The Use of Social Impact Measurements in Socially Entrepreneurial Organizations

A Quantitative Survey Study on Organizational Size

Albina Iljasov
Hillevi Eilard
Abstract

This thesis investigates the relationship between organizational size and the decision of using social impact measurements in Swedish socially entrepreneurial organizations, which include organizations and individuals that aim at solving social problems. Data is collected using an online survey, and three different ways to measure organizational size are used and tested whether they have a positive relationship with the use of social impact measurement. Organizational size is measured as the yearly turnover, available personnel and the number of members and participants in the organization. The association was analyzed through three different analytical methods, and we also present detailed descriptive statistics for the data sample. The results showed significant relationships between available staff as well as members and participants and the use of social impact measurements, while no significant relationship between the yearly turnover and the use of social impact measurements could be confirmed. We suggest a consideration of impact measurements when addressing sustainable organizational development as well as consideration of organizational size for local governments when introducing policies, funding and other support for socially entrepreneurial organizations.

Keywords: organizational size, social impact measurement, socially entrepreneurial organizations, performance measurements, evaluation

Words count: 18 726
# Table of Content

1. Introduction ................................................................................................................................. 1  
   1.1. Initiatives to address social needs .......................................................................................... 1  
   1.2. Performance evaluation and impact measurement ............................................................... 1  
   1.3. Socially entrepreneurial organizations .................................................................................. 2  
   1.4. The matter of size when measuring impact ......................................................................... 3  
   1.5. Problem formulation .............................................................................................................. 3  
   1.6. Purpose ................................................................................................................................. 4  
2. Methodology .................................................................................................................................. 5  
   2.1. Positivist approach and inference ......................................................................................... 5  
   2.2. Role of theory ......................................................................................................................... 5  
   2.3. Research design ..................................................................................................................... 6  
   2.4. Quantitative research approach .............................................................................................. 6  
3. Theoretical Framework .................................................................................................................. 7  
   3.1. Organizational evaluation ...................................................................................................... 7  
   3.1.1. Business sector ................................................................................................................. 7  
   3.1.2. Public sector ...................................................................................................................... 8  
   3.1.3. Civil sector .......................................................................................................................... 8  
   3.1.4. Socially entrepreneurial organizations .............................................................................. 8  
   3.2. Organizational size ................................................................................................................ 9  
   3.3. Conceptual framework .......................................................................................................... 10  
4. Methods ....................................................................................................................................... 11  
   4.1. Survey creation ......................................................................................................................... 11  
   4.2. Data collection ......................................................................................................................... 13  
   4.3. Data sample ............................................................................................................................. 14  
   4.4. Statistical analysis .................................................................................................................. 15  
   4.4.1. Model I ............................................................................................................................... 16  
   4.4.2. Model II ............................................................................................................................. 16  
   4.4.3. Model III ............................................................................................................................ 18  
   4.5. Ethical considerations and quality of research ...................................................................... 19  
   4.5.1. Ethical considerations ......................................................................................................... 19  
   4.5.2. Quality of the research ...................................................................................................... 19  
   4.5.2.1. Strengths .......................................................................................................................... 20
Table of Tables

Table 1: Independent variables ................................................................. 9
Table 2: Descriptive data ................................................................. 15
Table 3: Correlation ................................................................. 17
Table 4: Main challenges ................................................................. 25
Table 5: Main opportunities ................................................................. 25
Table 6: Results Model I H₈ ................................................................. 26
Table 7: Results Model I H₉ ................................................................. 27
Table 8: Results Model I H₊ ................................................................. 27
Table 9: Results Model II and III ................................................................. 28

Table of Figures

Figure 1: Sectors represented in our data sample ........................................... 22
Figure 2: Organizational age ........................................................................ 23
Figure 3: Initial funding ........................................................................ 23
Figure 4: Methods for performance evaluation ........................................... 24
Figure 5: Future use of social impact measurement ....................................... 25
1. Introduction

In the first chapter we give a short introduction to the topic of this thesis. Using relevant literature we discuss the current situation. We explain the used terminology and finally formulate the purpose of this study.

1.1. Initiatives to address social needs

New types of social challenges are continuously increasing in our fast developing society, and differ largely on their context, including the region, culture and its effect on the society (Nicholls, 2008). The failure to provide social and environmental services to address urgent needs leads to a raising number of new types of organizations and collaborative initiatives (Alvord, Brown & Letts, 2004; Fayolle & Matlay, 2010; Nicholls, 2008; Short, Moss & Lumpkin, 2009). The interest and scope of such organizations has increased significantly in recent years due to a perceived failure of the dominant corporate, liberal models governing consumption and welfare services today (Phillips, Lee, Ghobadian, O'Regan & James, 2015). With the added strain from a global, economic recession and increased pressure on public spending, the need for new, innovative solutions to solve social problems has grown stronger (Phillips et al., 2015; Short et al., 2009). This has resulted in a move by private organizations away from the more common CSR policies towards the development of corporate social innovation as a potential answer to pressing local and global needs, such as poverty, environmental destruction and social exclusion (Phillips et al., 2015). Due to what Alvord and colleagues (2004) describe as social market failures, these organizations and individuals identify opportunities to provide social goods and services to a needing community, and must balance the economical and social values in order to prevent a deficit on either one of the sides. This phenomenon has been described in the literature as the Window of Opportunity (Kingdon, 2003; Werther & Chandler, 2010). Social value and its creation are phenomena that have increasingly drawn attention from individuals and institutions at the local, national and global level to tackle social issues (Wilson, Silva & Ricardson, 2015). These actors are increasingly interested in investing in solutions for societal issues, such as improving social, environmental and socio-economic conditions of the society that can provide social as well as financial return for both the investors and the community (Wilson et al., 2015). However, while the interest in social investment is increasing and becoming a focus in the private and public sectors, consequently the question is arising on how to assess and measure the impact on society of such activities.

1.2. Performance evaluation and impact measurement

Evaluating organizational performance and thereby measuring the results is essential for all organizations that seek to improve and develop (Baron, 2011). It is important to have specific and measurable performance goals in an organization to be able to evaluate its performance (Caproni, 2011). The value of an accurate assessment of organizational performance can enhance internal governance, strategic planning and efficient resource allocation, which in turn are drivers for innovation and organizational effectiveness (Paton, 2003). Some researchers argue, based on organizational theories, that the organizational effectiveness and legitimacy are a construct within bounded rationality rather than objective reality (Meyer & Scott, 1992; Nicholls, 2008; Zammuto, 1984). Nicholls (2008) adds that “...legitimacy judgments represent organizational performance constructed via stakeholder voice...” (Nicholls, 2008:47). To improve innovativeness and effectiveness of an organization, it is therefore vital to address the right stakeholders and take their views into consideration (Werther & Chandler, 2010). A multi stakeholder perspective, as presented in Werther and Chandler (2010), includes organizational (internal to the organization), economical, and societal (external) stakeholders in the context of globalization and...
technology which enables an organization not only to identify all important stakeholders, but also to prioritize their respective needs and expectations. However, measuring performance is not the same as measuring impact; assessing social impact also means measuring what the organization accomplishes outside of their own organizational boundaries, and is in this way more challenging to quantify by measurements. Exploring ways to measure impact is not a new focus, but one that is becoming more of a concern as organizations look for new ways to improve their activities and demonstrate this impact to others, for example through CSR-work (Werther & Chandler, 2010). In part, this requires a shift in thinking from measuring merely outputs to additionally measuring outcomes, and the challenge is to find ways to measure elements related to social issues and not only the economic results. In addition, there is an interest from outside of organizations, a pressure from philanthropic funds and public service commissioners, to find ways of making their decisions on how resources are or should be allocated (Short et al., 2009). When moving from measuring organizational performance to measuring social impact, stakeholder consideration becomes increasingly important to provide guidance for internal stakeholders, managers, employees and shareholders (Kernot & McNeil, 2011). At the same time the impact can be presented to economical stakeholders, such as customers, beneficiaries, suppliers and competitors, on one hand for marketing purposes and on the other hand as inspiration (Kernot & McNeil, 2011). Societal stakeholders, such as governments, communities and other actors that might influence or be influenced by the organization, are important in the sense that they provide ground for any action the organization takes (Werther & Chandler, 2010). Providing tangible evidence that organizational activities have a social impact on society can be highly influential in terms of regulation settings, movements and social change. Further, social impact metrics are no longer managed independently by the organizations themselves, but defined and constructed together with relevant stakeholders (Costa & Pesci, 2016). This way the organization can show if it has reached the set goals and fulfilled its responsibilities towards clients and beneficiaries both in terms of results and impact.

1.3. Socially entrepreneurial organizations

Carroll’s (1991) pyramid of corporate social responsibility distinguishes between different kinds of organizations’ responsibilities towards the society. As a basis the economic responsibility towards owners and investors is followed by the legal responsibility to follow regulations and laws of the environment, and finally the ethical and the discretionary responsibilities to enhance the social standard of the community the organization operates in (Carroll, 1991). Werther and Chandler (2010) challenge this view by saying that the two later responsibilities of Carroll’s model are becoming a growing necessity for businesses today because of the fast changing environment and mindset (Werther & Chandler, 2010). The ethical and discretionary behavior can build the fundament for the economical and legal responsibilities and thus the success of the business (Werther & Chandler, 2010). Some organizations take this approach further to the core meaning of their existence and incorporate the ethical and discretionary responsibilities into their mission. In the literature, these are often described as socially entrepreneurial organizations (SEOs), including social entrepreneurs, social enterprises, and social businesses (Elkington, 2001; Emerson, 2003; Kaplan, 2002; Nicholls, 2008; Wilson et al., 2015; Zadek, 1998). These actors are described as organizations or individuals that operate at the intersection between innovation, market-change and sociality (Nicholls, 2008). In Sweden, none of these have any specific legal form yet (Alamaa, 2014: Defourny & Nyssens, 2008). Social entrepreneurship has previously been conceptualized in three different approaches. The first approach centers on social demands of vulnerable groups that are currently unmet (varies to a large extent based on the societal challenge addressed - hunger, poverty, human rights, etc.); the second approach is broader and targets challenges of society as a whole; and the third approach is concerned about systemic reforms of societal transformation (Short et al., 2009). Any final, valid definition for SEOs does not seem to exist. In this thesis we use the term socially entrepreneurial organizations to refer to social entrepreneurs, social enterprises and social businesses alike. We do not make the distinction between different types of organizations as they all have similar features, such as to do good for society, create public value, come about social or community change or develop sustainable community benefit. Meanwhile, they are business-like in operations but without primarily financial motives. They are aiming towards solving social problems or provide social services on the local, national or global level while taking any legal form or size (Fayolle & Matlay, 2010).
1.4. The matter of size when measuring impact

Like all other organizations, SEOs vary largely in organizational size, and while some assess social impact, others do not. A risk if only large organizations assess impact can be the complete loss of knowledge and innovation from smaller organizations. Evaluations in general can be regarded as a form of knowledge production, and therefore objective evaluations, both in terms of performance and impact, are one component of broader knowledge strategy and organizational learning (Nielsen, Lemire & Skov, 2011). This means that if no measuring of impact is done, no knowledge will be created regarding the ways to come about social change. This in turn can limit social innovation and creativity, also limiting social solutions to the existing ones. A related risk is decreased possibilities for funding opportunities without any evaluation or impact assessment (Kernot & McNeil, 2011). Potential funders and investors must choose among a large set of organizations which they have imperfect information about, and by measuring impact the organizations themselves can demonstrate to funders how valuable their work is and simultaneously compare their performance to others (Lecy, 2010). If organizations can show a social impact, funders and investors could then more easily invest into the most effective organizations and thereby disseminate a powerful incentive for adopting best practices (Lecy, 2010). Likewise, smaller organizations without possibilities to demonstrate an impact miss out on funding opportunities both from the public and private sector, potentially risking the whole organization’s existence. Further, the goal of evaluations in general is to continuously create and sustain an overall organizational process that from knowledge can create quality. Internal and external evaluations are required to make such a goal reality regardless of whether it is evaluating performance or impact, and thus measuring in general has become a valuable commodity for organizations in modern society (Baron, 2011). If no evaluation is done, there will be limited possibilities for development in that organizational field, and the organization will remain small without potential to grow. The literature shows that large organizations have a distinct advantage when it comes to evaluation of both results and impact due to extensive, available resources and opportunities for funding from numerous sources (Baron, 2011; Gallo & Christensen, 2011; Stanwick & Stanwick, 1998). Meanwhile, smaller organizations oftentimes may not have the resources or political influence to justify an evaluation staff or activity (Baron, 2011). No matter what the size of the organization, evaluations can be a productive activity that reaches beyond the political, economic, or social barriers to propel organizations to improve quality, efficiency, and effectiveness (Baron, 2011), making it maybe even more important for small organizations, and SEOs particularly, to assess and measure their social impact to gain the legitimacy and financial stability they need. As a final implication, if only large organizations measure impact, their assessment methods can become dominating in the field. Over time, it can create bias in the measurements as only bigger organizations use them and the tools become even more subjective rather than the opposite. This way, rigorous impact studies that validate large organizations’ activities become influential as they can generate field consensus and inform policy makers about a sector in the long run (Baron, 2011). If only larger organizations assess impact, it could thus be a problem for the development of reliable social impact measurements (SIM) as they also have to be versatile to fit many different organizations.

1.5. Problem formulation

As stated above, it is essential for SEOs to measure performance and social impact for internal development (such as guidance and inspiration) as well as external development (including acquiring legitimacy and attracting investors) (Caproni, 2011; Kernot & McNeil, 2011; Meyer & Scott, 1992; Nicholls, 2008; Paton, 2003; Werther & Chandler, 2010; Wilson et al., 2015; Zammuto, 1984). On the other hand, such assessments are commonly seen as subjective and not representative for organizational activities. In addition, they are often described as resource intensive and time consuming (Baron, 2011; Elkington, 2001; Emerson, 2003; Kernot & McNeil, 2011; Wilson et al., 2015). While it generally seems to be beneficial to measure social impact for organizations of all sizes that aim towards solving social problems or providing social services, some undertake such assessments and some do not. Gallo and Christensen (2011) examine a large number of firms and find out that organizational size is strongly connected
to sustainability related activities since such initiatives require resources in form of human and financial capital, and larger firms generally have more of them. Stanwick and Stanwick (1998) provide evidence that organizational size is indeed positively related to corporate social performance due to available resources on one hand and due to increased influence by multiple stakeholders on the other hand. Other studies show that size affects the leadership behavior needed for management innovation (Vaccaro, Jansen, Van den Bosch & Volberda, 2012). Most attention on organizational size has been drawn from the business sector; however it has also been shown that organizational size is an important predictor for innovations in nonprofit organizations (Jaskyte, 2013). Since organizational size seems to be an important factor for decision-making, evaluation and resource allocation (Baron, 2011; Tolbert & Hall, 2011), it is relevant to investigate if there is any relationship between the size of an organization and the use of social impact measurement.

1.6. Purpose

The purpose of this thesis is to explain how the use of social impact measurements in socially entrepreneurial organizations relates to organizational size.
2. Methodology

In this chapter we give a brief overview of our philosophy of science and state the warrant inference. Afterwards we shortly present the scope of our empirical research including the chosen research design, role of theory and the overall research approach.

2.1. Positivist approach and inference

Positivism has a broad interpretation in contemporary social sciences as it embraces any approach which applies scientific methods to a social and human phenomena conceived as belonging to some natural order (Hollis, 1994). In recent philosophy of science, positivism means that the theoretical knowledge can simply be reduced to an empirical observation so that one can predict what will be found in further observations (6 & Bellamy, 2012; Hollis, 1994). Based on a positivist ontological and epistemological assumption as defined by 6 and Bellamy (2012), we review collected empirical data and attempt at generalizing our inference as far as possible to broader context, as we believe that knowledge is transferable and comes from experiences independent of the learner. At the same time, we are aware that generalization across a field with little consensus and few clear definitions is difficult and hence risks the external validity of the study. However, it has previously been argued that the very nature of progress in academic research speaks against attempting at maximizing the external validity at any cost in the context of a single study (Calder, Phillips & Tybout, 1982). Thus, as any researcher, we desire to add knowledge regarding how the world works, and the reality we are investigating refers to whatever occurrence in universe that causes a certain phenomena (Hollis, 1994). If the strategy for doing this is sound, it can offer an explanation with enough probability to predict future happenings (Hollis, 1994). To avoid any biased generalization, we base our study on previous, peer-reviewed research, and test our explanatory inference on empirical data. The first chapter shows that some SEOs conduct such assessments and some do not, and based on the fact that organizational size is a decisive factor for decision making, we infer that the use of SIMs depends on organizational size.

2.2. Role of theory

In this thesis, we use a deductive approach to address our purpose. In accordance with our philosophy of science, a rather positivistic epistemology, we gather and organize knowledge that exists externally through quantitative collection of data and analytic methods. A deductive approach is therefore a common choice for positivists (6 & Bellamy, 2012). Conducting deductive research always holds a precise statement of expected findings based on what we know, or suspect to know (6 & Bellamy, 2012). It begins with stating a hypothesis. This hypothesis may stem from previous body of evidence, but it is not a criteria for deductive research (6 & Bellamy, 2012; Blaikie, 2003). In our case, we start with a practical observation and a potential explanation for why it occurs. We then proceed to deducing three hypotheses and testing whether our claim seems plausible by using appropriate, empirical data (Blaikie, 2003). Confirmation of the hypotheses can lead to conclusions regarding a certain phenomena, generalizations across units, settings and times, and the rejection of other, possible explanations than the one remaining at the end of the research (6 & Bellamy, 2012; Calder et al., 1982).
2.3. Research design

For our research we use a variable-oriented research design (VOR) for analyzing the relationship between variables chosen from a large number of observations (6 & Bellamy, 2012). Since we aim to carefully generalize our findings to a larger population of SEOs in Sweden, based on the findings from the sample, it is an appropriate design for our purpose. It is however always important to consider the sample of the population as it will be a major determinant for the quality of the study (Brown, 2006). We return to the sample issue in Chapter 4. The analysis of variable-oriented research is generally done statistically and determines the relationship of outcomes, or dependent variables, to particular variables that we believe are causing these, the explanatory (independent) variables (6 & Bellamy, 2012). VOR in general is often used to test such relationships, for example whether changes in one variable are systematically linked to changes in another. If such a relationship can be proven, theories and models can be verified, rejected or refined (6 & Bellamy, 2012). VOR is used in social sciences because it makes it possible to broaden the conclusions about social life and test general applications of theories (6 & Bellamy, 2012).

2.4. Quantitative research approach

There is a general distinction between qualitative and quantitative research; while the first one rather uses words, the later provides numerical proof (Blaikie, 2003). A common belief is that numbers are needed to ensure objectivity and provide reliability, but in fact quantitative data often starts with words which are later transformed into numbers. The data then goes through statistical manipulation, and is presented either in words, numbers, or both (Blaikie, 2003). For our research we choose a quantitative research approach, and we collect and analyze primary data using quantitative methods because it best suits our purpose of investigating a relationship between two different factors and it is in line with our methodological assumptions. The main purpose of this paper is to carefully try to explain how organizational size of SEOs in Sweden relates to the use of social impact measurements. For stronger representation of the target population, various organizations that differ in size, structure, development stage and sector are represented. By using a deductive and quantitative research approach, we create a framework anchored in current, academic literature when constructing the survey instrument, then collect, code and quantitatively analyze the empirical data. Based on our positivist methodological assumption, we then present the findings, discuss and interpret results.
3. Theoretical Framework

In this chapter we present our theoretical background in which we anchor this study. First, we provide an overview of organizational evaluation and the performance measurements used in different sectors. Then, we review the literature on organizational size and present the emerged dependent and explanatory variables of this study. We finally combine the two concepts in the third part of the chapter in a conceptual framework and state the hypotheses.

3.1. Organizational evaluation

To be able to evaluate if, or measure to what extent, organizations meet their performance and impact goals they first need to set their objectives. Defining goals can often be problematic since organizations have multiple goals that frequently contradict each other (Tolbert & Hall, 2011). This often results in organizations formulating official goals in a rather broad and abstract way to state the general purpose of the organization in annual reports and other publications. If the goals are abstract and broad, the assessment on how and when they are achieved becomes more complex. In addition, organizations also formulate operational goals internally to state what they are actually trying to achieve in a more detailed way (Perrow, 1961). Tolbert and Hall (2011) present three different approaches organizations use for performance orientation. The system-resource approach says that the main goal of an organization is to survive by acquiring and managing scarce resources; participant-satisfaction approach describes the orientation to attract contributions of members that are necessary for the organizational existence; and finally the stakeholder approach implies that to assess effectivity, the organization needs a referent (Tolbert & Hall, 2011). Either way, all kinds of organizations aim to evaluate their effectivity and performance, but the approaches can differ from organization to organization, from industry to industry and from sector to sector. In the following, we present how different sectors evaluate and measure their performance and what tools are commonly used for this.

3.1.1. Business sector

Evaluating performance is commonly done through the use of performance measurements, which is the assessment of an action that has been caused intentionally (Cordery & Sinclair, 2013; Payer-Langthaler & Hiebl, 2013). Originated from the business sector, performance measurements firstly concentrate on creating value for stakeholders, primarily for the organization’s owners or shareholders (Munir, Baird & Perera, 2013). In businesses, a traditional accounting system is preferred, such as the well-known bottom line (Kellen & Wolf, 2003). Behn (2003) argues that there are eight main purposes for the use of performance measurements, namely to evaluate, motivate, promote, budget, control, celebrate, learn and improve. Especially strategic management divisions use business performance measurements in order to evaluate and improve organizational performance (Venkatraman & Ramanujam, 1986). Business performance can be seen as a subset of the overall concept of organizational effectiveness. Business performance is dualistic, one part is the financial domain, hence how well the organization fulfills its economic goals, including return on equity and return on investment, as well as improved market value (Venkatraman & Ramanujam, 1986). The other part of business performance is indeed non-financial, the operational performance. This includes, among others, measuring the market-share, product quality and new product introductions, as well as technological efficiency (Venkatraman & Ramanujam, 1986). Strategic management objectives differ depending on the purpose of the managerial activity and therefore, performance measurements should be adjusted to the respective purpose of the action rather than a one-size-fits-all approach (Behn, 2003). The probably most common attempt at balancing the financial
and operational performance is the framework of *balanced scorecard*, which provides businesses with a set of broader measurements (Kaplan & Norton, 1996).

### 3.1.2. Public sector

In contrast, programs designed by governmental institutions often aim at various objectives with produced outcomes that compete in terms of created value, but at the same time there is a problem with existing evaluations only capturing one dimension of governmental performance (Oh & Bush, 2015). It has not yet been fully successful to apply the complexity of such measurements to the public sector (Ammons & Rivenbark, 2008). It is suggested that a universal approach is not realistic for the public sector, but rather a more representative evaluation can be achieved when the wide variety of performance measurements and contexts are considered (Costa & Pesci, 2016; Oh & Bush, 2015). The authors argue that existing performance measurements only partially cover the outcomes of governmental programs, and suggest that different aspects of performance measurement should exist within one of the three dimensions - value, orientation and data. These dimensions constitute the context, and therefore a representative performance measurement must be continuously considered from multiple perspectives, and adjusted according to specific settings (Oh & Bush, 2015). What is missing in the public sector performance measurements might in fact be an evaluation of the social impact achieved through different policies and practices that could encompass the outcomes of governmental programs that are not fully covered today.

### 3.1.3. Civil sector

In the civil sector, where SEOs are often placed, the ownership relation from the business sector does not exist and monetary value cannot or should not be distributed within the organization (Cordery & Sinclair, 2013). Huang and Hooper (2011) argue that the main reason for the civil sector to evaluate performance is to present their worthiness and effectiveness to potential or existing funders and beneficiaries, as well as to allocate resources and learn from feedback and comparisons to similar organizations. The accounting measures that are common for businesses are not the primary focus in the civil sector; funders rather want to see how well the organization delivers social goods and services, and if it generally learns from previous evaluations (Huang & Hooper, 2011). The balanced scorecard framework has been developed and adjusted for non-profits (and SEOs) and is one of the used SIM in the civil sector (Kaplan, 2002). The framework is considered to have advantages mostly for internal processes and mission guiding, but is rather limited and still not fit for comparisons (Nicholls, 2008).

### 3.1.4. Socially entrepreneurial organizations

Non-profit organizations have traditionally had an inherent interest in fostering social change to achieve societal improvements (Short et al., 2009). That’s why the civil sector is often called the third sector, and similarly why SEOs often have their rooting within the civil sector (Milbourne, 2009). However, many of them often come to face a tension between their social aims and demands of making financial profits to be able to keep providing some social good (Battilana, Sengul, Pache & Model, 2015; Battilana & Lee, 2014). As defined in the introduction part, we consider SEOs as organizations or individuals that tend to be business-like in operations, but without primarily financial motives. Although they are partially dependent on funding and external investments, they seek economic independence without compromising their social purpose (Fayolle & Matlay, 2010). At the same time, they provide products and services that governments fail to arrange, and often SEOs are in focus of policy makers who are trying to increase the quality of welfare while reducing the public expenditures (Cordery & Sinclair, 2013). We therefore place SEOs somewhere in between the three sectors, which results in complex performance measurements as a combination of all three. Another dimension specific for SEOs as the provider of social goods and services as suggested by current literature generally requires accountability (Bagnoli & Megali, 2011; Nicholls, 2008; Paton, 2003; Zadek, 1998). The issue of accountability concerns the relationship between the SEO and its beneficiaries and funders in form of trust and legitimacy, which can be built upon provision of transparent evidence-based results (Bagnoli & Megali, 2011). Performance measurements therefore become instrumental for SEO managements in terms of accountability.
and as the performance of SEOs oftentimes is expected to be a social impact, SEOs performance evaluations can be understood as the use of various social impact measurements. The use of such measurements to assess SEOs organizational performance becomes the dependent variable of our research.

3.2. Organizational size

Organizations are generally difficult to define, but common definitions include four main elements: size, particular goals to achieve, different roles, and an authority system (Tolbert & Hall, 2011). A great variety within each of these categories exists. The formal organizational structure has three dimensions that are still valid today: arrangements of organizational units, level of rules and standardized procedures as well as the level of autonomy in decision-making (Tolbert & Hall, 2011). In the closed-system approach as presented by Tolbert and Hall (2011), organizational size is next to technology and organizational culture, one of the key elements that shape the formal structure of organizations (Tolbert & Hall, 2011). Within these dimensions, size is the decisive factor which with ascent has great influence on organizational formal structure and internal processes. For example the bigger the organization, the more complex it becomes and therefore the need for more formalization arises; complexity thereon leads to decentralization of decision making due to increased control and coordination difficulties which again leads to increased formalization (Blau & Schoenherr, 1971). The question about what is the actual size of an organization then arises. Organizational size can be represented in at least four different concepts: 1) capacity, such as number of beds in a hospital or classrooms in schools; 2) staff available for the organization (this is the most common way to define size); 3) inputs, such as number of clients served, and outputs, for example sales volume; 4) discretionary resources, for businesses it is the amount available for investment (Kimberley, 1976). All four ways to define size has some disadvantages of only partially representing the size of an organization. For the first definition it is not given that the capacity is fully used, and therefore the numbers can be misleading. For the second approach, it needs to be clarified if part-time, temporary or volunteered personnel are included or not. The third and fourth definitions limit comparison to similar organizations and exclude some types. Evaluations highly depend on what measurements are used; different definitions deliver different results (Kimberley, 1976). In our study we base our independent variables on this idea, and operationalize according to the following: discretionary resources as the organization’s yearly turnover; staff as number of part- and full-time employees and volunteers, and inputs as number of members and participants in the organization. We do not include the capacity aspect because it is not relevant for this context since the range of goods and services provided is too wide to measure in this way and does not provide good ground for further analysis. These three size-measurements become the explanatory (independent) variables for our research.

Table 1: Independent variables. Suggestions from previous research on how to measure size with examples, and the operationalization and variable names for this study (based on Kimberley, 1976).

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Operationalization</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discretionary resources</strong></td>
<td>Resources available to the organization in form of wealth and net assets</td>
<td>Yearly turnover</td>
<td>size_turnover</td>
</tr>
<tr>
<td><strong>Personnel</strong></td>
<td>Staff available to the organization</td>
<td>Full-time employees</td>
<td>size_staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Part-time employees</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers</td>
<td></td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td>Number of clients served or inmates housed in a prison</td>
<td>Members and partici-</td>
<td>size_members</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pants</td>
<td></td>
</tr>
</tbody>
</table>
3.3. Conceptual framework

As presented above, all types of organizations evaluate and measure their performance in one way or another (Huang & Hooper, 2011; Munir, Baird & Perera, 2013; Oh & Bush, 2015; Tolbert & Hall, 2011). Depending on the sector and set goals, the organizations decide what they want to achieve and how it shall be done. For SEOs it is often done by measuring social impact in the community where they operate. The use of SIM is a way for SEOs to balance the performance measurements from all three sectors while adding a layer of legitimacy (Bagnoli & Megali, 2011; Nicholls, 2008; Paton, 2003; Zadek, 1998). There are various approaches on how to measure social impact developed in different areas of the world: Measuring Effectiveness from the UK (Leviner, 2006), PHINEO from Germany (Epkenhans, 2012), SIMPLE also from the UK (McLoughlin, Kaminski, Sodagar, Khan, Harris, Arnaudo & McBrearty, 2009) and many others (Elkington, 2001; Emerson, 2003; Kaplan, 2002; Nicholls, 2008; Zadek, 1998). Assessing social impact is difficult and the measurements need to be chosen with caution (Nicholls, 2008). However, some SEOs do undertake such assessments. Based on organizational theories that emphasize the importance of organizational size to have great effect on the formal structure as well as on the internal processes, we assume that the decision to use SIM to assess performance partially depends on the size of the organization. Many other factors can also influence the decision of using or not using SIM, one example could be whether the SEO receives public funding or not, since organizations or activities receiving governmental support need to provide evidence of their success to collect further funds (Huang & Hooper, 2011). Another factor can be the main field in which the SEO operates, for example initiatives that target a bigger part of the community and are expected to have an impact on society, such as organizations working in healthcare or labor integration. In contrast to SEOs that work mainly with animals, these organizations are maybe more likely to evaluate the social impact in order to legitimize their activities. In our research we will concentrate mainly on the size of the organization since there has been much research done on organizational size affecting the occurrence of activities related to evaluation, sustainability, corporate social performance and leadership (Gallo & Christensen, 2011; Jaskyte, 2013; Stanwick & Stanwick, 1998; Vaccaro et al., 2012). Therefore we hypothesize that organizational size is related to initiatives to measure social impact in SEO. Above we specify our independent and dependent variables and based on the presented assumption we formulate the following claim: the use social impact measurements increases with size of socially entrepreneurial organizations. From this we derive three different hypotheses based on Table 1:

\[ H_A: \text{The use of social impact measurements relates positively to yearly turnover in socially entrepreneurial organizations.} \]

\[ H_B: \text{The use of social impact measurements relates positively to number of available staff in socially entrepreneurial organizations.} \]

\[ H_C: \text{The use of social impact measurements relates positively to number of members and participants in socially entrepreneurial organizations.} \]
4. Methods

In the following chapter we detail how the survey for data collection was created and the rationale behind the asked questions. We then provide an overview on how we distributed the survey to collect relevant data, present the sample and describe our methods for the analysis. We finally argue for our ethical considerations, and the strengths and the limitations of this conducted research.

4.1. Survey creation

As suggested by Ejlertsson (2005), we began our data collection process with a thorough literature review (see Chapter 3) to identify the right terminology and existing problems in the field, but also to limit the scope of the study and keep it focused. Survey studies are suitable methods for research in many fields and levels, but to develop a quality instrument is more challenging than many novice researchers realize at first. Poorly designed surveys will produce unsatisfactory data that cannot answer the desired question (Passmore, Dobbie, Parchman & Tysinger, 2002). A good way to ensure the quality is to use existing questionnaires, as they have already been assessed for validity and reliability (Ejlertsson, 2005; Kitchenham & Pfleeger, 2002). Unfortunately, this was not possible in our case as we are doing research with an ambition to investigate the connection between organizational size and the use of SIMs which has to our knowledge not been done before. Despite the explorative topical component, the survey is constructed to describe and explain this association. The next task was to determine what language to use in the survey, and whether we should provide it in both English and Swedish in order to maximize the number of potential respondents. After careful considerations we chose to only formulate the questions in English; a detailed discussion of the advantages and disadvantages of this decision is provided in section 4.5.2. Quality of the research.

A survey asks the respondents to answer questions, so the starting point in designing a survey instrument should always be the survey's purpose and objectives. When writing this piece of information we considered components as suggested by previous research (Ejlertsson, 2005). The survey description needs to be brief but at the same time define the context of the survey, tell the respondents who is eligible to fill it in, what areas to cover, and what types of information we are looking for (Kitchenham & Pfleeger, 2002). It is particularly important to keep it short and concise. However, we expected our target population to have some prior knowledge of the areas, as they are in fact social entrepreneurs or working in socially entrepreneurial organizations, which allowed us to use a somewhat academic language. As a researcher you should however not assume that the respondents know the terminology you use, which is why we started the description by very briefly defining the terms socially entrepreneurial organizations and social impact measurements. This is to ensure that we have the same understanding of some basic concepts in order to minimize the risk of misinterpretation and internal loss of data (Kitchenham & Pfleeger, 2002).

As explained in Chapter 3, our dependent variable is the current use of SIM in SEOs. This variable is a self-reported, categorical question of whether the corresponding organization uses such measurements or not, giving the options of ‘No, never’, ‘Yes, once or twice’, ‘Yes, regularly’ or ‘I don’t know’. We also added a question about how the organizations evaluate their performance today to explore what kind of measurements are commonly used in SEOs in Sweden, if any at all. For our explanatory variables, we claim that organizational size has a positive association with the dependent variable. As stated above, organizational size is measured as 1) the yearly turnover; 2) the number of employees and volunteers, and 3) the number of members and participants within the organization. When formulating these questions for our survey instrument, we considered both options of open or closed questions. An open re-
Response format allows subjects to answer a question by typing their own response in text and, consequently, a question is closed when a respondent is asked to select an answer from a list of predefined alternatives (Kitchenham & Pfleeger, 2002; Passmore et al., 2002). By restricting the respondents' response options, closed questions are easier to analyze and are often preferred in surveys to ease the statistical analysis. However, if deciding to use a closed question option, we may need to include a ‘None of the above’ or ‘I don't know’ answer when asking categorical questions (Kitchenham & Pfleeger, 2002; Passmore et al., 2002). In addition, the type of questions we needed to ask would be challenging to create categories from, for example what is a suitable interval for a category asking for the organization’s yearly turnover? To avoid this issue and allow for more individual responses, we chose open questions for our explanatory variables. On the other hand, we chose a closed question design for the dependent variable; it is not asking the respondent to provide a number, but a report of some previous happening, and keeping it open would leave room for misinterpretation. A further decision when using closed categorical questions is whether the respondent can choose one or more responses to the question (Kitchenham & Pfleeger, 2002). Since we want the respondents to provide only one answer, we ensured that the categories are mutually exclusive (limited by the technical options provided by the survey platform). Answering open questions requires more time and effort from subjects than checking responses on a list and thus, the response rate may be lower and the data less complete than that obtained using a closed format (Passmore et al., 2002). But in our favor, the open questions were only asking about numbers therefore not too daunting or time consuming. Similarly, questions must be formulated in a way that respondents can answer them easily and accurately. It is also important to make sure that respondents have sufficient knowledge to answer the questions (Kitchenham & Pfleeger, 2002). In this case, the asked questions are on behalf of the organization the respondent represent, and the responses should be fairly easy to look up in case the respondent does not know the answer already.

Additionally to our dependent and explanatory variables we included a series of other questions. For example two questions to describe the context of the SEO, such as the area it operates in (e.g. education, environment, unemployment) which was a compulsory question allowing only one answer, and who are the main beneficiaries (e.g. unemployed, refugees, children) which was not compulsory and with the option of giving multiple answers. This allowed us to better understand our data sample and provide evidence that our results will represent not only organizations with similar mission and beneficiaries, but a very broad set of organizations.

Furthermore, it was important to keep in mind that the number of questions you can realistically ask in a survey depends on the amount of time respondents are willing to commit to it. In a self-administered questionnaire like ours it is possible to ask more questions, but we still needed to be sure that it is not considered too daunting. If it is, we risked getting answers that are inaccurate or incomplete (Kitchenham & Pfleeger, 2002). One way to reduce the time taken to complete a survey is to have standardized response formats, for example standardized to an ordinal scale (Kitchenham & Pfleeger, 2002). Apart from the dependent, independent and descriptive variables, we also wanted to describe what seems to be the major reason for assessing or not assessing social impact in Swedish SEOs. We asked one question respectively about the main challenges and opportunities of performing such assessments. These were formulated with a series of options that arose from reviewing current literature (Kernot & McNeil, 2011; Wilson et al., 2015), with an extra category of ‘Other’ for the respondents to specify if they did not feel their view was already represented as an option. For these questions we used a five-point Likert scale for frequency ratings of different alternatives importance ranging from 1 to 5, where 1 is ‘not at all important’ and 5 is ‘very important’ (Blaikie, 2003; Ejlertsson, 2005). A traditional Likert scale with this structure is commonly used in surveys, thus the format is familiar to respondents and it lends itself well to measuring constructs like attitudes and values like we did in this case. The number of scale points can vary; we chose five which is usually sufficient.

Some questions in our constructed survey were compulsory while others were not. The rationale behind each compulsory question is that it will give data that is important or even necessary for our research. The non-compulsory questions are rather interesting aspects that we could include or use in other ways that relates to our research, but that was not crucial for the results. In the end, we used a total number of 13 questions. We did not include a question about geographical location of SEOs mainly because we
wanted to limit the amount of items. We furthermore decided to view Sweden as a whole and would not have any relevant way to relate geographical data to our study since it is not a part of our hypotheses. Likewise, we believe it would be too broad to be able to interpret it in a meaningful way. After submitting the complete survey, respondents came to a page that thanked them for their participation and contribution to our study, and the option of registering for a copy of the study upon completion.

In the construction of this survey we have attempted to make the questions as clear as possible. Further, we refrained from using a large number of questions as long instruments tire respondents and can lower response rates (Passmore et al., 2002). After the first construction we performed a minor pilot study (Ejlertsson, 2005) where a few, people that were external to the study but easily available to us were asked to fill in the survey. After this, we studied the results and assessed how well the questions worked and if they were interpreted as we had intended. According to common practice we made some minor changes to the survey after the pilot study in order for it to be more functional and intuitive. These changes mainly concerned formulation matters in either the question descriptions or survey description, such as what the response is expected to contain or how we defined SEOs or SIMs. After these last edits, we were ready to launch. The complete questionnaire can be found in Appendix 1. The survey was done using Google Forms.

4.2. Data collection

Using common research methods like interviews or telephone surveys is often both an expensive and difficult task. The internet offers unprecedented research opportunities to overcome such barriers (Tijden, De Ruijter & De Ruijter, 2012). In this section, we discuss how we used the internet in this study, and how it can be used to meet the challenges of large-scale, continuous data collection from a great geographical base. Since the beginning of a more widespread internet access among the public the world has experienced a drastic expansion in number of web users, online services and applications, and every day millions of people use the internet to gather information, entertainment or to converse in online communities and social networks (Tijden et al., 2012). The internet can reach extremely large populations, as well as very rare populations, and it can provide data collection methods beyond traditional surveys (Tijden et al., 2012). A challenge however, is to recruit a sufficient number of web users with the right interest in these targeted organizations. One option is to identify websites or online communities aimed at specific topical issues, and ask them to either participate or help distribute the survey by linking it on their website (Tijden et al., 2012). Our distribution followed this pattern, and was done using internet by a form of distribution by interest method. It implies a method where a survey is distributed over different communication channels connected to a certain interest, for example when visiting a certain website or buying a particular magazine (Ejlertsson, 2005). In our case, we did this over Facebook (social media) and email. We approached different groups and organizations involved in social entrepreneurship and similar organizations and initiatives where actors can post messages and share a link to our survey. In some groups we could post this message ourselves, but for more formal organizational pages we asked for help from administrators to share it so that everyone who has liked their page, out of interest, could see our survey call (Ejlertsson, 2005). However, this only determines who can see the call for the survey, who then also chose to participate is beyond our control, and up to each respective individual. The included descriptive questions helped us then to identify the final sample of respondents.

In addition, we also had the condition that only people who are part of a SEO can participate in our survey. There is no way for us to ensure this is the case, though we constructed the questions in a way that if you were not eligible, you were not able to answer the questions properly. We also have no reason to suspect to that some answers were simulated since the survey was distributed by interest, and not openly shared with all Facebook-users. As mentioned earlier on, there is also some limitations concerning the language used as respondents had to be able to understand academic English, and also have access to a computer to be able to respond. Hence, the total population (Ejlertsson, 2005) is SEOs operating in Sweden, and the target population of the survey on the other hand was SEOs in Sweden with access to a computer, and who are able to understand the language in the survey.
Since there is no consensus or vast body of evidence within the field, there is so far no real possibility to determine how big the total population of SEOs in Sweden actually is, and consequently how big share of the population our sample represents. In order to generalize findings, the sample must be representative for the entire population which is commonly achieved through a randomized selection to avoid any upward or downward bias (Ejlertsson, 2005). Randomization is generally the best sampling method but very rarely possible, and unfortunately nor was it in our case. We instead chose the method of self-selection as described above. Previous studies have shown that the quality of data does not differ significantly between web-based and hard-copy surveys, providing support for the feasibility of using a web-survey format (Kiernan, Kiernan, Oyler & Gilles, 2005; McCabe, Boyd, Young, Crawford & Pope, 2005). In general, web surveys have relatively poor response rates, likely because they are in most cases self-selective (Converse, Wolfe, Huang & Oswald, 2008). We were aware of this fact, and in addition to the web survey we approached a large number of SEOs over email (closer to 350 organizations), distributing the survey to a more targeted group than what was possible on the online platform. Contacts details to these 350 organizations were found on a web platform created by Sofisam, which is a collaboration between three Swedish authorities (the Swedish Agency for Economic and Regional Growth, the Swedish Social Insurance Agency and the Swedish Unemployment Office). The platform contains a large share of all Swedish work-integrating, social enterprises (in Swedish: arbetsintegrerande sociala företag). However, this too is based on voluntary registration and did not enable us to reach any total population. Self-administered surveys distributed by email in this way are not expensive or difficult to use, and provide privacy and anonymity to respondents (Passmore et al., 2002). However, such surveys can typically yield numerous unusable or incomplete responses and may require multiple mailings to obtain a response rate high enough to generalize the data gathered to the whole target population (Passmore et al., 2002). In our case, the internal loss of data due to such problems were however very low, and to obtain a satisfying response rate we also sent out a reminder email one week after the first emailing. The survey accepted responses for a duration of two weeks, and even though we received emails from additional respondents after the closing date, we did not allow for any more responses once the survey had closed.

Our survey software (Google Forms) is an online method that instantly imports the data into an Excel sheet and both counts and sorts the responses. Although, we had to recode some of our variables to enable statistical analysis, which is explained in section 4.3 Data sample and 4.4. Statistical analysis.

4.3. Data sample

From the Facebook posts and approximately 350 email requests we had gotten a number of 67 respondents when the survey closed down. In order to be able to use our dependent variable as intended, we had to recode it to a dummy variable rather than a categorical as otherwise the number of responses in each category would have been unsatisfactory. A dummy variable means using binary categories or options rather than the four categories initially used as response options for the question: ‘Have you undertaken any assessment of your organization’s social impact?’ In addition, we got three different responses in the field for ‘Other’. Thus, we had to organize our data differently to get a binary variable. The ‘Other’ responses were easily sorted into the existing categories (for example ‘once every fiscal year’ were put into the response category ‘Yes, regularly’). We then eliminated the respondents who said ‘I don’t know’ (n=11), as these are irrelevant for our analysis. This number can be considered as internal loss of data. For the remaining (n=56), we coded the data according to this “No, never” became ‘0’ and both ‘Yes, once or twice’ and ‘Yes, regularly’ became ‘1’. Thus, the interpretation of the variable is that if the organization has not undertaken any assessment of their social impact, it is represented by a 0, and if it has ever undertaken any such assessments, the organization is represented by a 1. The total sample from this point is n=56.

For our independent variables we also had to perform some coding prior to conducting the statistical analysis. When respondents gave an answer that ranged from one number to another (such as 40-45), we consistently chose the larger number. There might for example be some vacancies not represented in currently employed staff number or other reasons for why they chose a ranging response. Similarly,
if they responded a number with decimals, we always rounded up rather than according to common mathematics. For example, if they had 2.25 employees the number was rounded up to three as there are three individuals working in the organization. As explained in Chapter 3, we consider both full-time and part-time employees in this measurement of size and did not restrict answers merely to full-time, however this adjustment only concerned five observations. The rationale behind both these decisions were mainly to be consistent in our choices, and we decided to choose max value. Further, some organizations indicated both a number of ‘clients’ and another number of ‘members’ for the question “How many members and participants do you have in your organization?”. In this case we used only members and disregarded the number of ‘clients’. For the variable of yearly turnover, both a no profit situation and a negative profit (loss) are represented as ‘0’. In cases where respondents stated several annual turnovers we chose the most recent year (2015). As shown in Table 2, some additional internal loss of data was present in the independent variables, and the total number of observations differs slightly between variables. A much more detailed descriptive analysis of our data will be provided in the next chapter, section 5.1 Descriptive results.

### Table 2: Descriptive data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total n</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>evaluation*</td>
<td>56</td>
<td>0 (n=30)</td>
<td>1 (n=26)</td>
<td>0.46</td>
<td>-</td>
</tr>
<tr>
<td>size_turnover</td>
<td>52</td>
<td>0</td>
<td>15 000 000</td>
<td>2 445 365.39</td>
<td>1 000 000</td>
</tr>
<tr>
<td>size_staff</td>
<td>56</td>
<td>0</td>
<td>1 000</td>
<td>77.98</td>
<td>9.5</td>
</tr>
<tr>
<td>size_members</td>
<td>54</td>
<td>3</td>
<td>8 000</td>
<td>352.44</td>
<td>20</td>
</tr>
</tbody>
</table>

4.4. Statistical analysis

All three hypotheses in this study, H_A, H_B and H_C, are stating a positive relationship. There is a connection between sample size and the ability to estimate population parameters, or to be confident about a particular relationship in a subsample. If the set significance level is equaled or below in the analysis, the estimator is said to be significant. That is, the estimated value is expected to be present in the total population from which the sample was drawn. Generally, we can be more confident about results from large samples, and have to be more cautious when small samples are used (Blaikie, 2003). However, the case against the broad use of statistical significance testing has been developed by many opponents (Carver, 1978; Johnson, 1999). The discussion of statistical significance uses the argument that as the world is not perfect, we must allow for some errors in how confident we are of our estimated results. While we may be pretty sure we know what the factors are, we can never be 100% sure - there is always a chance, no matter how small, that we might be wrong (Blaikie, 2003; Carver, 1978). The statistical significance of an estimator is represented by a p-value, and the significance level can be made as small as one wishes. The use of a fixed significance level of say 0.05 promotes the seemingly random distinction between a significant finding if p = 0.049, and a nonsignificant finding if p = 0.051. Thus, if an important finding is statistically non-significant, the researcher will conclude that the sample was insufficient and call for further research, an action criticized by some for being of schizophrenic nature (Johnson, 1999). It has even been claimed that all forms of statistical significance testing should be abandoned in favor of other measures such as effect estimates (Carver, 1978; Carver, 1993). In general, a significance level of 0.01, 0.05 or 0.10 is commonly used (Blaikie, 2003; Dougherty, 2011). If p-values are larger than 0.05, for example 0.12, it still means that we can with a certainty of 88% say that the hypothesis is probable.

Despite this above discussion we use, as much research suggests (Blaikie, 2003; Dougherty, 2011; Johnson, 1999), a significance level of 0.05 as determinant for rejection or confirmation of our hypotheses.
However, we specify for all models to what level the results are significant, and even if not met at the 0.05 level still discuss its relevance. All analyses were performed in the statistical software Stata (version 12).

4.4.1. Model I
When doing hypothesis testing with two different sample groups, a t-test is a common approach. It compares the difference in means between two different groups assuming normally distributed and continuous dependent variable (Blaikie, 2003). However, as our dependent variable is binary, this is not a good approach for us. Therefore, we instead conducted a Chi-square test. Applied researchers have employed this test for more than one hundred years, and some claim there is a call for greater consideration of foundational techniques like the Chi-square test, even before newer, more advanced methods arguing that the Chi-square test remains an important and useful method to analyze categorical data (Sharpe, 2015). In fact, a Chi-square test is probably the most commonly used test as it can be used for any distribution of nominal (also known as categorical) data (Gilbert & Prion, 2016). This means variables with characteristics that cannot be quantified, because they are instead used to separate units or individuals into groups or categories, rather than each item having an own, numerical value (Gilbert & Prion, 2016). A Chi-square test can be calculated using any number of categories, but the simplest example compares two variables in two categories creating a 2x2 table. Sharpe (2015) suggests one should avoid a Chi-square contingency table greater than 2x2 if possible, and so we followed this advice. After collecting and coding our categorical data, a Chi-square test is a simple technique to find out if more cases fell into one category than another, what is called a goodness of fit (Sharpe, 2015). A cell-by-cell comparison of observed frequencies and percentages helps us to better understand the nature of the evidence, and is provided in the next chapter (5.2 Statistical results).

It is important to have a somewhat equal distribution of observations between groups. Our groups were created from the dummy variable, 0=have not assessed any social impact, and 1=have assessed social impact at least one time. For example, if the group who did not measure social impact consisted of 90% of the respondents, the outcome would be attenuated such that it mathematically cannot approach the maximum value regardless of the strength of the relationship between the variables (Sharpe, 2015). In our case the non-evaluating and evaluating group are similar in size, which makes our data suitable for a Chi-square test. In any study where a Chi-square test is used, the categories must be established in a logically defensible and reliable manner (Sharpe, 2015). As our independent variables are in fact non-categorical after data collection, we grouped them appropriately at the median for each variable (size_staff=9.5 size_members=20, size_turnover=1 000 000).

4.4.2. Model II
After this first hypotheses testing, we moved on to performing a simple linear regression for each variable. We began by estimating an Ordinary Least Squares regression which assumes a linear relationship between two or more variables, in our case only two. It is used to develop estimates for coefficients that seek to quantify the relationship when a dependent variable is assumed to be determined by another variable, normally by estimating a difference in means (Dougherty, 2011). One problem with the mean as a measurement is that the tails of a distribution can dominate its value entirely, making the mean value inordinately influenced by only one or a few extremely large or small observations (Tukey, 1977). Thus, if a measurement is intended to reflect what a typical phenomenon is like, the mean can fail (Tukey, 1977). Extreme values like this are called outliers, commonly considered to be a data point that is far outside the norm for a specific variable (Osborne & Overbay, 2004). A high variance or the presence of outliers can lead to inflated error rates and substantial distortions of parameter and statistic estimates, resulting in a higher risk of falsely rejecting the hypothesis (Osborne & Overbay, 2004). One method to adjust for the presence of extreme outliers in the data is to exclude those particular observations from the analysis completely to determine whether they play a significant role for the explanation value of the model. Previous studies has shown a strong beneficial effect of such a technique, as accuracy tended
to increase significantly and errors of inference tended to drop substantially once outliers were completely removed (Osborne & Overbay, 2004). After careful consideration, we decided to do so in the variable where such a value seemed particularly problematic (size_turnover). This action lowered the mean value with over 500,000, and the median with exactly 100,000 (in Table 2, this observation is already excluded).

After this, there were however still some presence of very large values, particularly for the variables size_members and size_turnover as can be seen in Table 2 by the fact that the median value differs largely from the mean value. A popular measure of the dispersion of the probability distribution is the standard errors, a method to measure the dispersion of distributions around the mean (Blaikie, 2003; Dougherty, 2011). The lower the standard errors, the more concentrated is the data to the mean, larger standards errors on the contrary indicates that the data is spread wider from the mean (Blaikie 2003).

Another strategy for dealing with the problem of a few very large values is to give less weight to the values in the tails of the distribution and pay more attention to the values near the center (Tukey, 1977). Due to the nature of our data distribution, we therefore chose to use robust standard errors to correct for some of the variation caused by extreme values present in our explanatory variables. Using robust procedures such as robust standards errors is a common method when researchers want to protect their data from being distorted by the presence of some extreme values or large variance. Such techniques accommodate the extreme values in the data set at no serious inconvenience, making the data robust against the presence outliers (Osborne & Overbay, 2004).

We must also consider the possibility that our size variables are correlated. If so, we would not be able to isolate any effect to any particular variable of size, as they would be correlated to each other and thereby clouding the results. Table 3 presents the linear association of all variables. Correlation indicates the extent to which different units such as individuals or organizations occupy the same relative position on two variables (Lee Rogers & Nicewander, 1988). To test the correlation we use a common correlation coefficient, Pearson’s \( r \), which is limited to values between -1 and 1 (Blaikie, 2003). A value of 1 indicates a perfect positive correlation between two variables, meanwhile -1 is the perfect negative correlation. A correlation coefficient of 0 shows no correlated relationship between the two variables (Blaikie, 2003). A simple scheme for Pearson’s \( r \) indicates that a coefficient of approximately 0.1 is considered low, 0.3 medium and 0.5 as rather high (Blaikie, 2003). In our case almost all linear relationships are between 0.1 and 0.3, which is considered as low to medium. The only coefficient that reaches a value over 0.5, which is generally considered high, is the relationship between the explanatory variables size_members and size_staff. This represents a correlation between the number of members and the number of staff, implying that we can expect the effects of these variables to correlate to some extent. For this reason, we chose to perform three individual regression models instead of a multiple regression combining all three explanatory variables.

<table>
<thead>
<tr>
<th>Table 3: Correlation. The correlation matrix shows the correlation values for dependent and independent variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td>evaluation</td>
</tr>
<tr>
<td>evaluation</td>
</tr>
<tr>
<td>size_turnover</td>
</tr>
<tr>
<td>size_staff</td>
</tr>
<tr>
<td>size_members</td>
</tr>
</tbody>
</table>

To return to the method of Ordinary Least Squares, it implies an equation based on some underlying assumptions. Suppose a single linear relationship exists between a number of variables according to \( Y = \alpha + \beta X + \varepsilon \). The variables of the hypothesized relationship are called \( X \) and \( Y \), \( \varepsilon \) is a random error term with constant variance, and \( \alpha \) and \( \beta \) are constants to be determined by estimation. Assuming the X-
variables are independent of the random error term, the best linear unbiased estimates of $\alpha$ and $\beta$ are given by the method of least squares (Cochrane & Orcutt, 1949; Dougherty, 2011). Based on this, we derived our first equations for Model II:

$$Y_{org} = \alpha_{org} + \beta X_{size\_turnover, org} + \epsilon_{org} \quad (1)$$

$$Y_{org} = \alpha_{org} + \beta X_{size\_staff, org} + \epsilon_{org} \quad (2)$$

$$Y_{org} = \alpha_{org} + \beta X_{size\_members, org} + \epsilon_{org} \quad (3)$$

where $Y$ is the dependent dummy variable of whether the organization has assessed their organization’s social impact. The subscript $org$ implies that $Y$ depends on the organization, and takes an individual value of 0 or 1 depending on which observation we intend. $\alpha$ is the intercept, and $\beta$ our coefficient to be estimated that will determine the relationship between $Y$ and $X$. As our dependent variable is binary, the resulting beta coefficient can be interpreted on the probability scale, or transformed into a percentage point difference by multiplying its value by 100. $X$ in turn is understood as our independent variable that should explain why $Y$ takes a certain value. As explained earlier, it is measured as total yearly turnover ($X_{size\_turnover}$), number of staff ($X_{size\_staff}$), and number of members or participants within the organization ($X_{size\_members}$). This number too depends on the individual observation and is therefore followed by the subscript $org$. The last term is the error term which is also depending on each individual observation.

4.4.3. Model III

Lastly, we conducted a third model, a manipulation of the data and regressions from Model II in an attempt to simplify the interpretation of the results from the previous findings. What we did for Model III was a log-transformation of the independent variables. Logarithmically transforming variables in a regression model is often used to manage situations where a non-linear relationship exists between the independent and dependent variables, and therefore suits our case well. Using the logarithm of one or more variables instead of the un-logged form makes the effective relationship non-linear, while still preserving the linear model (Benoit, 2011). Logarithmic transformation is a convenient way of transforming highly skewed variables into something, approximately, more normal. In fact, there is a distribution called the log-normal distribution defined as a distribution whose logarithm is normally distributed, but whose untransformed scale is skewed (Benoit, 2011). Log-transformation is thus also a common method to adjust for large outliers in data, for example heavy tails. This is because it takes the natural logarithm of all single values, which homogenizes variance across the entire range of sampled data (Basu, Manning & Mullahy, 2004; Manning & Mullahy, 2001). However, it is not possible for observations with a value of 0 (as you cannot take the natural logarithm of 0), and thus all zero observations are excluded in Model III. These are however very few (the total number of observations for each analysis is presented in Table 9). To test Model III, we first generated three new variables from the previous $size\_turnover$, $size\_staff$ and $size\_members$, and got the new, log-transformed $ln\_turnover$, $ln\_staff$ and $ln\_members$. After this, we created the same model but with the new variables, resulting in the following log-linear equations for Model III:
\[ Y_{\text{org}} = \alpha_{\text{org}} + \beta X_{\ln \text{turnover, org}} + \epsilon_{\text{org}} \]  
\[ Y_{\text{org}} = \alpha_{\text{org}} + \beta X_{\ln \text{staff, org}} + \epsilon_{\text{org}} \]  
\[ Y_{\text{org}} = \alpha_{\text{org}} + \beta X_{\ln \text{members, org}} + \epsilon_{\text{org}} \]  

A log-transformation is not always suitable for small data samples like ours because it is highly challenging to interpret the results as they become very small. In the linear regression for Model II explained above, the coefficient \( \beta \) is interpreted as the change in \( Y \) for a one-unit change in \( X \). No additional interpretation is required beyond the estimate \( \beta \). This literal interpretation is the same even when variables have been logarithmically transformed, but instead a change of one-unit on the log scale. However, to enable a more intuitive interpretation it usually makes sense to interpret the changes in percentage change, rather than in log-units (Benoit, 2011). In a linear-log model like Model III, the literal interpretation of the estimated coefficient \( \beta \) is that a one-unit increase in \( \ln X \) will produce an expected change in \( Y \) of \( \beta \) units. To instead get a percentage change in \( X \), we can use the following: the expected change in \( Y \) associated with a \( Z \)-% change in \( X \) can be calculated by \( \beta \cdot \log([100+z]/100) \). Thus, to work out the expected change associated with a 10% increase in \( X \) we therefore multiply \( \beta \) by \( \log(110/100) = \log(1.1) =0.095 \). Multiplied by 100, this gives the expected percentage change in \( Y \) when \( X \) increases by 10% (Benoit, 2011). This procedure generates a percentage change in likelihood of using SIM when \( X \) increases with 10%, as will be detailed in section 5.2. Statistical results.

4.5. Ethical considerations and quality of research

Generally in social science research when entering some kind of relationship with people, either studying their behavior or asking them questions, researchers’ values and responsibilities towards the studied subjects are of great importance, and therefore ethical aspects need to be considered in all research (Esaiasson, 2007; Silverman, 2010). At the same time, researchers need to be aware of possible limitations of the conducted study, and weigh strengths and weaknesses against each other (6 & Bellamy, 2012; Silverman, 2010). Our ethical considerations and the quality aspects of this research are presented below.

4.5.1. Ethical considerations

As this study is done anonymously and further does not contain any sensitive or vulnerable data that in any way risk the integrity of respondents, and nor is it possible to connect the responses with any particular individual or organization, we did not take any extra measures to protect the information. All respondents were informed in the survey description about the purpose of the study and their answers, and gave consent before submitting (Esaiasson, 2007; Silverman, 2010). Responses were not collected using log in details, but done fully anonymously. Due to the anonymity of the online survey, the concern that responses will reflect what the respondents perceive as the ‘true’ or ‘best’ answer diminishes, and preserves privacy (Elias, Lacetera & Macis, 2015). We saw no reason to provide a secure web address, or take any other measures based on ethical considerations. In addition, all respondents were offered the chance to sign up for a copy of the thesis upon completion.

4.5.2. Quality of the research

Two essential aspects of conducted research are reliability and validity for the chosen research design and methods to collect and analyze the data. They contain both limitations and strengths of the study.
By being aware of the limitations, researchers can take measures to increase both validity and reliability of the research, and by being aware of the strengths, they can better argue for the value of their results and findings (6 & Bellamy, 2012; Silverman, 2010).

4.5.2.1. Strengths

One strength of a survey study is that it enables a much bigger sample than with for example interviews. In our case, the geographical context is the whole country, and the use of a survey and a technological distribution method enables a bigger geographical sample than with other methods (Ejlertsson, 2005). Some of the benefits of a web format survey include a broader access to respondents, also to those who might be generally difficult to reach due to geography, a desire for anonymity or reduced pressure to provide answers desirable by society, all while making it easier for the researchers to collect data (Joni- kas, Kisk, Grey, Hamilton & McNulty, 2010). In addition, the information is not given under any time pressure, which allows respondents to think through each question and reflect upon it if required. All respondents answered the exact same questions which improves the strength of this study in comparison to other collection methods, where the formulation and stating of questions might differ between researchers and occasions (Ejlertsson, 2005). Reliability is a measure for how consistently we measured the data. To test reliability of a study, it is common to refer to its replicability, meaning that if someone else would conduct the same study using same measures, sample and methods, the same results would be obtained (6 & Bellamy, 2012), and thus we have detailed each action and decision as carefully as possible. Generally in variable oriented research it is easier to control reliability since statistical measures are rather free from researcher’s subjectivity (6 & Bellamy, 2012). In addition, one alternative for measuring reliability is the internal consistency, which is the extent to which some variables are correlated to each other (Ejlertsson, 2005). We have tested this correlation of variables earlier on in this chapter, and could determine that it did not seem problematically high. The reliability of our analytic methods can therefore be considered rather high.

Measurement validity refers to how accurate the measurements measure what they are meant to (6 & Bellamy, 2012). Since we based our selection of variables on the theoretical framework provided in Chapter 3 and distributed our survey by interest method, we can with confidence claim that our conceptual validity is satisfying, an aspect of measurement validity that shows if the measures are in line with the specified theoretical framework (6 & Bellamy, 2012). Criterion validity refers to similar measures that are accepted as accurate in the research community (6 & Bellamy, 2012). To strengthen our criterion validity we reviewed the literature on organizational size and performance measurement in other sectors. It is very seldom possible to measure criterion validity, but it is generally considered as the coherence between an instrument (question) and the criterion (Ejlertsson, 2005). As our questions are in most cases closed, or only requiring a very short, specific answer, we do however deem the criterion validity relatively high. Similarly, we also consider content validity relatively high, as we use measures of size previously derived from other research (Ejlertsson, 2005). This is in accordance with our positivist methodological assumption, and therefore we consider it as one of our strengths in this research. The relevance of external validity is highly debated in research literature; some saying that generalization is all that matters and some saying that it depends on the interest of the researcher. If the goal is of theoretical or empirical nature with no intention of generalizing, the external validity is of little concern, and is even in many cases sacrificed for the greater statistical power. One does not require external validity for any single research study, but rather as the progress of adding experience from one study to another to prove or disprove earlier theories (Calder et al., 1982; Shadish, Cook & Campbell, 2002). We feel cautiously confident our results are generalizable to the total population of SEOs in Sweden, and they could be valid for a much bigger set of organizations than the ones included in our sample.

4.5.2.2. Limitations

Even though surveys as a collection method enable a bigger sample than other methods, the potential loss of answers needs to be considered. First of all, the answering of questions relies on memory and honesty of the respondents (this is not a limitation specific to our study) (Passmore et al., 2003). Another, more relevant limitation is how well respondents can understand the asked questions. This can be affected by the level of reading skills, respondent’s culture and language and many other factors (Passmore et al., 2003). In our case, this limitation will mainly concern the external loss of data, i.e., the risk that potential respondents will not participate because they do not fully understand either the purpose or the
questions (Ejlertsson, 2005). Internal loss on the other hand is when a respondent has avoided answering one or several questions, either by mistake or on purpose (Ejlertsson, 2005). By nature, this type of loss is easier to detect. We have attempted to minimize the internal loss of data by formulating clear and short questions, and also by not making all questions compulsory in case respondents lack some vital information for one or two questions it will not jeopardize the respondent’s participation. Any such missing data were not replaced in any way (Ejlertsson, 2005). Missing data on the dependent variable (n=11) were excluded completely, and whatever losses remain were mainly external. It is further possible that the main part of external loss of data was due to language difficulties. The option to provide the survey in both English and Swedish to correct for external losses was considered, but later rejected as certain terminology might be lost and consequently a difficulty of interpreting different results from differently constructed surveys.

Some of the general limitations of web surveys include among others concerns about self-selection bias in the sample as well as the validity of the collected data, but these can also occur in mail and personal surveys (Jonikas et al., 2010). Another weakness is the self-reported nature of the collected data as we cannot guarantee that respondents accurately represent their organizations. Additionally, with the anonymity design of the survey follows that we cannot assure that the same people did not complete the survey several times (Jonikas et al., 2010), however, we consider the risk low. In addition, there are some issues regarding the external validity of the study, which is to what extent it is possible to generalize the results (Calder et al., 1982). Since our findings are based on a non-random sample of SEOs, with access to internet and understanding of English, we must cautiously interpret the results since they do not represent all SEOs in Sweden which limits our possibilities to generalize the results across time and context. Further, the subsample used for our analysis included only n=56, which is relatively small in quantitative studies (Blaikie, 2003). Thus, there may be systematic differences between the sample and other target areas that could impair the ability to generalize our findings to all SEOs. Although, we believe our criteria is relevant for examining such a problem, and aimed at getting as many respondents as possible during the short time frame. Furthermore, generalizations of the results to communities that differ greatly from the Swedish context, whether nationally or internationally should be done with great caution. We cannot generalize these results to other countries since the results are conditional on the Swedish context and conditions Swedish SEOs are operating under.
In this chapter we first present the descriptive statistics of our sample in pie charts and tables for better visual understanding. We then provide the results from the statistical analysis of Model I, II and III, including coefficients, p-values and standard errors.

5. Results

5.1. Descriptive results

Through the distribution by interest-method we reached SEOs from various sectors with very broad and versatile missions. The entire sample’s (n=67) sector identity is presented in the pie chart below (Figure 1). Almost half of all respondents (45%) are working in areas targeting work integration and unemployment or other reasons for social exclusion. Their core activities are in several instances phrased as long-term unemployment, integration difficulties, physical or mental disabilities or other social problems such as drug addiction. These organizations are aimed at creating work opportunities and pathways to employment for groups who have particular difficulties in entering or re-entering the labor market. Another big part of Swedish SEOs work, according to our sample, is with the disabled community (15%). Also trade is well represented (12%) which includes second hand stores, sustainable manufacturing and catering services. Sustainable urban development (9%) and education (6%), including objectives such as working with children, mentorship positions or networking are also represented. Environment (3%), animal welfare (2%) and LGBTQ (1%) have the smallest shares in the sample. In the category Other (7%), we collected organizations that did not obviously fit into any of the other categories. Examples of these are health and elderly care as well as working with drug abusers or people on long-term sick leaves.

In the second question (see Appendix 1), we asked respondents about the organization’s main beneficiaries, most organizations (over 60%) named unemployed, the local community, and society in general. This shows that SEOs do expect to have a more general social impact on the community they operate in that exceeds the specific organizational objectives, and strengthens the idea that this performance too should be measured.

Figure 1: Sectors represented in our data sample. Main area of organizational activity chosen by each respondent, and calculated using the responses from Survey Question 1 (see Appendix 1).

Question 7 in our survey (see Appendix 1) asked the respondents about the age of their organization by stating the organization’s starting year. From that we calculated the age of the organization and clustered...
it into appropriate categories, which are presented in Figure 2. Nearly two thirds (64%) of all participating SEOs (n=67) are seven years or younger. 26% of all organizations were established either a few years before or a few years after the beginning of this century. In total, only 10% are older than 19 years old. As already mentioned in Chapter 1, SEOs is a rather young occurrence in Sweden, which is also the case in our sample.

Figure 2: Organizational age. The age of SEOs clustered in categories, and shares calculated using the responses from Survey Question 7 (see Appendix 1).

Furthermore, we asked the entire sample (n=67) of how they initially financed their organizational activities (Question 8, Appendix 1). Almost half (48%) of the respondents used public funding to start up their ideas. Public funds are often provided by local or national governments for ideas that somehow are expected to have an impact on the community, increase the quality of life or provide welfare, meanwhile reducing public expenditures (Cordery & Sinclair, 2013). Runner up is private funding, which makes up almost a quarter (23%) of all initial funding, followed by own savings (19%) and participant fees, including memberships and crowd-funding, that have the smallest share of only 10%.

Figure 3: Initial funding. The type of funding or financing SEOs used at the initial stage. Shares calculated using responses from Survey Question 8 (see Appendix 1).
In addition, we also asked the respondents how they evaluate the performance of their organization to this date (Question 9, Appendix 1). Out of all respondents, 17% do not evaluate their activities in any way, and the remaining share mentioned one or up to three different ways they assess their organizational performance. The most common ways to do so is through participant surveys (23%), conducting interviews and collecting feedback (19%), and counting the number of participants and members (17%).

11% of respondents use some existing SIM, such as SROI and Social IMPact measurement for Local Economies (SIMPLE). Other methods (13%) to evaluate present activities are mainly expressed as done through economic assessments, observations of individual development and summarizing activities in annual reports, however, this does not sufficiently describe any social impact.

Importance of different challenges and opportunities of using SIMs are presented in Table 4 and Table 5 below. The respondents could for this question rate the importance of each challenge and opportunity from 1 (not at all important) to 5 (very important). ‘I don’t know’ responses were immediately excluded for an easier interpretation of the results, and are not included in the presented average ratings. The highest rated challenge of using SIMs as viewed by the organizations (Table 4) is that it is a resource intensive process (average rating 3.74), followed by a lack of understanding (3.45) on how such assessments work and its rather subjective approach (3.43). Other challenges, such as difficult data collection (3.30), not being comparable to other organizations (3.13) or not a good representation of their activities (2.82) are rated as the least important aspects, but still score an average above 2.5 (median). The highest rated opportunity of using SIM is according to respondents the creation of awareness internally as well as externally (3.98) and the internal development (3.94) of the organization. Marketing purposes (3.88) and improving performance (3.84) are further benefits of assessing social impact that score relatively high on importance. To secure funding (3.50) is fifth place, followed by the chance to replicate the idea (3.44) and internal allocation of resources (3.29). All opportunities are rated above 3. In general the opportunities were rated higher than the challenges of measuring social impact. The highest rated challenge (resource intensive process with 3.74) has lower average rating than the biggest opportunity (create awareness with 3.98). The lowest rated challenge (not representative measurements with 2.82) is less relevant than the lowest rated opportunity (allocation of resources with 3.29). This shows that the respondents clearly see the benefits of using SIM and assessing their organization’s social impact and accept these as more or equally important as the challenges that appear with the application of such tools. We will discuss this further in the next chapter.
Table 4: Main challenges. Average rating of relevance for each challenges of using social impact measurements calculated using Question 11 from Survey (see Appendix 1).

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Average rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource intensive</td>
<td>3.74</td>
</tr>
<tr>
<td>Lack of understanding</td>
<td>3.45</td>
</tr>
<tr>
<td>Subjective evaluation</td>
<td>3.43</td>
</tr>
<tr>
<td>Data collection is difficult</td>
<td>3.30</td>
</tr>
<tr>
<td>Not comparable</td>
<td>3.13</td>
</tr>
<tr>
<td>Not representative</td>
<td>2.82</td>
</tr>
</tbody>
</table>

Table 5: Main opportunities. Average rating of relevance for each opportunity of using social impact measurements calculated using Question 12 from Survey (see Appendix 1).

<table>
<thead>
<tr>
<th>Opportunity</th>
<th>Average rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create awareness internally and externally</td>
<td>3.98</td>
</tr>
<tr>
<td>Internal development</td>
<td>3.94</td>
</tr>
<tr>
<td>Marketing purposes</td>
<td>3.88</td>
</tr>
<tr>
<td>Improving performance</td>
<td>3.84</td>
</tr>
<tr>
<td>Finding funding</td>
<td>3.50</td>
</tr>
<tr>
<td>Replication of organizational idea</td>
<td>3.44</td>
</tr>
<tr>
<td>Allocation of resources</td>
<td>3.29</td>
</tr>
</tbody>
</table>

Figure 5: Future use of social impact measurement. Self-rated forecast of the organization carrying out social impact assessments in the future. Calculated using responses from Survey Question 13 (see Appendix 1).
At last, (Question 13, Appendix 1) we wanted to see the general attitude of Swedish SEOs towards a future use of a SIM. Over 54% think that they will use such assessments in the future, 39% of respondents report they do not know, and only 7% say they are not planning to use SIMs in the future.

5.2. Statistical results

This section presents the findings from conducted Chi-square test, Ordinary Least Squares regression analysis and regression analysis after log-transforming the independent variables. For each model we discuss the significance level and estimated coefficient, and finally we confirm or reject the hypotheses for each model.

5.2.1. Model I

In the tables below, frequency counts and percentages are presented for each Chi-square test with corresponding $p$-values. The column evaluation=0 is the group who reported that they had not used any methods to assess social impact, and consequently evaluation=1 is the group who reported that they have used such methods. The organizations were as explained in the previous chapter divided into small and large segments by median, and this value is also displayed for each test. In the first test (Table 6), we tested whether the size of staff seemed to have any impact on the probability for assessing an organization’s social impact. From the 28 organizations coded as small (below median), 64% did not assess any social impact, and just over 35% reported that they did. This is a fairly large difference considering the total sample, 30 organizations in total did not assess social impact, and 18 of these are considered small. Looking at the group of larger organizations (above median), the difference is somewhat smaller, almost 43% do not evaluate and just over 57% do. However, the results are statistically non-significant at the 5%-level, and should hence be interpreted with caution. It is however very close to being significant at the 10%-level, and as discussed in the previous chapter this means that we can with a 89% certainty say that our hypothesis seems plausible, and that there is an association with number of staff and a use of SIMs.

<table>
<thead>
<tr>
<th>size_staff</th>
<th>evaluation=0</th>
<th>evaluation=1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>18</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>64.29%</td>
<td>35.71%</td>
<td>100%</td>
</tr>
<tr>
<td>large</td>
<td>12</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>42.86%</td>
<td>57.14%</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>26</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>53.57%</td>
<td>46.43%</td>
<td>100%</td>
</tr>
</tbody>
</table>

*Median value=9.5, p-value: 0.108 (non-significant)*

For the following two tests of the number of members and yearly turnover (Table 7 and 8), frequency counts and percentages are presented as well. However, the differences between groups are very small and not close to being statistically significant. Thus, we will refrain from interpreting these results as an association with a $p$-value of close to 50% could be considered just as likely as an association by chance. What we can say, is that there seems to be a similar pattern of larger organizations in the evaluating group, and more of the smaller organizations in the non-evaluating group.
Table 7: Results Model I H₀. Frequencies and percentages of each respective group of conducting a Chi-square test testing the association between number of members and the use of social impact measurements.

<table>
<thead>
<tr>
<th>size_members</th>
<th>evaluation=0</th>
<th>evaluation=1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>19</td>
<td>14</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>57.58%</td>
<td>42.42%</td>
<td>100%</td>
</tr>
<tr>
<td>large</td>
<td>11</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>47.83%</td>
<td>52.17%</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>26</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>53.57%</td>
<td>46.43%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Median value=20, p-value: 0.471 (non-significant)

Table 8: Results Model I H₀. Frequencies and percentages of each respective group of conducting a Chi-square test testing the association between yearly turnover and the use of social impact measurements.

<table>
<thead>
<tr>
<th>size_turnover</th>
<th>evaluation=0</th>
<th>evaluation=1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>16</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>59.26%</td>
<td>40.74%</td>
<td>100%</td>
</tr>
<tr>
<td>large</td>
<td>14</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>54.55%</td>
<td>45.45%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Median value=1 000 000, p-value: 0.491 (non-significant)

5.2.2. Model II and III
The results for Model II and III are presented in Table 9. The results for Model II, as introduced in the chapter above including robust standard errors, are presented in the top part of the table. The constant (α) is not directly related to the independent variable but is dependent on the organization and shows where the curve intercepts the Y axis. It reaches values between 0.423 to 0.453 and is significant for p<0.01. The intercept is however quite irrelevant to interpret, but is that if X=0, this would be the corresponding Y-value, so that if size=0 the value of the constant is the probability of using SIMs. Therefore, we immediately move on to the coefficient (β), which is directly related to the independent variable. Both results for the variable size_staff and size_members are statistically significant at our chosen significance level of p<0.05. The coefficient for the variable size_turnover is however not significant, and therefore irrelevant to interpret in this form. The estimator for size_staff indicates that for each extra person of staff, the probability for using SIMs increases with 0.0005 units, which is maybe easier understood as a 0.05 percentage point increase. The estimated coefficient for size_members indicates that for each extra member, the probability that the organization uses SIMs increases with 0.0001, or 0.01 percentage points. These results are hard to understand in this state as they are very small, and we therefore log-transformed each independent variable as presented in Chapter 4. These results are also presented in Table 9. However, on the behalf of an easier interpretation, we lost some significance. In Model III none of the results are significant at the 5% level, but on the other hand we have already established that there seems to be an association of sort between use of SIM and size of staff and size of members, and a logarithmic transformation allows us to better interpret the data. After testing Model III,
we performed some further manipulation of data as explained in the previous chapter to be able to present a result of a change higher than one unit. These results show that if the yearly turnover increases by 10%, it generates the increase in likelihood of using SIM by 0.19 percentage points, however this association has never shown a significant result and we therefore entirely reject the corresponding hypothesis. Similarly, if the number of staff increases by 10%, it generates an increase in likelihood of using SIM by 0.67 percentage points. If the number of members increases by 10% it generates the increase in likelihood of using SIM by 0.43 percentage points. However, it needs to be noted that also the p-values for size_staff and size_members in this model are too high to be considered significant at our chosen level. In our case the results for R² tell us that variable size_turnover only accounts for 1.3% of the variance in the model. For size_staff and size_members 4.4% and 4.9% of the variance are accounted for by the model. This basically means that the prediction of future outcomes based on this data set is rather limited.

Table 9: Results Model II and III. The table shows the results from the performed analysis. The asterisk tells us about the significance of the test, *: p<0.10, **: p<0.05, ***: p<0.01. Standard errors are presented within parentheses.

<table>
<thead>
<tr>
<th>Model I</th>
<th>turnover</th>
<th>staff</th>
<th>members</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>0.4521***</td>
<td>0.4232***</td>
<td>0.4313***</td>
</tr>
<tr>
<td>(0.085)</td>
<td>(0.072)</td>
<td>(0.071)</td>
<td></td>
</tr>
<tr>
<td>coefficient</td>
<td>-4.01e-09</td>
<td>0.0005**</td>
<td>0.0001***</td>
</tr>
<tr>
<td>(1.89e-08)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.013</td>
<td>0.044</td>
<td>0.049</td>
</tr>
<tr>
<td>Number of observations</td>
<td>53</td>
<td>56</td>
<td>54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model III</th>
<th>turnover</th>
<th>staff</th>
<th>members</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>0.1650</td>
<td>0.2690**</td>
<td>0.3072**</td>
</tr>
<tr>
<td>(0.703)</td>
<td>(0.132)</td>
<td>(0.131)</td>
<td></td>
</tr>
<tr>
<td>coefficient</td>
<td>0.0204</td>
<td>0.0700*</td>
<td>0.0453</td>
</tr>
<tr>
<td>(0.0504)</td>
<td>(0.0411)</td>
<td>(0.0317)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.0113</td>
<td>0.0487</td>
<td>0.0311</td>
</tr>
<tr>
<td>Number of observations</td>
<td>50</td>
<td>55</td>
<td>54</td>
</tr>
</tbody>
</table>

5.3. Summarized results

Based on the presented findings, we below summarize the results in regards to the stated hypotheses with a significance level of 5%.

Based on the conducted Chi-square test in Model I, we reject H_A, H_B and H_C.

H_A: The use social impact measurements relates positively to yearly turnover of socially entrepreneurial organizations → not confirmed

H_B: The use of social impact measurements relates positively to number of available staff of socially entrepreneurial organizations → not confirmed

H_C: The use of social impact measurements relates positively to number of members and participants of socially entrepreneurial organizations → not confirmed
Based on the regression analysis performed in Model II, we reject $H_A$ and accept $H_B$ and $H_C$.

$H_A$: The use social impact measurements relates positively to yearly turnover of socially entrepreneurial organizations $\rightarrow$ not confirmed

$H_B$: The use of social impact measurements relates positively to number of available staff of socially entrepreneurial organizations $\rightarrow$ confirmed

$H_C$: The use of social impact measurements relates positively to number of members and participants of socially entrepreneurial organizations $\rightarrow$ confirmed

Based on the log-transformed regression analysis in Model III, we reject $H_A$, $H_B$ and $H_C$.

$H_A$: The use social impact measurements relates positively to yearly turnover of socially entrepreneurial organizations $\rightarrow$ not confirmed

$H_B$: The use of social impact measurements relates positively to number of available staff of socially entrepreneurial organizations $\rightarrow$ not confirmed

$H_C$: The use of social impact measurements relates positively to number of members and participants of socially entrepreneurial organizations $\rightarrow$ not confirmed

To conclude we can to some extent confirm our claim that the use of SIM is positively related to organizational size.
6. Discussion

In the following part we interpret our findings from above to address the research problem stated in the introduction part. We also discuss the practical implications for sustainable organizational development of SEOs and policy makers.

6.1. The decision to assess social impact

Based on the results from our three models it can be said that there seem to be an established relationship between the size of a SEO measured by number of staff as well as by number of members, and the use of SIM with a statistically significant certainty. However, the effect estimates are very small, therefore it is still difficult to speak about any strong relationship, and nor can we talk about causality at all. One possible explanation for the relationship between organizational size in term of people and the use of SIM is that it is a resource intensive process, which requires high ability in staff in order to conduct them. At the same time, not only the amount of employees and/or volunteers is relevant but also the number of members is important as they must participate in the data collection through filling out surveys, giving feedback etc. to enable the measuring. The association between size measured by yearly turnover and the use of SIM on the other hand, is not significant in any of the three models and the corresponding hypothesis is therefore entirely rejected. The fact that we cannot detect any significant relationship between these factors was surprising at first, but a potential explanation could be that a high yearly turnover can imply that the SEO works closely with the private sector. Such a relationship can impair difficulties with maintaining a social purpose at the core, and therefore the SEO tends to measure performance more in economic terms rather than assess any potential social impact (Defourny & Nyssens, 2010). In addition, yearly turnover is not the same as organizational profit - profit is calculated by subtracting the costs from the yearly turnover. The yearly turnover therefore does not represent the amount that is to be reinvested in some social good or the amount available for other activities, such as measuring social impact (which is generally a costly activity).

In general, there seems to be some established relationship between the use of SIM and organizational size among SEOs, meaning that the smaller the SEO, the smaller the chance of using SIM. Small SEOs often struggle to survive financially, and therefore rarely have the time or resources to evaluate performance in a more advanced way (Alamaa, 2014; Baron, 2011). This was also a concern put forward in the survey where respondents defended their reason for not assessing social impact with the argument that they are instead focusing on surviving on a day to day basis. Their primary concern is to pay salaries and acquire necessary equipment to keep providing social goods and services, and all other tasks become inferior to this. Bigger SEOs on the other hand tend to use SIM more, presumably due to greater availability of needed instruments in form of staff, finances and time (Gallo & Christensen, 2011;Stanwick & Stanwick; 1998). Larger organizations are oftentimes also more influenced by multiple stakeholders (e.g. other SEOs, NGOs, government, business partners etc.), and since social impact metrics are no longer managed by the organizations independently but rather constructed together with relevant stakeholders (Costa & Pesci, 2016), there could be a bigger pressure on larger organizations to provide tangible evidence and legitimize their activities by assessing social impact.

As shown in our study, SEOs in Sweden represent very diverse sectors and target different beneficiaries. A large proportion of them are aimed at labor market integration or local community development in general. This means that most SEOs have a built-in intention to enhance quality of life and have a social impact on the (local) society. At the same time, SEOs in general is a rather new phenomena in Sweden and is still in a nascent stage, which is evidenced by Figure 2 showing a large majority of our sampled SEOs are only 12 years or younger. So far, Sweden has only two main initiatives for Social Impact
Investment (Leksell Social Ventures and Impact Invest Scandinavia) for SEOs to receive funding based on their social impact evaluation. Another discussed issue is that Sweden still does not have a specific legal form for social enterprises (Alamaa, 2014; Defourny & Nyssens, 2008). The dominant organizational form for Swedish SEOs is still as a non-profit organization or the so-called cooperative or social cooperative (Alamaa, 2014; Defourny & Nyssens, 2008). These do not sufficiently represent the scope of many SEOs, and constrain them in some organizational activities. In line with our philosophy of science our aim is to generalize our research findings to a broader context. With our data, we cannot generalize to other countries than Sweden due to the specific contextual and ideological standpoint, for example a very well developed welfare state (Burström, 2015; Larsson, Letell & Thörn, 2012). As our study has shown, almost half of all Swedish SEOs financed their activities by public funding in the initial phase. In a welfare state, national and local governments are more likely to invest in initiatives and practices that contribute to raising the quality of social welfare and societal change (Cordery & Sinclair, 2013). However, we feel semi-confident to generalize our findings to the total population of Swedish SEOs and possibly as a pattern in other Scandinavian countries with similar contexts.

As presented in the theoretical framework, all types of organizations need to evaluate their activities and operations for several reasons and in different ways. In fact, according to our sample 83% of SEOs are somehow evaluating their organizational performance. Currently SEOs use both quantitative (such as participant surveys and counting the number of participants) as well as qualitative methods (such as interviews) to evaluate performance. Although, this does not necessarily mean that they measure any social impact of their activities. Balancing the evaluation requirements from the private, public and civil sector is already a challenging task, and additionally a layer of legitimacy needs to be considered when evaluating performance to assure evidence-based social impact assessment. Such assessments are according to our sampled SEOs very resource intensive and require both available staff and finances. Another challenge acknowledged by our sampled SEOs is the lack of understanding on how such assessments are carried out and where to seek external, evaluative support. The two main initiatives for Social Impact Investment are both located in Stockholm and one of them (Impact Invest Scandinavia) is aimed at Nordic countries in general and therefore primarily provides information in English. During our data collection process we received emails from several SEOs saying that due to their low understanding of English, they were not able to participate in the survey. This implication could drastically reduce the amount of available support and information for SEOs in Sweden concerning investment for social impact and measuring its return. Other challenges such as the existing measurements being rather subjective, not representative or comparable to other organizations, and difficulties with data collection can further influence the decision to measure social impact. However, while being aware of the challenges of SIM, the SEOs at the same time weigh the opportunities that come with such assessments as more important. They acknowledge both internal opportunities (such as creating awareness, internal development and allocation of resources) as well as external opportunities (such as marketing purposes and improved funding possibilities). Furthermore, most SEOs claim they are planning to use SIM in the future to evaluate performance and present more tangible evidence for their activities. A cautious interpretation of this self-reported prediction could mean that while the organization grows in size and acquires both more resources and staff they will move towards evaluating performance and social impact, confirming our claim that the use of SIM seems to depend on organizational size.

Our study to some extent confirms the relationship between organizational size and use of SIM, although the established effect is very low and several small SEOs do evaluate impact while larger ones do not as can be seen in Table 6. This could have several reasons which lie beyond the scope of our data. However, some explanations appear more plausible to us. For example, Swedish SEOs often receive public funding and support from the government, and therefore often have to provide tangible evidence to collect further funds, which require improvements in their evaluating capacity for measuring social impact (Alamaa, 2014). This fact could contribute to the integration of SIM even in smaller organizations as they too receive public funding, and thus are required to show an impact. Consequently, it is probable that if a large SEO has already reached an evaluative capacity and size, it has to some extent gained legitimacy and trust from beneficiaries, which can mean that the pressure to assess impact to gain further ground decreases. Performance evaluation can then be done by simpler methods that require fewer resources and are sufficient enough to present to the main stakeholders. Thus, it could be possible
that larger SEOs experience less pressure from funders to evaluate social impact since they can sustain financially through their own activities and thereby are less depending on external funding. These can be potential explanations for why the shown relationship is so small, and does not apply to all organizations.

6.2. Sustainable organizational development

According to our positivist methodological assumptions, we furthermore want to connect our findings to practical implications for sustainability that could contribute to both research and practice in the field (6 & Bellamy, 2012). One aspect concerning sustainability in organizations is organizational learning. Previous research has claimed that for organizations to be sustainable in operations and results, they need to become learning organizations (Argyris, 1977; Crossan, Lane & White, 1999; Ellström, 2001). According to Ellström (2001) organizational learning happens in two major modes, adaptive and developmental learning. The modes are based on a consideration of three work-learning situations: the tasks to be performed, the methods to be used, and finally the results to be achieved. As evaluation in general can be regarded as a form of knowledge production, objective evaluation is part of a broader knowledge strategy and organizational learning process (Nielsen et al., 2011). Organizational learning is a dynamic process that enables the required organizational change. For example the strategic renewal from Crossan and colleagues (1999) represents the combination of continuity and change which should encompass the entire organization. An organization that recognizes and manages the tension between exploration (acquiring new learning) and exploitation (using what has already been learnt) is a learning organization (Crossan et al., 1999). Organizational learning can thus be seen as a mean to provide a competitive advantage which enables the organization to sustain and develop (Crossan et al., 1999). Appropriate evaluations of social impact can provide an opportunity for both small and big organizations to become learning organizations, which should be desirable if they are hoping to be prolific and sustainable.

In addition, one of the key advantages of measuring social impact is the potential for funding opportunities. Economical together with environmental and social aspects build the ground for how sustainability is understood today, oftentimes termed as the Triple bottom line (Elkington, 1997). Funders and investors can, when faced with tangible evidence, choose the best applied practices for further investment to achieve economical sustainability (Lecy, 2010). Without performance evaluation, organizations have very limited possibilities for development and face the risk of remaining small without potential to grow (Baron, 2011). Similarly, if only larger organizations measure social impact, it could negatively affect the development of reliable measurements as they need to fit many different organizations. It could become a problem if bigger organizations were to dominate the field and develop biased measurements that only capture impact specific to their organization (Baron, 2011).

Thus, the concluding results of this study can in a sustainability perspective be useful for policy makers when supporting organizations by providing both knowledge and financial support, as well as when developing legislation and practices. If a society or community is lacking a certain social good or service, and an organization can correct for this and accomplish societal change for their stakeholders, it can too be expected to have social impact. The expected organizational performance then can be measured by assessing social impact. Based on our study, it does seem like there is a relationship between the decision to evaluate and organizational size. However, all SEOs need to measure their performance regardless of size since social impact and societal change is generally a desired outcome.
7. Conclusion

In the following chapter we summarize the findings, answer the purpose of this thesis, and further suggest a few future developments of social impact measurements.

7.1. Relationship between organizational size and social impact measurements

From our research, we can conclude that while there is no significant relationship between the organizational size measured as yearly turnover and the use of social impact measurements, our data suggests a relationship between size measured as the number of staff as well as measured by the number of members and use of social impact measurements. By increasing the size measured as number of staff, which includes both full- and part-time employees and available volunteers, by one unit the likelihood of using SIM to evaluate performance increases by 0.05 percentage points. At the same time by increasing the number of members and participants by one unit, the likelihood of using SIM increases by 0.01 percentage points. Even if this association is very small, it is statistically significant in our data set. Based on this we partially confirm our claim that the use of SIM seems to be related to organizational size, and we herewith answer the purpose of our thesis and provide an explanation of how organizational size in terms of people relates to the use of SIM. The most evident reasons for the fact that larger SEOs are more likely to evaluate performance by using SIM is that they have more available human and financial resources and are pressured by a larger number of stakeholders to provide tangible evidence of impact, meanwhile smaller SEOs with less resources are primarily fighting for their mere existence (Alamaa, 2014; Costa & Pesci, 2016; Gallo & Christensen, 2011; Stanwick & Stanwick, 1998). The rather small effect of organizational size on use of SIM can be explained by our rather small data sample which limits the statistical possibilities, but also by the fact that both small and big SEOs experience the need to assess social impact and at the same time see the challenges of using such tools. Therefore SEOs have to decide for themselves either to evaluate in this way or use simpler methods that do not fully capture their underlying ambition of achieving a social impact.

7.2. Future research and practical implications

To correct for some of the limitations inherent to our study we suggest that further research is conducted, and in this section we briefly detail some suggestions. First of all, we suggest that similar studies are done in a larger context with additional, more extensive data in order to determine what factors serve as incentives and barriers for the use of social impact measurements and other suitable performance evaluations of SEOs. As the number of SEOs increases, the larger becomes the need to evaluate and assess best practices and how such organizations can become more sustainable. We further suggest studies on how the type of funding can affect both the probability of using SIM and potential differences in the way they assess social impact, for example depending on whether they have received public funding or funds from business angels. In addition, we suggest more research on the field of social impact measurements that sufficiently measure an impact regardless of context or size, and that can serve as reference points and comparisons. Research should also develop common solid definitions for better comparisons and categorizations to evolve hybrid models from different combinations of instruments for improved evaluations of performance and impact. This has previously been emphasized by the OECD, claiming that there is an urgent need for “…new tools, including results-based financing, outcome-based approaches, market-based solutions and different forms of public-private partnerships…” (Wilson et al., 2015:17) to enable SEOs to evaluate their impact in a satisfactory way and at the same time stay within
a reasonable budget. Kernot and McNeil (2011) has previously proposed that relevant information about social impact measurements should be clearly embraced and disseminated by local governments to raise awareness and motivate organizations to assess impact (Kernot & McNeil, 2011). Further, in the Swedish context an official legal form for social enterprises needs to be introduced in the near future to address legal and institutional questions for practitioners in the field. To shift the mindset of the institutionalized, traditional markets as the only option for investments is a challenging process. Many organizations and individuals are working on promoting the social values among investors and society in general (e.g. OECD, NESTA, Social Ventures etc.), but this transformation might require a paradigm shift before they can be properly integrated, what Nicholls (2010) describes as *institutional transformation of social investment*. An ideal scenario is the *institutionalization of social investments* which can generate a systemic change across all investment parties, and that people would change their consumption behavior through value-driven rationality, and instead demand an economy built on virtues of fairness (Nicholls, 2010).
Reference List


Appendix 1

The Use of Social Impact Measurements in Socially Entrepreneurial Organizations

This survey is done as a part of our Master Thesis at Malmö University in the program Leadership for Sustainability. Our study is aimed at organizations in Sweden that target social issues, while being businesslike in operations without having a primarily financial motive. Such organizations are generally called socially entrepreneurial organizations. This study will investigate if socially entrepreneurial organizations measure their contribution to society and a sustainable development through the use of Social Impact Measurements. Social Impact is defined as a long-term change occurring in communities or societies as a direct result of a taken action, and Social Impact Measurements are any methods or tools for assessing this type of impact.

The questionnaire will only take about 3-5 minutes. By answering you agree to participate and give full consent. All answers will be kept anonymous and confidential, and only used for the purpose of this study.

If you have any further questions or would like a copy of the results upon the study's completion, please contact either albina.iljasov@gmail.com or hillevi.eilard@gmail.com

*Required

1. In which sector is your organization operating? *
   
   Mark only one option.

   - Unemployment
   - Women and gender issues
   - LGBTQ
   - Disabled
   - Education
   - Children
   - Social exclusion
   - Trade / second hand
   - Sustainable urban development
   - Environment
   - Animal welfare/protection
   - Networking/Mentorship
   - Other:

2. Who do you see as your main beneficiaries?
   
   Tick all that apply.

   - Unemployed
   - Immigrants
   - Refugees
   - Children
   - Youths
   - Women
   - Society
   - Local community
   - Homeless
3. How many people are employed in your organization? *
Include both part time and full time. Please try to answer as detailed as possible, however, if you do not know exact numbers, try to give an approximate answer.

4. How many people volunteer in your organization? *
How many people do you have working or helping out without financial compensation in your organization?

5. How many members and participants do you have? *
If you have a membership condition, how many members did you have during 2015? If you do not have a membership condition or register, estimate how many people were active in your organization as participants during 2015. Please try to answers as detailed as possible, however, if you do not know try to give an approximate answer.

6. How big is your yearly turnover? *
In SEK or other preferred currency (please indicate which one). Please try to answers as detailed as possible, however, if you do not know try to give an approximate answer.

7. When did you start up your organization?
Indicate the starting year for your organization (no exact date needed).

8. How did you finance your organization in the beginning?
Tick all that apply.

☐ Own savings
☐ Public funding
☐ Private funding
☐ Participant fees
☐ Other:

9. How do you evaluate the success of your organization today? *
Tick all that apply.

☐ We don't really evaluate our activities
☐ By interviews
☐ Participant surveys
☐ By numbers of participants and/or members
☐ By using existing Social Impact Measurements (for example Social Return On Investment, Social IMPact measurement for Local Economies, Demonstrating Value etc)
☐ Other:

10. Have you undertaken any assessment of your organization's social impact? *
Mark only one option.

☐ No, never
☐ Yes, once or twice
☐ Yes, regularly (please specify below how often)
☐ I don't know
11. What do you think are the main challenges of using Social Impact Measurements?
Please rate each challenge between 1-5, where 1 is not at all challenging and 5 is very challenging aspect. 
*Mark only one number per row.*

<table>
<thead>
<tr>
<th>Challenge</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of understanding on how to do it</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collection difficulties (integrity issues, willingness to participate)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource intensive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not comparable with other organizations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not representative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you choose "Other", please indicate what you see as an opportunity:

12. What do you think are the opportunities that come with using such measurements?
Please rank each opportunity 15, where 1 is not at all considered an opportunity of using Social Impact measurements, and 5 is what you think would benefit your organization the most. 
*Mark only one number per row.*

<table>
<thead>
<tr>
<th>Opportunity</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marketing purposes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improving performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal development / getting feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allocation of resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create awareness internally and externally</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replication of the idea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If you choose "Other", please indicate what you see as an opportunity:

13. Are you planning to use any Social Impact Measurement tools in the future? *
Mark only one option.

☐ Yes, I think we will
☐ No, I don't think we will
☐ I don't know

Thank you for your participation! It was very helpful for our work.