DRUID Working Paper No. 12-12

Two's Company: Human Capital Composition and Performance of

Entrepreneurial Pairs

By

Alex Coad and Bram Timmermans

Danish Research Unit for Industrial Dynamics

www.druid.dk

Two's Company: Human Capital Composition and Performance of Entrepreneurial Pairs

Alex Coad
SPRU – University of Sussex and
Aalborg University
Department of Business and Management
E-mail: acoad@business.aau.dk

Bram Timmermans
Aalborg University
Department of Business and Management
Fibigerstræde 4, DK-9220 Aalborg Oe
E-mail: bram@business.aau.dk

October 31, 2012

Abstract:
We study the effects of diverse team composition on the survival and growth of new ventures using the Danish Linked Employer-Employee database. To get cleaner measures of diverse team composition, we focus on entrepreneurial dyads, and also investigate the asymmetric effects of team composition by distinguishing between the `primary' and the `secondary' founder. We complement existing work by showing that heterogeneity in team composition is affected by the asymmetric hierarchical structure within the team, and that a unidimensional diversity indicator (which is usually applied) fails to capture a number of performance effects of heterogeneous team composition. Ventures with a STEM-educated primary founder and a Business-educated secondary founder have high employment growth, while the opposite combination (Business first, STEM second) has low employment growth. Pairs of younger individuals have lower survival chances but higher employment growth. Family firms have lower employment growth, especially when formed with your mother.

www.druid.dk

www.druid.dk

1   Introduction

From the late 1980s onwards, we have observed a gradual shift from treating entrepreneurship as an act of one lone individual towards entrepreneurship as a collective activity (Cooney, 2005; Harper, 2008).[1] Consequently, researchers within the field of entrepreneurship started to investigate the composition of these entrepreneurial teams (e.g. Ruef et al., 2003). In close relation to identifying this composition, there is also an interest in investigating whether composition affects the performance of these teams, which varies from member entry and exit to growth and survival, and if so what the nature of this relationship is. Studies have focused on various compositional measures of teams; for example, industry experience (Delmar and Shane, 2006), entrepreneurial experience (Ucbasaran et al., 2003; Delmar and Shane, 2006), and prior affiliations of team members (Beckman et al., 2007). However, inspired by the upper echelon theories on diversity in top management teams (Hambrick and Mason, 1984), there has been an increased focus on diversity in entrepreneurial teams arguing that the heterogeneity of these teams affects how they work together, which ultimately will affect their performance (Gartner, 1985; Roure and Maidique, 1986; Ensley et al., 1998). Indeed, Ensley et al. (2002) argue that "the richest and most interesting studies of TMTs are likely to involve new ventures" (p. 381).

Not surprisingly, and in accordance with studies on top management teams, the impact of diversity is inconclusive. Studies have not specified whether, how and why team diversity positively or negatively affects the performance of start-ups. This can attributed to: (i) the assumption that the same level of diversity might have a different impact on various performance indicators, i.e. diversity might be good for firm growth but bad for firm survival, (ii) that the approach in investigating team composition is too ambitious by constructing an overall (scalar) diversity measure, and (iii) studies investigate entrepreneurial teams of different sizes, thereby introducing an undesirable level of complexity, making it even more difficult to estimate the impact of diversity on performance. In this paper, we will focus on a subset of entrepreneurial teams – dyads. This approach will provide us with cleaner measure of heterogeneous team composition, and a focus on dyads is a theoretically meaningful way of analyzing entrepreneurial teams (Harper, 2008). In addition, it is one of the most common forms of entrepreneurial team size (Ruef et al., 2003), which is supported by many studies on entrepreneurial teams report

---

[1]Decades earlier, researchers already started to promote the notion of entrepreneurial teams but the work of Kamm et al. (1990) and Gartner et al. (1994) called for a more systemic approach in studying the performance of entrepreneurial teams (Ruef, 2010).

averages sizes between two and three members (Eisenhardt and Schoonhoven, 1990; Ucbasaran et al., 2003; Chowdhury, 2005; Clarysse et al., 2007) indicating the bulk of the distribution is represented by two-person teams.

To investigate the issue, we use the Danish Integrated Database for Labour Market Research (IDA) to identify these entrepreneurial pairs. This database provides detailed information on the demographic characteristics of individuals, e.g. age, gender, education, and the dynamics of organizations, i.e. birth, growth and exit of firms, which allows us to analyse the relations that exist between the demographic characteristic of the entrepreneurial team and new venture performance. The direction of the diversity is determined by the position of the individuals in the firm where a higher ranked individual, based on ownership and occupation code, is considered to be the *primus motor* of the start-up. We select a sample of 3777 entrepreneurial dyads in the Danish private sector in the period 1999-2003 and follow these start-ups for five years after founding.

We contribute to the literature in a number of ways. We investigate the effect of diversity on performance using an especially rich dataset that contains details on a number of variables including educational background and family ties. While much previous work has focused on small samples, we provide, as suggested by Vanaelst et al. (2006), representative large-sample evidence using detailed administrative data. We investigate the performance of new businesses in terms of both survival and employment growth. While previous work has grouped together ventures of different ages, we observe new ventures from their first year of business (as indicated by their date of official registration). In response to calls for diversity research to focus more on dynamic effects (Horwitz and Horwitz, 2007), we exploit our longitudinal data to consider lagged effects of diversity (that is, the effects of heterogeneous combinations of start-up and pre-start characteristics on five-year performance). Furthermore, we contribute to the literature by moving on from assuming that power relations are symmetric between team members – we distinguish between the primary and secondary founder, and investigate which characteristics matter for each of the two founders.

The preliminary analysis indicates that, when focusing on entrepreneurial pairs, there is indeed a difference in performance depending on the direction of this diversity. With regards to education, the best performing firms are not composed of similar individuals. Ventures with a STEM-educated primary founder and a Business-educated secondary founder enjoy relatively high employment growth, while interestingly enough the oppo-

site combination (Business first, STEM second) has low employment growth. Pairs of younger individuals have lower survival chances but higher employment growth. Performance of mixed-race and mixed-gender ventures depends upon the identity of the primary founder. Family firms have equal survival chances but lower employment growth – consistent with suggestions that they persist for an unnecessarily long period of time.

The remainder of the paper is structured as follows. We review the related literature in Section 2 and formulate some propositions. Our methodology is described in Section 3. We present our data in Section 4. Section 5 contains our analysis, where we begin with non-parametric representations of team composition and performance before moving on to parametric regressions. Section 6 contains a synthetic discussion of our findings and revisits our propositions. Section 7 concludes.

2   Background

2.1   Related literature

Issues on team diversity are not a new phenomenon; on the contrary, a survey of the literature indicates that there exists a long tradition in linking the diverse composition of teams with their performance (see, e.g., Williams and O'Reilly (1998) and Horwitz (2005) for a literature review). However, a closer inspection of these studies reveals that the interest is traditionally based upon teams in larger organizational settings, e.g. top management and product development teams (Murray, 1989; Bantel and Jackson, 1989; Ancona and Caldwell, 1992; Pelled, 1996; Dahlin et al., 2005). More recently, studies that investigate the diverse composition of entrepreneurial teams have emerged and an increase in the number of such studies is visible. This steady increase runs parallel with the increased focus on entrepreneurial teams in general (Cooney, 2005; Harper, 2008).

Studies that investigate the composition of entrepreneurial teams (e.g. Baron et al., 1999; Ruef et al., 2003; Steffens et al., 2011; Kaiser and Müller, 2012) show that entrepreneurial teams are mainly characterized by homophily, at least regarding gender, ethnicity and occupation (more visible characteristics), while we can observe more heterogeneity in terms of functionality and status.[2] The homophily in these teams can be explained by the social selection mechanism behind recruitment that is often driven by interpersonal attraction (Forbes et al., 2006); not only because these teams rely on social networks

---

[2]The same also holds for other organizational units that rely on voluntary participation (McPherson et al., 2001).

(Aldrich and Langton, 1998; Aldrich and Ruef, 2006), which are homogeneous (McPherson et al., 2001), but also based on the other recruitment channels. The underlying rationale is that interpersonal attraction based on the demographic attributes will cause less (personal) trouble in start-ups (Beckman et al., 2007); consequently, the limited resources will be used to deal with the liability issues that start-ups face.

In contrast to the above-mentioned perspectives on the importance of homophily, there are studies that stress the positive impact of diversity on performance as a result of the unique set of skills, abilities and knowledge that are brought into the team (Cox and Blake, 1991; Hambrick et al., 1996; Williams and O'Reilly, 1998; Horwitz, 2005). This line of argument is similar to other approaches within management theory, in particular the resource-based view of the firm, which argues that a heterogeneous resource composition, including human resources, determines a firm's competitive advantage (Barney, 1991). Within the upper echelon studies on top management teams, which have their origin in Hambrick and Mason (1984), it is widely accepted that it is important that these teams collectively possess the skills that are necessary to run a successful business (Beckman et al., 2007). The majority of studies on entrepreneurial teams share this perspective as their superior performance compared to solo entrepreneurs is believed to be driven by the access to various forms of human capital and the presence of different perspectives (Kamm et al., 1990; Eisenhardt and Schoonhoven, 1990; Watson et al., 1995). Nevertheless, most studies on these entrepreneurial teams focus on human capital theory and look at overall team characteristics (e.g. average level of education or length of experiences) to explain performance (Beckman et al., 2007). This approach is very helpful in explaining the performance of individual entrepreneurs (Davidsson and Honig, 2003) but fails to capture the impact of the diversity in skills that are present in a collective. By adopting an organizational demography approach, we can consider both the average characteristics of the human resources in ventures and the differences between the human resources (Beckman et al., 2007).

The above-mentioned theoretical approaches provide sound but contradictory arguments on the potential effect of team diversity on team performance. It is therefore not surprising that empirical studies have found both positive, negative and non-significant effects of diversity in entrepreneurial teams. This inconsistency is illustrated by the selection of studies on diversity in entrepreneurial teams and the impact on various performance indicators in Table 1. The conflicting results found in previous work on small samples motivates our interest in a large-scale analysis. Another, rather unexplored perspective

on why there is so much ambiguity of diversity in entrepreneurial teams might be because of the rather reductionist way in which studies have treated team heterogeneity. In the remainder of this paper we will address this issue empirically by investigating how a diverse composition in a particular subset of entrepreneurial teams – dyads – affects new venture performance.

Table 1: Studies on Entrepreneurial Teams and Diversity

Study	Dependent Variable	Teams	Age	Gender	Education	Industry and/or Work Experience	Functional Background	Skill Composition
Amason et al. (2006)	Sales growth (LN/HN)†	174	ns (-/+)		ns (ns/-)		ns (ns/-)	
	Profitability (LN/HN)†		ns (ns/-)		ns (ns/-)		ns (ns/-)	
	Market Performance (LN/HN)†		ns (ns/-)		ns (ns/-)		ns (ns/-)	
Beckman (2006)	Exploration strategy	170						+
	Exploitation strategy							ns
	Speed to market							ns
	Firm growth							+
Beckman et al. (2007)	Ability to attract VC	161						+
	Succeed an IPO							+
Chandler et al. (2004)	Entry of team members	124		ns	+	+		
	Exit of team members			ns	ns	+		
Chowdhury (2005)	team effectiveness	79	ns	ns			ns	
Clarysse et al. (2007)	Team member addition	140	ns				ns	
Eisenhardt and Schoonhoven (1990)	Organizational growth	92				+		
Ensley et al. (1998)	Sales growth	88						-
	Profitability							ns
	Revenues							-
Foo et al. (2005)	External evaluation of ideas	154	-	ns	+			
Steffens et al. (2011)	Short term performance	202	ns	ns				
	Long run performance		+	ns				
Ubasaran et al. (2003)	Entry of team members	92					ns	
	Exit of team members						ns	
Vanaelst et al. (2006)	Team development‡	10			(.)			(.)
Watson et al. (2003)	Profit	175		ns	ns			
	Growth			ns	ns			

Note: we only report the most common demographic and human capital dependent variables. Other measures used but not included in the table are diversity in religious and political background, and various constructs, e.g.: conflict, task related diversity, non-task related diversity etc.
† LN=low novelty, HN=high novelty
‡ case study

2.2 Development of propositions

The previous literature has generally formulated hypotheses in terms of how diversity in one particular dimension (e.g. age, gender, race, prior professional affiliations) affects the performance of the firm. For example, Pelled et al. (1999) distinguish between task conflict and emotional conflict, while Foo et al. (2005) write in terms of task-based and non-task based diversity. Task-based diversity can have beneficial effects on performance if different perspectives lead to superior problem-solving capabilities (Østergaard et al., 2011), while non-task related diversity will offer no such advantages (Foo et al., 2005). In our analysis, we focus primarily on a number of human capital variables – that is, age, education (level and type), and prior industry experience. Diversity in age can have advantages if energetic youth can be combined with the wisdom that accompanies age. Diversity in education type can lead to a broader set of available skills and benefits of specialization. We do not expect any benefits from diversity in education level, however, because individuals will need a common level of education in order to effectively communicate their diverse perspectives. Diversity in prior industry experience is not expected to be an advantage, although it may be that it is sufficient if one of the partners has prior industry experience. We also control for diverse combinations in terms of gender, nationality, and marital status, although we do not hypothesize any particular advantages for these variables because we classify them as non-task-related characteristics.

Our main theoretical focus, however, is on developing some 'propositions' to loosely guide our empirical investigations, that will be used to evaluate the validity of our empirical approach. To begin with, we deliberately distinguish between the 'primary' and 'secondary' founding entrepreneur in our analysis, and suggest that the effect of diversity on performance is not invariant to which individual has which characteristics. For example, it may be that entrepreneurial pairs need one brash, energetic young individual to take the leading role, with an older and wiser individual acting as a 'guiding hand.' It may also be that the primary founder needs to have sound technical knowledge of the product, while benefitting from commercial advice from a supporting partner.

Proposition 1 *Structures of power and authority within teams are not symmetric, and the 'direction' of heterogeneity moderates the effect of team composition on performance*

We also take a non-standard approach to measuring heterogeneity of team composition, because we suspect that the standard practice of reducing heterogeneity to a single

summary scalar index of diversity leads to the loss of considerable information on team composition. Consider the variable age: first of all, we suspect that age has a non-linear effect on performance (from the liability of youth to the 'golden age' to senescence). A second drawback is that it is likely that 10 years difference in age matters more when the two founders are on average 25 years old than when they are both on average 60 years old.

Therefore we posit:

Proposition 2 *Team composition cannot easily be reduced to a single summary scalar index of diversity, because many interesting effects will remain hidden*

Another feature of our paper is that we have two performance indicators: survival and employment growth. While each of these indicators is associated with firm performance, they shed light on different facets of performance. We prefer growth as an indicator of success, because some firms may survive and persist even if they experience poor performance (the so-called 'living dead').

Proposition 3 *Heterogeneity of team composition has different effects for survival and growth*

3   Method

In the majority of studies on team diversity, diversity is defined as a function of differences among team members with respect to a common attribute. Consequently, diversity is often regarded as a unit-level compositional construct (Harrison and Klein, 2007). Overall, diversity on these attributes can be measured on three dimensions: variety, balance, and disparity (Stirling, 2001; Harrison and Klein, 2007). Variety takes into account the number of categories within a certain attribute where more categories result in higher diversity. With balance, the shares of the specific category are measured and a more equal balance between categories results in a higher degree of diversity. Disparity refers to the distance between the outer boundaries of the various categories within one characteristic. Harrison and Klein (2007) distinguished between separation and disparity where the first relates to horizontal differences (diversity based on opinions or expertise), and the latter on vertical differences (diversity based on hierarchy or power). To study this diversity, we will follow the methods proposed in the existing work on

team diversity. The majority of these studies have used the techniques of organizational demography. This means that the level of diversity is measured based on observable demographic characteristics, where demography is defined as: "the composition, in terms of basic attributes such as age, sex, educational level, length of service, race and so forth of the social unit under study" (Pfeffer, 1983, p. 303).

Such an empirical strategy leads to several challenges when investigating the impact of diversity on the performance of the team. First, researchers create a unidimensional summary indicator of diversity for each attribute; this approach does not take into account that diversity might in reality have a 'directional' character, in the sense that the value of an individual's characteristics is moderated by their position in the hierarchy. Second, there is the challenge on how to find a concise representation of the high dimensionality, i.e. large teams have more nodes leading to a higher level of complexity.

3.1   Focus on pairs only

To keep the dimensionality manageable, we focus on entrepreneurial pairs. Focusing on these dyads is a theoretically meaningful way of simplifying the analysis of entrepreneurial teams (Harper, 2008). With pairs, there is only one possible relationship in which heterogeneity can be measured – that is, the relationship of A to B. With triads, one may look at the heterogeneity between A and B, or A and C, or B and C; and the analysis of heterogeneity becomes even more complex with four or more founders. Another main reason why we focus on pairs is that, contrary to other studies that investigate entrepreneurial team performance, we consider that entrepreneurial teams of different sizes are qualitatively different. In pairs, for example, there is always the tension of a head-on conflict, and disputes are resolved essentially through the mechanism of 'my word against yours.' In keeping with insights from geometry (that is, the stability of triangular structures), an entrepreneurial team of three founders will have more stability as the dynamics of majority rule is more flexible, with each individual taking turns as the swing voter and arbiter, and being able to move from side to side to form new majority coalitions with one of the two others. With teams of four individuals, there may be a tendency to split into rival groups (of pairs) within the team, for individuals to seek strong pair-bonds within the team, or for minority views to acquiesce relatively easily. In short, there may be nonlinearities between number of team members and the nature of diversity within the team, because integers can be seen as being qualitatively different (Schimmel, 1994). Teams of different sizes have fundamentally different opportunities

for specialization, that do not scale up with team size in a linear way. To keep our observations as comparable as possible, we focus only on the most numerous team-size, which is the team of two individuals.

The final reason to focus only on entrepreneurial dyads is grounded in this form is a very common size, if not the most common, in the entrepreneurial team literature. Many existing studies report, including the majority of the studies reported in Table 1, average sizes between two and three members (Eisenhardt and Schoonhoven, 1990; Ucbasaran et al., 2003; Chowdhury, 2005; Clarysse et al., 2007; Beckman, 2006; Beckman et al., 2007; Steffens et al., 2011). Despite of the existence of larger teams, particular in more high-tech areas, the bulk of the size distribution corresponds to two-person teams.

3.2 Quantifying diversity

Table 2: Indicators of diversity used in the literature

Indicator	Formula	Types of variables	Examples
Coefficient of variation	$c_v = \frac{\sigma}{\mu}$	Continuous	Pelled et al. (1999, p11); Foo et al. (2005, p393)
Herfindahl-Hirschman index	$H = 1 - \sum_{l=1}^{l}(P_i)^2$	Categorical	Pelled et al. (1999, p11); Foo et al. (2005, p393); Beckman et al. (2007, p157)
Shannon index	$H = - \sum_{l=1}^{l} P_i(lnP_i)$	Categorical	Pelled et al. (1999, p16); Ucbasaran et al. (2003, p116); Beckman et al. (2007, p156)

Notes: The Shannon index is referred to as Teachman's index in Pelled et al. (1999, p16) and Ucbasaran et al. (2003, p116). The Herfindahl-Hirschman Index is referred to as Blau's heterogeneity index in Pelled et al. (1999, p11).

Table 2 summarizes the most common indicators of diversity used in the literature. We will argue that these measures of diversity have a number of drawbacks. First of all, the numerical value of such an index may have no intuitive interpretation. Second, we may be interested in asymmetric roles (due to power structures in a hierarchy) for individuals i and j, instead of assuming the two to be interchangeable. Third, the benefits of diversity may vary across the distribution of x (for example, being ten years younger may be more important if your partner is 30 than if your partner is 60). This will be difficult to quantify without making the results difficult to interpret. Therefore, instead of trying to quantify diversity, we instead aim to present information on diverse team compositions in the most accessible way possible.

In our view, the standard scalar indicators of diversity suffer from problems related to extreme reductionist simplification. For example, a team of two men and one women

is treated as having an identical gender composition as a team of two women and one man, or even a team of four women and two men (because firm size is seldom interacted with the diversity indices). The maximum possible amount of diversity also depends on the group size (e.g. the maximum score for gender diversity in a team of three is not the same as the maximum score for a team of four). With regards to information on educational background, standard diversity measures provide information on the number of different backgrounds but they remain mute on *which* backgrounds are represented. For example, when looking at the diversity of educational backgrounds, a team where everyone has a STEM background is indistinguishable (to the econometrician) from a team where everyone has a business background (both teams would have zero diversity). To deal with these problems, we develop a less parametric approach to investigating diversity and performance, by representing heterogeneity in terms of coordinates in an n-dimensional Euclidean disparity space (Stirling, 2007).

3.3   Empirical strategy

We begin with some non-parametric illustrative statistics of the performance of pairs to give the reader an intuitive grasp of the composition of teams and their performance outcomes. Anticipating that most readers will not have a 3D printer, and furthermore in recognition of the fact that the human brain is not well adapted to considering graphs containing three or more dimensions (which would be problematic if we had more than two team members), we plot the two founders on two axes and report the outcome in the resulting two-dimensional disparity space, using contour plots and cross-tabulations.

We then complement our 'raw' non-parametric results with parametric regressions, that have the advantage of allowing us to include control variables. In our parametric regressions, we prefer not to collapse information on diversity into a single summary diversity index, because this might not have a ready or 'intuitive' interpretation. Instead, we include a dummy variable for each category of combinations of partners. This gives us a different problem – that of having to include a large number of dummy variables for each pair-wise combination of characteristics. To deal with this latter issue, we adopt a 'stepwise' regression approach, whereby we repeat our regressions in iterative progression, at each step removing the least significant variable, and proceed until all of the remaining explanatory variables are above a minimum threshold level of significance.

4   Data

To investigate whether the direction of the employment diversity affects the performance of the new venture, we make use of the information gathered from Danish government registers. This database, which is maintained by Statistics Denmark, is known under the name Danish Integrated Database for Labor Market Research (from now on referred to by its Danish acronym IDA). IDA is suitable for the analysis as its longitudinal characteristic allows us to follow individuals, establishments and firms over time. As a result, firm dynamics (birth, death and growth rate of firms) and the employment history of the active labor force can be identified. The database holds information on various demographic characteristics, such as gender, age, country of origin, type and level of education, which university the individuals attended, occupation and work experience. Because these individuals can be matched to a firm at any given year, it is possible to observe the team composition of the start-up.[3]

4.1   Start-ups, Entrepreneurial Pairs, and Directionality

To conduct the various analyses, we created a sample of all start-ups in the period 1999 to 2003 where we exclude all start-ups in the primary and public sector.[4] Table H.2 in the Appendix shows a more detailed industry breakdown of ventures in our sample. The motivation for selecting the time-period is two-fold. First, we want to be able to use the growth in sales as one of the firm growth measures; due to the break in the data between 1998 and 1999 it is problematic to include start-ups founded prior to 1999. Second, we want to follow the start-up for a number of years after founding to identify whether they survive and to establish their growth rates. The current dataset has data up to 2008, which allows us to follow each start-up for up to at least five years after founding.

To select our sample of start-ups it is important to identify the founding year. To do so, we use information on the firm's founding date from the company register in combination with the plant and firm identification number. We identify a start-up as a one-plant firm with no prior firm and plant identification number, which is in line with Dahl and Reichstein (2007). Furthermore, to select genuinely new firms, we exclude all start-ups that are the result of a separation or merger of previously existing plants. Based on the

---

[3]See Timmermans (2010) for a more detailed description of the database.

[4]Start-ups that are not within the 15 and 75 two-digit level NACE code are excluded. Within these two two-digit codes there is one classification, 40 to 45 (energy), that is a mix of both public and private firms, which also will be omitted.

above-mentioned selection criteria we identify 12,861 start-ups in the period 1999-2003.

To identify the disparity we need to identify the persons that are involved in the start-up in the year of founding. These persons are identified by merging two datasets: i.e. (i) the entrepreneurship database, which provides detailed information on who is the owner of the start-up; and (ii) the employee dataset that provides information on a person's primary workplace. We add all these individuals to identify the size of the start-up in the year of founding. We are aware that a firm changes its human resource composition, but the motivation for choosing the human resources in the first year is: (i) the observation that most firms start small and hardly change in size during their lifetime (Aldrich and Ruef, 2006); (ii) the initial resource profile can be used to predict start-up performance, including failure (Cooper et al., 1994); and (iii) founder characteristics (Boeker, 1989), early hiring decisions (Baron et al., 1999), and strategies at start-up (Romanelli and Tushman, 1994) have lasting consequences for new organizations.

As explained in Section 3 we will focus only on two-person start-ups.[5] Similarly to the studies reported in Section 3, entrepreneurial pairs is the most common team size. This selection criteria will decrease the sample to 4002 new ventures. More descriptive statistics on the overall sample, including the performance indicators and demographic variables are presented in Table H.1 in the Appendix.

Since we investigate the directionality of the disparity, we need to assign a *primus motor* (individual i) for each two person start-up. To do so, we conduct several steps to find this individual, which is a combination of ownership, occupation code and the length of the attachment to the workplace. The majority of the our primary individuals are identified as the owner of the new venture, i.e. 2487. In 183 pairs the primary individual is identified as a director of the business and 1070 cases both are registered as an employee but then the highest employee rank in combination with the number of days of employment enables us to identify a primary individual. For 225 cases it is not possible to identify a direction and these cases are excluded from our sample; consequently, we end up with a sample of 3777 entrepreneurial pairs. Afterwards, we identify the disparity of this pair in terms of age, education (both degree and discipline) and industry experience, as well as other common diversity indicators (gender, nationality, and civil status).

---

[5]We only restrict our sample where two individuals have a primary attachment to the new venture in the founding year. Other individuals might be connected to this new venture but this venture is not their primary employment. Note also that it is possible that a team of two founders takes their first employee within the first year.

14

4.2 Survival and Growth

As mentioned above, we investigate the impact of this directional disparity on firm performance. The first performance measure is firm survival – due to the unique identification number associated to firms and plants we can follow the status of these organizational units in all years up to a change in this identification number.[6] A change in this identification number is always connected to a variable that indicates the status; when this status is identified as being a closure we consider it to be a non-survivor. In reality, firms might re-enter into the same or in different industries; however, for analytical purposes we will not consider this as a new entry and these firms do not re-appear in the sample. In addition to the closure of a business a firm might continue in another form, e.g. as a result of a merger or acquisition. We will treat these firms as survivors but these observations will be censored due to the structural change of these firms. This is the case for 114 firms in the sample. In total, 1256 firms (that is, 33.33%) survive up to the fifth year.

We also investigate the impact of team composition on employment growth. We measure growth in terms of number of employees,[7] and track the employment growth of the firm after 5 years. It is straightforward for us to measure the employment growth of our firms, because they all start with two individuals – we need only consider the number of employees in year 5.[8]

---

[6]Timmermans (2010) discusses in more detail when firms and plants change identifier. In short, IDA follows a person-oriented approach towards change. Consequently, an establishments identification number remains the same from one year to the other whenever one of the following criteria is fulfilled: (i) a plant has the same owner and is active in the same industry; (ii) a plant has the same owner and the same labor force; or (iii) a plant has the same labor force and is located on the same address or is active in the same industry.

[7]Measured in terms of an employee headcount that takes place on a date in November of each year, as chosen by the government administration.

[8]Although we also have data on sales, value-added and profits, in this paper we focus on growth measured in terms of employees. This is easier to count, it is a relatively transparent indicator, it makes for meaningful comparisons across sectors and years, and furthermore it makes it easy for us to calculate growth because all firms have the same initial size (i.e. 2 founders). We repeated our analysis by looking at growth of sales, value-added and profits, and got similar results, but for the sake of space they are not reported here.

5 Analysis

5.1 Non-parametric analysis

In the following subsections we will present the non-parametric analysis that provides us with a first indication of the effect of team composition on new venture performance. We will start with the impact of age, followed by education (both the level and the type of education), and industry experience.

5.1.1 Age

We start with presenting some non-parametric analysis on the age of the founders. As presented in Table H.1 in the Appendix, the average age of the individuals in the sample is around 35 years of age.

Figure 1 shows the survival rates of entrepreneurial pairs, conditional on the age of individuals i and j. On the whole, there appears to be a rather uniform pattern – the survival of startups seems to be constant across the distribution of ages of i and j. That said, we observe a slightly longer survival of firms depending on the age of the entrepreneurial team, i.e. where either i or j is old.

Figure 2 shows the employment growth outcomes associated with different partnership combinations according to age. A first observation is that the best performing ventures, in terms of employment growth, are those where the primary founder has an age of around 20, while the secondary partner has an age of around 30. This suggests that both partners should be relatively young, to cope energetically with the workload of starting a new venture, although the secondary founder should be noticeably older than the first. Hence, some diversity in age can be valuable. Other regions associated with high employment creation are also visible, such as when age_i=45 and age_j=40, or when age_i=55 and age_j=30. A second observation is that job creation generally seems to decrease with age of both the primary and the secondary partner, although the relationship is not smooth or linear. To the extent that the best outcomes are not on the 45° line (the 'diagonal'), Figure 2 provides early evidence that diversity in age can be beneficial.
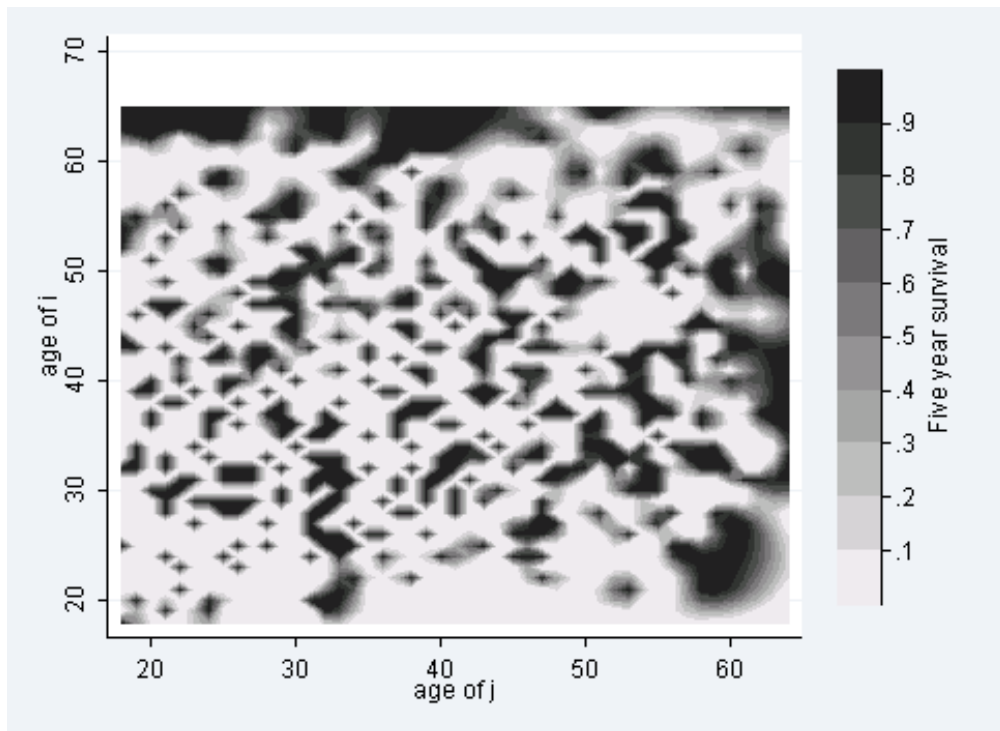
Figure 1: Contour plot of the survival of entrepreneurial pairs. z-axis: average survival after 5 years. Contour plot produced using thin-plate-spline interpolation.
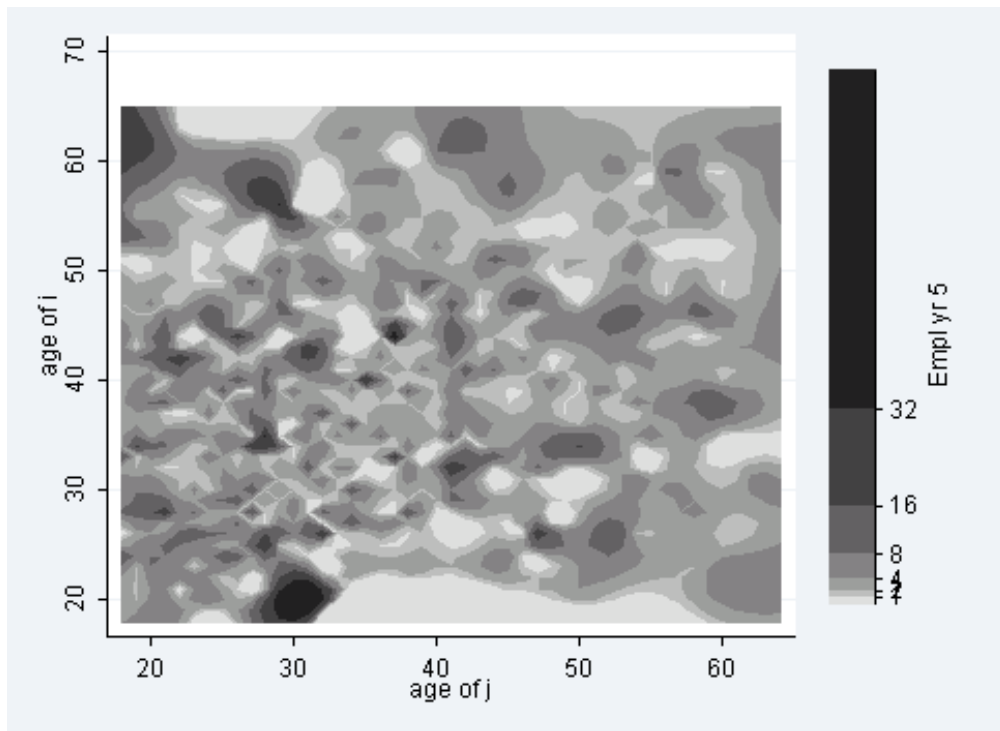
Figure 2: Contour plot of the outcomes associated with entrepreneurial pairs. z-axis: employment after 5 years, measured in terms of number of employees at the date of annual compulsory registration (in November of each year). Contour plot produced using thin-plate-spline interpolation.

			j							j		
		1	2	3	4	5			1	2	3	4	5
	1	0.241	0.327	0.353	0.286	0.227		1	3.909	4.677	4.111	3.833	5.000
	2	0.319	0.441	0.457	0.385	0.350		2	4.411	4.224	4.952	4.500	7.714
i	3	0.373	0.520	0.409	0.235	0.125	i	3	4.760	5.654	5.111	9.750	2.000
	4	0.341	0.448	0.444	0.565	0.353		4	4.696	3.821	6.875	4.269	5.333
	5	0.241	0.333	0.273	0.154	0.625		5	3.846	4.923	7.000	6.250	10.080

Figure 3: Performance after 5 years, by education level, for individuals i (primary founder) and j. Cells with above-median values are highlighted. Left: survival; right: means of numbers of employees.

5.1.2 Education

In our analysis of the impact of directional disparity in education, we would like to take into account both the level of education and the type of education. The structure of the education variable in IDA allows us to identify the level of education (based on the first two-digits of an eight-digit code) and the discipline being taught (digit three and four of the eight-digit education code).[9] We begin by considering education level before moving on to education type.

Education level   In preparing the sample for the non-parametric analysis on education, we have merged the different education codes in five education level dummies, i.e.: for all up to (and including) high-school (1); vocational training (2); Vocational short cycle education (3); Bachelor (4), which includes professional and academic bachelor; Post graduate (5), which includes Master and PhD graduates. For this analysis we also drop individuals for who we do not know the education they obtained.

In Table 3 we present the survival rates of the start-ups based on the combination of education levels of the founders and the directionality of these education levels. We observe that those with the least education (i.e. public/primary school) generally seem to have the lowest survival rates. Paradoxically, the highest education category also has lower survival rates, presumably because highly-educated entrepreneurs have attractive outside options (Gimeno et al., 1997). We also observe that diversity in education level is not necessarily an encumbrance, because high survival rates are also observed away from the diagonal.

Table 3 (right) shows the number of employees after 5 years by education level. It

---

[9]For more information on our education variable, see: http://www.dst.dk/Statistik/dokumentation/ times3/emnegruppe/emne/variabel.aspx?sysrid=372481&timespath=19|1013| .

		j			
		≤ Voc Tr	STEM	Business	Other
i	≤ Voc Tr	0.318	0.408	0.488	0.263
	STEM	0.465	0.556	0.571	0.542
	Business	0.397	0.222	0.412	0.231
	Other	0.311	0.000	0.500	0.345

		j			
		≤ Voc Tr	STEM	Business	Other
i	≤ Voc Tr	4.263	4.897	5.400	4.371
	STEM	5.414	6.600	17.500	4.231
	Business	6.074	12.500	4.857	14.667
	Other	3.842	.	3.750	6.684

Figure 4: Performance after 5 years, by education type, for individuals i (primary founder) and j. Cells with above-median values are highlighted. Left: survival; right: means of numbers of employees.

would appear that post-entry growth is low when $Educ_i = 1$ or $Educ_j = 1$ (that is up to and including high school) and relatively high when $Educ_i$ or $Educ_j$ are equal to 3 (vocational short cycle education) or 5 (post graduate education).

Education type From the level of education we try to identify whether the type of education matters for survival and growth. The education types are divided in four categories. One type are all the programmes in vocational training and below (≤ Voc Tr) and three where the educational programmes above this level have been divided in: degrees within science technology, engineering and mathematics (STEM), business related degrees (Business); and other degrees (Other). Figure 4 contains a number of interesting results, among which some evidence that the direction of diversity matters. The best performing teams (in terms of survival and growth) occur when $EducType_i$=STEM. {$EducType_i$=STEM, $EducType_j$ = Business} has a high survival rate (0.571) while {$EducType_i$ = Business, $EducType_j$ = STEM} has a low survival rate (0.222). t-tests with unequal variance reveal that this difference is significant at the 10% level (p-value = 0.0942). This pattern is also visible in the right panel of Figure 4, which pertains to growth. The highest employment growth (mean of 17.5 employees after 5 years) is associated with {$EducType_i$=STEM, $EducType_j$ = Business}; while the employment growth associated with {$EducType_i$=Business, $EducType_j$ = STEM} is lower (but the difference is not significant). This suggests that STEM and Business education backgrounds complement each other in complex ways, in line with our propositions.

5.1.3 Industry experience

Results for prior industry experience are shown in Figure 5. A first observation is that not all combinations of industry experience are observed. The worst outcome for survival

		j	
		None	2-digit	4-digit
	None	0.265	0.311	.
i	2-digit	0.331	0.428	.
	4-digit	0.412	.	0.429

		j	
		None	2-digit	4-digit
	None	4.798	4.197	.
i	2-digit	5.652	4.618	.
	4-digit	4.395	.	4.149

Figure 5: Performance after 5 years, by industry experience, for individuals i (primary founder) and j. Cells with above-median values are highlighted. Left: survival; right: means of numbers of employees.

is when neither i nor j have any industry experience – pairwise t-tests with respect to the baseline case of $\{indexp_i = indexp_i = 0\}$ show that the differences are all highly significant. The performance outcomes associated with industry experience for i at the 2-digit level are not considerably lower than those obtained for experience at the 4-digit level, which suggests that experience at the 2-digit level may be sufficient.[10] For growth, high employment growth can occur even if j= no prior industry experience. It seems that it is not necessary for both founders to have prior industry experience in order to grow.

5.2 Regressions

The non parametric analysis we presented above give an indication of the role of team composition on new venture performance. To study this effect in more detail, we apply other estimation techniques to control for other factors that might explain new venture performance. To do so, we estimate the following regression equation:

$$
\begin{aligned}
Y_k = & \ \beta_0 + \beta_1 \cdot \sum_{i=1...3}\sum_{j=1...3} AgeGroup_{ij} + \beta_2 \cdot \sum_{i=0,11}\sum_{j=0,11} Education_{ij} \\
& + \beta_3 \cdot \sum_{i=0,2}\sum_{j=0,2} IndustryExp_{ij} + \beta_4 \sum_{i=0,1}\sum_{j=0,1} Gender_{ij} \\
& + \beta_5 \cdot \sum_{i=0,1}\sum_{j=0,1} Danish_{ij} + \beta_6 \cdot \sum_{i=0,1}\sum_{j=0,1} Married_{ij} \\
& + \beta_7 \cdot VentureCharacteristics_k + \varepsilon_k
\end{aligned} \tag{1}
$$

---

[10]t-tests reveal that there are no significant differences between the outcomes for i when industry experience is measured at the 2-digit or 4-digit level, with the exception of survival when $indexp_j=0$ (that is, the mean values 0.412 and 0.331 are significantly different at the 5% level).

The unit of observation is the performance of venture k, and the explanatory variables include a constant term β_0, characteristics of the two founders i and j, as well as some venture-specific controls. A closer analysis of non-linearities and interdependencies affecting team composition is enabled by mapping the two-dimensional disparity space with a set of dummy variables for each combination of characteristics for individuals i and j.

We recoded the age categories by dividing age in three separate classes, i.e. less or equal to 30; age between 31-45 years, and 45 years and over. This gave us $3 \times 3 = 9$ possible configurations for i and j which are represented by 8 dummy variables (with the omitted base dummy corresponding to a team of young partners). In each case, the omitted dummy variable (corresponding to the base case) is the combination of lowest values for i and j.

We recoded our education variables to take into account the interdependence of education level and education type. Those with the lowest educational qualifications have not had the opportunity to specialize, and therefore the types of education refer only to those above a minimum level of education. To take this into account, we recoded our education variables $Education_{ij}$ such that i and j can take the following values: 1 for all up to (and including) high-school; 2 for vocational training; 3_1, 3_2 and 3_3 for vocational oriented short-cycle education programmes that specialize in either STEM, Business or other (respectively); 4_1, 4_2 and 4_3 for undergraduate (both academic and professional bachelor degrees), that specializes in either STEM, Business or other, respectively, and 5_1, 5_2 and 5_3 graduate and PhD education that specializes in either STEM, Business or other, respectively.

Industry experience is calculated with respect to the individual's work experience in the previous 5 years. Individuals can either have no prior experience, industry experience at the 2-digit NACE industry class, or industry experience at the 4-digit level (following Dahl and Sorenson, 2012).

$VentureCharacteristics_k$ includes a set of control variables. The first control variables we include are industry controls. In some industries, such as manufacturing, we have only a few firms present. To deal with this, we regroup some sectors together, following the Eurostat industry classification scheme for manufacturing sectors.[11] We also have

---

[11]The Eurostat manufacturing industry classification scheme has the following four categories: high tech, medium-high tech, medium tech and low tech. More details can

few firms in two-digit NACE sectors 65 and 67 (banking, insurance, etc) and so we merge these sectors together with 66 (life insurance, pensions, etc) to generate a new industry group which corresponds to the Eurostat definition of "Knowledge-intensive financial services."[12] Second, the entrepreneurial pairs might be based on family relationships. As this relationship can influence the performance of the firms in different ways we included four dummies making a distinction whether the entrepreneurial pairs are spouses, siblings, father and son/daughter, or mother and son/daughter. The inability of previous work to control for spousal relations has in fact been identified as a weakness of previous work (Hellerstedt et al., 2007). Family firms account for around 20% of our sample. Third, similar industry experience is an important factor that explains new venture performance, in particular survival (Dahl and Reichstein, 2007). To control for this factor we created two variables that indicate whether the two entrepreneurs have common industry experience in the previous 5 years.[13] Fourth, we introduce cohort dummies, which correspond to the year (1999-2003) in which the firm was founded. Finally, to control for the regional dimension, we created a set of five region dummies that correspond to the five Danish administrative regions.[14]

Equation (1) corresponds to a cross-sectional regression setup, where we explain performance at time $t + 5$ as a function of characteristics at startup (time t). When the dependent variable is survival, we apply a logit regression model (Jenkins, 1995).[15] When the dependent variable is employment growth, we measure this by taking the (natural logarithm of) number of employees after five years. Indeed, a meaningful indicator of the growth of new ventures is their size at the end of the period of observation (Eisenhardt and Schoonhoven, 1990; Storey, 1994; Colombo and Grilli, 2005). Although final size confounds the effects of initial size and post-entry growth, nonetheless all the firms in our sample have the same initial size of two employees.

Altogether, we will have an unusually rich set of variables in our initial regressions – too many to fit in a conventional results table. To maintain clarity, we adopt a stepwise algorithm for iterative removal of the least significant variable, and then repeating the

---

be found at http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:High-tech_class ification_of_manufacturing_industries.

[12]See http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/hrst_st_esms_an9.pdf.

[13]That is, whether one or both have worked in the same 4-digit NACE industry class.

[14]These regions are: Capital Region of Denmark, Zealand, North Denmark, Southern Denmark, and Central Denmark.

[15]An additional advantage of logit regression in our context is that it can be implemented inside our stepwise regression algorithm.

regression until only the most significant variables remain. We take the full model in Equation (1) as a starting point before stepwise removal of the least significant explanatory variable. We stop when the least significant explanatory variable is significant at the 15% level. As a result, it can be expected that our stepwise algorithm will give a different set of significant variables for different dependent variables and for different subsamples. This implies that it is not possible for us to report all of our results as adjacent columns in the same regression table – instead each regression of Equation (1) will be reported in a separate results table.

5.2.1 Survival

Table 3 contains the regression results for survival after 5 years. Teams that include old founders tend to have a higher survival rate (compared to the base case of two young founders). The two best performing age combinations (in terms of survival) are when the secondary founder is in the 45 years and over category.

Education seems to help survival, because all of the significant education dummies are positive (with respect to the omitted baseline case corresponding to two founders with minimal education). Interestingly enough, many (but not all) of the significant education dummies correspond to symmetric configurations where both founders are in the same education category. (However, this symmetry with respect to education is not found for employment growth, as we shall see). Industry experience has a positive effect on survival – because all of the significant dummies are positive with respect to the omitted baseline category of no experience for both founders.

In terms of ethnicity, we observe that teams of two Danes have the highest expected survival. Higher survival is observed for ventures where both the founders are in a registered partnership. However, there are no significant differences in the survival of family businesses.

5.2.2 Employment growth

Table 4 contains the regression results for number of employees after 5 years. A first observation is that we have a larger number of significant variables when employment growth is the dependent variable.[16]

---

[16] A practical implication of the difficulty of finding variables that predict survival, but not growth, is that it will not be easy to apply a two-stage Heckman selection model (which requires the existence of

For age, we observe results that contrast to our findings for survival – the base case (youngest age category for both founders) has the highest expected employment growth, judging by the fact that all of our significant dummies are negative. In fact, teams of two old founders have the coefficient with the largest magnitude, which indicates the lowest employment growth. Combining our results for survival and growth, firms with older pairs demonstrate stronger persistence, but firms composed of young founders demonstrate higher growth (conditional on survival). This mirrors the well-known result that younger *firms* grow faster (i.e. firm age measured as years since start-up; at the business-level rather than at the level of individuals).

With regards to education, we have a mixed bag of results, considering that some education dummies are positive and others are negative. Furthermore, many are not significant at the usual 5% significance level. We observe that $\{i = 3\_1, j = 3\_2\}$ has a large positive effect on employment growth, while the opposite configuration, $\{i = 3\_2, j = 3\_1\}$, has a large negative effect on employment growth. This echoes our earlier findings from Figure 4 (right). This asymmetry between STEM and Business education is not always observed however, the cases of post-graduate education levels, $\{i = 5\_1, j = 5\_2\}$ and $\{i = 5\_2, j = 5\_1\}$ are the combinations with the highest employment growth. More generally, teams where the primary founder has a post-graduate education often have higher employment growth.

With regards to industry experience, it is puzzling to see that *ceteris paribus* the few significant dummies are all negative. It is not clear why businesses without prior industry experience would not outperform those without such knowledge capital. Although we include industry dummies in our linear regression framework, it could be that more complex interactions of industry, experience, and employment growth are driving these findings.

Regarding ethnicity, the highest-growth ventures are those led by a Dane (although it matters little whether the second founder is a Dane or a foreigner). Family firms consistently have negative coefficients – whether we consider businesses formed with the father, with the spouse, or with the mother. This likely reflects the particularities of the 'business plan' of family firms, which puts more emphasis on guaranteeing a relaxed family lifestyle rather than the pursuit of commercial ambitions. The lowest-growth businesses are formed with one's mother.

variables that predict survival but not growth).

Another interesting finding is that the R^2 from the growth equation is higher than what is usually found for growth regressions (Coad, 2009, Table 7.1), perhaps because of the level of detail in our explanatory variables, and perhaps because we have a homogenous subset of firms (pairs) for which growth is measured over 5 years instead of annually. Moreover, the growth regression R^2 is much higher than the psuedo-R^2 obtained from the survival equation.[17]

5.2.3 Robustness analysis

To deal with the risk that our stepwise procedure will remove potentially valuable variables, we compared our stepwise estimates with those of the full model (all variables included) to see if there were any major discrepancies; however no such discrepancies were found. We also changed the time period over which performance was measured – instead of focusing on the five years after entry, we also investigated three-year and four-year periods for our two performance indicators (survival and employment growth). This gave similar results.

It has recently been argued that taking survival as a binary variable means that exit is underspecified, because no distinction is made between successful exits and failures (Wennberg et al., 2010). Some firms in our dataset underwent acquisitions – these firms are identified as survivors in the year prior to the merger and then excluded. Furthermore, to address concerns that firm exit might correspond to closure of a successful business for retirement reasons, we removed those individuals aged 55 or above (cf. Coad et al., 2012; Dahl and Sorenson, 2012) and repeated the analysis. Our main results remained unchanged.

Alternative performance indicators were considered – namely, growth of sales, value-added, and profit (results available upon request). These results are presented in the appendix, in Tables H.3, H.4, and H.5. Results for the growth of sales and value added are similar to those obtained for employment growth. With regards to growth of profits – we observe that age has no significant effect, and that education in general, and business education in particular, has stronger benefits than when performance is measured in terms of employment growth.

---

[17]Note however that an OLS R^2 and a logit pseudo-R^2 are not strictly comparable. A comparison of the Nagelkerke R^2 statistics for both equations (obtained from estimation of the full model in equation (1), without stepwise elimination of insignificant coefficients, yields estimates of the Nagelkerke R^2 of 0.152 for the survival regression and 0.217 for the growth regression.

Table 3: Stepwise logit regression of equation (1), where the binary dependent variable is survival after 5 years.

	Coefficient	Robust Std. Error	z-stat
Age group dummies			
agegrp_dummy_21	0.255	0.108	2.36
agegrp_dummy_22	0.224	0.106	2.12
agegrp_dummy_23	0.336	0.153	2.20
agegrp_dummy_32	0.236	0.155	1.53
agegrp_dummy_33	0.348	0.157	2.21
Education			
education_dummy_1__3_2	0.837	0.535	1.56
education_dummy_2__2	0.307	0.096	3.18
education_dummy_3_1__2	0.834	0.278	3.00
education_dummy_4_1__2	0.598	0.415	1.44
education_dummy_4_1__4_1	1.544	0.724	2.13
education_dummy_4_3__2	0.473	0.281	1.68
education_dummy_4_3__4_3	0.904	0.408	2.22
education_dummy_4_3__5_2	2.042	0.956	2.14
education_dummy_5_1__5_1	2.106	0.588	3.58
education_dummy_5_3__5_3	1.623	1.102	1.47
Industry experience			
indexp_dummy_1_0	0.215	0.144	1.50
indexp_dummy_1_1	0.622	0.164	3.80
indexp_dummy_2_0	0.697	0.106	6.60
indexp_dummy_2_2	0.632	0.111	5.69
Non-task characteristics			
dane_dummy_11	0.621	0.116	5.34
marr_dummy_01	0.209	0.122	1.71
marr_dummy_10	0.310	0.099	3.13
marr_dummy_11	0.506	0.107	4.75
Industry dummies	yes		
Region dummies	yes		
Year dummies	yes		
β_0: Constant term	-2.146	0.183	-11.70
Obs	3604		
Pseudo-R^2	0.0808		

Table 4: Stepwise OLS regression of equation (1), where the dependent variable is (log of) the number of employees after 5 years. Robust standard errors obtained from the Huber/White/Sandwich estimator.

	Coefficient	Robust Std. Error	t-stat
Age group dummies			
agegrp_dummy_12	-0.236	0.132	-1.79
agegrp_dummy_13	-0.271	0.139	-1.95
agegrp_dummy_21	-0.239	0.082	-2.90
agegrp_dummy_22	-0.400	0.081	-4.93
agegrp_dummy_23	-0.492	0.098	-5.01
agegrp_dummy_31	-0.460	0.124	-3.71
agegrp_dummy_32	-0.487	0.109	-4.48
agegrp_dummy_33	-0.715	0.099	-7.19
Education			
education_dummy_1__3_3	-0.651	0.360	-1.81
education_dummy_2__4_1	0.377	0.146	2.57
education_dummy_3_1__3_1	0.601	0.106	5.67
education_dummy_3_1__3_2	1.364	0.145	9.38
education_dummy_3_1__3_3	0.342	0.115	2.97
education_dummy_3_1__4_3	-0.856	0.225	-3.81
education_dummy_3_1__5_3	-0.807	0.203	-3.98
education_dummy_3_2__3_1	-1.200	0.162	-7.42
education_dummy_3_3_2	0.634	0.290	2.18
education_dummy_3_3__3_2	-0.575	0.171	-3.36
education_dummy_4_1__1	0.555	0.271	2.05
education_dummy_4_1__3_2	1.443	0.151	9.55
education_dummy_4_3__3_3	-1.529	0.102	-14.97
education_dummy_4_3__5_2	-0.583	0.195	-2.98
education_dummy_5_1__3_3	0.581	0.225	2.58
education_dummy_5_1__4_1	-0.540	0.119	-4.53
education_dummy_5_1__4_3	0.907	0.114	7.94
education_dummy_5_1__5_1	0.455	0.254	1.79
education_dummy_5_1__5_2	1.374	0.552	2.49
education_dummy_5_2__2	0.653	0.249	2.62
education_dummy_5_2__4_3	0.842	0.127	6.61
education_dummy_5_2__5_1	2.027	0.103	19.67
education_dummy_5_2__5_3	-0.377	0.128	-2.94
education_dummy_5_3__3_3	1.024	0.099	10.37
education_dummy_5_3__5_2	0.440	0.122	3.61
Industry experience			
indexp_dummy_0_1	-0.108	0.064	-1.68
indexp_dummy_2_2	-0.143	0.063	-2.29
Non-task characteristics			
dane_dummy_10	0.283	0.140	2.02
dane_dummy_11	0.255	0.099	2.56
marr_dummy_10	0.124	0.060	2.06
marr_dummy_11	0.115	0.061	1.89
Family-firm dummies			
mom	-0.568	0.107	-5.3
spous	-0.323	0.082	-3.96
dad	-0.218	0.120	-1.82
Industry dummies	yes		
Region dummies	yes		
Year dummies	yes		
β_0: Constant term	3.010	0.220	13.68
Number of obs	1207		
R^2	0.174		

28

We also took an alternative employment growth indicator. In our baseline analysis, we measured employment taking the number of employees (November headcount). To check the robustness of our measure, however, we also considered employment (growth) measured using the Full-Time Equivalent (FTE; which nonetheless had more missing observations than our main employment variable). All in all, the results are very similar.

To check that our identification of 'primary' and 'secondary' entrepreneur is valid, we repeated the analysis on a subsample of businesses where we could be more confident that our attribution of founders as 'primary' and 'secondary' was meaningful. In this case, we only select those firms where the primary founder is registered as an owner and the secondary founder as a middle and lower ranked employee. This would get a stronger disparity between the different founders. The results from this analysis are not very significant (presumably because of a lower number of observations) but generally are in accord with our main findings.

For comparability with previous work, we also estimated Cox and Heckman regression models. This required reformatting our data into a panel with observations at an annual frequency. Cox regressions gave similar findings for the role of age and industry experience, although our findings for the education dummies were weaker. We also estimated a Heckman selection model for growth (conditional on survival). An advantage of applying a Heckman model in our case is that we have comprehensive coverage of all two-founder startups (even though some exit before the end of the 5-year period). A drawback of the Heckman model, however, is that the first-stage equation requires the inclusion of variables that affect survival (but not growth). As can be seen from our results tables 3 and 4, there are more variables that predict growth than survival. Nonetheless, lagged growth has been shown to be a strong predictor of survival, but weakly related (if at all) to subsequent growth (Coad et al., 2012), and so we took lagged growth to explain survival but not growth. The non-selection hazard (that is, the inverse of the Mills Ratio) was significant and negative, but in the Heckman model most of our explanatory variables became insignificant (no doubt because explaining annual growth is more difficult than explaining growth over a five-year period).

6   Discussion

In this section we will seek to 'digest' our findings by referring to our three propositions.

Proposition 1 stated that the effects of diversity on the outcomes of new businesses were heavily moderated by the 'position' or 'status' within the hierarchy. We find considerable support for this hypothesis because our results were far from 'symmetric' in a number of cases. This suggests that beneficial characteristics of the primary founder are not necessarily those that would best befit the secondary founder. With regards to age, growth tends to be higher if the primary founder is younger than the second. With regards to type of education, we obtained a mixed set of results, although businesses with a commercially-minded individual playing a secondary role performed better in terms of survival and growth than when a commercially-minded individual was the primary founder. More generally, our results for education type were far from symmetric. With some of our other variables, however, symmetry in characteristics space was associated with better outcomes (such as two Danes as founders; or where both founders are married (positive effects for firm survival); or two founders with low education having the worst survival chances).

Proposition 2 stated that the effects of diversity were non-linear and complex and could not easily be represented using a linear unidimensional indicator. We observed that the 'optimal' position in characteristics space was not monotonically increasing – for example, low education was associated with low survival, but there was little dependence of survival on education above a certain threshold. We also observed that the 'optimal' position in characteristics space depended on the characteristics of the partner – a powerful illustration of this idea is that, controlling for other factors, a configuration of education types $\{i = 3\_1, j = 3\_2\}$ had high expected employment growth, while the inverse configuration $\{i = 3\_2, j = 3\_1\}$ yielded a negative coefficient (with respect to the base case of minimal education), and furthermore, changing the characteristics of the 'second fiddle' turned the coefficient from strongly positive to negative (compare e.g. employment growth for $\{i = 5\_1, j = 4\_1\}$ with $\{i = 5\_1, j = 4\_3\}$). Finally, another problem is that diversity will probably interact with firm size (this was not examined here because all businesses in our sample have the same start-up size of 2 individuals).

Taken together, our support for Propositions 1 and 2 provides justification for our new methodology, which has identified effects of team composition on performance that could not have been uncovered using the standard diversity indicators.

Proposition 3 predicted that diversity has different effects on survival and employment growth (even though these two could be considered as indicators of firm performance).

In the case of family firms, we observe that they generally have an average performance for survival (because we observe no significant coefficients apart from a positive coefficient for spouses), although family firms are associated with slower employment growth. Regarding our employment growth regressions, we observe the largest negative effects for firms founded with mothers, then fathers, then spouses. This is consistent with the notion that family firms are under pressure to keep the family 'tradition' alive (perhaps even in the face of prolonged poor performance), although they do not seek employees either through a mistrust of 'outsiders' or an aversion to the perceived risks or growth. Similarly, firms composed of older founders have better survival rates, but lower employment growth. Pairs of young founders have the highest employment growth. This could be because pairs of older founders do not want to take risks or over-exert themselves, but would prefer to 'coast along' before retirement. Younger pairs seem to be more willing to 'experiment' in their businesses, having higher exit hazards but often experiencing faster employment growth.

Limitations A number of caveats of our analysis can be mentioned. One open question is how valid our findings our for different institutional contexts in other countries. Denmark has a fluid labour market that has been described as being about as fluid as that of the US (Bingley and Westergaard-Nielsen, 2003), although there are generous state unemployment benefits for those who do lose their jobs. It is not clear, therefore, how easily our findings regarding employment growth can be generalized to countries with tighter job protection laws.

As with many other studies on entrepreneurial teams, this paper suffers from endogeneity. Studies such as ours that deal with voluntary participation in teams do not observe exactly what makes the individuals form teams. Therefore, our results are not intended for guiding how individuals should be 'fused' together as teams. Self-selection matters in establishing a team. We cannot identify the underlying mechanisms on why teams are formed, but given that entrepreneurial teams rely on voluntary participation there is likely to be a subjective motivation on the formation of the team, which is probably influenced by factors that remain unobserved to the econometrician. Our results do not purport to prescribe 'arranged marriages' in the sense that our results are not sufficient to say how one individual should be 'fused' to another individual they don't know, in order to make the ideal entrepreneurial team. Instead, we merely observe start-ups that have formed themselves at time t, and observe regularities concerning the outcomes associated with these start-ups at time $t + 5$. We have no information on the business

opportunity being exploited, or the quality of the business idea. It may, of course, be the case that higher-quality opportunities make it easier to team up with higher-skilled individuals.

Another issue that needs to be addressed is the founder selection and identification. Since we rely on official tax records, we can only identify individuals that have a formal attachment to the start-up, an identification that is made only once a year. In addition, it might be that the secondary individual only later in the year joined the business; however, as mentioned in Section 3, early hiring decisions affect new venture performance (Baron et al., 1999). Furthermore, the distinction between primary and secondary partner is crucial. Some teams may be symmetric (some of these cases ($n = 225$) have been identified and dropped; however it is possible that power relations in a new venture go counter to the legalistic foundations of their relationship, in the same way that (according to the caricature) a CEO may be subservient to his secretary. Further work on this topic would be needed. In particular, our work would benefit from being complemented by case studies and qualitative research to shed further insights on these themes.

7   Conclusion

This study on 3777 entrepreneurial pairs, and the amount of detail that is provided by the Danish register data on these pairs, provide interesting insights in how team composition affects performance. In particular it places question marks on the way diversity is treated in the various studies that exist on the topic. First, we provided evidence that the effects of diversity are moderated by the hierarchy that exist in the firm. Second, diversity is clearly not a linear and unidimensional indicator. This calls for an overall re-evaluation of the existing approaches to investigating diversity of team composition. Third, diversity has a different impact on different performance measures.

The findings reported in this paper will also be of interest to the following:

- Academics and scholars interested in the effects of team composition and diversity on venture performance.

- Angel investors who are interested in the outcomes associated with observed entrepreneurial partnerships.

- University entrepreneurship promotion schemes will be interested in our results

concerning the diversity of education subjects. For example, our finding that pairs of educational profiles consisting of STEM and business do best, may have implications for encouraging collaboration between e.g. engineering departments and business schools.

- Policy makers interested in offering assistance to potentially high-impact new ventures will be interested in information that helps identify which ventures will grow faster.

- Policy makers seeking to have a more efficient entrepreneurship policy should observe that family firms do worse in our case, and should perhaps rethink the specific benefits these firms get. For example, it is not clear why, in the UK, family firms get implicit subsidies (such as relief from inheritance tax) even though they are observed to be noticeably unproductive (Bloom and Van Reenen, 2010).

- Entrepreneurs who are interested in choosing a partner for their new business idea, from a pre-selected list of possible candidates. However, it should be clear that our results provide large-sample regularities rather than deterministic causal effects. Furthermore, the explanatory power of our regressions was not higher than about 20%. Entrepreneurs have richer information about their set of potential partners, and have detailed knowledge about the specificities of their business opportunities, personalities, strengths/weaknesses, technical challenges, etc.

Our results offer insights concerning the stereotype (to be found in venture capital circles or the university spinout literature) that startups have good technical ideas but poor business/commercial skills (and hence need VC business guidance to succeed). For example, Wennberg et al. (2011, p. 1138) write that "an important imperative to assist USOs in building viable teams that have the requisite commercial experience to succeed." We observed that it is better, in terms of employment growth, to be configured with STEM first and business second, than to have a business-educated founder first and STEM second. This hints that there may be problems if the focus is on commercial aspects, with the technical side taking a back seat. Our analysis provides tentative evidence that while commercial skills are important, they should not dominate the technical aspects. Commercial viability should, perhaps, be seen as a constraint to be satisfied rather than the primary aim of the new venture.

Finally, we would like to provide some suggestions for further work. First, we consider that there is still plenty of opportunity for finding richer quantitative tools for analyzing

diverse entrepreneurial teams. It seems slightly ironic to us that it is frequently acknowledged that diversity is a 'double-edged sword' and often yields mixed results, and yet researchers generally compress the numerous dimensions of diversity into a single indicator and then calculate the 'average effect' through standard regressions. We would like to see more 'diversity' in quantitative research into the role of diversity in teams. For example, future work could try to decompose the two edges of the 'sword' to investigate which factors affect conflict more than creativity (that is, distinguishing between the 'gross' and the 'net' costs and benefits of diversity). Second, it would be interesting to see if the degree of diversity in an entrepreneurial team affects the likelihood that the founder will stay with the firm in later years. Although there does exist literature on team member exit (Ucbasaran et al., 2003; Chandler et al., 2004; Hellerstedt et al., 2007), the time span of these studies are limited.

References

Aldrich, H. E. and Langton, N. (1998). Human resource management practices and organizational life cycles. In Reynolds, P., Bygrave, W. D., Davidsson, P., Gartner, W. B., and Carter, N. M., editors, *Frontiers of Entrepreneurship Research*, pages 349–357. Babson Park, MA: Babson College, Center for Entrepreneurial Studies.

Aldrich, H. E. and Ruef, M. (2006). *Organizations Evolving*. SAGE publications ltd, second edition.

Amason, A. C., Shrader, R. C., and Thompson, G. H. (2006). Newness and novelty: Relating top management team composition to new venture performance. *Journal of Business Venturing*, 21:125–148.

Ancona, D. G. and Caldwell, D. F. (1992). Demography and design: Predictors of new product team performance. *Organization Science*, 3(3):321–341.

Bantel, K. and Jackson, S. (1989). Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal*, 10:107–124.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1):99–120.

Baron, J., Hannan, M., and Burton, M. (1999). Building the iron cage: Determinants of managerial intensity in the early years of organizations. *American Sociological Review*, 64(4):527–547.

Beckman, C. M. (2006). The influence of founding team company affilitations on firm behavior. *The Academy of Management Journal*, 49(4):741–758.

Beckman, C. M., Burton, M., and O'Reilly, C. (2007). Early teams: The impact of teams demography on vc financing and going public. *Journal of Business Venturing*, 22(147-173).

Bingley, P. and Westergaard-Nielsen, N. (2003). Returns to tenure, firm-specific human capital and worker heterogeneity. *International Journal of Manpower*, 24(7):774–788.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1):203–224.

Boeker, W. (1989). The development and institutionalization of subunit power in organizations. *Administrative Science Quarterly*, 34(3):388–410.

Chandler, G. N., Honig, B., and Wiklund, J. (2004). Antecedents, moderators, and performance consequences of membership change in new venture teams. *Journal of Business Venturing*, 20:705–725.

Chowdhury, S. (2005). Demographic diversity for building an effective entrepreneurial team: is it important? *Journal of Business Venturing*, 20:727–746.

Clarysse, B., Knockaert, M., and Lockett, A. (2007). Outside board members in high tech start-ups. *Small Business Economics*, 29(3):243–259.

Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence.* Edward Elgar Publishing.

Coad, A., Frankish, J., Roberts, R., and Storey, D. J. (2012). Growth Paths and Survival Chances: An Application of Gambler's Ruin Theory. *Journal of Business Venturing*, forthcoming.

Colombo, M. and Grilli, L. (2005). Founders' human capital and the growth of new technology-based firms: A competence-based view. *Research Policy*, 34(6):795–816.

Cooney, T. M. (2005). Editorial: What is an entrepreneurial team. *International Small Business Journal*, 23(3):226–235.

Cooper, A. C., Gimeno-Gascon, F. J., and Woo, C. Y. (1994). Initial human and financial capital as predictors of new venture performance. *Journal of Business Venturing*, 9:371–395.

Cox, T. H. and Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *The Executive*, 5(3):45–56.

Dahl, M. S. and Reichstein, T. (2007). Are you experienced? prior experience an the survival of new organizations. *Industry and Innovation*, 14(5):497–511.

Dahl, M. S. and Sorenson, O. (2012). Home sweet home: Entrepreneurs' location choices and the performance of their ventures. *Management Science*, 58(6):1059–1071.

Dahlin, K., Weingart, L., and Hinds, P. (2005). Team diversity and information use. *The Academy of Management Journal*, 48(6):1107–1123.

Davidsson, P. and Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing*, 18:301–331.

Delmar, F. and Shane, S. (2006). Does experience matter? the effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, 4(3):215–247.

Eisenhardt, K. and Schoonhoven, C. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, 35(3):504–529.

Ensley, M., Pearson, A., and Amason, A. (2002). Understanding the dynamics of new venture top management teams: cohesion, conflict, and new venture performance. *Journal of Business Venturing*, 17:365–386.

Ensley, M. D., Carland, J. W., and Carland, J. C. (1998). The effect of entrepreneurial team skill heterogeneity and functional diversity on new venture performance. *Journal of Business and Entrepreneurship*, 10(1):1–15.

Foo, M. D., Wong, P. K., and Ong, A. (2005). Do others think you have a viable business idea? team diversity and judges' evaluation of ideas in a business plan competition. *Journal of Business Venturing*, 20:385–402.

Forbes, D., Borchert, P., and Zellmer-Bruhn, M. (2006). Entrepreneurial team formation: An exploration of new member addition. *Entrepreneurship Theory and Practice*, pages 225–246.

Gartner, W. B. (1985). A conceptual framework for descriing the phenomenon of new venture creation. *The Academy of Management Review*, 10(4):696–706.

Gartner, W. B., Shaver, K. G., Gatewood, E., and Katz, J. A. (1994). Finding the entrepreneur in entrepreneurship. *Entrepreneurship Theory and Practice*, Spring 1994:5–9.

Gimeno, J., Folta, T., Cooper, A., and Woo, C. (1997). Survival of the fittest? entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, pages 750–783.

Hambrick, D., Cho, T., and Chen, M. (1996). The influence of top management team heterogeneity on firms' competitive moves. *Administrative Science Quarterly*, 41:659–684.

Hambrick, D. and Mason, P. (1984). Upper echelon: The organization as a reflection of its top managers. *The Academy of Management Review*, 9:193–206.

Harper, D. (2008). Towards a theory of entrepreneurial teams. *Journal of Business Venturing*, 23(6):613–626.

Harrison, D. and Klein, K. (2007). What's the difference? diversity constructs as separation, variety, or disparity in organizations. *The Academy of Management Review*, 32(4):1199–1228.

Hellerstedt, K., Aldrich, H., and Wiklund, J. (2007). The impact of past performance on the exit of team members in young firms: The role of team composition. Frontiers of Entrepreneurship Research 27 (11), article 2.

Horwitz, S. (2005). The compositional impact of team diversity on performance: Theoretical considerations. *Human Resource Development Review*, 4(2):219–245.

Horwitz, S. and Horwitz, I. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, 6:987–1015.

Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57(1):129–136.

Kaiser, U. and Müller, B. (2012). Team heterogeneity and its development over time. *MIMEO*.

Kamm, J., Shuman, J. C., Seeger, J., and Nurick, A. J. (1990). Entrepreneurial teams in new venture creation: A research agenda. *Entrepreneurship Theory and Practice*, Summer 1990:12.

McPherson, M., Smith-Lovin, L., and Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.

Murray, A. (1989). Top management heterogeneity and firm performance. *Strategic Management Journal*, 10:125–141.

Østergaard, C., Timmermans, B., and Kristinsson, K. (2011). Does a different view create something new? the effect of employee diversity on innovation. *Research Policy*, 40(3):500–509.

Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. *Organization Science*, 7(6):615–631.

Pelled, L. H., Eisenhardt, K., and Xin, K. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, 44(1):1–28.

Pfeffer, J. (1983). Organizational demography. In Staw, B. and Cummings, L., editors, *Research in Organizational Behavior*, volume 5, pages 299–357. Greenwhich, CT: JAI press.

Romanelli, E. and Tushman, M. (1994). Organizational transformation as punctuated equilibrium: an empirical test. *Academy of Management Journal*, 37(5):1141–1166.

Roure, J. and Maidique, M. (1986). Linking prefunding factors and high-technology venture success: An exploratory study. *Journal of Business Venturing*, 1:295–306.

Ruef, M. (2010). Entrepreneurial groups. In Landström, H. and Lohrke, F., editors, *Historical Foundations of Entrepreneurship Research*. Edward Elgar Publishing, Aldershot, UK.

Ruef, M., Aldrich, H., and Carter, N. (2003). The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review*, 68(2):195–222.

Schimmel, A. (1994). *The Mystery of Numbers*. Oxford University Press.

Steffens, P., Terjesen, S., and Davidsson, P. (2011). Birds of a feather get lost together: new venture team composition and performance. 10.1007/s11187-011-9358-z.

Stirling, A. (2001). A general framework for analyzing diversity in science, technology and society. *SPRU Working Papers 156*.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15):707–719.

Storey, D. (1994). The role of legal status in influencing bank financing and new firm growth. *Applied Economics*, 26(2):129–136.

Timmermans, B. (2010). The danish integrated database for labor market research: Towards demystification for the english speaking audience. *Druid Working Paper Series*, 10-16.

Ucbasaran, D., Lockett, A., Wright, M., and Westhead, P. (2003). Entrepreneurial founder teams: Factors associated with member entry and exit. *Entrepreneurship Theory and Practice*, Winter 2003:1042–2587.

Vanaelst, I., Clarysse, B., Wright, M., Lockett, A., Moray, N., and S'Jegers, R. (2006). Entrepreneurial team development in academic spinouts: An examination of team heterogeneity. *Entrepreneurship Theory and Practice*, 30(2):249–271.

Watson, W., Ponthieu, L., and Critelli, J. (1995). Team interpersonal process effectiveness in venture partnerships and its connection to perceived success. *Journal of Business Venturing*, 10:393–411.

Watson, W. E., Stewart, W. H., and BarNir, A. (2003). The effects of human capital, organizational demography, and interpersonal processes on venture partner perceptions of firm profit and growth. *Journal of Business Venturing*, 18:145–164.

Wennberg, K., Wiklund, J., DeTienne, D., and Cardon, M. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing*, 25(4):361–375.

Wennberg, K., Wiklund, J., and Wright, M. (2011). The effectiveness of university knowledge spillovers: Performance differences between university spinoffs and corporate spinoffs. *Research Policy*, 40(8):1128–1143.

Williams, K. and O'Reilly, C. (1998). Demograhpy and diversity in organizations: A review of 40 years of research. In Staw, B. and Cummings, L., editors, *Research in Organizational Behavior*, pages 77–140. JAI Press, Greenwich.

8 Appendix

Table H.1: Summary statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
age i	3777	36.967	10.263	15	91
age j	3777	33.615	11.941	15	79
male i	3777	0.700	0.458	0	1
male j	3777	0.617	0.486	0	1
dane i	3777	0.881	0.324	0	1
dane j	3777	0.867	0.340	0	1
educ_level (i)	3678	1.997	1.097	1	5
educ_level (j)	3639	1.722	0.977	1	5
educ_type (i)	3777	0.387	0.900	0	3
educ_type (j)	3777	0.260	0.770	0	3
survival	3777	0.333	0.471	0	1
empl (t=5), November headcount	1259	4.563	5.310	0	88
empl (t=5), FTE	1259	3.223	3.919	0	60
turnover (100k DKK) (t=5)	892	50.597	114.642	0.960	2007.530
family firm: brother	3777	0.032	0.175	0	1
family firm: spouse	3777	0.073	0.261	0	1
family firm: mother	3777	0.030	0.170	0	1
family firm: father	3777	0.049	0.216	0	1
family firm	3777	0.184	0.387	0	1
2-digit industry experience (i)	3777	0.443	0.497	0	1
2-digit industry experience (j)	3777	0.387	0.487	0	1
4-digit industry experience (i)	3777	0.317	0.465	0	1
4-digit industry experience (j)	3777	0.262	0.440	0	1
same_school	3777	0.045	0.207	0	1
Copenhagen	3777	0.449	0.497	0	1

Table H.2: Frequency of ventures disaggregated by 2-digit NACE sectors.

NACE02	Frequency	Percent
15	22	0.58
17	5	0.13
18	8	0.21
19	1	0.03
20	7	0.19
21	1	0.03
22	21	0.56
24	3	0.08
25	2	0.05
26	6	0.16
27	3	0.08
28	46	1.22
29	16	0.42
30	4	0.11
31	10	0.26
32	3	0.08
33	4	0.11
34	1	0.03
35	8	0.21
36	8	0.21
37	1	0.03
45	552	14.61
50	94	2.49
51	256	6.78
52	762	20.17
55	850	22.5
60	223	5.9
61	7	0.19
63	24	0.64
64	25	0.66
65	6	0.16
66	1	0.03
67	7	0.19
70	57	1.51
71	23	0.61
72	216	5.72
73	9	0.24
74	485	12.84
	3777	100.00

Table H.3: Stepwise OLS regression of equation (1), where the dependent variable is log of sales after 5 years. Robust standard errors obtained from the Huber/White/Sandwich estimator.

	Coefficient	Robust Std. Error	t-stat
Age			
agegrp_dummy_22	-0.121	0.072	-1.68
agegrp_dummy_33	-0.430	0.136	-3.16
Education			
education_dummy_1__3_1	0.344	0.180	1.92
education_dummy_1__5_1	0.679	0.367	1.85
education_dummy_2__3_2	1.476	0.180	8.20
education_dummy_2__3_3	0.831	0.471	1.76
education_dummy_2__4_1	0.602	0.145	4.16
education_dummy_2__5_1	-1.207	0.209	-5.77
education_dummy_2__5_2	1.177	0.182	6.48
education_dummy_3_1__3_1	0.477	0.075	6.34
education_dummy_3_1__3_2	0.743	0.099	7.51
education_dummy_3_1__3_3	0.901	0.155	5.79
education_dummy_3_1__4_3	1.121	0.619	1.81
education_dummy_3_2__3_1	-0.696	0.155	-4.51
education_dummy_4_1__3_1	-0.671	0.158	-4.25
education_dummy_4_1__4_1	0.775	0.244	3.18
education_dummy_4_1__5_1	-5.105	0.109	-46.73
education_dummy_4_3__3_2	-1.241	0.156	-7.94
education_dummy_4_3__3_3	-2.227	0.154	-14.44
education_dummy_5_1__3_3	-1.752	0.140	-12.48
education_dummy_5_1__4_3	0.564	0.105	5.39
education_dummy_5_1__5_2	2.862	0.187	15.27
education_dummy_5_2__2	1.202	0.347	3.47
education_dummy_5_2__4_3	1.026	0.176	5.82
education_dummy_5_2__5_1	1.766	0.110	16.06
education_dummy_5_2__5_2	1.539	0.150	10.25
Non-task characteristics			
man_dummy_10	0.261	0.088	2.97
man_dummy_11	0.271	0.081	3.37
dane_dummy_11	0.199	0.084	2.36
mom	-0.670	0.164	-4.09
spous	-0.276	0.108	-2.57
dad	-0.444	0.169	-2.63
Industry dummies	yes		
Region dummies	yes		
Year dummies	yes		
β_0: Constant term	14.727	0.111	132.95
Obs	863		
R^2	0.328		

Table H.4: Stepwise OLS regression of equation (1), where the dependent variable is log of profit after 5 years. Robust standard errors obtained from the Huber/White/Sandwich estimator.

	Coefficient	Robust Std. Error	t-stat
Age dummies			
Education dummies			
education_dummy_1__3_1	-1.030	0.438	-2.35
education_dummy_1__3_3	2.206	0.111	19.87
education_dummy_1__4_1	0.931	0.331	2.82
education_dummy_1__5_2	0.264	0.137	1.93
education_dummy_2__3_2	0.396	0.106	3.72
education_dummy_2__5_2	2.412	0.278	8.66
education_dummy_3_1__3_1	0.694	0.106	6.58
education_dummy_3_1__3_3	2.591	0.195	13.27
education_dummy_3_1__4_3	0.715	0.363	1.97
education_dummy_3_3__1	0.895	0.124	7.24
education_dummy_4_1__3_1	-5.695	0.112	-50.86
education_dummy_4_3__3_2	0.501	0.195	2.57
education_dummy_4_3__5_2	1.064	0.092	11.58
education_dummy_5_1__4_3	0.518	0.144	3.60
education_dummy_5_1__5_2	4.001	0.110	36.34
education_dummy_5_2__2	1.485	0.287	5.18
education_dummy_5_2__4_3	2.758	0.092	29.94
education_dummy_5_2__5_1	1.557	0.110	14.14
education_dummy_5_3__2	1.314	0.297	4.43
education_dummy_5_3__5_3	1.143	0.124	9.24
Industry experience			
indexp_dummy_2_0	0.234	0.101	2.30
indexp_dummy_2_2	0.247	0.101	2.44
Non-task characteristics			
man_dummy_10	0.226	0.133	1.70
man_dummy_11	0.303	0.107	2.84
marr_dummy_10	0.322	0.105	3.08
marr_dummy_11	0.253	0.107	2.37
mom	-0.629	0.312	-2.02
Industry dummies	yes		
Region dummies	yes		
Year dummies	yes		
β_0: Constant term	11.744	0.116	101.68
Obs	762		
R^2	0.202		

Table H.5: Stepwise OLS regression of equation (1), where the dependent variable is log of Value Added after 5 years. Robust standard errors obtained from the Huber/White/Sandwich estimator.

	Coefficient	Robust Std. Error	t-stat
agegrp_dummy_13	0.228	0.149	1.53
agegrp_dummy_22	-0.128	0.075	-1.72
agegrp_dummy_33	-0.298	0.123	-2.42
Education dummies			
education_dummy_1__3_1	0.434	0.144	3.00
education_dummy_2__3_2	1.533	0.186	8.26
education_dummy_2__4_1	0.565	0.203	2.78
education_dummy_2__5_1	-0.435	0.198	-2.20
education_dummy_2__5_2	1.745	0.242	7.20
education_dummy_3_1__3_1	0.683	0.098	6.93
education_dummy_3_1__3_2	1.048	0.124	8.42
education_dummy_3_1__3_3	1.249	0.124	10.06
education_dummy_3_2__3_1	-0.535	0.139	-3.86
education_dummy_3_2__3_2	0.258	0.116	2.22
education_dummy_3_3__1	0.285	0.154	1.85
education_dummy_4_1__2	-0.328	0.157	-2.09
education_dummy_4_1__3_1	-0.441	0.170	-2.59
education_dummy_4_1__4_1	0.677	0.302	2.24
education_dummy_4_3__3_3	-1.080	0.128	-8.40
education_dummy_5_1__4_1	-0.229	0.135	-1.69
education_dummy_5_1__4_3	0.704	0.142	4.94
education_dummy_5_1__5_2	3.283	0.204	16.1
education_dummy_5_2__2	1.122	0.168	6.69
education_dummy_5_2__4_3	1.421	0.182	7.82
education_dummy_5_2__5_1	1.839	0.133	13.78
education_dummy_5_2__5_2	0.788	0.112	7.02
education_dummy_5_3__2	0.411	0.229	1.79
education_dummy_5_3__5_3	0.516	0.121	4.28
Non-task characteristics			
man_dummy_10	0.256	0.096	2.68
man_dummy_11	0.200	0.089	2.25
dane_dummy_11	0.193	0.104	1.84
dad	-0.563	0.206	-2.73
mom	-0.685	0.175	-3.93
spous	-0.235	0.115	-2.04
Industry dummies	yes		
Region dummies	yes		
Year dummies	yes		
β_0: Constant term	13.471	0.198	68.16
Obs	964		
R^2	0.2391		