CATÓLICA LISBON
BVSINESS & ECONOMICS

The influence of a start-up process on the entrepreneurs' emotions, deduced by their Twitter accounts

Matteo Maria Gambardella

Dissertation written under the supervision of Francesco Castellaneta

Dissertation submitted in partial fulfilment of requirements for the MSc in Management with specialization in Strategy and Entrepreneurship, at the Universidade Católica Portuguesa, January 2017.

In this paper, entrepreneurship has been analysed along a temporal range, stressing the influence that being founded has on an entrepreneur's emotions.
Starting from a vast base of Twitter accounts, through the text analysis, have been extracted ratios concerning the presence within the Tweets of a positive and negative emotions. Those ratios have been calculated equating the amount of specific words referring to that emotions on the total amount of words tweeted every year by the entrepreneurs. Those people are mostly from US, and the investors founding their ideas undertook many different investment strategies. The main focus lays within the effect that the overall founding strategy has on the entrepreneurs' emotions.
The purpose is comparing those emotional ratios with the verification of the founding process. This aims to test how the emotions are affected by the fact that the start-up has been financed, or it has not received any funds. The results have been analysed through Stata, yielding interest findings shown along the whole paper.

Nesta dissertação, o empreendedorismo foi analisado dentro de um intervalo de tempo, de modo a fazer entender a maneira como o financiamento influencia as emoções dos empreendedores. Como ponto de partida, vários textos de uma vasta base de contas de Twitter foram analisados com o objectivo de construir rácios que destaquem a presença de emoções positivas e negativas nos Tweets. Estes rácios foram calculados equacionando o montante específico de palavras referentes a este tipo de emoções sobre o montante total de palavras trocadas nos Tweets todos os anos pelos empreendedores. A amostra de empreendedores é maioritariamente dos EUA e os investidores que financiaram as suas ideias utilizaram várias estratégias diferentes de investimento.
O propósito é comparar estes "rácios emocionais" com os seus processos de financiamento. Assim sendo, o objetivo é testar como é que as emoções são afetadas pelo facto da start-up ser financiada ou não receber qualquer tipo de fundos. Os resultados foram analisados através do programa Stata e são responsáveis por várias conclusões interessantes que são discutidas ao longo da dissertação.

ACKNOWLEDGMENTS

This work has been possible through the help and sustainment of many people. Firstly, my supervisor Francesco Castellaneta, who helped me along the whole work, since September, with constant feedbacks and a terrific patience. Another huge thanks must be said to Amulya and Prof. Stefano Brusoni, who allowed me join them last summer and to access the data gathered by their ETH team in Zurich. My family has been the main reason why I am here, now, with a Master's degree. Without my parents', Luca and Daria, as well as my sister Marta's support I would not be here. Their encouraging words and actions mixed with their immense love pushed me to go through many situations in my life, and to obtain important achievements, so "grazie mille" is the least they deserve. I also would like to thank my friends, my Italian crew Matteo, Lorenzo, Andrea, Claudio, Stefano, Davide, Chicco e Simone who have been and are going to be an incredible source of happiness and joy in my life. A mention is also necessary for the wonderful people I met in this year and a half in Portugal, they made this experience unique. In those last months, I also met Ilaria, who has become a not-substitutable person in my life and thanks to whom I felt love again. Eventually, I would like to address this whole work to my grandmother Alba, who left us this summer. Even though she has never read something in English, I am sure she would have been really proud of his nephew. She has been a wonderful person who is going to be forever in my hearth.

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

Are emotions affected by an entrepreneurial process? This is the main question this paper aims to answer, providing a point of view which embeds both the emotional and entrepreneurial side linked with a founding process. The main issue is understanding what happens after being founded, emotion-wise. In fact, many paper have been discussing entrepreneurs' emotions (Drnovsek, 2005) (Cardon, 2012), but it has not been treated what is the process to become an entrepreneur's influence on the emotions. This helps to have a clearer idea of what are the consequences of entrepreneurship on someone's feelings, stepping out from the pure monetary side. The peculiarity of the research is the tool adopted to gather the information about the emotion: the social media Twitter. Entrepreneurship has been stressed under different aspects, but there is a lack of information about it linked with emotions and social media. The most common studies deal with social entrepreneurship, which is more likely to be involving emotions (Toyah L. Miller, 2012), (Goss, 2014), (Arend, 2012). Through the whole work, there are many positions similar or personally interesting articles/papers, which will support the different information provided, as well as will back up my assumptions. In this context, the focus is not on the profitability of the company, but on the very first moments of its "life". The attention is given to the birth of the firm, in the moment when the inventors notice the idea and commit themselves to it through money. The final step of this process can end up in different ways, generally defined under the category of founding (Lueger, 2012). Hence, in this paper founding is considered as the final outcome of the entrepreneurial process (Shepherd D. A., 2013). An entrepreneur, aiming to be founded, has to go through many stages before presenting his product to the investors. The feedback received from those people is going to bring the company to be founded or not (successful or a failure). This last step perfectly matches with the emotions enhanced by it, being that either positive or negative (Lueger, 2012). Are those emotions in line with a succesul or disastrous founding process? Are they directly linked? Successful founding implies an increasing of positive emotions, and vice versa? How the emotions answer to the presence of founding is the focal point of this research. More precisely, it is not taken under analysis which emotions are generated, but if the emotions are, more broadly, positive or negative. What are the most prevailing feelings? This is the purpose of the paper, understanding the relationship of founding and emotions, without considering the other thousands of factors that are evaluated before triggering an investment process. But, are we supposed to understand the emotions behind this process? In this paper, the social media, Twitter, have been adopted to gather information about the personal inner features of their users. Twitter is the instrument adopted to come up with the entrepreneurs' emotions. An interesting topic that could have been stressed is if the presence of a strong network can enable better outcomes, under both a financial and emotional side. For instance, there are many interesting papers about solo entrepreneurs (De Klerk, 2002), which display how an entrepreneurial progression can be valuable even though run by only one head. But, in this paper the team size as well as the other factors intervening in the process have been put aside. In fact, the only factor taken in consideration is the nature of the emotions and from those, it has been possible identifying the research questions:

- **RQ1:** Does a successful founding process increase the level of an entrepreneur's positive emotions?
- **RQ2:** Does a successful founding process decrease the level of an entrepreneur's negative emotions?

1.1 The structure

The work consists in an academic dissertation aiming to, as already said, to inspect the effect of start-up founding on the entrepreneurs' emotions. To help the readers to go through the paper, a brief summary about its structure is here presented.

Chapter 2: The literature review focuses initially on the previous works that tackled this topic, showing that this has been treated formerly. This short paragraph introduces to the three main variables taken in consideration along the whole work: Twitter, Founding and Emotions. Twitter has been adopted, through the LIWC text analysis, to gather all the information about the Emotions. They have been investigated for the whole research long, identifying how they react to the founding process in a start-up and have been split into positive and negative. Founding regards the process that takes a company to get financially sustainable, and it has been stressed its influence on the entrepreneur’s emotions indeed. Eventually, the last chapter aims to combine the information about Twitter, Founding and Emotions with the objective of providing a framework within inserting those three concepts. The reasoning behind this will deliver the hypotheses that constitute the base of the analysis (Chapter 4)

Chapter 3: This part dives more deeply into the numerical part of the paper. It focuses on the way the data have been gathered and clustered to obtain the final database, on which the analysis has been developed afterwards.

Chapter 4: This shows the actual analysis, presenting some code strings from the software Stata. This part answers to the hypotheses which have been identified taking in consideration the all former contents of the paper. It yields the final results, which gave room to further reasoning

Chapter 5: This is the last relevant chapter of the paper, as far as it concerns contents. In fact, it sums up the outcomes of the analysis and explains what could be the next steps, following this paper.

2. LITERATURE REVIEW

2.1 Previous works

The main question, the research pretends to find an answer for, is how an entrepreneurial process affects the entrepreneurs' emotions, understood by a meticulous analysis of their Twitter accounts, and the consequences of these on the "founding strategy adopted" of the company (temporal based). This question has emerged as the main one due to the absence of a clear and narrowed analysis about this topic. Entrepreneurship and emotions have been covered in many papers but the scope has never been pointing out the consequence of entrepreneurship on emotions. Most of the literature about this combination tackles the influence of emotions in the entrepreneurial development. Mateja Drnovsek, Melissa S. Cardon, Joakim Vincent and Jagdip Singh in 2005 (Drnovsek, 2005) wrote an interesting paper about entrepreneurial passion as the main driver for the development and success of a company, together with a large explanation on how the emotions develop along the entrepreneurial process. Both those aspects have been recalled in this research. About time, the date when the founding took place has been identified and merged with emotions identified in that precise year. Entrepreneurial passion has been presented under the pure emotional aspect, showing whether the emotions are positive or negative. This paper gives the chance do dive deeper into the temporal topic within the analysis. In another paper is reported the entrepreneurs' ability in identifying the different commercial occasions related to their business (Keh, Foo, & Lim, 2002), stressed also by Maw-Der Foo in 2009, who explained how the evaluation of emotions influences how the risk is perceived and the preferences related to this. In this work is presented the distinction between state and trait emotions: the first ones refer to the individual capability of feeling certain emotions, while state emotions indicate situations where only certain emotions are taken in consideration. This is what it is going to be tested: the emotions' behaviour when the founding process takes place. Those emotions have been gathered through Twitter. Conducting a text analysis on entrepreneurs' Twitter accounts has allowed to collocate the emotions in a temporal framework where the emotions have been gathered along the founding process, and this has been proven to be effective (Fischer, 2014).

As briefly said above, the whole work's base is the text analysis made on thousands of Twitter accounts both privates' and companies'. This has been possible with LIWC (Linguistic Inquiry and Word Count). LIWC collects 2,300 English words which refer to almost 80% of the terms adopted in daily dialogues and point out the emotional perceptions behind these. The stop-words in the tweets have been removed to count the amount of words which can be identified as positive and negative. Those two categories refer to "positive emotions" and "negative emotions" within LIWC's dictionary. The two sections have been merged to calculate the LIWC score, which recalls Kramer's score (Kramer 2010).

$$Sentiment_i^{WC} = \frac{p_i - \mu_p}{\sigma_p} - \frac{n_i - \mu_n}{\sigma_n}$$

Figure 1. Kramer's score (2010)

In order to have clearer idea about this formula, the different variables have to be explained briefly:

- pi (ni) is the ratio of positive (negative) words referring to the user i
- μp (μn) is the ratio of positive (negative) words for the total number of users
- σp (σn) is the standard deviation referring respectively to positive and negative emotions

The distribution of positive and negative English terms consists in the normalization method developed through means and standard deviations. This method has been adopted by Jon Quercia, Daniele Capra, Licia Crowfort in 2012 to deduct the nature of the emotions (positive or negative) form people's Twitter accounts (Quercia, 2012) . Furthermore, apart from many articles tackling this topic I have many data available gathered by Ph.D. students of ETH Zürich (D- MTEC). To verify whether the companies have been founded and how much has been the amount obtained they adopted the data available on Crunchbase[1], mostly for what concerns the founding rounds. In fact, the founding process is the main driver for understanding the performance of a start-up, meaning that a larger amount of money raised as well as several founding rounds, point out the appeal of the company toward the investors.

[1] https://www.crunchbase.com/#/home/index

They are disposed to disburse money for companies which can provide likely high returns, even though Shikhar Ghosh from Harvard declared that the failure rate for start-ups in the US oscillates between 30 and 40%. Of the remaining 60%, 40% returns the initial investment and the 20% booms and succeeds (Gage, 2012). He also declared that another problem is timing, which can directly influence the opportunity of a company to be founded or not, considering also its exit strategy: becoming public or being acquired. This stresses the main question of my research, identifying the reaching of the founding and throughout which strategy as one of the main outputs to whoever investors bases his venture process. It also mentions the difference between entrepreneurial and personal failure, which Ghosh explains as the stubbornness to avoid the entrepreneurial failure. This last one is considered by the author as a constructive failure, while that personal is able to destroy a career. How should I consider the entrepreneurial failure within my study? Always recalling what Ghosh said, also company that failed can create business opportunities under the network point of view, binding venture capitalists with entrepreneurs who have seen their companies already flourishing. Furthermore, he mentioned how in Silicon Valley, the failure of your company can be also seen as a “badge of honour”. Including a negative exit strategy, such as failure, is a useful predictor for the overall results of my work. In addition to the positive or negative nature of the failure, it would be interesting to include whether this is personal or entrepreneurial, but with the data we dispose it is impossible to come up with such details (Gage, 2012). It could also be taken in consideration the investment risk, which strongly affect the company’s – investors’ returns, but this a way too broad topic to be included within this research. Agnes Virlics (Virlics, 2013) shows the process behind investment decisions and the risks generated from them. It stresses two important features regarding the risk itself: the objective component within the investment and the subjective factor of the investment decision making. As already mentioned, this topic requires a whole work about, and this would deviate from mine.

This brief introduction aims to clarify the two sides of the methodology: the research and the analysis. Both of those need literature references able to back them up and prove the validity of the data I am using. The research combines mostly literature references with available data obtained by other people’s work. The process through which these data have been gathered will constitute the research method, while the analysis method regard how this data will be processed to show the eventually show the research's final outcomes. The whole set of data I dispose has been kindly given us by the D-MTEC department of ETH Zurich, after a summer internship I did there. This implies that the whole explanation of the data gathering process is not referring to us, but to them. So, all these are secondary data, that I have been permitted to access by the ETH University based in Zurich. The references regarding the data gathering, about all the different datasets (Twitter, Founding and Emotions) refer directly to a paper published in 2015 by three people within the MTEC department of ETH University in Zurich (Tata, 2015)

To give a clearer and more precise idea about the whole data gathering method, it is necessary to distinguish the data as well. In fact, there are different datasets regarding diverse parts of the research. Twitter, founding and emotions are the three main categories, which need to be discussed separately.

2.2 Twitter

Twitter data regard all the tweets posted by the entrepreneurs on the social platform, as well as their personal information. The easiness to access those data together with their easily interpretable message template, allows many more users to adopt microblogging services (Haewoon Kwak, 2010). Usually, microblogs' authors tell about different topics, ranging from personal issues to discussions about current trends (Haewoon Kwak, 2010). The reliability of the Twitter data has been also confirmed as far as it concerns the entrepreneurs' thoughts spontaneously (Fischer, 2014) and, especially if the user is a chain-publisher, all the tweets posted shape a precise longitudinal set of data about their thinking (Tata, 2015). In the ETH paper it has been explained how the dataset that I am going to take in consideration has been shaped. Twitter data of public domain have been collected, dealing both with the founder's private and company's account (Tata, 2015). The focus about the data collection is around the Californian area and uses a set of 21048 Twitter accounts from previous research (A. Abisheva, 2014) , extracted from a large dataset of Twitter users in the US. The accounts were selected per their location as disclosed on their profile, matching places in California by using Yahoo! Placemaker. The oldest tweet in this dataset was in 2006 and the latest tweet in mid-2013. In total, 2880 company accounts have been selected, whose around 80% had personal Twitter accounts about their founder(s) (Tata, 2015). This means that there is no distinction between solo and companies founded by more than one individual. We only focused on the personal Twitter accounts of the founders to have a clear measure of their expressed thoughts. To give consistency to our data we took in consideration accounts presenting at least 2000 words. The boundaries of the Twitter API did not permit to gather as many information as available, it has not been possible to collect more than 3200 Tweets per person indeed (Tata, 2015). Furthermore, it has been possible to obtain each person's total number of Tweets since joining Twitter, which recalls the longitudinal concept mentioned above. This yielded a sample of 3796 entrepreneurs. In fact, due to the quantity constraints, the tweets show also the frequency of people's posting. An assiduous Twitter user will have a shorter temporal gap to analyse, while a seldom user is taken in consideration for a longer time frame. It is also worth mentioning the fact that entrepreneurs approach the tweets exactly as they approach a normal conversation, making them feel free to share what concerns both them and their companies with their followers (Tata, 2015). Recent work by Lee *et al.* (2016) has shown that what a CEO tweets reflects a transparent picture of his personal features (J. M. Lee, 2016). The spontaneity of Twitter data is very high, in fact, even though some contents are automatically banned by the social network itself, they are way more direct than other types of information (Tata, 2015). Finally, due to the availability of Twitter data over time, it enables multiple observations for each individual. Twitter data have been adopted in many research contexts: understanding the mood of the stock market (J. Bollen, 2010), to manage crises in organizations (F. Schultz, 2011), to predict election results (Andranik Tumasjan, 2010), and to assess how users view an organization (C. Tan, 2011) (Tata, 2015).

2.3 Founding

This entire knowledge about the financing trial has been extracted from Crunchbase (see above), which allowed to observe the whole investors' steps toward the companies. Crunchbase is an online crowdsourced database providing information about start-ups, especially tech ones. Its active users base is formed by almost 50,000 people, proving the validity of its findings. It has been pointed out the information's reliability that databases like Crunchbase have been gaining in the research context

(Arora A, 2012) (Tata, 2015) (Anne L.J. Ter Wal, 2016). It is noticeable that also from this dataset emerges the temporality of the data, meaning the temporal scale along which the whole research has been developed. The time range goes from March 2008 to June 2014, Crunchbase's information are displayed per companies, without stressing the individuality within them, anyway they show the number of founders. The data are presented into the different founding rounds, gathering information about the number of rounds and their date, information presented in the final database as cumulative ones, such that there is the total amount invested along the whole set of rounds. In order to obtain a homogenous sample especially under the temporal point of view, and to focus on an environment listing many tech start-ups newly founded, the analysis' sample has been mainly extracted from the San Francisco Bay area (Tata, 2015). In 2013, P. Delevett showed the importance of this area regarding the start-ups' context, which indeed received around 46% of the whole national VC founding (P.Delevett, 2013). A research conducted in 2002 by USA Today highlighted how the San Francisco Bay area was the first in the list concerning the start-ups' scenario (Graham, 2012). The main reasons for this are the availability of talent, relatively low costs, and the possibility to raise money (Graham, 2012).

The data have been downloaded by the ETH team ((Tata, 2015) on 6th June 2014 using Crunchbase's Excel exports (Crunchbase, 2014). This final spreadsheet (crunchbase_monthly_export_2014) presents several sheets that progressively narrow the scope from the whole founding event to the single founding rounds. In fact, it initially stresses the companies as a unique entity, which received a certain amount of money in X rounds, afterwards the scenario changes. As shortly mentioned before, there is presented every single round. This comes with other information apart from the amount of money, like the type of founding received (VC, seed, convertible note, debt financing) and the time period of the financing, starting from the year, the quarter, the month and eventually the day. Another sheet presents further information about the investors, diving deeper in the individuals/companies/founds that invested in the company. The information, even though vast, is very precise: it is possible to deduce the country, region and the city of the company financed as well as the investing company, as far as this is not an individual or there is not another bigger found owning the one acting as the final investor. In this case, the spreadsheet matches every company with the investor's financer, either he is another company or an individual (e.g. Robert Branson), along with the investor's market (consulting, investment management, agriculture, marketing and so on). The date and the kind of founding, useful for doing a cross check with the rounds' sheet, are also mentioned for each investment. Lastly, it is presented the total quantity founded by every investor to the companies. This information is not divided per investor, but per company, giving us the chance to then group all the companies financed by a single investor. In fact, many of them invested in diverse companies presented: for instance, Richard Branson financed 10 different firms.

2.4 Emotions

This is the most difficult part about the data gathering, since all of them have to be interpreted after having been collected. In fact, the emotions result from extracting information throughout LIWC (see research question paragraph). This software is used to compute the level to which both cognitions and emotions are identified within a text. To calculate this number, it adopts the frequency of words and of phrases linkable with a specific word category. For example, the category referring to the past will list words such as "could have", "did", "had" etc. The whole set of measures reported in the database has been calculated and normalized in line with the length of the tweet, such that the information reported are directly proportional to the entire number of tweets that have been analysed before (Tata, 2015). This process avoids the measures to be biased by either the "tweeting" frequency or the number of words tweeted (Tata, 2015). LIWC has been recently adopted to cipher the CEO's temporal orientation (Nadkarni, 2014). The results yielded a high level of reliability of LIWC for calculating temporal measures. This is why LIWC has been as the main tool to come up with the most relevant variables within the dataset, especially about the entrepreneurial process (Tata, 2015). Entrepreneurship is referred to an "emotional journey" (Cardon, 2012). When creating companies, entrepreneurs are exposed to factors such as high uncertainty, work overload, managing individuals and capital (Baron, 2013). Thus, it is commonly assumed that entrepreneurs experience higher levels of stress and negative emotions. Negative emotions are not only generated by situations internal to the company and entrepreneurial development, but also from the interaction with "social capital". This concept regards "the direct participation of at least two parties on a recurrent basis with a shared focus on the exchange of resources" (Goss, 2014). Recent studies focus on the influence that emotions have on individual preferences or behaviours, without mentioning how they affect the relationships among the various economic players (Bandelj, 2009). A definition of emotion is: “An emotion is a reaction to a stimulus that has consequences for cognitions and behaviours” (Frijda, 1988). So, which role do negative emotions have within entrepreneurship’s social influence? The role of negative emotions when social interactions are studied is to work as a Negative emotions have salience in the study of social interactions, functioning as a memento about the attention to be given to interactions (Larsen, 1989), pushing toward conformity processes (Taylor, 1991), and eventually to enhance relationships’ preservation or to create a more homogeneous resources’ distribution (Morris, 2000). The relevance for the entrepreneurs is such a high level, that it can consist in either success or failure (Goss, 2014). This recalls the founding approaches, main topic within our research. Another important feature regarding the driver and passion behind entrepreneurship is passion (Smilor, 1997). This term must not be equivocated, in fact passion means “a deeply felt or strongly held emotion, such as hope, pride, anger, frustration” (Nair, 2003) and others. On one hand, passion has been also mentioned as a lasting emotional state, easily associable with enthusiasm, joy or zeal (Smilor, 1997). On the other hand, other scholars identified emotions as a consequence deriving from a precise event, like a company’s failure (Shepherd D. A., 2013). Due to the importance held by both the temporary and enduring feelings, this prevents them from being overall consistent with each other. The interesting concepts deriving from those two concepts is “sweat equity. This boundary pushes entrepreneurs to attribute business success to themselves, as well as business fallacies are identified as their own failures or mistakes (Shepherd D. A., 2003). This is an important aspect that has to be added to my research. Passion seen as a bucket filled with emotions directly linked to the entrepreneurial success or failure, enables a deeper understanding of what an entrepreneur feels during the whole business process. Yet, as Shane and Venkataraman (Shane & Venkataram, 2000) tell us: “Entrepreneurial thoughts and behaviours, and by extension emotions, are not stable characteristics that differentiate some people from others across all situations”. But, the individual must be taken in consideration with his venture, due to the linkage between them, that results in entrepreneurial success (Shook, 2003). Therefore, it is worthier considering emotions in entrepreneurship as the whole mechanism behind the process and not only as the feeling involving a single character.

2.5 The Twitter-Founding-Emotions “mix”

This chapter aims to explain the linkage between the three concepts analysed above, to recognise the most efficient way to create this binding and eventually to come up with the hypotheses to be tested in the analysis.

Many entrepreneurs are using Twitter to enhance their business activities and to widen their personal network (Haewoon Kwak, 2010). Therefore, Twitter is the perfect place to find information about what those people feel (Fischer, 2014). Everyone’s feelings are related with their life events, from the daily live situations to precise and life-changing events. In entrepreneurship, there is a specific term related with the emotions linked with this concept: entrepreneurial passion (Cardon, 2012). Many studies have been conducted to analyse in more detail this concept. The most relevant model regarding affective experience is the “circumplex model” of affect (Remington, 2000). This model identifies an underneath structure which places the affective emotional situations on the borders of a circle (Remington, 2000), which is represented by two different dimensions: core affect and valence. The pattern followed by entrepreneurial emotions strongly recalls those two dimensions. That is exactly what this research aims to do, identifying how those emotions are affected by entrepreneurship, intended as a start-up process. As mentioned by this model, a part of the underlying structure of the emotions is pleasure-displeasure, which are meant in this paper as positive and negative emotions. Until which extent the level of those emotions is affected by entrepreneurship? This question will be answered through the analysis.

How can the founding distress the emotions of an entrepreneur? Money has a huge influence on people, especially on their psychological dimension (Vohs K. D., 2006). Founding is the most important and relevant event in the early life of a start-up and its occurring strongly affects the involved people’s life as well as their feelings (Shepherd D. A., 2013). That is why in this paper, founding is considered as the main factor regarding the influence on the emotions, in order to understand whether positive or negative ones prevail, and how. This equation has founding as the fundamental variable, even though the term is not specific for a special founding process, but it refers more generally to every kind of money stream within the start-ups (Weiblen, 2015). The amount founded is also not taken in consideration, in fact the data regarding the founding size were uncomplete and incompatible with the analysis intended to be developed. As far as concern the founding process, there are many information available, but the focus has been directed to solely an event: whether there has been founding or not. Therefore, the most appropriate variable, mentioned in the following chapters, is a dummy variable pointing out whether founding took place or it did not. This variable allows to immediately understand the linkage between the founding process and the nature of the emotions behind it (Foo, 2011). The other set of information regarding founding is more investors-oriented. It tells about the country, investment strategy, investors' personal information, all information that could be used for developing an analysis for instance regarding the influence of the investors' features on the start-up development.

As mentioned at the beginning of the chapter, Twitter is the most suitable platform to identify this information (Qiu, 2012), and with Crunchbase's data, has been possible merging the whole amount of information and to come up with the final database through which the analysis has been developed. Twitter's function has been to provide the emotional aspect that has been inspected along the whole work, using the text analysis. This is a common practice, and the results' validity has been proven to be highly suitable for this kind of research (Summer, 2012). The ratios concerning the emotions refer to two general categories, positive and negative emotions, which are intended as the two emotional outcomes deriving from the entrepreneurial process. The information provided by Twitter and that available through Crunchbase aim to identify the two main hypotheses detected in the following analysis. They regard the influence of the financial process on the emotions, considering them as merely positive and negative and zooming into the way the founding affects them.

- **H1:** The positive emotions increase when there is founding
- **H2:** The negative emotions decrease when there is founding

The following chapter will show how the data have been collected and clustered, in order to reach the analysis with the data here explained and then merged within a unique database.

3. METHODOLOGY AND DATA COLLECTION

1.1 The research

The whole work resides on the influence that an entrepreneur's business life has on his emotions. The model I mean to develop, aims to put the emotions as DV, attesting that many factors can influence them, intervening both on their nature and on their changes. This has been possible throughout the coding of their Twitter accounts. Collecting all the information has generated a data base ordered for categories, according to the different terms encountered in the Tweets. For instance, under the voice *anxiety*, there are many terms such as *afraid, discomfort, dwell, overwhelm* etc. All of them refer to the root of the word, and all the compounded words as well as the adjective and so on, refer back to this category. There are diverse databases taken in consideration, but the variables measured at this point of the analysis are included in *twitdat_annual_140826_2b*. This set of variables presents more than 35,000 rows, divided per user whose account is split along the amount of years considered by the quantity of tweets. As stated before, the years' range depends on the tweeting frequency. In fact, some Twitter accounts have data available from 2006 to 2014, like *barbralee*: this user's tweets have not been very frequent and she started using Twitter early, enhancing us to have a wide temporal gap (9 years) to analyse. Every column of the database presents different variables extracted through the text analysis (LIWC), tackling different topics depending on the analysis' scope. In fact, the data gathered regard text words, emotions, time, general activities and Twitter numbers as well. All this information has been obtained using a table listing all the words that were fitting within the definition of the category. Within the file *"LIWC_CODE"*, all these terms have been mentioned in alphabetical order respectively for every word then adopted in *twitdat_annual_140826_2b*. This suits perfectly the initial phase of the analysis, whose aim is to identify possible relations between all these variables grouped per Twitter' s ID. Furthermore, this also allows to compute the impact of people's characteristics on the different communicational attitudes they adopt. For instance, some studies tried to identify the gender from well-written (Pennacchiotti, 2011), traditional text (Herring, 2010) (Singh, 2001), blogs (Burger, 2010), reviews (Otterbacher, 2010), e-mail (Garera, 2007), user search queries (Jones, 2007); and, for the first time, Twitter (Rao, 2010). To understand in more detail the Twitter users' stream and their network structure, the Twitter users' personal features have been gathered and analysed very precisely (Pennacchiotti, 2011). "Show me how you tweet and I will tell who you are" is a sentence that Java and colleagues (Java, 2007) made come true. They identified users who are information seekers (seldom post) and information providers (the opposite and URLs in their posts). It has been affirmed that, differently from Java et al.'s assumptions, users' profiling is given by linguistic features (Rao, 2010). The linguistic features, in particular whether concerning a specific topic, have been proven to be reliable by the analysis' results. The contents of the social network can also be explicit, and even if their collection is pretty demanding under the financial point of view, they are worth it particularly when the people targeted are celebrities, or popular people in general, whose presence on Twitter is very high (Pennacchiotti, 2011). In this case though, we consider Twitter categories not for their popularity but type. In fact, we are targeting entrepreneurs' Twitter accounts, whose importance is about what they post.

The research considers the "Twitter-based social interactions", coding adopted in the research held by Fischer and Reuber in 2011 (Fischer, 2014). The paper adopts an "inductive, theory-building" methodology, which enhances the development of definitions about the impact on entrepreneurs' effectuation process when using Twitter. The paper showed that interactions through Twitter can lead to effectual cognitions, even if the abuse of this kind of interaction can generate an effectual disturb. Fischer and Reuber identify also the "perceived time affordability" as one element to deduct the degree of social interaction adopted by an entrepreneur trough Twitter. For my research, the perceived time affordability is the main factor under the temporal gap point of view, and not under the frequency of publishing on the social media. Another important aspect emerging from Fisher and Reuber's paper are the people targeted: entrepreneurs (Fischer, 2014). They investigate how they got involved in Twitter, its influence on their business and their thoughts about being both producers and consumers of this service. The mean adopted to gather this information is interviews. This method provides a higher quality of data, both under the information and additional insights due to the face-to-face situation, but they are

costly and do not allow to have a large sample size (Wyse, 2014). In our case, the dimension of the sample is essential to address as many entrepreneurs as possible, stressing the largest number of fields, types of founding (JVs, VCs and so on), as well as emotions intervening during the process. The more unbiased the sample is, the better is for the final outcome of the work.
I am always working on *twitdat_annual_140826_2b*. The size of the sample must be in line with the rest of the work, that is why the initial database has been drastically reduced. Many variables included are too specific and/or do not point out meanings important for the work. Without providing a useful source of information for the development of the analysis. additionally, the temporal gap analysed must be coherent with the whole entrepreneurial process, denoting that the founding process has to reach an end able to show the exit strategy of the company. In this case, it is worth starting examining another database: *yearlyfirmvariables*. This dataset is important for financial data's analysis. In fact, it reports the data divided per each founding round. For each round, as mentioned previously, there are data about the founder, the type of founding, the amount founded and the time gap between the different rounds. Several data are not available (NA), that is why there have been applied filters to consider only the consistent and relevant information. The dataset presents cells differently coloured, which show the NA cells for every founding round. It results that the white cells (due to cells' formatting) are the only having all rounds' data available (until the 4^{th}). Further in the investigation, it is important understanding what is the least number of rounds to consider, and whether it is necessary a constraint about a minimum amount of money collected. Another solution, that is the most likely to be adopted, is presenting all the founding rounds under the same variable, in such a way it is removed the bias generated by the different number of rounds per company and the different amount raised. In fact, working on a single financial variable enables to go more deeply into the emotional aspects.

3.2 The dataset

The data have been clustered in a final dataset "*StataFile.*xlsx", which presents many details about several users. The two main databases adopted to come up with this data are *twitdat_annual_140826_2b"*and *yearlyfirmvariables*. The most evident problem in merging those two sets of information is their different structure and amount of data. The first database presents a different temporal gap among the users which, as mentioned already above, depends on the frequency of tweets per each user. This is a problem related with the possibility to group the different accounts. In fact, there are not criteria to obtain clearly distinct groups of Twitter accounts and it is even more complicated to deal with it due to the very vast amount of data. *Yearlyfirmvariables* presents a different Twitter account in every row of the spreadsheet. The total amount of user IDs is way lower than in *"witdat_annual_140826_2b*, in fact *yearlyfirmvariables* presents almost 2,500 accounts, where almost 500 do not present any information (NA). Due to the complete lack of data in those rows, I removed them to work only with the accounts presenting at least one founding round.

These two databases have then been incorporated within the final one, adopted for further developing the analysis. The final database has been obtained by conditional formatting the two datasets and identifying the Twitter accounts present in both. The work has required more time than expected, due to the different way of presenting the names of the users. In fact, a manual selection had to be done before of proceeding with the creation of the last database. A part from that, this cluster of data presents both the founding details and the ratios of the emotions deriving from the tweets. Initially, it presents a general overview of the positive and negative emotions, shown as *posemo* and *negemo,* extracted from *twitdat_annual_140826_2b*. Both those variables, as all the others in the database, have been created throughout a yearly tweets' text analysis based on the presence of certain words able to define the category itself. These categories have been validated in previous studies with self-reported measures and have been shown to capture inter-individual differences excellently (Timmons, 2015). *Posemo*'s dictionary consists of 407 words such *happy, glad* and *success*, while that of *negemo* is formed by 506 words, which can be categorized in *anger, sadness* and *anxiety*. These three sub categories are in *twitdat_annual_140826_2b* too. LIWC's validity has been proven by a study about the differences between Christians and atheists, developed following the text analysis (Ritter, 2013). LIWC categories about positive and negative emotions have also been shown to be relatively independent (Quercia, 2012).How have been all those "emotional" variables been calculated? Every year (*t1*=2008, *t7*=2014) each Twitter

ID has *x* words. Accessing the dictionary referring to each variable, LIWC gathers *y* words referring to that precise category and the result is a ratio which represents the percentage of *y* in *x* from *t1* to *t7*. The words are cross-variables, denotating that some words can refer to more than one variable. That is why the results have been normalized according to the text length. After merging the two databases have been identified 1,592 different IDs, each of them presenting the same years' range, even though a part of the IDs do not report data for this whole time period. For instance, the first ID is *0xdata* (see *figure3*), and it is present from 2008 to 2014, meaning that this ID's data have been reported for the all seven years in *twitdat_annual_140286_2b*. Nevertheless, also the second ID *1000memories* has been listed along the same seven years period, but the 0 in *posemo* and *negemo* mean that in the years 2008, 2009 and 2014, where the ratio is 0, those values have not been analysed because not present (see *figure3*). This permits to identify the development of emotions along the founding process and compare it with the users who have not had the chance, or have not caught the opportunity, to get founded. Furthermore, this structure (cross-section/ time series data) must be presented through a panel data, where the IDs are observed along the years they have been monitored together with the emotions' (*posemo, negemo*) ratio, even though it is important reminding that the years' range is from 2008 to 2014 to have a consistent panel data. Those data are presented with a FE (Fixed Effects) model that enables to analyse only the variables that change over time. The effect of the features neutral to time has been removed by the FE (Fixed Effects), such that is easier coming up with the effect that the predictors have on the variable considered as a outcome. Are the founding procedures directly linked with the emotional status of the entrepreneur? If yes, how does it affect them? The analysis aims to identify how the founding process within an entrepreneurial journey affects the entrepreneurs' emotions. This hypothesis is tested adopting the variables *posemo*, *negemo* and a dummy variable able to show whether the founding happened: *DPF* (Dummy Post Founding). *DPF* is equal to 1 whether there has been founding, to 0 if the founding has not been done and -1 if there are not data about it. The other variables included in the data set, a part from the *ID, Year, DPF, posemo and negemo* are *country, investmentsize, investmenttype* and *founding data*. The variable *country* symbols where the investment has been made and it is divided in *US* (United States) or NA (not available) and so it is not consistent for the purpose of the analysis. In fact, this variable is not changing along the years. The same happens for founding data, which identifies whether the founding data are available or not, being *DPF* = 1 or =0 for *foundingdata= Founding* and *DPF*= -1 for *foundingdata= Unknown founding* data. The others two variables regard the investment done from external actors in the start-up. *Investmentsize* shows the different ways of founding adopted by the investors: *angel, convertible_note, debt_financing, grant, non_equity_assistance, private_equity, product_crowdfounding, seed, undisclosed* and *venture*. This variable presents also some NA data, that must be excluded from the set of information about the investment type. *Investmentsize* shows the total amount invested in the start up during the whole founding process. This number has been reported though the whole-time gap analysed, meaning that it is exactly the same for each year. Both this variable and *ID* have been grouped, such that along the different years, the data are not repeated. The analysis has been developed through the software Stata.

4. RESULTS' ANALYSIS

As already mentioned, the main objective is to identify how the start-up process affects the entrepreneurs' emotions. The two hypotheses mentioned before are going to be tested in this part.

To verify the hypotheses, and to answer the main research question, it has been adopted a panel data with fixed effects model. Stata, as already mentioned, has been adopted to run the regression and to come up with the tests' results.

The dataset "*StataFile.xlsx*" has been imported in Stata and saved as "*Regression_data.dta*". This is the file through which the analysis has been developed and used along it.

The data must be reviewed to make the analysis possible, especially because the model adopted has strict requirements. That is why I had to get rid of the negative values of *DPF* (*DPF* = -1), in fact those data could not be processed in the analysis and they have not significance in the analysis:

drop if dpf==-1 – to remove the values of *DPF* equal to -1

As anticipated before, *ID* must be grouped to avoid the presence of same data in the data set. This process has been feasible throughout the creation of a new variable, which include the IDs groups just mentioned above: *newid*

egen newid = group (id) – to group the variable *ID* in *newid*

This formula allows to shape the information within the panel data, so that is possible to run the analysis and eventually to yield the results.
Before, the string variables must be converted into numeric variables, and the function can also be reversely applied.
Another set of data which needs attention is *Investmenttype,* even though this is not a relevant variable for the analysis. In fact, it is a string variable and it cannot be used in a factor analysis.
This variable presents some errors (NA) though, which must be erased, before creating
keep if investmenttype=="angel" | *investmenttype=="convertible_note"* | *investmenttype=="debt_financing"* | *investmenttype=="grant"* | *investmenttype=="non_equity_assistance"* | *investmenttype=="private_equity"* | *investmenttype=="product_crowdfounding"* | *investmenttype=="seed"* | *investmenttype=="undisclosed"* | *investmenttype=="venture"* – this formula keeps only *investmenttype*'s values which show a category of investment, removing the NA values.
After having clarified those issues, the analysis can begin, from stating through a formula the intention to use panel data:
Xtset newid – Now the data set I am working on is a panel data
Xtset panelvar defines the data set to be a panel, and the observations' order is not important. The usual panel data formula in Stata is: *xtset panelvar time* var, but the *timevar* is not necessary because the time range is repeated for every single ID and Stata does not allow to consider it as a *timevar;* "repeated time values within panel" (Stata' automatic description). In fact, in this case, *newid* (see above) is the group of IDs and within that there are the IDs listed in each year, from 2008 to 2014.
The model to be set must follow, as already said, a FE (Fixed Effects) model within panel data. The DV (Dependent Variable) varies depending on what is the purpose of the analysis: analysing the behaviour either of the positive or negative emotions toward the presence or not of the founding (*DPF=1* or *DPF=0*).

xtreg posemo i.dpf i.year fe cluster (newid)
In this first regression *posemo* has been identified as the DV, and it is shown along the year range 2008 – 2014, for *dpf* assuming value of either 1 or 0. This regression is presented with the fixed effects model, which removes constant term, and the *cluster* function that clusters the standard errors indeed.

Figure4 shows the strong influence of *dpf* on *posemo*. In fact, the level of positive emotions increases of 2.612 points when *dpf=1*. This recalls H1 (Hypothesis 1), proving that this is true: the presence of founding increase the positive feelings an entrepreneur has. This result is an important signal for the research, proving that the founding process, if happening, enhances the entrepreneurs' positive sentiments. A financial strain caused by a job loss generates physical and mental illness that is facilitated by the reduction of perceived personal control (Vohs K. D., 2006). Nevertheless, in this scenario the prospective must be swapped.

Focusing on the emotions, in this case as consequences of a worth-to-remember event, and on the event, itself, can the above mentioned "job loss" be considered as a failure regarding the founding process? Yes, in fact an entrepreneur aiming to trigger a new business has as the task to make it profitable/successful. The fiasco of not receiving founds is going to prevent him to achieve that, creating a job loss. So, directing the analysis toward the negative emotional aspect, it is important testing the influence of a "job gain" on it. Through Stata it has been run a regression similar to that whose results are reported in *figure4*, but with the negative emotions (*negemo)* set as DV. The model is the same proposed before, but with a different DV:

xtreg negemo i.dpf i.year fe cluster (newid)
In *figure5* it is noticeable how the presence of founding affects way less *negemo* compared to the former analysis regarding the positive emotions (*posemo*). Nevertheless, founding affects positively also the negative emotions, which is something relevant for the final results. The negative emotions increase of 0.308 points when *dpf=1*, meaning that H2 must be rejected, in fact negative emotions do not decrease but slightly increase. This hypothesis (H2) was the "natural" consequence of H1, assuming that an increase in the positive emotions would have led to the sinking of that negative. As shown, that is not happened, even though the founding's

influence on negative emotions is almost eight times lower than that manifested by the positive emotions.

To have a clearer idea about founding process' impact on the emotions, being that negative or positive, in Stata it has been created a ratio comparing those two variables:

gen ratio_posemo_negemo = posemo/negemo

Like the former regressions, also in this case, the formula is the same but with the ratio as the DV:

xtreg ratio_posemo_negemo i.dpf i.year fe cluster (newid)

This last regression aims to test on the founding on the connexion between the positive and negative emotions, yielding the result of 2.104. This number shows that the influence exerted by the founding process (*dpf=1)*, independently from its nature, has a far stronger effect on the positive emotions than on the negative ones.

In the three regressions run before there is no point in diving more deeply in the coefficients' results, in fact those are linked with every single year from 2008 to 2014, and are used only to position the emotions within the funding period. Furthermore, due to the fact that not all the companies have been funded in the same year, whether they have actually been founded, it is useless proceeding further with the coefficients' analysis. The years contributed at the creation of the panel data, but their coefficients are not influencing the result taken in consideration for computational purposes.

5. MAIN CONCLUSIONS AND FUTURE RESEARCH

The outcomes of the analysis show that the start-process affects the entrepreneurs' emotions, as well as that Twitter is a valid instrument to gather this information ((Fischer, 2014). These two have been the main points of the research, highlighting important details about how the social media can be adopted to forecast feelings and relational ties (Gillbert, 2009) and pointing out the impact of entrepreneurial patterns on the emotions (Shepherd D. A., 2013). The findings have been surprising, especially as far as they concern the negative feelings. In fact, the positive side of a human being is strictly related with achievements (Corradino, 2016) and this emerges also in the LIWC dictionary adopted for calculating the *posemo* ratio, where there are terms such as commitment, determined, glory, improve, profit, values and wealth which identify a strong presence of a high striving level for achieving objectives. This is positively related with the emotions, bringing out the high correlation between the obtainment of the founding and the positive feelings behind it. Nevertheless, negative reactions identify a different trend, actually opposite to the one expected (see H2, pg. 32). In fact, the negative emotions increase when the founding takes place. Is this reasonable? Always looking at the LIWC dictionary adopted for the ratios' calculations, emerge important details from the *negemo*'s words: agitation, anxiety, loss, stress, struggle. Those variables are linked with the founding process and it is reasonable noticing this trend in the level of negative emotions along with the financial aspect, start-up wise ((Vohs K. D., 2006). Furthermore, the ratio between *posemo* and *negemo* highlights the different weights those two variables have. The difference between positive and negative emotions is noticeable and leads to the confirmation of H1 (first hypothesis) of the analysis. These are few of the terms recalling the direct link between money (founding) and emotions (Vohs K. D., 2006). Those findings allow to answer another important set of questions, which are related with the social media. In 2016, being active on those platforms leads to develop a wide and strong network that enhances relationships under many points of view (Coppersmith, 2014). This whole research has been developed from Twitter, enabling to come up with findings close and at the same time far from the entire social-monetary aspect. In other words, Twitter enabled to integrate "soft" aspects, such as words, within a way larger and deeper vision about entrepreneurship (Fischer, 2014). Large datasets, used to develop this analysis, enhance the development of many other works related to the entrepreneurial pattern, from scratch to the actual company. The data present all the different founding tranches that the company collected, as well as other rations about precise emotions and contexts within a person's life (work, leisure, friends and so on). It would be interesting understanding the impact of an "entrepreneurial adventure" on the entrepreneurs' daily life, and on his relationships too. Furthermore, as already said, there are also information the start-ups' team members and the investors who financed those companies. This leaves room for analysing the team aspects behind a new enterprise, pointing out how the creation of this together with its following steps are affected or also affect the team members and their interpersonal relationships. The attention given to entrepreneurship could be tackled by many other different sides, instead of getting stuck on the pure financial aspect of it. The personality of an entrepreneur could be an interesting starting idea.

6. APPENDIX

ID	Year	DPF	posemo	negemo	Country	Investment size	Investment type	Funding data
0xdata	2008	-1	2.48	1.72	NA	0	NA	Unknown funding data
0xdata	2009	-1	2.93	1.99	NA	0	NA	Unknown funding data
0xdata	2010	-1	2.89	2.03	NA	0	NA	Unknown funding data
0xdata	2011	-1	2.72	2.03	NA	0	NA	Unknown funding data
0xdata	2012	-1	3.37	1.73	NA	0	NA	Unknown funding data
0xdata	2013	-1	3.63	1.89	NA	0	NA	Unknown funding data
0xdata	2014	-1	3.88	2.49	NA	0	NA	Unknown funding data
1000memories	2008	0	0	0	US	2535000	seed	Funding
1000memories	2009	0	0	0	US	2535000	seed	Funding
1000memories	2010	1	8.53	0.68	US	2535000	seed	Funding
1000memories	2011	1	7.3	0.61	US	2535000	seed	Funding
1000memories	2012	1	9.47	0.24	US	2535000	seed	Funding
1000memories	2013	1	6.62	0	US	2535000	seed	Funding
1000memories	2014	1	0	0	US	2535000	seed	Funding

Figure 2. Dataset's extract.

```
. xtreg posemo i.dpf i.year, fe cluster (newid)

Fixed-effects (within) regression               Number of obs     =      7,441
Group variable: newid                           Number of groups  =      1,057

R-sq:                                           Obs per group:
     within  = 0.3158                                         min =          7
     between = 0.0228                                         avg =        7.0
     overall = 0.2487                                         max =         14

                                                F(7,1056)         =     408.81
corr(u_i, Xb)  = -0.0792                        Prob > F          =     0.0000

                            (Std. Err. adjusted for 1,057 clusters in newid)
------------------------------------------------------------------------------
             |               Robust
      posemo |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       1.dpf |   2.612375   .1548765    16.87   0.000     2.308474    2.916276
             |
        year |
       2009  |   .4145062   .0707895     5.86   0.000     .2756021    .5534104
       2010  |   .8396062   .0941776     8.92   0.000     .6548097    1.024403
       2011  |   1.683729   .1288662    13.07   0.000     1.430866    1.936592
       2012  |   2.449451   .1786499    13.71   0.000     2.098902         2.8
       2013  |    2.67666   .1767615    15.14   0.000     2.329816    3.023503
       2014  |   2.073756   .1814125    11.43   0.000     1.717786    2.429726
             |
       _cons |   -.111237   .0599337    -1.86   0.064    -.2288396    .0063657
-------------+----------------------------------------------------------------
     sigma_u |  1.7035881
     sigma_e |  3.0497223
         rho |  .23782742   (fraction of variance due to u_i)
```

Figure 3. panel data FE with DV=posemo

```
. xtreg negemo i.dpf i.year, fe cluster (newid)

Fixed-effects (within) regression               Number of obs     =      7,441
Group variable: newid                           Number of groups  =      1,057

R-sq:                                           Obs per group:
     within  = 0.1479                                         min =          7
     between = 0.0205                                         avg =        7.0
     overall = 0.1172                                         max =         14

                                                F(7,1056)         =     164.65
corr(u_i, Xb)  = -0.0155                        Prob > F          =     0.0000

                            (Std. Err. adjusted for 1,057 clusters in newid)
```

negemo	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
1.dpf	.307867	.0336601	9.15	0.000	.2418187	.3739153
year						
2009	.083467	.0198919	4.20	0.000	.0444349	.1224991
2010	.1366159	.0265255	5.15	0.000	.0845672	.1886646
2011	.2657928	.0322465	8.24	0.000	.2025182	.3290674
2012	.3813182	.0391216	9.75	0.000	.3045533	.4580832
2013	.4536536	.0412377	11.00	0.000	.3727365	.5345707
2014	.3965141	.0431912	9.18	0.000	.3117637	.4812645
_cons	.004281	.0139127	0.31	0.758	-.0230187	.0315807
sigma_u	.39187714					
sigma_e	.70760127					
rho	.23471707	(fraction of variance due to u_i)				

Figure 4. panel data FE with DV=negemo

```
. xtreg ratio_posemo_negemo i.dpf i.year, fe cluster (newid)

Fixed-effects (within) regression               Number of obs     =      2,937
Group variable: newid                           Number of groups  =        975

R-sq:                                           Obs per group:
     within  = 0.0100                                         min =          1
     between = 0.0015                                         avg =        3.0
     overall = 0.0042                                         max =          8

                                                F(7,974)          =       2.04
corr(u_i, Xb)  = -0.0381                        Prob > F          =     0.0474

                              (Std. Err. adjusted for 975 clusters in newid)
------------------------------------------------------------------------------
             |               Robust
ratio_pose~o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       1.dpf |   2.104479   .7867655     2.67   0.008     .5605288     3.64843
             |
        year |
       2009  |  -.4246615   1.051296    -0.40   0.686    -2.487728    1.638405
       2010  |   1.063331   1.070633     0.99   0.321    -1.037681    3.164343
       2011  |   .3503811   1.178844     0.30   0.766    -1.962985    2.663747
       2012  |   -.606026    1.16643    -0.52   0.603    -2.895032     1.68298
       2013  |  -1.819891   1.282026    -1.42   0.156    -4.335743    .6959605
       2014  |  -1.887146   1.323633    -1.43   0.154    -4.484647    .7103551
             |
       _cons |   8.523096   1.043601     8.17   0.000     6.475131    10.57106
-------------+----------------------------------------------------------------
     sigma_u |  8.5978186
     sigma_e |  9.2628144
         rho |  .46281906   (fraction of variance due to u_i)
```

Figure 5. panel data FE with DV=ratio_posemo_negemo

126 Posemo			127 Negemo		
accept	freed*	partie*	abandon*	enrag*	maddening
accepta*	freeing	party*	abuse*	envie*	madder
accepted	freely	passion*	abusi*	envious	maddest
accepting	freeness	peace*	ache*	envy*	maniac*
accepts	freer	perfect*	aching	evil*	masochis*
active*	frees*	play	advers*	excruciat*	melanchol*
admir*	friend*	played	afraid	exhaust*	mess
ador*	fun	playful*	aggravat*	fail*	messy
advantag*	funn*	playing	aggress*	fake	miser*
adventur*	genero*	plays	agitat*	fatal*	miss
affection*	gentle	pleasant*	agoniz*	fatigu*	missed

agree	gentler	please*	agony	fault*	misses
agreeab*	gentlest	pleasing	alarm*	fear	missing
agreed	gently	pleasur*	alone	feared	mistak*
agreeing	giggl*	popular*	anger*	fearful*	mock
agreement*	giver*	positiv*	angr*	fearing	mocked
agrees	giving	prais*	anguish*	fears	mocker*
alright*	glad	precious*	annoy*	feroc*	mocking
amaz*	gladly	prettie*	antagoni*	feud*	mocks
amor*	glamor*	pretty	anxi*	fiery	molest*
amus*	glamour*	pride	apath*	fight*	mooch*
aok	glori*	privileg*	appall*	fired	moodi*
appreciat*	glory	prize*	apprehens*	flunk*	moody
assur*	good	profit*	argh*	foe*	moron*
attachment*	goodness	promis*	argu*	fool*	mourn*
attract*	gorgeous*	proud*	arrogan*	forbid*	murder*
award*	grace	radian*	asham*	fought	nag*
awesome	graced	readiness	assault*	frantic*	nast*
beaut*	graceful*	ready	asshole*	freak*	needy
beloved	graces	reassur*	attack*	fright*	neglect*
benefic*	graci*	relax*	aversi*	frustrat*	nerd*
benefit	grand	relief	avoid*	fuck	nervous*
benefits	grande*	reliev*	awful	fucked*	neurotic*
benefitt*	gratef*	resolv*	awkward*	fucker*	numb*
benevolen*	grati*	respect	bad	fuckin*	obnoxious*
benign*	great	revigor*	bashful*	fucks	obsess*
best	grin	reward*	bastard*	fume*	offence*
better	grinn*	rich*	battl*	fuming	offend*
bless*	grins	ROFL	beaten	furious*	offens*
bold*	ha	romanc*	bitch*	fury	outrag*
bonus*	haha*	romantic*	bitter*	geek*	overwhelm*
brave*	handsom*	safe*	blam*	gloom*	pain
bright*	happi*	satisf*	bore*	goddam*	pained
brillian*	happy	save	boring	gossip*	painf*
calm*	harmless*	scrumptious*	bother*	grave*	paining
care	harmon*	secur*	broke	greed*	pains
cared	heartfelt	sentimental*	brutal*	grief	panic*
carefree	heartwarm*	share	burden*	griev*	paranoi*
careful*	heaven*	shared	careless*	grim*	pathetic*
cares	heh*	shares	cheat*	gross*	peculiar*
caring	helper*	sharing	complain*	grouch*	perver*

casual	helpful*	silli*	confront*	grr*	pessimis*
casually	helping	silly	confus*	guilt*	petrif*
certain*	helps	sincer*	contempt*	harass*	pettie*
challeng*	hero*	smart*	contradic*	harm	petty*
champ*	hilarious	smil*	crap	harmed	phobi*
charit*	hoho*	sociab*	crappy	harmful*	piss*
charm*	honest*	soulmate*	craz*	harming	piti*
cheer*	honor*	special	cried	harms	pity*
cherish*	honour*	splend*	cries	hate	poison*
chuckl*	hope	strength*	critical	hated	prejudic*
clever*	hoped	strong*	critici*	hateful*	pressur*
comed*	hopeful	succeed*	crude*	hater*	prick*
comfort*	hopefully	success*	cruel*	hates	problem*
commitment*	hopefulness	sunnier	crushed	hating	protest
compassion*	hopes	sunniest	cry	hatred	protested
compliment*	hoping	sunny	crying	heartbreak*	protesting
confidence	hug	sunshin*	cunt*	heartbroke*	puk*
confident	hugg*	super	cut	heartless*	punish*
confidently	hugs	superior*	cynic*	hell	rage*
considerate	humor*	support	damag*	hellish	raging
contented*	humour*	supported	damn*	helpless*	rancid*
contentment	hurra*	supporter*	danger*	hesita*	rape*
convinc*	ideal*	supporting	daze*	homesick*	raping
cool	importan*	supportive*	decay*	hopeless*	rapist*
courag*	impress*	supports	defeat*	horr*	rebel*
create*	improve*	suprem*	defect*	hostil*	reek*
creati*	improving	sure*	defenc*	humiliat*	regret*
credit*	incentive*	surpris*	defens*	hurt*	reject*
cute*	innocen*	sweet	degrad*	idiot	reluctan*
cutie*	inspir*	sweetheart*	depress*	ignor*	remorse*
daring	intell*	sweetie*	depriv*	immoral*	repress*
darlin*	interest*	sweetly	despair*	impatien*	resent*
dear*	invigor*	sweetness*	desperat*	impersonal	resign*
definite	joke*	sweets	despis*	impolite*	restless*
definitely	joking	talent*	destroy*	inadequa*	revenge*
delectabl*	joll*	tehe	destruct*	indecis*	ridicul*
delicate*	joy*	tender*	devastat*	ineffect*	rigid*
delicious*	keen*	terrific*	devil*	inferior*	risk*
deligh*	kidding	thank	difficult*	inhib*	rotten
determina*	kind	thanked	disadvantage*	insecur*	rude*

determined	kindly	thankf*	disagree*	insincer*	ruin*
devot*	kindn*	thanks	disappoint*	insult*	sad
digni*	kiss*	thoughtful*	disaster*	interrup*	sadde*
divin*	laidback	thrill*	discomfort*	intimidat*	sadly
dynam*	laugh*	toleran*	discourag*	irrational*	sadness
eager*	libert*	tranquil*	disgust*	irrita*	sarcas*
ease*	like	treasur*	dishearten*	isolat*	savage*
easie*	likeab*	treat	disillusion*	jaded	scare*
easily	liked	triumph*	dislike	jealous*	scaring
easiness	likes	true	disliked	jerk	scary
easing	liking	trueness	dislikes	jerked	sceptic*
easy*	livel*	truer	disliking	jerks	scream*
ecsta*	LMAO	truest	dismay*	kill*	screw*
efficien*	LOL	truly	dissatisf*	lame*	selfish*
elegan*	love	trust*	distract*	lazie*	serious
encourag*	loved	truth*	distraught	lazy	seriously
energ*	lovely	useful*	distress*	liabilit*	seriousness
engag*	lover*	valuabl*	distrust*	liar*	severe*
enjoy*	loves	value	disturb*	lied	shake*
entertain*	loving*	valued	domina*	lies	shaki*
enthus*	loyal*	values	doom*	lone*	shaky
excel*	luck	valuing	dork*	longing*	shame*
excit*	lucked	vigor*	doubt*	lose	shit*
fab	lucki*	vigour*	dread*	loser*	shock*
fabulous*	lucks	virtue*	dull*	loses	shook
faith*	lucky	virtuo*	dumb*	losing	shy*
fantastic*	madly	vital*	dump*	loss*	sicken*
favor*	magnific*	warm*	dwell*	lost	sin
favour*	merit*	wealth*	egotis*	lous*	sinister
fearless*	merr*	welcom*	embarrass*	low*	sins
festiv*	neat*	well*	emotional	luckless*	skeptic*
fiesta*	nice*	win	empt*	ludicrous*	slut*
fine	nurtur*	winn*	enemie*	lying	smother*
flatter*	ok	wins	enemy*	mad	smug*
flawless*	okay	wisdom			
flexib*	okays	wise*			
flirt*	oks	won			
fond	openminded*	wonderf*			
fondly	openness	worship*			
fondness	opportun*	worthwhile			

forgave	optimal*	wow*	
forgiv*	optimi*	yay	
free	original	yays	
free*	outgoing		
freeb*	painl*		
	palatabl*		
	paradise		

Figure 6. posemo and negemo ratios. Source: LIWC_CODE.xlsx

7. BIBLIOGRAPHY

A. Abisheva, V. K. (2014). Who watches (and shares) what on YouTube? And when? Using Twitter to understand YouTube viewership. *Association for Computing Machinery*.

Adcock AB, L. M. (2013). *CS 224W Final Report Group 37.* Stanford University.

Ahmetoglu, G. L. (2011). EQ-nomics: Understanding the relationship between individual differences in Trait Emotional Intelligence and entrepreneurship. *Personality and Indivdual Differences*, (8): 1028-1033.

Alexy OT, B. J. (2012). Social capital of venture capitalists and start-up funding. . *Small Business Economics*, 835-851.

Andranik Tumasjan, T. O. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.*

Anne L.J. Ter Wal, O. A. (2016). The Best of Both Worlds: The Beneifts of Open-specialized and Closed-diverse Syndication Networks for New Ventures' Success. *Administrative Science Quarterly* , 406 - 407.

Arend, R. J. (2012). A Heart-Mind-Opportunity Nexus: Distinguishing Social Entrepreneurship for Entrepreneurs. *ACAD Manage Review*.

Arora A, N. A. (2012). Insecure advantage? Markets for technology and the value of resources for entrepreneurial ventures. *Strategic Management Journal*, 33(3): 231-251.

Bandelj, N. (2009). Emotions in economic action and interaction. *Theroy and Society*.

Baron, R. R. (2013). Why Entrepreneurs Often Experience Low, Not High, Levels of Stress The Joint Effects of Selection and Psychological Capital. . *Journal of management*.

Barrick, M. R. (1991). The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Pshycology*, 1-26.

Behling, O. (1998). Employee selection: will Intelligence and Conscientiousness do the job? *The Academy of Management Executive*, 12 (1): 77-86.

Block J, S. P. (2009). What is the effect of the financial crisis on venture capital financing? Empirical evidence from US Internet start-ups. . *Venture Capital 11(4)*, 295-309.

Burger, J. H. (2010). An exploration of observable features related to blogger age. *Computational apporaches to analyzing weblogs: Papers from the 2006 AAAI Spring Symposium*, 710-718.

C. Tan, L. L. (2011). User-Level Sentiment Analysis Incorporating Social. *ACM (Association for Comouting Machinery) Journal*.

Caligiuri, P. T. (2004). International assignee selection and cross-cultural training and development. In I. B. Günter K. Stahl, *Handbook of Research in International Human Resource Management* (pp. 321-342).

Cardon, M. M. (2012). Exploring the heart: Entrepreneurial emotion is a hot topic. *Entrepreneurship Theory and Practice*, 1-10.

Coppersmith, G. D. (2014). Quantifying Mental Health Signals in Twitter. *Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, (pp. 51-60).

Corradino, C. F. (2016). Positive emotions and academic achievement. *Applied psycology OPUS, NYU Steinhardt*.

De Klerk, G. J. (2002). The driving force behind entrepreneurship: An exploratory perspective. *Rencontres de St-Gall, Radical change in the world - will SMEs soar or crash?; Umbruch der Welt - KMU vor Hohenflug oder Absturz?* Zurich, Switzerland.

Elfenbein, H. A. (2007). Emotion in organizations: A review and theoretical integration. *Academy of Management AnnalsAcademy of Management Annals*, 1. 371 - 457.

F. Schultz, S. U. (2011). Is the medium the message? Perceptions of and reactions to crisis communication via twitter, blogs and traditional media. *Public Relations Review*.

Fairlie, R. W. (2013). *Kauffman Index of Entrepreneurial Activity.* Kansas City: Ewing Marion Kauffman Foundation.

Fischer, E. R. (2014). Online entrepreneurial communication: Mitigating uncertainty and increasing differentiation via Twitter. *Journal of Business Venturing*.

Foo, M. (2011). Emotions and entrepreneurial opportunity evaluation. *Entrepreneurship: theory and practice*, 375-393.

Forgas, J. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117, 39-66. .

Frijda, N. (1988). The Laws of Emotion.

Furnham, A. Z. (2005-2006). The relationship between psycometric and self-estimated intelligence, creativity, personality and academic achievement. *Imagination, Cognition and Personality*, 25(2), 119-145.

Gage, D. (2012). The Venture Capital Secret: 3 Out of 4 Start-Ups Fail. *The Wall Street Journal*.

Garera, N. Y. (2007). Modeling latent bio- graphic attributes in conversational genres. *Conference on Information and Knowledge Management.*

Gillbert, E. K. (2009). Predicting Tie Strength With Social Media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 211-220). Boston.

Goss, R. a. (2014). The Role of Negative Emotions in the Social Processes of Entrepreneurship: Power Rituals and Shame-Related Appeasement Behaviors. *Entrepreneurship: theory and practice*.

Graham, J. (2012). Top cities for technology start-ups. *http://usatoday30.usatoday.com/tech/columnist/talkingtech/story/2012-08-22/top-tech-startup-cities/57220670/1*.

Haewoon Kwak, C. L. (2010). What is Twitter, a Social Network or a News Media? *The International World Wide Web Conference Committee (IW3C2)*.

Herring, S. P. (2010). Gender and genre variations in weblogs. *Journal of Sociolinguistics*, 710-718.

J. Bollen, H. M. (2010). Twitter mood predicts the stock market. *Arxiv.org*.

J. M. Lee, B. H. (n.d.). Are founder CEOs more overconfident than professional CEOs? Evidence from S&P 1500 companies.

Java, A. S. (2007). Why we tweet: Understanding microblogging usage and communities. *9th WebKDD and 1st SNA-KDD 2007 workshop on web mining and social network analysis*, (pp. 12-15). San Jose, California.

Jones, R. K. (2007). I Know What you Did Last Summer - Query Logs and User Privacy. *Conference on Information and Knowledge Management.*

Kuratko, D. F. (2009). Strategic entrepreneurship: Exploring different perspectives of an emerging concept. *ET&P*.

Larsen, R. J. (1989). Extraversion, neuroticism, and susceptibility to positive and negative mood induction procedures. *Personal & Individual References*, 10, 1221-1228.

Lerner, J. S. (2000). Beyong valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion*, 14(4), 473-493. .

Leutner, F. A.-P. (2014). The relationship between the entrepreneurial personality and the Big Five personality traits. *Personality and Individual differences*, (63): 58-63.

Levine, R. (2008). A geography of time: On tempo, culture, and the pace of life. *Basic Books*.

Lueger, A. K. (2012). Predicting founding success and new venture survival: a longitudinal nascent entrepreneurship approach. *Journal of entreprise culture*.

Mason, M. K. (2016). What is the Average Number of New Businesses That Start Up Worldwide Each Day?

Mateja Drnovsek, M. S. (2005). The nature and experience of entrepreneurial passion. *Academic of Management Review*.

Morris, M. &. (2000). How emotions work: The social functions of emotional expression in negotiations. . *Research in Organizational Behavior*.

Nadkarni, C. J. (2014). Bridging yesterday, today, and tomorrow: CEO temporal focus, environmental dynamism, and rate of new product introduction. *Academy of Management Journal*.

O'Boyle, E. H. (2010). The relationship between Emotional Intelligence and Job Performance. A meta-analysis. *Journal of Organizational Behavior*, 10, 1002.

Otterbacher, J. (2010). Inferring gender of movie reviewers: Exploiting writing style, content and metadata. *CIKM.*

P.Delevett. (2013). Silicon Valley continues to outstrip rest of country in tech investing, and it's not even close.

Pennacchiotti, M. &. (2011). A machine learning approach to Twitter-user classification. *Fifth International AAAI Conference on Weblogs and Social Media*, (pp. 281-288). Sunnyvale.

Qiu, L. L. (2012). You are what you tweet: Personality expression and perception on Twitter. *Journal of research in personality*, 710-718.

Quercia, D. C. (2012). The Social World of Twitter: Topics, Geography, and Emotions. *Sixth International AAAI Conference on Weblogs and Social Media*.

Rao, D. D. (2010). Classifying Latent User Attributes in Twitter. *SMUC (Search and Mining User-generated Contents)*, (pp. 710-718).

Rauch, A. F. (2007). Let’s put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners’ personality traits, business creation, and success. *European Journal of Work and Organizational Psychology*, 16, 353-385.

Remington, N. A. (2000). Reexamining the circumplex model of affect. *Journal of personality and social psycology*, 286-300.

Ritter, R. S. (2013, June 18). *Happy Tweets: Christians Are Happier, More Socially Connected, and Less Analytical Than Atheists on Twitter.* Retrieved from Sage Publications: http://journals.sagepub.com/doi/abs/10.1177/1948550613492345

Russell, J. A. (1999). Structure of Self-Reported Current Affect: Integration and Beyond. *Journal of Personality and Social Psycology*, 77(3), 600-619.

Russell, J. A. (2003). Core affect and the psycological construction of emotions. *Psycological Review*, 110(1), 145-172.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, Vol 82(1).

Shane, S. V. (2000). The promise of entrepreneurship as a field of research. *Academic of Management Journal*, 25, 217-226.

Shane, S., & Venkataram, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review*, 217-226.

Shepherd, D. A. (2003). Learning from business failure: propositions about the grief recovery process for the 'self-employed'. *Academy of Management Review*, 318-329.

Shepherd, D. A. (2013). Project failures arising from corporate entrepreneurship: Impact of multiple project failures on employees' accumulated emotions, learning, and motivation. *Journal of Porduct Innovation Management*.

Shook, J. R. (2003). Entrepreneurship and values in a democratic and pragmatic economics: commentary on 'A transactional view of entrepreneurship'. *Journal of Economic Methodology*, 10 (2): 181-190.

Singh, S. (2001). A pilot study on gender differences in con- versational speech on lexical richness measures. *Literary and Linguistic Computing*.

Smilor, R. W. (1997). Entrepreneurship in the Next Century: where will Venture Capitalists find their next pearl? *The Journal of Private Equity*, 53-58.

Summer, C. B. (2012). Predicting Dark Triad Personality Traits from Twitter usage and a linguistic analysis of Tweets. *11th International Conference on Machine Learning and Applications, ICMLA 2012*, (pp. 386-393).

Taylor, E. B. (1991). A review of local adaptation in Salmonidac, with particular reference to Pacific and Atlantic salmon. *Aquaculture*, 185-207.

Timmons, A. C. (2015). Physiological linkage in couples and its implications for individual and interpersonal functioning: A literature review. *Journal of Family Psycology*, 29(5), 720-731.

Toyah L. Miller, M. G. (2012). Venturing for others with heart and head: how compassion encourages social entrepreneurship. *Academy of Management Review*.

Van Rooy, D. L. (2004). Emotional Intelligence. A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behaviour*, 65, 71-95.

Virlics, A. (2013). Invmestment Decision Making and risk. *International Economic Conference of Sibiu 2013 Post Crisis Economy: Challenges and Opportunities, IECS 2013*.

Vohs, K. D. (2006). The psychological consequences of money. *Science*, 1154-1156.

Vohs, K. D. (2006). The psychological consequences of money. *Science*, 1154-1156.

Weiblen, T. C. (2015). Engaging with startups to engage corporate innovation. *California Management Review*, 66-90.

Wright, T. A. (1998). Emotional exhaustion as a predictor of job performance and voluntary turnover. *Journal of Applied Psycology*, 486-493.

Wyse, S. E. (2014, October 15). *Advantages and Disadvantages of Face-to-Face Data Collection.* Retrieved from SnapSurveys: http://www.snapsurveys.com/blog/advantages-disadvantages-facetoface-data-collection/

Xiang G, Z. Z. (2012). A Supervised Approach to Predict Company Acquisition with Factual and Topic Features Using Profiles and News Articles on TechCrunch. *ICWSM.*

Zhao, H. S. (2006). The Big Five Personality dimensions and entrepreneurial status: a meta-analytical review. *The Journal of Applied Psychology*, 91, 259-271.

Zhao, H. S. (2010). The relationship of personality to entrepreneurial intentions and performance: a meta-analytic review. *Journal of Management*.