment decisions, eventually coupled with the belief that outcome information about rejected proposals is irrelevant, may lower the motivation to actively search for such (potentially costly) information.

The biased sampling of business plans creates a situation in which the set of plans that VCs encounter cannot be considered to be a representative sample of the business plans in the population. The psychological literature is rich in examples demonstrating that performance is

severely impaired due to such a sample selection bias (see Brunswik, 1955; Fiedler, 2000; Dhimi et al., 2004).

To see how this bias affects investment results, we also created a condition in which the simulated VCs were only able to include in their learning sample those business plans they invested in (note that this reduction of feedback was only implemented after the fund started, but it did not affect the experience base that VCs were equipped with before they started to make decisions). The results are displayed in Figure 2. As the comparison between this figure and that of the standard condition shows, all strategies perform worse if the feedback was selective – be it because the size of the learning sample was smaller, or because the learning sample was biased, or both.

Even though the order of the four strategies with respect to their profit was the same for both learning conditions, there are some striking differences. The strategy that was most affected by restricting the learning sample was CART: As in the standard condition, the slope starts out negative, but when information on rejected business plans was lacking, this strategy was not able to enter an informed phase in which the constructed trees could make substantial profits, and, as a consequence, also not able to compensate for the losses made during its learning phase.

While in the standard condition, Equal Weighting has a slope of about .3 during its informed phase (see also Table 1, ROI_{51-75}), when feedback was selective this slope is reduced by roughly 20% to about .24. Note that slopes of .3 and .24 correspond to 86.6% and 82.7% successful outcomes among business plans that received funding, respectively, which should be compared to the a priori success probability of .2. For Logistic Regression the slope was reduced from about .25 by roughly 28% to about .18, corresponding to 83.3% and 78.7% of successful outcomes among all investment decisions. Thus, not only did Equal Weighting have a higher

slope than Logistic Regression in both feedback conditions, it also suffered less when outcome information on rejected plans no longer entered the learning sample; that is, it was more robust when sample selection bias was introduced. To put the absolute effect between unselected and selected feedback received during the informed phase in context, note that a drop in profits of .06 (Equal Weighting; from .3 to .24) and .07 (Logistic Regression; from .25 to .18) is still more than 10% of the profit for a successful deal, which was .5 investment units.

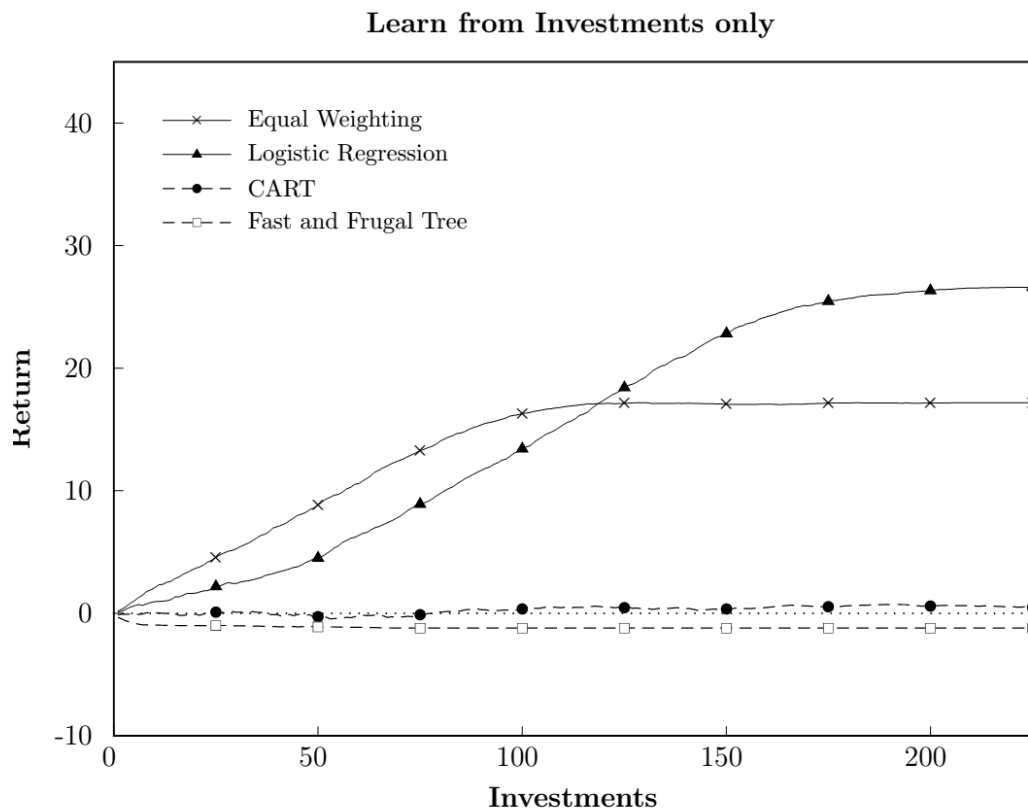


Figure 2: Return after k investments while learning from investments only

The drop in performance due to selectivity of outcome feedback for three of the four strategies can also be seen when considering ROI_{all} . For Equal Weighting, Logistic Regression,

CART, and the Fast and Frugal Tree, (ROI_{all}) was .21, .16, .00 and -.44, respectively, when feedback was selective. Comparing these values to the last column of Table 1 (where feedback was given for all plans), the drop in performance due to this manipulation was roughly .06, .05, .11 for the first three strategies. The Fast and Frugal Tree improved by roughly .08, while the ROI_{all} stayed negative.

4.3. The Cue Weighting Structure

For the standard condition, we used arithmetic cue weights when generating business plans. As explained in the Setup section, we also generated business plans for which the cues were equally predictive (equal weights), or for which their weights adhered to a geometric distribution (see Table 2). We reran the standard condition except that we used business plans generated with the arithmetic and geometric cue weighting structure. The performance of the strategies in the two new environments is shown in Figure 3. It is evident that the relative success of the four strategies depends on the environment in which they are tested.

In the Arithmetic Weights Environment, and especially in the Equal Weights Environment, Equal Weighting was the best strategy (in terms of average return on investment during the informed phase). In the Geometric Weights Environment, Equal Weighting was the worst performing strategy, while in contrast, the Fast and Frugal Tree, which had failed in the Equal Weights Environment, now performed best, almost converging against the theoretically optimal slope of .5. CART, which had shown the worst performance during the learning phase and the second worst performance during the informed phase in the other two environments, showed virtually the same (good) performance as Logistic Regression in the Geometric Weights Environment.

For the Equal Weighting strategy, the average number of investments did not vary much between the environments (94.4, 92.9, and 82.7 for Equal Weights, Arithmetic Weights and Geometric Weights, respectively). In contrast, both for the Logistic Regression (127.8, 157.7, and 181.6, respectively) and for CART (126.0, 147.1, and 182.0), this number increased as the skewness of the distribution of cue weights increased (note that a higher skew made it easier to identify what the predictive cues in the environment were). The difference was most pronounced for the Fast and Frugal Tree which made, on average, 2.5 investments in the Equal Weights Environment, 2.0 in the Arithmetic Weights Environment, and 103.8 in the Geometric Weights Environment.

The cue weighting structure turns out to be a crucial factor in the study of the ecological rationality of strategies: a strategy tended to perform well if its architecture matched the properties of the environment (here, the cue weighting structure). Equal Weighting excelled in an environment in which the cues were equally predictive, but in the Geometric Weights Environment in which the distribution of cue weights is highly skewed, this strategy was punished for using a non-matching weighting structure. Treating some cues as more important than others (as Fast and Frugal Trees do in the extreme) makes sense only if some cues are indeed more important than others in the respective environment. Therefore, it showed the opposite pattern compared to Equal Weighting. Martignon et al. (2008, Result 2) have shown that each linear threshold model with non-compensatory weights shows the same decision behavior as a corresponding Fast and Frugal Tree using the same (or a subset of the) cues used in the linear model. Therefore, the Fast and Frugal Tree can match the Geometric Weights Environment as the latter can be described as a linear threshold classifier with non-compensatory weights.

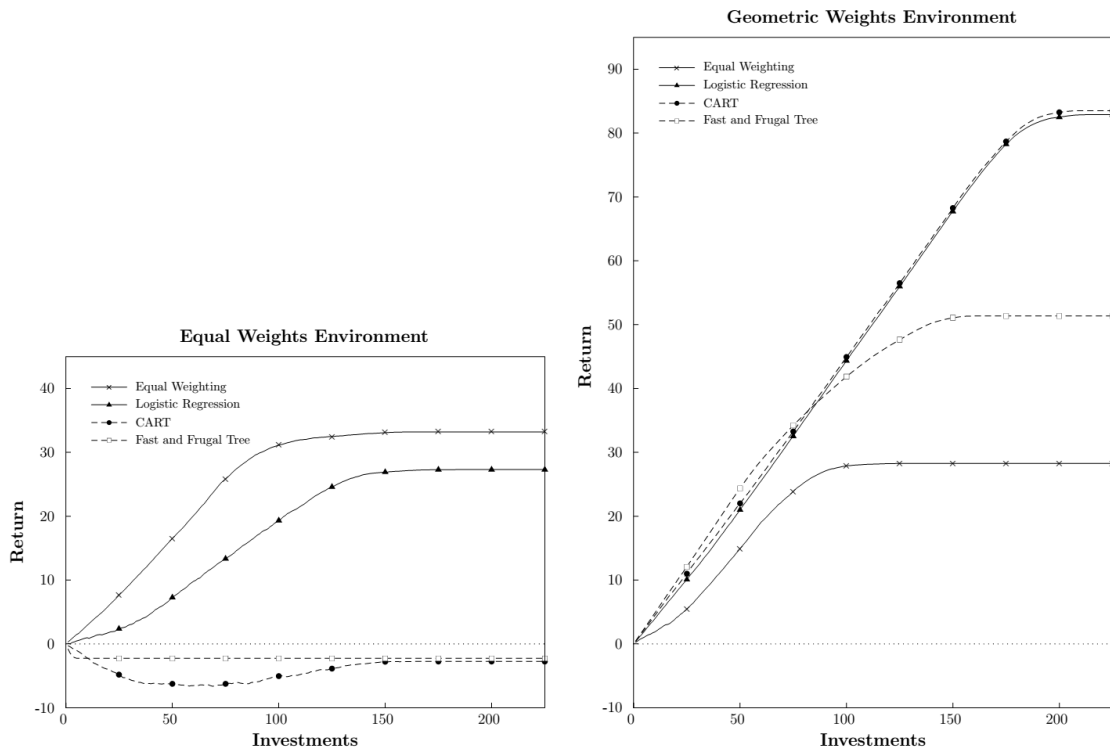


Figure 3: Return after k investments for the Equal Weights and the Geometric Weights Environment

These two examples were extreme in the sense that the two strategies had an extreme architectural bias (Equal Weighting is biased towards flat weighting, and the Fast and Frugal Tree is biased towards a steep hierarchy) and the two environments had extreme properties (flat and highly skewed cue weighting structures). How did the strategies perform in the Arithmetic Weights Environment whose cue weighting structure was in between that of the other two environments (see Table 2)? Consistent with the title Dawes (1979) used for his classic paper (“The Robust Beauty of Improper Linear Models in Decision Making”), Equal Weighting was robust enough to fare well when the cue weights followed an arithmetic structure – and more than this, it was even the best performing strategy. In contrast, this cue weighting structure was

not skewed enough to allow the Fast and Frugal Tree to realize its potential. As Table 1 shows, VCs using this strategy could not construct a tree that led to a substantial number of investments and those few investments obtained, on average, a negative return.

In contrast to Equal Weighting and the Fast and Frugal Tree, the other two strategies were, at least in principle, flexible enough to adjust to a wide range of environmental properties. One could therefore argue that in the Geometric Environment, CART should be able to match the performance of the simpler Fast and Frugal Tree, since the structure of the tree constructed by the latter strategy is within CART's search space. However, this is not necessarily true as the larger size of CART's search space requires a longer learning phase to find the structure of this tree (and even with an unlimited learning phase this is not guaranteed). Given the parameters we used in our simulations, even though CART's performance came close to that of the Fast and Frugal Tree, it did not reach it (the slope in the informed phase was about 7% lower). CART also has the potential to construct trees that match the performance of Equal Weighting. As it turned out, however, it struggled in the Arithmetic Weights Environment and even more so in the Equal Weights Environment. This can be explained by the fact that the trees necessary to capture the cue weighting structures of these environments need to be much larger than those for the Geometric Weights Environment – and even CART's prolonged learning phase for those two environments was not sufficient to construct a competitive complex tree.

In a similar vein, Logistic Regression had the potential to perform well in all three environments. It owes its flexibility to the fact that it estimates cue weights without any restrictions. Indeed, there is no environment in which VCs using this strategy are heavily outperformed. Nevertheless, there is not a single environment that allows Logistic Regression to win the competition against all three other strategies. The lack of a stronger bias towards a

particular cue weighting structure and the impact of sampling error during the learning phase prevents the strategy from finding the cue weights that are optimal for the environment, and hence, from reaching the performance that could theoretically be achieved.

4.4. The Number of Cues

In the standard condition, a business plan was described by 8 cue values. We manipulated this number and set it to be either 6, 8 (standard), 10, or 12. The manipulation of the number of cues did not result in a change of the ordering of the strategies, except for the fact that CART's performance decreased (compared to that of the other strategies) as the number of cues increased. Figure 4 shows how each of the four strategies was affected by this manipulation.

The Equal Weighting strategy performed almost the same in the four cue number conditions, except that it made more investments as the number of cues increased. For Logistic Regression, in contrast, an interesting pattern can be observed: during the learning phase, the VCs using this strategy made better decisions when the number of cues was smaller, whereas this relationship was reversed when considering the informed phase. In addition to the resulting crossing of the return curves, it can be observed that the learning phase was longer and the number of investments was higher when more cues were available. For CART, the curves do not cross and in each phase, this strategy both made fewer investments and performed better when the number of cues was smaller. Finally, the inability of the Fast and Frugal Tree to cope with an Arithmetic Weights Environment (that we described above for the standard condition) was robust against this manipulation, except that performance was worse for smaller numbers of cues.

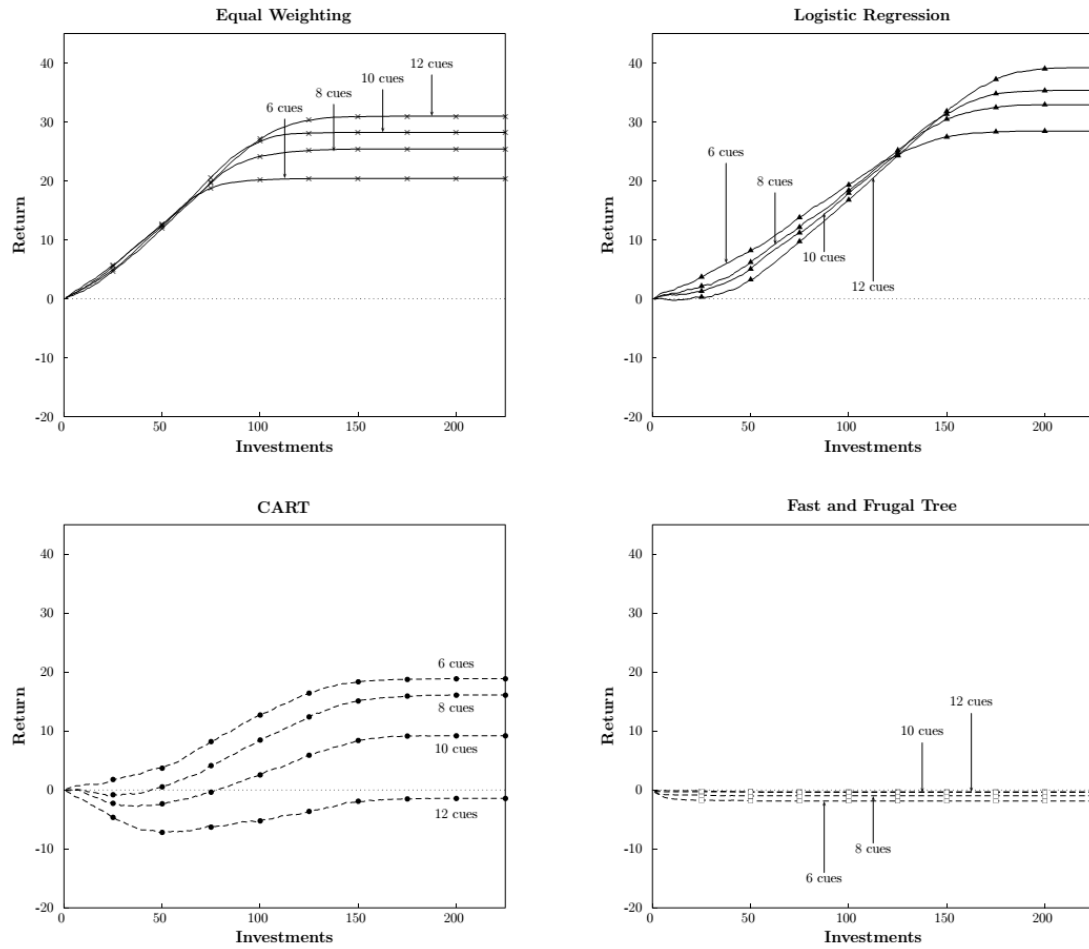


Figure 4: Return after k investments for different numbers of cues

These results again demonstrate how important it is to consider strategies and environments jointly (Todd et al., 2012). The performance of a strategy can only be determined in a given environment and strategies' performance can be affected differentially by changes of environmental features. Whereas the number of cues (almost) did not affect Equal Weighting and the Fast and Frugal Tree, CART and Logistic Regression performed differently in the various cue number conditions. For CART, the effect of the number of cues is readily explained: the strategy obviously found it easier to construct a well performing tree when the number of cues (and, hence, the universe of possible trees) was smaller. For Logistic Regression, it is easier to estimate

the parameters when the number of cues is smaller (hence the shorter learning phase for a smaller numbers of cues), but once the strategy is in the informed phase and estimates are relatively stable, the strategy benefits if more information can be processed.

5. Discussion and Conclusion

One of our central simulation results was that the Equal Weighting strategy was competitive with more complex investment decision strategies. This finding is consistent with a general pattern found across a wide range of tasks and domains: Under specific conditions, simple strategies can outperform complex strategies, presumably not despite of their simplicity but because of it (Martignon & Hoffrage, 2002). Specifically, their success when deciding upon new business plans can largely be attributed to the robustness of their parameter estimates (see also Dawes, 1979). Complex strategies have an advantage when it comes to fitting known data. This was not the situation, however, in the present set of simulations, neither is it a relevant condition for VCs who will have to make investment decisions for new business plans outside the learning sample, in particular, if this learning sample is small. In such a situation, simple strategies that are adapted to the environmental structure will prevail. For instance, when Armstrong & Graefe (2011) compared forecasting methods in the context of the last 29 U.S. presidential elections, they found that an equal weighting model based on 59 biographical variables of presidential candidates (which they called “index method”) outperformed – in cross-validation – Gallup polls, prediction markets, and three respectable econometric models. Moreover, they found that a model based on this bio-index was competitive with seven econometric models in forecasting voting shares (see also Graefe, this issue).

In a similar vein, Gigerenzer & Brighton (2009) showed for small learning samples that a lexicographic decision strategy outperformed several more complex strategies in a wide range of conditions (see also Brighton & Gigerenzer, this issue), von Helversen & Rieskamp (2008) demonstrated that a simple strategy which used equal weighting of predictors outperformed multiple regression in an estimation task, and DeMiguel et al. (2009) demonstrated that the simple $1/N$ heuristic that allocated wealth equally across N assets could not be consistently outperformed by any of the optimizing allocation model that they used as benchmarks. Goldstein & Gigerenzer (2009) offer a more extensive evaluation of simple forecasting strategies. They demonstrate the successful performance of simple prediction rules in tasks as diverse as portfolio selection, the prediction of future customer purchase activity and future high-value customers, the forecasting of tennis and soccer results, and the forensic practice of geographic crime profiling.

Our results are in line with these findings: The performance of the simple Equal Weighting strategy was robust across environments in the sense that it either outperformed the other strategies (in Arithmetic and Equal Weights Environments) or did not lag much behind the winner (in a Geometric Weights Environment). VCs who do not know the statistical structure of the environment are thus well-advised to use this simple and robust strategy as a starting point and may, after a substantial number of investments followed by outcome information, consult and carefully evaluate the experience base to check whether the millions or billions they have at their disposal should better be invested by adapting a different plan for making investment decisions.

Our findings are also consistent with the notion of ecological rationality (Todd & Gigerenzer, 2007), and demonstrate how important it is to consider strategies and environments

jointly: (i) The performance of a strategy can only be determined within the context of a given environment and (ii) strategies' performance can be differentially affected by changes of environmental features. Building on these observations, VCs are well advised to learn more about their environment. Different VCs are likely to operate in environments with different statistical structures, in other words, there will be variance across time, regions, markets and fund sizes. While our simulations used parameter ranges that were informed by the real world, VCs need to know which parameter constellation is relevant for them in particular.

Consequently, VCs may gain a competitive advantage from systematically studying their decision making context, for instance, by using the information stored in their archives, or, even better, by creating protocols that allow for convenient forms of data analysis. In this learning process VCs should, however, consider that learning from outcomes of their own investments alone may lead to biased parameter estimates for the decision strategies, as our simulations have shown – simply because the amount of decision making experience, and learning as a result of the same, by any VC is too limited to enable them to estimate cue weights with enough precision to apply the model. The consideration of their own investments should thus be complemented by securing other information sources to obtain a more complete representation of the statistical properties of their environment.

While the vast majority of research on VC decision making has focused more on the VCs' assessment of specific cues, our focus was on the performance of forecasting and deal selection strategies that differ with respect to how they process cues. Similarly, Zacharakis & Meyer (2000) used two bootstrapping models to predict the success of ventures and compared the performance of these models to that of actual VCs (see also Armstrong, 2001). Their first model used cue weights that were fitted to data from an earlier study and their second model

weighted the cues equally. The authors found that the equally weighted bootstrap model was clearly better at selecting successful ventures than both the alternative model and the actual VCs. They conclude that bootstrap models are useful decision aids for VCs – a conclusion that we fully endorse (see also Zacharakis & Meyer, 1998; Shepherd & Zacharakis, 2002).

Our Equal Weighting strategy outperformed more complex strategies in two of the three types of environments we analyzed. Zacharakis and Meyer's bootstrap model with weights that have been fitted to earlier data corresponds to the Logistic Regression strategy in our simulation with regression weights that have been fitted to the VC's experience base. Thus, in both studies the Equal Weighting strategy is more robust and hence more useful than the regression model when it comes to forecasting the success of previously unseen business plans.

Similarly, Åstebro & Elhedhli (2006) found that a simple conjunctive model with separate counts for good and bad cues also turned out to be extremely successful when modeling experts' forecasts of early-stage ventures' performance. They concluded that "simple decision heuristics can perform well in a natural and very difficult decision-making context" (p. 395).

In the real world, VCs may provide value-added activities to their portfolio companies that will enhance their performance (Sapienza et al., 1994; Cumming et al., 2005). Also, during times of increased VC investment activity or in the case of an exceptionally promising proposal, VCs are often forced to compete for the opportunity to invest in selected deals (Gompers & Lerner, 2001; Lockett & Wright, 2001). While we have not explicitly considered these factors so far, our simulations do allow for their inclusion. In the first instance, when evaluating a deal a VC will be assessing the need for, as well as their ability to provide, any necessary hands-on involvement or other value-added activities. As such, the perceived potential for the VC to provide support for a deal beyond the financial investment can be one of the cues that are

evaluated by a VC. Secondly, the actions of another VC interested in a potential deal, either observed directly by the focal VC or reported by the entrepreneur or a third party, can be one of the cues that are used in our simulations.

The decision strategies we have studied are described in a way that will assist future researchers further the scope of empirical testing in the VC setting. The literature to date has provided many valuable insights on the content and process aspects of the core VC activities but has predominantly provided descriptive accounts that offer limited input to the VC community. Our work builds upon this body of research and evaluates the VC investment decision task through a more prescriptive approach. Moreover, in our simulations the term VC may represent an individual VC or a team of VCs within a firm so we hope that our findings will provide the basis for more successful investment decision strategies, at both the individual and firm levels. Finally, we believe that this paper can serve as a basis for more longitudinal research that focuses on the intersection of the weighting of informational cues, the forecasting and decision making process, and VC performance and learning.

Extending beyond the VC context, our simulations are equally relevant to angel investment networks, the managers of corporate venture funds and private equity funds. Although each of these investors has their own unique preferences and investment objectives, they all seek profitable deals and draw upon multiple informational cues when evaluating potential investment opportunities. Additionally, we believe that our results will inform others who engage in evaluation and selection routines such as human resource professionals or any manager responsible for selecting job candidates. Given the robustness of the pattern obtained in our own simulation and in the related research discussed above, we predict that simple strategies for forecasting and for decision making will prevail in these domains as well.

Appendix A: The cue weighting structures

5.1. Eight cues

Whether a business plan with a given cue profile has a positive or negative outcome value ($f(\mathbf{C}_i) = 1$ vs. $f(\mathbf{C}_i) = -1$) is determined by comparing the value of

$$u(\mathbf{C}_i) = \sum_{j=1}^k w_j c_{i,j} + \varepsilon, \tag{1}$$

where w_i are cue weights as specified below and q is a random variable (whose values are drawn from a uniform distribution over $[-e; e]$, where the parameter constant e represents the amount of uncertainty), with a threshold θ :

$$f(\mathbf{C}_i) = \begin{cases} 1 & , u(\mathbf{C}_i) \geq \theta \\ -1 & , u(\mathbf{C}_i) < \theta. \end{cases} \tag{2}$$

The probability for each cue profile to be associated with a positive outcome value is therefore determined as follows:

$$p_s(\mathbf{C}_i) = \begin{cases} 1 & , \sum_{j=1}^k w_j c_{i,j} - e > \theta \\ .5 + \frac{\sum_{j=1}^k w_j c_{i,j} - \theta}{2e} & , \sum_{j=1}^k w_j c_{i,j} - e \leq \theta \leq \sum_{j=1}^k w_j c_{i,j} + e \\ 0 & , \sum_{j=1}^k w_j c_{i,j} + e < \theta \end{cases} \tag{3}$$

Table 2: Parameters for the Standard Condition.

	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	θ	e
Equal Weights	1	1	1	1	1	1	1	1	4.9	1.25
Arithmetic Weights	1.7	1.5	1.3	1.1	0.9	0.7	0.5	0.3	4.9	1.4
Geometric Weights	4	2	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	1.249	0.55

The weights w_j in Equation 1 together with the threshold and error define the relative importance of cues. The ordered set of weights conforms to one of three different types of sequences, namely, constant, arithmetic, and geometric. Given that a constant sequence of weights corresponds to a set of equal weights, we will henceforth use the following terminology for the three types of cue weighting structures:

1. Equal: The weights are all set to $w_i = 1$.
2. Arithmetic: The weights are determined by $w_i = s - (i - 1) \frac{2s - 2}{k - 1}$. The parameter $1 < s < 2$ determines the highest weight (s) and the lowest weight ($2 - s$).
3. Geometric: The weights are determined by $w_i = s_1 \cdot s_2^{i-1}$. The parameter $s_1 > 0$ determines the largest weight, the parameter $0 < s_2 < 1$ determines the ratio between subsequent weights.

5.2. Different numbers of cues

For different numbers of cues, the weights of the cues had to be adapted in order to still be able to satisfy the constraints imposed by the cue weighting structure (for the Arithmetic

Weights Environment $w_i = s - (i - 1) \frac{2s - 2}{k - 1}$). The maximum and the minimum cue weight as

well as the size of the error component were kept constant, only the threshold was increased by one for every two cues that were added. The resulting cue weighting structures are displayed in Table 3 (for the Arithmetic Weights Environment).

Table 3: Parameters for cue number variations (rounded to two decimal places).

	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}	θ	e
6 cues	1.70	1.42	1.14	0.86	0.58	0.30	-	-	-	-	-	-	3.9	1.4
8 cues	1.70	1.50	1.30	1.10	0.90	0.70	0.50	0.30	-	-	-	-	4.9	1.4
10 cues	1.70	1.54	1.39	1.23	1.08	0.92	0.77	0.61	0.46	0.30	-	-	5.9	1.4
12 cues	1.70	1.57	1.45	1.32	1.19	1.06	0.94	0.81	0.68	0.55	0.43	0.30	6.9	1.4

Appendix B: Calculation of occurrence probabilities

The following procedure was used to determine occurrence probabilities for cue profiles in a way that guaranteed the predetermined overall success probability π_s . In a first step, each profile was assigned to one of three classes ($\mathbf{C}_{\pi_s}^+$, $\mathbf{C}_{\pi_s}^0$, or $\mathbf{C}_{\pi_s}^-$) as a result of the comparison between its individual success probability $p_s(\mathbf{C}_i)$ and the predetermined overall success probability, so that $\mathbf{C}_{\pi_s}^+ = \{\mathbf{C}_i \mid p_s(\mathbf{C}_i) > \pi_s\}$, $\mathbf{C}_{\pi_s}^0 = \{\mathbf{C}_i \mid p_s(\mathbf{C}_i) = \pi_s\}$, and $\mathbf{C}_{\pi_s}^- = \{\mathbf{C}_i \mid p_s(\mathbf{C}_i) < \pi_s\}$.

Now, let $n_h = |\mathbf{C}_{\pi_s}^+|$, $n_l = |\mathbf{C}_{\pi_s}^-|$ and $n_0 = |\mathbf{C}_{\pi_s}^0|$. We can then state the average probability of success for profiles with probabilities above and below the predetermined overall probability as $\overline{p_s^h} = \frac{1}{n_h} \sum_{i \in \{x \mid \mathbf{C}_x \in \mathbf{C}_{\pi_s}^+\}} p_s(\mathbf{C}_i)$ and $\overline{p_s^l} = \frac{1}{n_l} \sum_{i \in \{x \mid \mathbf{C}_x \in \mathbf{C}_{\pi_s}^-\}} p_s(\mathbf{C}_i)$, respectively. The occurrence probabilities for cue profiles are then determined by

$$p_o(\mathbf{C}_i) = \begin{cases} \frac{n_h + n_l}{n_h(n_h + n_l + n_0)} \frac{\pi_s - \overline{p_c^l}}{\overline{p_s^h} - \overline{p_s^l}}, & \mathbf{C}_i \in \mathbf{C}_{\pi_s}^+ \\ \frac{n_h + n_l}{n_l(n_h + n_l + n_0)} \frac{\overline{p_s^h} - \pi_s}{\overline{p_s^h} - \overline{p_s^l}}, & \mathbf{C}_i \in \mathbf{C}_{\pi_s}^-, \\ \frac{1}{n_h + n_l + n_0}, & \mathbf{C}_i \in \mathbf{C}_{\pi_s}^0 \end{cases} \quad (4)$$

ensuring that the overall success probability for a set of business plans with these occurrence probabilities is at the predetermined level.

Appendix C: Implementation of Algorithms

Implementation of Equal Weighting

Equal Weighting is based on the number of cues whose values predict success.

Specifically, the learning sample is first used to determine, for each cue, which cue value, if any, is associated with a higher percentage of successful plans. Equal Weighting then counts, for a given cue profile \mathbf{C}_i , the number $w(\mathbf{C}_i)$ of cues whose values point to success. All observed cue profiles that share the same $w(\mathbf{C}_i)$ are grouped into one class, for which it is determined whether it is better to invest or not. This, of course, depends on the financial consequences of investment decisions, specifically on the outcomes for successes after decisions to invest (o_{is}), failures after decisions to invest (o_{if}), successes after decisions not to invest (o_{ns}), and failures after decisions not to invest (o_{nf}).

The fraction $\tau = \frac{o_{nf} - o_{if}}{o_{nf} - o_{if} + o_{is} - o_{ns}}$ represents the relationship of the consequences

incurred by decisions concerning plans that fail and plans that succeed: If a plan fails, the difference between not investing and investing into the plan is $o_{nf} - o_{if}$. If a plan succeeds, the

difference between investing and not investing is $o_{is} - o_{ns}$. These differences reflect the financial impact of a correct versus false decision for unsuccessful and successful plans, respectively. The fraction τ will have values between 0 and 1, with this value being close to 1 when the impact of investment decisions (invest versus not invest) for successful plans is negligible when compared to the corresponding impact for unsuccessful plans. Conversely, the value of τ is close to 0 when the impact of investment decisions for unsuccessful plans is negligible compared to the corresponding impact for successful plans. A value of .5 represents a situation where the two impacts of decisions for successful and unsuccessful plans are the same. In this situation, a VC can gain as much from investing in a successful plan relative to not investing as the VC can lose from investing in an unsuccessful plan relative to not investing. If Equal Weighting is used in this case, then it is sufficient to determine the fraction of successful and unsuccessful business plans in memory within the class defined by $w(C_i)$ and follow the majority in the sense that an investment is made, whenever the majority of plans within this class is successful. In the general case, whether an investment is made or not is based on

$$D(C_i) = \begin{cases} 1, & \frac{|\{C_j \in L \mid w(C_j) = w(C_i) \wedge f(C_j) = 1\}|}{|\{C_j \in L \mid w(C_j) = w(C_i)\}|} > \tau \\ -1, & \frac{|\{C_j \in L \mid w(C_j) = w(C_i) \wedge f(C_j) = 1\}|}{|\{C_j \in L \mid w(C_j) = w(C_i)\}|} < \tau. \\ 0, & \frac{|\{C_j \in L \mid w(C_j) = w(C_i) \wedge f(C_j) = 1\}|}{|\{C_j \in L \mid w(C_j) = w(C_i)\}|} = \tau \end{cases} \quad (5)$$

Whenever $D(C_i) = 1$, the decision is made to invest in C_i and one investment unit is immediately invested. When $D(C_i) = -1$, the decision is made not to invest in the business plan. Finally, when $D(C_i) = 0$, one of the two decisions is selected randomly with equal probabilities. This procedure ensures that, at least for the set of cue profiles in L with identical value of w , the

most profitable decision (invest or not invest) is chosen if it exists (based on the assumption that the business plans in memory are representative of the population, of course).

Implementation of Logistic Regression

The dependent variable is not directly regressed onto the predictors, rather the probability of being associated with a positive outcome value (p_i) is estimated for each profile. Because p_i can only take values within the interval $[0;1]$, p_i is replaced by the so-called link function

$logit(p_i) = \log\left(\frac{p_i}{1-p_i}\right)$ which can take values in the interval $]-\infty;+\infty[$. This log-odds-ratio is, in

turn, regressed onto the predictors in the form $logit(p_i) = \beta_0 + \sum_{j=1}^k c_{ji}\beta_j$.

In this sense, a *Logistic Regression* involves nonetheless a *linear* combination of predictor values. Unlike in OLS regression, the predictors are estimated via a maximum likelihood estimation algorithm (see Agresti, 2002). In the present simulation, Matlab's "glmfit" procedure was used for this task. The probability of success for a given business plan that is estimated by Logistic Regression (\hat{p}_i) is then compared to the ratio of potential losses and gains, similar to the procedure described above for Equal Weighting:

$$D(\mathbf{C}_i) = \begin{cases} 1, & \hat{p}_i > \tau \\ -1, & \hat{p}_i < \tau \\ 0, & \hat{p}_i = \tau \end{cases} \quad (6)$$

Again, the decision to invest or not to invest is based solely on $D(\mathbf{C}_i)$, following the same procedure that is explained above for Equal Weighting.

Implementation of CART

CART (Breiman et al., 1984) is an algorithm to develop binary trees (i.e., trees in which each node has either two or zero child nodes) without a priori constraints pertaining to the tree

structure. CART uses a binary recursive partitioning algorithm for creating a binary decision-tree structure. The process starts by creating the root node that contains the total set of business plans of L . If and only if this number of plans is above a predefined threshold (in our implementation, 10), two child nodes are created from this root node. Each of the (dichotomous) attributes that describe a plan can be used to assign a plan to one of the child nodes. CART chooses the attribute that maximizes the homogeneity of the resulting subgroups with respect to the outcome (technically, this is achieved by minimizing the GINI's diversity index). Each child node that results from this process will, in turn, be used as a new parent node, and the procedure described above will be applied to it and to any further child nodes. This procedure continues until no further splits are possible, which is the case if all end nodes either contain less than 10 cases, or contain business plans with only one outcome. A given end node is associated with the decision (namely, to invest or not) that maximizes the profit that results from applying this decision to all plans that are contained in this node. In our simulation, Matlab's standard "treefit" procedure was used to implement CART.

Implementation of Fast and Frugal Trees

Fast and frugal trees are a class of binary classification trees that are structurally constrained in the following way (see Martignon et al., 2008, for a formal treatment of fast and frugal trees). Following the root node there is exactly one end node at every level, except for the last level, which has two end nodes. Therefore, a tree with m levels (including the root node) has $2m-1$ nodes. On each level, exactly one cue is considered. Depending on the cue value, either an end node is reached or information processing continues down the tree by looking up information on another cue. If the cue value leads to an end node, then an investment decision is made, be it to invest or not to invest. Thus, fast and frugal trees exemplify *one-reason decision*

making. Moreover, it is a non-compensatory strategy because once an end node is reached, no further information is processed and the decision associated with this end node cannot be changed (see Martignon & Hoffrage, 2002).

In a simulation involving 30 empirical data sets, Martignon et al. (2008) compared various algorithms to construct such fast and frugal trees. In the present paper, a variant of the winner in their competition is used, henceforth referred to as the *Fast and Frugal Tree*. Its structure is constrained such that the decisions associated with the end nodes on subsequent levels alternate between investing and non-investing. Within this structure, cues are ordered according to their usefulness to either correctly identify successful or unsuccessful business plans – depending on whether the current level demands that an investment or non-investment decision should be made. This structure helps to protect against all cases being classified in the same way. In some environments enforcing such a balance may exact a price, specifically, if the base rate of successful plans is either very high or very low, or if the two misclassification costs are highly dissimilar. To mitigate such disadvantages, the procedure described above is modified as follows: If the base rates of successful and unsuccessful plans are unequal, then a bias is introduced that allows to have several end nodes of the same type in sequence.² In addition, the misclassification costs and the observed diagnosticity of cue values determine which cues cannot be utilized as a

² Let f_{pos} represent the fraction of successful plans in the learning sample. If $f_{pos} > .5$

then the first $\left\lceil \log_2 \frac{f_{pos}}{1-f_{pos}} \right\rceil + 1$ end nodes are linked to decisions to invest, and if $f_{pos} < .5$ then

the $\left\lceil \log_2 \frac{1-f_{pos}}{f_{pos}} \right\rceil + 1$ end nodes are linked to decisions not to invest.

basis for positive and negative decisions.³ Even though this restriction can lead to a pruned tree or, in a more extreme case, to the same default decision for all plans, they prohibit the use of cues that make a monetary loss highly probable if decisions are based solely on a single cue value.

³ Any cue, for which the percentage of successful plans among plans with a positive cue value in the learning sample is lower than the cost ratio τ is excluded from the set of cues that could form the basis for a decision to invest. Conversely, any cue for which the percentage of unsuccessful plans among plans with a negative cue value in the learning sample lower than $1 - \tau$ is excluded from the set of cues that could form the basis for a decision not to invest.

References

- Agresti, A. (2002). *Categorical Data Analysis*. Hoboken, NJ: Wiley.
- Armstrong, J. S. (2001). Judgmental bootstrapping: Inferring experts' rules for forecasting. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 171–192). Boston, MA: Kluwer Academic Publishers.
- Armstrong, J. S., & Graefe, A. (2011). Predicting elections from biographical information About candidates: a test of the index method. *Journal of Business Research*, *64*, 699–706.
- Åstebro, T., & Elhedhli, S. (2006). The effectiveness of simple decision heuristics: Forecasting commercial success for early-stage ventures. *Management Science*, *52*, 395–409.
- Bergemann, D., & Hege, U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance*, *22*, 703–735.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.
- Brighton, H., & Gigerenzer, G. (this issue). The bias bias. *Journal of Business Research*.
- Bruno, A. V., & Tyebjee, T. T. (1985). The entrepreneur's search for capital. *Journal of Business Venturing*, *1*, 61–74.
- Bruno, A. V., & Tyebjee, T. T. (1986). The destinies of rejected venture capital deals. *Sloan Management Review*, *27*, 43–53.
- Brunswik, E. (1955). In defense of probabilistic functionalism: A reply. *Psychological Review*, *62*, 236–242.
- Burgess, E. (1939). *Predicting success or failure in marriage*. New York: Prentice-Hall.

- Burgman, M. A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L., & Twardy, C. (2011). Expert status and performance. *PLoS ONE*, *6*(7), e22998. doi:10.1371/journal.pone.0022998.
- Chan, Y.-S., Siegel, D., & Thakor, A. V. (1990). Learning, corporate control and performance requirements in venture capital contracts. *International Economic Review*, *31*, 365–381.
- Chrisman, J. J., Bauerschmidt, A., & Hofer, C. W. (1989). The determinants of new venture performance: An extended model. *Entrepreneurship Theory and Practice*, *23*, 5–30.
- Cumming, D., Fleming, G., & Suchard, J. (2005). Venture capitalist value-added activities, fundraising and drawdowns. *Journal of Banking & Finance*, *29*, 295–331.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 97–118). Oxford University Press, USA.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.
- De Clercq, D., & Sapienza, H. J. (2005). When do venture capital firms learn from their portfolio companies? *Entrepreneurship Theory and Practice*, *29*, 517–535.
- DeMiguel, V., Garlapp, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953.
- Denrell, J. (2003). Vicarious learning, undersampling of failure, and the myths of management. *Organization Science*, *14*, 227–243.
- Dhimi, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*, 959–988.

- Dimov, D., & De Clercq, D. (2006). Venture capital investment strategy and portfolio failure rate: A longitudinal study. *Entrepreneurship Theory and Practice*, 30, 207–223.
- Dixon, R. (1991). Venture capitalists and the appraisal of investments. *Omega*, 19, 333–344.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, 85, 395–416.
- Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, 107, 659–676.
- Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (2008). Venture capitalists' evaluations of start-up teams: Trade-offs, knock-out criteria, and the impact of VC experience. *Entrepreneurship Theory and Practice*, 32, 459–483.
- Fried, V. H., & Hisrich, R. D. (1994). Toward a model of venture capital investment decision making. *Financial Management*, 23, 28–37.
- Gifford, S. (1997). Limited attention and the role of the venture capitalist. *Journal of Business Venturing*, 12, 459–482.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristics: Why biased minds make better inferences. *Topics in Cognitive Science*, 1, 107–143.
- Goldstein, D. G., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting*, 25, 760–772.
- Gompers, P., & Lerner, J. (1999). *The venture capital cycle*. Cambridge, MA: MIT Press.
- Gompers, P., & Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives*, 15, 145–168.
- Gompers, P. A. (1998). Venture capital growing pains: Should the market diet? *Journal of Banking & Finance*, 22, 1089–1104.

- Gorman, M., & Sahlman, W. A. (1989). What do venture capitalists do? *Journal of Business Venturing, 4*, 231–248.
- Graefe, A. (this issue). Improving forecasts using equally weighted predictors. *Journal of Business Research*.
- Gupta, A. K., & Sapienza, H. J. (1992). Determinants of venture capital firms' preferences regarding the industry diversity and geographic scope of their investments. *Journal of Business Venturing, 7*, 347–362.
- Hall, J., & Hofer, C. W. (1993). Venture capitalists' decision criteria in new venture evaluation. *Journal of Business Venturing, 8*, 25–42.
- Hammond, K. R. (1966). Probabilistic functionalism: Egon Brunswik's integration of the history, theory, and method of psychology. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik* (pp. 15–80). New York: Holt, Rinehart and Winston Inc.
- Harvey, N., & Fischer, I. (2005). Development of experience-based judgment and decision making: The role of outcome feedback. In T. Betsch, & S. Haber stroh (Eds.), *The routines of decision making* (pp. 119–137). Mahwah, NJ: Lawrence Erlbaum.
- Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review, 112*, 494–508.
- Hisrich, R. D., & Jankowicz, A. D. (1990). Intuition in venture capital decisions: An exploratory study using a new technique. *Journal of Business Venturing, 5*, 49–62.
- Hoffman, C. A. (1972). *The venture capital investment process: A particular aspect of regional economic development*. Ph.D. thesis University of Texas Austin, TX.
- Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review, 114*, 733–758.

- Kaplan, S. N., & Strömberg, P. (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *Journal of Finance*, *59*, 2177–2210.
- Khan, A. M. (1987). Assessing venture capital investments with noncompensatory behavioral decision models. *Journal of Business Venturing*, *2*, 193–205.
- Lockett, A., & Wright, M. (2001). The syndication of venture capital investments. *Omega*, *29*, 375–390.
- MacMillan, I. C., Siegel, R., & Subba Narasimha, P. (1985). Criteria used by venture capitalists to evaluate new venture proposals. *Journal of Business Venturing*, *1*, 119–128.
- Maier, J. B., II, & Walker, D. A. (1987). The role of venture capital in financing small business. *Journal of Business Venturing*, *2*, 207–214.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29–71.
- Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology*, *52*, 352–361.
- Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman, & L. Macchi (Eds.), *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making* (pp. 189–211). Chichester: Wiley.
- Mason, C. M., & Harrison, R. T. (2002). Is it worth it? The rates of return from informal venture capital investments. *Journal of Business Venturing*, *17*, 211–236.
- Meggison, W., & Weiss, K. (1991). Venture capitalist certification in initial public offerings. *Journal of Finance*, *46*, 879–903.

- Muzyka, D., Birley, S., & Leleux, B. (1996). Trade-offs in the investment decisions of European venture capitalists. *Journal of Business Venturing, 11*, 273–287.
- Petty, J. S., & Gruber, M. (2011). “In pursuit of the real deal”: A longitudinal study of VC decision making. *Journal of Business Venturing, 26*, 172–188.
- Riquelme, H., & Rickards, T. (1992). Hybrid conjoint analysis: An estimation probe in new venture decisions. *Journal of Business Venturing, 7*, 505–518.
- Robinson, R. B. J. (1987). Emerging strategies in the venture capital industry. *Journal of Business Venturing, 2*, 53–77.
- Sahlman, W. A. (1990). The structure and governance of venture-capital organizations. *Journal of Financial Economics, 27*(2), 473–521.
- Sapienza, H. J., Amason, A. C., & Manigart, S. (1994). The level and nature of venture capitalist involvement in their portfolio companies: *A study of three European countries. Managerial Finance, 20*, 3–17.
- Shepherd, D. A. (1999a). Venture capitalists’ assessment of new venture survival. *Management Science, 45*, 621–632.
- Shepherd, D. A. (1999b). Venture capitalists’ introspection: A comparison of “in use” and “espoused” decision policies. *Journal of Small Business Management, 37*, 76–87.
- Shepherd, D. A., Armstrong, M. J., & Lévesque, M. (2005). Allocation of attention within venture capital firms. *European Journal of Operational Research, 163*, 545–564.
- Shepherd, D. A., & Zacharakis, A. (2002). Venture capitalists’ expertise: A call for research into decision aids and cognitive feedback. *Journal of Business Venturing, 17*, 1–20.

- Shepherd, D. A., Zacharakis, A., & Baron, R. A. (2003). VCs' decision processes: Evidence suggesting more experience may not always be better. *Journal of Business Venturing, 18*, 381–401.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.
- Sorenson, O., & Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology, 106*, 1546–1588.
- Tetlock, P. E. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton University Press.
- Todd, P., Gigerenzer, G., & the ABC Research Group (2012). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
- Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science, 16*, 167–171.
- Tyebjee, T. T., & Bruno, A. V. (1984). A model of venture capitalist investment activity. *Management Science, 30*, 1051–1066.
- von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General, 137*, 73–96.
- Zacharakis, A., & Shepherd, D. A. (2005). A non-additive decision-aid for venture capitalists' investment decisions. *European Journal of Operational Research, 162*, 673–689.
- Zacharakis, A. L., & Meyer, G. D. (1998). A lack of insight: Do venture capitalists really understand their own decision process? *Journal of Business Venturing, 13*, 57–76.
- Zacharakis, A. L., & Meyer, G. D. (2000). The potential of actuarial decision models: Can they improve the VC investment decision? *Journal of Business Venturing, 15*, 323–346.