



Outcome-Based Contracting metrics, targets and incentives: The case of Portuguese Social Impact Bonds

Francisco Anselmo
152118323

Dissertation written under the supervision of Professor Francesco Longo.

Dissertation submitted in partial fulfilment of requirements for the MSc in International Management with major in Entrepreneurship and Innovation, at the Universidade Católica Portuguesa and for the MSc of Economics and Management of Innovation and Technology, at Università Luigi Bocconi.

6th of January, 2021

ABSTRACT

Title: *Outcome-Based Contracting metrics, targets, and incentives: the case of Portuguese Social Impact Bonds*

Author: Francisco Martinho Bragança Catarino Anselmo

Keywords: social impact bonds; outcome-based contracting; social impact investment; social service delivery; outcome metrics

Over the past decade, Social Impact Bonds (SIBs) emerged as an innovative outcome-based financing solution for social programs. Despite the general benefits associated with the mechanism, there are concerns regarding the appropriate metrics and evaluation methods to use. The analysis has considered two major aspects which may affect SIB success and diffusion: the definition of outcome metrics and its underlying incentives. To shed some light on these issues, six contracts were considered and discussed based on interviews with key stakeholders, providing an analysis of the challenges in the implementation of outcome-based mechanisms, mostly SIBs, in Portugal. The main findings highlight: (1) the need to develop further incentives for all stakeholders through more sophisticated metrics; (2) to decrease the administrative complexity of outcome reporting processes; and (3) to formally structure a potential integration of the intervention's methodologies and insights into public policy, supported by more robust evaluations. Overall, the dissertation confirms the main advantages and struggles faced by practitioners worldwide, with additional operational challenges resulting from the unique nature of the Portuguese SIB ecosystem. In particular, the existing constraints associated with the use of European Funds by the public outcome payer.

SUMÁRIO

Título: *Métricas, metas e incentivos em contratualização por resultados: o caso de Títulos de Impacto Social em Portugal*

Autor: Francisco Martinho Bragança Catarino Anselmo

Palavras-Chave: títulos de impacto social; contratualização por resultados; impacto social; investimento social; métricas de pagamento por resultados

Na última década, os Títulos de Impacto Social (TIS) surgiram como uma solução de financiamento inovadora para programas sociais mediante a contratualização por resultados. Apesar dos benefícios atribuídos ao mecanismo, ainda existem questões em relação às métricas e métodos de avaliação adequados. Desta forma, a análise focou-se maioritariamente em dois aspetos que podem afetar o sucesso e a difusão dos TIS: a definição e avaliação das métricas de resultado e os seus incentivos subjacentes. De forma a clarificar estas questões, seis estudos de caso foram considerados e discutidos com base em entrevistas com intervenientes, apresentando uma análise dos desafios na implementação de mecanismos de contratualização por resultados, maioritariamente TIS, em Portugal. As principais conclusões indicam: (1) a necessidade de desenvolver incentivos para todas as partes envolvidas através do uso de métricas mais sofisticadas; (2) diminuir a complexidade administrativa dos processos de apresentação e validação de resultados; e (3) estruturar formalmente uma potencial integração das metodologias e aprendizagens das intervenções em políticas públicas, suportada por métodos de avaliação mais robustos. De maneira geral, esta dissertação apresenta um alinhamento com as principais vantagens e dificuldades enfrentadas pelo mecanismo TIS a nível internacional, com desafios operacionais adicionais resultantes da natureza particular do ecossistema Português. Em particular, as limitações existentes associadas ao uso de fundos Europeus no pagamento de resultados.

TABLE OF CONTENTS

1. INTRODUCTION	4
2. LITERATURE REVIEW	7
2.1 OBC: THE CASE OF SIBs	7
2.1.1 <i>Where It Happens More Often?</i>	8
2.1.2 <i>Traditional SIB Design</i>	9
2.1.3 <i>Main Drivers and Barriers for OBC and SIB Use</i>	11
2.2 OUTCOME MEASUREMENT	12
2.2.1 <i>Metrics</i>	13
2.2.2 <i>Monitoring and Measurement Evaluation</i>	15
2.3 UNDERLYING INCENTIVES	16
2.3.1 <i>Public OBC</i>	17
2.3.2 <i>Perverse Incentives</i>	18
2.3.3 <i>Financial and Reputational Risk</i>	20
2.4 PORTUGUESE SIBs	22
3. METHODOLOGY	24
3.1 RESEARCH PHILOSOPHY	24
3.2 RESEARCH STRATEGY AND DESIGN	24
3.3 CASE STUDIES	26
3.4 THE DATA COLLECTION PROCESS	28
3.4.1 <i>Primary Data</i>	28
3.4.2 <i>Secondary Data</i>	29
3.5 DATA QUALITY AND ETHICAL CONSIDERATIONS	30
4. CASE STUDY FINDINGS	31
4.1 MEASUREMENT	31
4.1.1 <i>Metrics and Targets</i>	31
4.1.2 <i>Monitoring</i>	35
4.1.3 <i>Measurement and Impact Evaluation</i>	37
4.2 UNDERLYING INCENTIVES	39
4.2.1 <i>Perverse Incentives</i>	39
4.2.2 <i>Binary Thresholds</i>	41
4.2.3 <i>Financial, Social and Reputational Incentives</i>	43
4.2.4 <i>Expense Reporting and Investor Risk</i>	48
4.2.5 <i>Investor and Provider Diversity</i>	53
4.3 PUBLIC OBC	57
4.3.1 <i>Public Monitoring</i>	57
4.3.2 <i>Communication</i>	59
4.3.3 <i>Integration into Public Policy</i>	60

5. DISCUSSION.....	63
6. CONCLUSION.....	67
7. REFLECTION.....	70
7.1 SUGGESTIONS.....	70
7.2 LIMITATIONS.....	71
7.3 FUTURE RESEARCH.....	72
8. REFERENCE LIST.....	73
9. APPENDIX.....	83
9.1 INTERVIEW GUIDE.....	83
9.2 ORGANIZATIONS DESCRIPTION.....	85
9.2.1 <i>Portugal Social Innovation (PSI)</i>	85
9.2.2 <i>Maze Impact</i>	85
9.2.3 <i>Calouste Gulbenkian Foundation</i>	85
9.2.4 <i>TESE (Association for Development)</i>	86
9.2.5 <i>ASSOP</i>	86
9.2.6 <i>Movimento Defesa da Vida (MVD)</i>	86

LIST OF FIGURES

FIGURE 1: IMPACT BONDS PER SECTOR 9
FIGURE 2: THE PETERBOROUGH SOCIAL IMPACT BOND..... 10
FIGURE 3: HIERARCHY OF EVIDENCE QUALITY..... 15
FIGURE 4: INTERVIEWEES SUMMARY.....28

LIST OF TABLES

TABLE 1: CASES SUMMARY26

1. INTRODUCTION

Over the past decades, public administrations worldwide, particularly in Western countries, have experienced organizational reform, namely through the introduction of the New Public Management (NPM) trend (Hood, 1991). NPM broadly involved an attempt to implement management techniques and values from the private sector into public services (Haynes, 2003; Mulgan *et al.*, 2011; Liebman, 2011), especially relevant in a time when governments continued to search for means of reducing public sector budgets in the aftermath of the 2008 crisis (Andrews & Van de Walle, 2012). One of the critical elements of this paradigm shift focused on government settings moving from efficiency and inputs to broader goals and outcomes (Moore, 1997; O'Flynn, 2007).

The economic and financial crisis also contributed to an aggravation of social problems concerning social cohesion, unemployment, homelessness, healthcare, education, societal aging, and immigration. The rise of these so-called "Wicked" problems presented new multi-level, multi-actor, and multi-sectoral challenges (Borgonovi *et al.*, 2019). Such challenges require a higher engagement from civil society and major stakeholders in the decision-making process (Bryson *et al.*, 2014; Lovan *et al.*, 2017).

On the supply-side, recent attempts by public administrations to re-engineer their contracting processes to outcome-based go in line with the emergence of social impact finance, pay-for results, and outcome-based financing models (Arena *et al.*, 2016). Financial markets witnessed an increase of socially conscious investors, willing to intentionally pursue social impact alongside the financial return (Bugg-Levine & Emerson, 2011). In this context, the Social Impact Bond (SIB) model, as a form of social impact investing, started to be considered as a possible response to the problem of funding service delivery and an instrument capable of fostering efficiency, effectiveness, and innovation (Fox & Albertson, 2011). Policymakers in several countries claimed for increased use of SIBs and highlighted the need to create an adequate ecosystem to enhance the use of these instruments (G8, 2014).

Looking at the demand-side, the resulting need to transform the public approach to the delivery of social services incentivized policymakers to consider third sector and private organizations as potentially more efficient outsourcing partners than the established public sector providers (Allen, 2009; Millar, 2012; Arena *et al.*, 2016). Social Impact Bonds,

aligned with NPM values (Hood, 1991), can provide a solution by linking the traditionally separate private sector and social services; and go even further by promoting collaboration between public, private, non-profit and voluntary sectors (Fraser *et al.*, 2016; Mulgan *et al.* 2011; Cohen, 2011).

However, despite the popularity rise of the SIB model, its potential benefits for the public and third sector are also weighted against the challenges stakeholders face in shaping this instrument (Chiodo, 2015). Besides the lack of provider competition, access to data, and commissioner organizational capacity, one of the most recurring scale barriers identified by SIB stakeholders worldwide concern measurement design and evaluation methodologies (Gustafsson-Wright, 2020a). On the one hand, the public sector reports difficulties in measuring performance and accountability (Martikke & Moxham, 2010). On the other hand, social sector organizations denounce a low involvement from the public authority in the specification of services and associated risks, constraints on target setting, and related monitoring procedures (Chiodo, 2015). A common challenge, however, concerns the ability of simple metrics, typically more suitable to the SIB mechanism, to measure the successful impacts of interventions (Gustafsson-Wright, 2020a). Finally, the novelty of the approach also implies increased political and media attention which translates into higher reputational risk for stakeholders (Gustafsson-Wright & Osborn, 2020a).

This thesis attempts to explore what are the primary motivations for stakeholders to engage in such mechanisms and also identify the main enabling factors, in particular with regards to outcome measurement and its underlying incentives, for SIBs to be established as a feasible instrument to improve the purchasing and delivery of social services.

As identified in the report developed by Brookings, a better understanding of the driving forces or barriers to SIB development in certain countries could help provide insights into the different paths to create a competitive market for outcomes (Gustafsson-Wright, 2020a). While it is often difficult to replicate entire interventions, some aspects of the SIB model, namely outcome metrics, were identified as potentially more suitable for replication (Ecorys, 2019).

With this in mind, the study's primary goal is to shed light on the metric and threshold selection in outcome-based contracting (OBC) mechanisms in Portugal, through a particular

focus on SIBs, and to understand the impact of the underlying incentives of these indicators on stakeholders and the overall success of the interventions.

Overall, six OBC case studies were evaluated through semi-structured interviews and secondary data: The first SIB to ever be launched in Portugal and the only one outside Portugal Social Innovation (PSI) scope; the three SIBs included in the first wave by PSI which are currently entering the final stage; a recent SIB launched in 2019 with PSI; and finally, a Joint Action Plan Pilot, with a common provider as one of the evaluated SIBs, which allowed for a direct comparison between SIBs and other outcome-based financing mechanisms. There were a total of 8 interviewees across 6 organizations: two public entities; two service providers; one social investor; and one specialized intermediary.

The study is driven by the following research questions: 1) How are SIB's outcome-based contracts structured? 2) Which metrics and measurement processes do they use? 3) Are they related to incentives? 4) Which are the main barriers and drivers that may explain a higher use of such mechanisms in specific contexts?

From the four research questions, the literature's scope and structure was defined, which translated into the subsequent framework for the interviews, interview findings, and discussion.

2. LITERATURE REVIEW

The purpose of this chapter is to lay the theoretical foundation of this research. The structure starts on a broad level by evaluating the traditional SIB model and becomes more specific into the respective topics of research as it proceeds.

2.1 OBC: The Case of SIBs

Historically, most social service contracting between governments and providers has been basing payments on inputs. More recently, several financing and contracting mechanisms, of which impact bonds are one example, have contributed to shifting payments from inputs to the output and outcome side (Gustafsson-Wright, 2020b).

Contract's terms can be broadly based on the service's inputs, outputs, outcomes, or a combination. Inputs are the resources ("things") used to deliver the service, the labor, skills, and materials combined in the provisioning process. Outputs are the actual results of the service itself (Brown *et al.*, 2016). An outcome represents an objective a program wants to achieve - not what the program itself produced (the output), but the consequences of services in the local areas of competence (Sorci, 2019).

Outcome-based contracting (OBC), where provider's payment is dependent on the achievement of specific outcomes, is taking on increasing significance as a commissioning model in public services reform. Other terms that include outcome-based contracts include payment by results (PbR), performance-based contracts, payment for performance, and finally, social impact bonds (SIBs). The latter is also known as pay for success in the U.S., and social benefit bonds in Australia (Farr, 2015). A Social Impact Bond can introduce a new type of public-private partnership to provide social services (Jackson, 2013; Nicholls & Tomkinson, 2013), as an extension of payment for performance in public services (Lagarde *et al.*, 2013). From a public reform perspective, through OBC, the SIB model is able to foster entrepreneurship in the public sector (Osborne & Gaebler, 1992) while adopting private sector practices such as a higher exposure to competitive tendering and the application of private sector-influenced audit systems (Power, 1999; Fraser *et al.*, 2016).

In practice, a SIB is a financing mechanism aimed to fund preventive interventions relying on an outcome-based contract (Chiodo, 2015). Through a shift from process to outcome metrics, SIBs adopt an "evidence-based" approach where the goal is to understand what works (Deering, 2014). The focus on prevention is associated with the potential "savings" from public service budgets by improving outcomes (e.g. lower recidivism rates will accrue savings in police, courts, prison, probation, etc.) (Fraser *et al.*, 2016). Some impact bonds even feature outcome metrics structured around these future avoided costs (Rudd *et al.*, 2013). The ability to prevent complex social problems and foster innovation are particularly relevant in the current climate, with significant cuts in public spending (Griffiths & Meinecke, 2014; Dodd & Moody, 2011; Jackson, 2013). Advocates of the SIB model highlight the potential for SIBs to transfer risk from the public to the private sector (Cabinet Office, 2013). In fact, most of the sectors and services, which are frequently targeted by SIBs, are also sectors where there is a high potential return for the Government by avoiding social expenditures (e.g. prison recidivism, homelessness) (Gustafsson-Wright & Osborn, 2020b).

2.1.1 Where It Happens More Often?

Ever since the first SIB was launched at Peterborough Prison in the U.K. in 2010 (Disley *et al.*, 2011), in either high, middle, and low-income countries, efforts have been underway to scale outcome-based financing more broadly by consolidating the instrument as a viable funding and investing solution for a range of intractable societal challenges (Gustafsson-Wright, 2020a).

According to reports by Brookings, as of July 2020, 194 impact bonds were contracted in 33 countries across six sectors, representing over \$421 million in upfront investment in social services and \$460 million in total outcome funding committed. The majority of deals concentrate in just a few countries; the number of contracts in the U.K. (69), the U.S. (26), the Netherlands (15), Portugal (13), and Australia (10) make up 69 percent of the total number of launched impact bonds. In terms of developing countries, India has the most contracted impact bonds (Gustafsson-Wright, 2020a; Gustafsson-Wright & Osborn, 2020b).

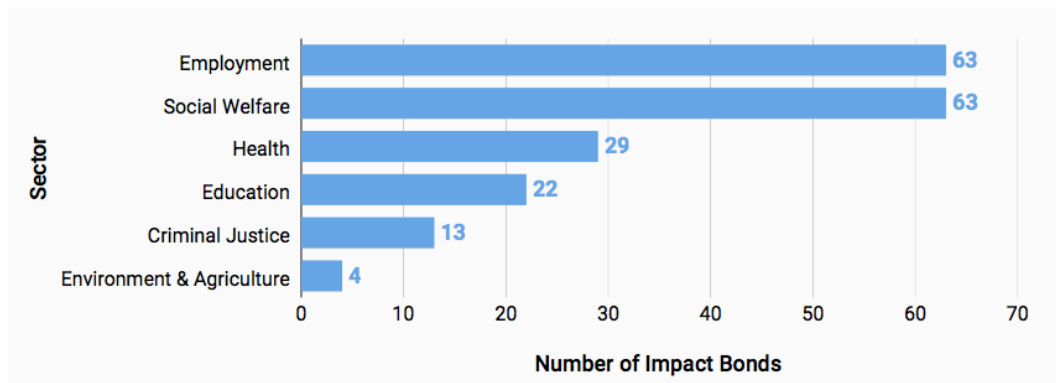


Figure 1: Impact Bonds per Sector (Gustafsson-Wright, 2020a)

SIBs focus not only on the prevention but also on the reduction of complex social issues such as improving skills, reducing youth unemployment or reducing loneliness amongst the elderly (Tan *et al.* 2015; Fraser *et al.*, 2016). The most impacted sectors include social welfare (63) and employment (63), as shown in Figure 1. In developing countries, health and employment have been the dominant sectors, with five impact bonds each (Gustafsson-Wright, 2020a).

When looking at the investment side, it is relevant to consider that SIBs cannot and do not aim to substitute traditional methods to finance social projects (Fox & Albertson, 2012; Liebman, 2011; Moynagh, 2010; Mulgan *et al.*, 2011). Impact bonds still remain a small proportion of overall government spending (Gustafsson-Wright, 2020a). In the United Kingdom, where most impact bonds were launched, according to estimates by Bridges Fund Management (2019), SIBs represent just GBP20 million of an overall budget of GBP730 billion.

Despite the rise in the adoption of the SIB mechanism worldwide, there are still questions regarding its development stage. There is no definite consensus in the literature if the innovative mechanism is still in the early adoption phase, and massive diffusion is yet to come, or if it simply constitutes a niche phenomenon (Arena *et al.*, 2016).

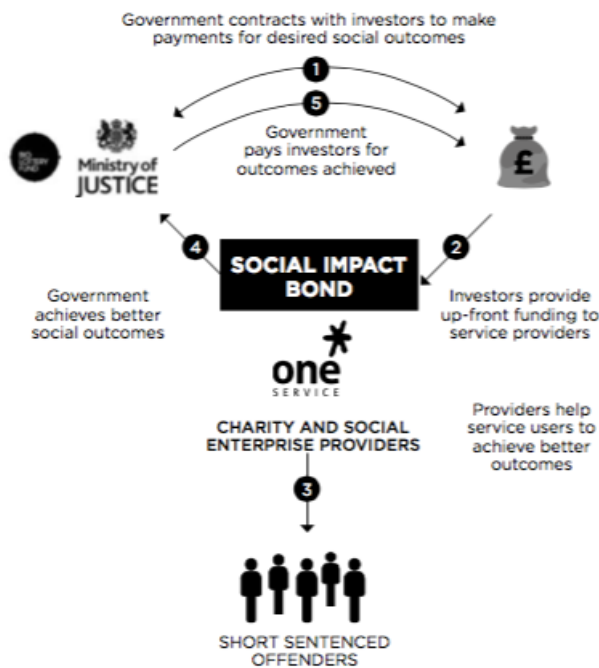
2.1.2 Traditional SIB Design

At its most fundamental, a SIB is an outcome-based contract between a public sector commissioner and investors (Social Finance, 2015a). Despite the term, Social Impact Bonds are not strictly bonds as the return is contingent, much like equity, on the out-turn of projects being financed (Fraser *et al.*, 2016; Warner, 2013). They can perhaps be better described as

hybrid instruments including elements of both equity and debt (Bolton & Savell, 2010; Liebman, 2011).

SIBs typically involve four different parties: the service providers who deliver the service; a government commissioner responsible for ensuring relevant services are made available to the intervention's target; external investors who cover upfront costs in exchange for a potential repayment plus return; and intermediaries, usually specialized in the SIB instrument or the intervention area, facilitating and managing the project planning and delivery (Fraser *et al.*, 2016).

In the SIB model, investors provide an upfront amount to finance a certain social intervention aimed to reach an outcome that is of interest to a government commissioner. If the provided



services achieve the pre-agreed social benefits, traditionally¹, the public commissioner pays back the nominal principal to the investor plus a return for the financial risks they took. If the outcomes are not met, investors lose their investment (Chiodo, 2015). In this way SIBs enable risk sharing between the two sectors. An example of a typical SIBs structure is presented in Figure 2.

Figure 2: The Peterborough Social Impact Bond (Social Finance, 2015a)

The prototypical SIB structure differs from standard payment-by-results models in three major aspects. First, the provider organization is not required to take on delivery risk; second, external capital is used to fund the service; and lastly, the investors are able to hold providers accountable (Social Finance, 2015a).

¹ A notable exception will be discussed in chapter 4. Case Study Findings, for the case of Portugal Social Innovation (PSI).

2.1.3 Main Drivers and Barriers for OBC and SIB Use

A better understanding of the driving forces and barriers to SIB development within certain contexts can provide valuable insights into the different paths to create a competitive market for outcomes (Gustafsson-Wright, 2020a).

Public sector leaders have historically tended to use mostly inputs and outputs as the measures of success. However, while such commissioning approaches might be considered more straightforward to measure, they still present several limitations (Social Finance, 2015a). First, traditional contracting by inputs or outputs risks providers cutting costs without accountability for the quality of service provided. Second, input or output oriented contracts tend to carefully specify activities at the time of contracting, leaving little room for adaptation and flexibility (Gustafsson-Wright *et al.*, 2020). And finally, management resources end up being overly spent on procedure compliance (Fox & Albertson, 2011).

On the other hand, outcome-based contracts, as is the case with SIBs, attempt to build a monitoring and evaluation culture that can provide a clearer definition of the aims of the programme, focused incentives to achieve them and higher freedom to personalize the intervention and foster innovation in delivery (Social Finance, 2015a; Leventhal, 2012; Jackson, 2013; Clark *et al.*, 2014; Gustafsson-Wright *et al.*, 2020). In addition, as the SIB model does not depend on the public sector for covering up-front costs of service provision, it allows, not only a reduction in government risk, but also an alleviation of the problem of public funding constraints for service providers, in particular third sector organizations (Fox & Albertson, 2011; Mulgan *et al.*, 2011; Warner, 2013; Chiodo, 2015).

Through the perspective of public sector reform, by transferring the risk from the state to external investors, interventions that might have been considered too experimental for traditional public funding also have the potential to benefit from this instrument (Fraser *et al.*, 2016). From a private sector perspective, the SIB model could attract private funding by allowing investors to align their financial interests with social impact through social entrepreneurship (Cohen, 2011; Liebman, 2011; Moore *et al.*, 2012). This may also allow investors to improve their perception and image in the eyes of the public (Barajas *et al.*, 2014). Proponents of the social entrepreneurship movement include these advantages as part of a broader trend of socially responsible business practices worldwide (Porter & Kramer, 2006; Bugg-Levine & Emerson, 2011).

Regarding its impact on the entire ecosystem, the SIB model is able to reshape the relationships between the stakeholders engaged in the commissioning of social services (Nicholls & Tomkinson, 2013; Palandijan & Hughes, 2014), and contribute to align the incentives of multiple stakeholders - government commissioners, third sector organizations, private organizations, financial intermediaries, and investors (Charlton *et al.*, 2013; Nicholls & Tomkinson, 2013; Chiodo, 2015). According to Gustafsson-Wright *et al.* (2020) report, improved stakeholder collaboration was highlighted as one of the most prominent ecosystem effects of the SIB model.

However, as identified by Fraser *et al.* (2016), in contrast to the generally positive literature on the public and private sector narrative, there are still cautions regarding the appropriateness of private sector values and practices to public services (Fraser *et al.*, 2016). It is argued by some that SIBs could represent an extension of neo-liberal logic in public policy (Warner, 2012, McHugh *et al.*, 2013, Sinclair *et al.*, 2014). This could potentially have a negative impact by incentivizing non-profit or philanthropic organizations to weight financial interests over their social impact missions (McHugh *et al.*, 2013). Lastly, as will be further discussed in section 2.2.2 Monitoring and Measurement Evaluation, there is still few empirical proof that contracting a service through a SIB model would in fact be more efficient and cost-saving than a traditional publicly provided service (Mullins *et al.*, 2011). The lack of rigorous comparison data and longitudinal evaluations still hinder the attribution of improvements in performance management and sustained impact to the SIB model (Gustafsson-Wright *et al.*, 2020).

For social investment and SIBs in particular to surpass their niche status there needs to be a higher focus on standardization of impact measurement, communication and risk assessment procedures (Antadze & Westley, 2012; Gustafsson-Wright, 2020a).

2.2 Outcome Measurement

A crucial stage in the development of SIBs is the selection of outcome metrics, thresholds and the measurement method used to evaluate success (Gustafsson-Wright *et al.*, 2020). As highlighted in the Social Finance (2015a) report, when it comes to outcome measurement, turning theory into practice can present a range of technical challenges.

Whilst there is a broad agreement in the literature regarding the potential benefits of the instrument as part of a broader shift to outcomes-focused measurement (Fraser *et al.*, 2016), SIBs are often reported as too expensive to design and implement (Gustafsson-Wright & Osborn, 2020a). The process of constructing outcome measures and metrics often requires a trade-off between trying to capture the overall objectives of the project, and measuring what it is possible to quantify (Social Finance, 2015a). Additionally, SIB negotiations usually present difficulties to agree on the appropriate measurement targets, time-scales for the outcome monitoring, and the role of each party in outcome monitoring (Tan *et al.*, 2015; Disley *et al.*, 2011). The underdevelopment of social impact measurement culture implies that SIB developers cannot fully rely on established methodologies and metrics. The lack of a dominant design hinders the development of SIBs, since suitable metrics and indicators must sometimes be created from scratch (Arena *et al.*, 2016).

A more complete understanding of the metrics, measurement process, and underlying incentives was identified as a key enabler for the development and scalability of the SIB model worldwide (Social Finance, 2015a). The following sections attempt to summarize existing research on these topics.

2.2.1 Metrics

For optimal OBC results, metrics should be meaningful (predictive of long-term improvements in participants), measurable in a specific context, and malleable (achievable within a reasonable time horizon) (Gustafsson-Wright *et al.*, 2015).

When developing a SIB, one or more outcome metrics will be established. The process of identifying the appropriate outcome is often challenging. While in some cases the objectives of the project are easily quantifiable (e.g., employment), in other instances, outcome selection can be difficult as the programme might have few measures that can be practically used or that directly reflect improvements in the target population (e.g. improve community cohesion) (Poister, 2003; Banke-Thomas *et al.*, 2015; Social Finance, 2015a). The more intangible and subjective the outcome, the more likely social benefits may be overlooked from an analysis of social impact (Polonsky & Grau, 2008, 2011).

The risk and return preferences of investors can also have a big impact on the outcome metric definition. A distinction should be made between "niche" pro-social investors willing to take higher risks and lower returns and those investors looking for social investments with higher yields and guaranteed returns (Fraser *et al.*, 2016). As an example, risk-averse SIB investors may push for repayments to also be partially tied to outputs, as opposed to more complex outcomes, as it is easier and more straightforward to measure (Gustafsson-Wright, 2020b).

When metrics include a combination of outputs and outcomes, this often occurs over the life of the project, with the outputs and possibly some outcomes representing intermediate measurements and proxies for the true outcome of interest (Cabinet Office, 2013; Gustafsson-Wright, 2020b). Proxy measures are predictors designed to track the short-term impact of the necessary policies to achieve the bottom-line result the program seeks in the long-term (Williams, 2019). In many cases, intermediate outcomes and proxy measures can also be used to track efficiencies and cost effectiveness of program activities as well as provide valuable information on "what works" and "for whom" (Gustafsson-Wright, 2020b; Segal, 2002). This can be particularly useful in SIBs by alleviating the trade-off between the cost-reduction of shorter time scales and a higher certainty of sustained impact indicated by larger measurement periods (Social Finance, 2015a). A traditional example is the use of an intermediate outcome indicator of employability as a predictor of long-term labour market performance (Martini *et al.*, 2017). When it comes to "end outcomes", these represent the result of agency programs (outputs and the associated intermediate outcome's impact) on bottom-line results produced (Segal, 2002). Outcome matrices detailing tangible outcome measures (listed by broader social interest area) have been developed by social impact investors to help social sector organizations to identify and measure outcomes (Social Finance, 2015a).

Finally, a crucial consideration when selecting outcomes is how to define success (Social Finance, 2015a). When it comes to the metrics related to payments, traditional metrics include: binary thresholds, outputs or outcomes achieved per capita, per the percentage of the total beneficiary population, or per the distance traveled by either an individual or the aggregate participant group (Gustafsson-Wright, 2020b). While a binary measure may provide a more absolute interpretation of success, it may also mean that improvements in individual's circumstances are not recognized. Overall, the end focus should be to ensure the root causes of the problems are addressed, this can be done through, for example, the use of improved or progressive outcomes (Social Finance, 2015a).

2.2.2 Monitoring and Measurement Evaluation

In the case of SIBs, payments can be based on the results of periodic evaluation and/or ongoing monitoring. While monitoring is the process of tracking what is happening day-to-day within an intervention (e.g. performance data; and auditing vendor activities in the field), evaluation is a more formalized assessment of the extent to which an intervention has met its objectives at designated points in time (e.g., beginning and end) (Brown & Potoski, 2003a; Gustafsson-Wright, 2020b).

In most cases, following commissioning, the outcomes linked to payment are measured so that the effectiveness of the approach can be judged. This typically involves carefully identifying the characteristics of the group of participants in the programme, and a source of comparison data – typically through a control group or baseline data (Cabinet Office, 2013; Social Finance, 2015a). The ultimate purpose of the outcome metric is to demonstrate the level of impact the intervention has (Gustafsson-Wright, 2020a). In the world of government evaluation, this is typically referred to as “additionality” which equates the net impact after allowances have been made for what would have happened had the intervention not been used (McPherson & McDonald, 2010; Social Finance, 2015a). There is a range of ways of demonstrating additionality of outcomes achieved through measurement. As a reference, Cabinet Office (2013) has developed a hierarchy of evidence quality, as shown in Figure 3.

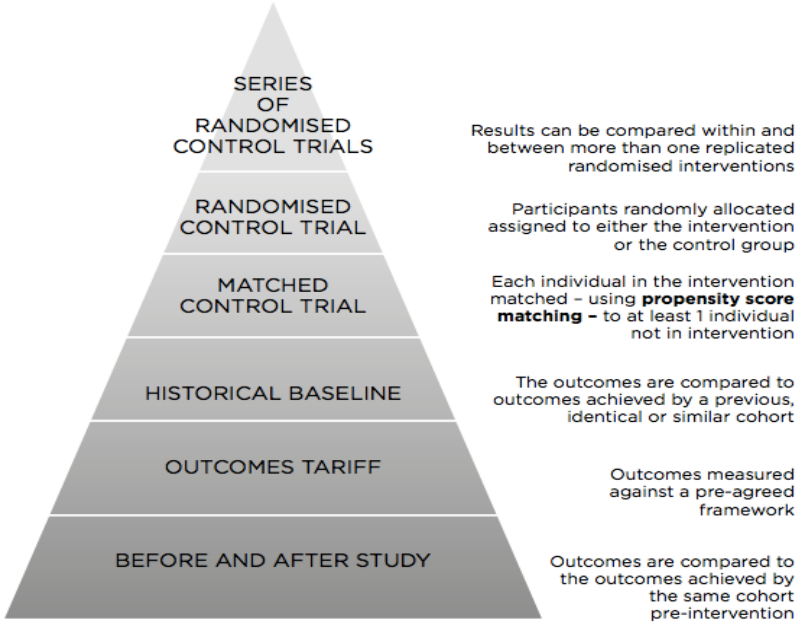


Figure 3: Hierarchy of evidence quality (Cabinet Office, 2013)

Looking at the Figure 3, at the bottom, we find the most straightforward to structure and manage, with the downside of being the most susceptible to perverse incentives and inaccurate impact assessments. At the top of the hierarchy we can find the most robust methods, but also the most expensive and time-consuming. As an example, recent empirical studies of U.K. SIBs suggested difficulties in developing robust counterfactual groups in most outcome measurement evaluations (Tan *et al.*, 2015; Fraser *et al.*, 2016). It should be noted however, that most SIBs target a small number of individuals, which usually results in high relative costs as many evaluation costs are fixed (Sturla *et al.*, 2018). An additional factor when evaluating the impacts of an intervention is the associated measurement period. The time scale selected should ideally leave enough room to make an assessment of whether positive outcomes have been sustained, while balancing the desire for a short repayment period (Social Finance, 2015a; Cabinet Office, 2013).

In practice, choosing the appropriate measurement system will be based on a range of considerations, such as cost, pragmatism and the quality of existing evidence available (Social Finance, 2015a). The negative impact of such factors might explain the lack of empirical evidence on results-based instruments, such as SIBs, delivering better outcomes than traditional reimbursement methods applied to public services (Lagarde *et al.*, 2013). This evidence gap is especially relevant for public commissioners since the lack of counterfactual evidence still raises questions surrounding the decrease in government's operational risks (Gustafsson-Wright *et al.*, 2020). Finally, it may also be beneficial for the parties, particularly for government bodies or commissioners, to evaluate the impact of the intervention on other outcomes than those related to payment (Berlin, 2016). This can be referred to as explanatory information (Segal, 2002). Often the product of process evaluation, these data provide context within a service area and give program managers an opportunity to give policymakers a better perspective on what happened (Cabinet Office, 2013).

2.3 Underlying Incentives

Once an outcome measure has been selected, the next stage is to consider the practicalities of measuring these outcomes by thinking through issues such as aligned incentives and the financial and reputational risks associated with outcomes-based contracting (Social Finance, 2015a).

2.3.1 Public OBC

The inherent diversity of public services hinders the alignment of incentives and motivations across policy makers, investors, service providers, service users and citizens. The co-creation of value in public services may include specific moral and ethical issues, and value could be perceived differently depending on stakeholder's position (Farr, 2015).

Overall, impact bond development can be significantly impacted by government experience, and interest in capacity building around the instrument itself (Gustafsson-Wright, 2020a). As in the case of Social Impact Bonds in the U.K., the central government has played a key role in commissioning OBC instruments, with a higher policy interest in particular sectors (Ecorys, 2019). When contracting with social business organizations, public authorities can finance these organizations directly, by subcontracting some social services or participating as a partner in carrying out a social programme (Bengo *et al.*, 2016). Governments should develop *evaluation capacity*, described as the capacity to evaluate the contractor's performance in terms of the service quality and impact generated by social programmes (Brown & Potoski, 2003b; Hyndman & McMahon, 2011; G8, 2014). Additionally, specific considerations such as longer measurement periods may allow commissioners to better assess and pay for the sustained impact of the intervention (Social Finance, 2015a; Cabinet Office, 2013).

Commissioners in a SIB should ideally focus on the social value of the interventions; however, their decision to engage may also weight in budgetary considerations. While some stakeholders believe governments embrace SIBs based on the social impact and learnings they can deliver, regardless of the potential realization of financial savings (Cabinet Office, 2013), evidence appears to suggest otherwise. The vast majority of policy areas that are targeted frequently by impact bonds, such as reductions in prison recidivism, homelessness, and preventable health conditions, imply high returns on government investment through avoided social expenditures (Fraser *et al.*, 2016). In fact, this is often presented as a key argument in support of impact bonds, in addition to social benefits and long-term outcomes achievement. Some impact bond projects even feature outcome metrics structured specifically around calculations of these avoided future costs (Rudd *et al.*, 2013; Gustafsson-Wright & Osborn, 2020a).

Lastly, despite the potential benefits of the SIB model on public expenditure, according to the cautionary narrative identified in the literature by Fraser *et al.* (2016), there are still inherent problems when calculating governmental risks in SIB contracts. First, the public sector often reports difficulties in measuring performance and accountability with outcomes based performance (Martikke & Moxham, 2010). Second, there is a reported lack of longitudinal evaluations to assess the sustainability of impacts beyond their completion (Gustafsson-Wright *et al.*, 2020). Finally, while the introduction of private financing appears to alleviate the public sector's financial risks, avoiding payment can hardly be perceived as a political "win" (Gustafsson-Wright *et al.*, 2020). In case of a potential failure in addressing targeted social problems, reputational risks ultimately end up falling within the scope of public commissioners (Fraser *et al.*, 2016).

2.3.2 Perverse Incentives

A cousin of transaction costs theory, principal-agent theory, argues that the root problems and risks in contractual situations begin with information asymmetries and goal incongruence between principals and agents (Miller, 1992; Brown & Potoski, 2003a). Despite the fact that, in social impact bonds, the financial and operational risk is shifted away from the service provider to an investor, these issues still persist (Gustafsson-Wright & Osborn, 2020b).

In general, one of the main difficulties associated with information asymmetries in services is that service quality is difficult to measure - and often much harder than measuring the quality of a good (Jensen & Stonecash, 2005). Service quality assessments may require subjective judgment rather than mere accumulation of data (Domberger & Jensen, 1997). Furthermore, the relationship between inputs, outputs and outcomes is in most cases uncertain, which could incentivize stakeholders to manipulate outcomes measures (Jackson, 2013; Liebman, 2011).

One way of managing these risks is through aligning incentives and motivations towards the outcomes by encouraging collaboration across all parties involved (Ng *et al.*, 2009). The specification of an outcome is particularly important as both service providers and users need to be actively motivated to achieve such outcomes (Farr, 2015). SIBs' outcome-based contracts intend to introduce greater accountability between commissioners and providers by setting clear expectations since the beginning (Stoesz, 2014). The stakeholders involved in the SIB planning process should consider the behavioral impact the initial outcome selection

might have on service providers, and look for ways of mitigating any conflicts or detrimental incentives (Social Finance, 2015a).

As the multitasking approach developed by Holmstrom & Milgrom (1991) suggests, in instances where an agent is required to perform a number of different tasks, effort will be allocated to the task that is most easily measured and therefore rewarded. Focusing on the implications of OBC for service provision, the measurement of outcomes can lead delivery organizations to focus resources on members of the target population whose outcomes determine payments because they are easier to measure or more likely to succeed. This is commonly referred to as "cherry picking" or "creaming" (Fox & Albertson, 2011; Disley & Rubin, 2014; McHugh *et al.*, 2013; Farr, 2015). On the flip side, service users with more complex needs may be given worse service provision as providers might be discouraged to focus effort on those unlikely to achieve further outcomes (ICF Consulting, 2019). This is commonly referred to as "parking" where "harder-to-help" target groups are neglected to some extent (Farr, 2015; Hudson *et al.*, 2010; Shutes & Taylor, 2014).

In the case of SIBs, when compared to traditional payment-by-results systems, there is less evidence of creaming and parking practices (Disley & Rubin 2014; DCLG 2014, 2015). This difference illustrates how the nature of contracting systems can affect the dynamics of service delivery (Farr, 2015). SIBs benefit, in most cases, from investors taking the risk for the delivery of outcomes, which reduces the risk of behaviors such as parking or cherry picking, because provider payments will not be tied to success of individual service users. The underlying philosophy of this approach is that the resources of service providers should be focused on the impact of the intervention and not on optimizing payment (Social Finance, 2015a).

Additional methods for encouraging service providers to work equally with all beneficiaries in the SIB target group, require careful metric considerations. First, the selection of the thresholds on which payments are based is critical to both the successful achievement of outcomes and to the credibility of the results. While low thresholds might favor investor risk, it also decreases the financial incentive for obtaining outcome evidences beyond the required threshold (Maze, 2020a). On the other hand, if thresholds are set too high, this could lead to perverse incentives such as cherry picking and parking, unless properly mitigated by SIB design (Gustafsson-Wright & Osborn, 2020b).

Second, impact bond designers can also account for perverse incentives by including metrics and bonuses which link repayment to the performance of the entire targeted population; the level of progress toward the outcome; or the weight of specific beneficiary characteristics (Gustafsson-Wright, 2020b; Disley *et al.*, 2015). For skills pilots in Canada, for example, the repayment calculation weighted different participant characteristics such as gender and baseline literacy level (Gustafsson-Wright & Osborn, 2020b).

Finally, if the target population and outcome metrics are not carefully specified in the SIB contract, the potential for perverse incentives such as creaming tends to increase (Fraser *et al.*, 2016). In fact, social sector organizations often highlight a low involvement from the public authority in the specification of services and associated risks (Chiodo, 2015). An accurate profiling of service-users who will ultimately benefit from the programme can have an important preventive impact on stakeholder incentives. The design of the selection mechanism should include an independent referral mechanism, ideally with clear objective criteria so that providers do not have undue influence on the service users engaged in the programme (Social Finance, 2015a).

2.3.3 Financial and Reputational Risk

When drafting contracts, exchange rules specify what the buyer and seller need to do to execute the exchange. Exchange rules have an important consequence on who bears contract's risks (Lamothe & Lamothe, 2012; Whitford & Miller, 2007). The most common exchange rule is compensation, the price for the product or service. The two most common types of payment rules are fixed-price contracts, which set compensation on the seller's outputs or outcomes; and cost-reimbursement contracts, which set compensation on inputs. Under a cost-reimbursement approach, the payer has the incentive to clearly specify service characteristics, since they will ultimately drive the price and prevent the seller from adding unnecessary features to the service (Brown *et al.*, 2016).

One suggested approach to manage the party's incentives and risks is to increase funding "conditionality" - the obligation of contracted providers to achieve outcomes as a condition of funding (Shutes & Taylor, 2014). However, while some research highlights new opportunities in conditionality of funding (Miller, 2013), others suggest a negative effect on provider

diversity. Particularly, some research suggests an adverse effect on smaller, voluntary sector organizations with additional costs and risks (Shutes & Taylor, 2014; NCVO, 2014). Overall, as outcomes-based contracting is still an historically recent approach, there is long-term risk associated with the unpredictability of future outcome payments, which could lead to higher costs in the future (Audit Commission, 2012; Farr, 2015).

In the specific case of SIBs, there is also a reported uncertainty regarding the calculation of risk amongst the different parties (Fraser *et al.*, 2016). SIB financial payments can either be directly tied to outputs and outcomes (fixed priced contract) or indirectly, through reimbursement set on inputs, but still contingent on the achievement of pre-defined outcomes². In theory, since private investors are the ones providing the upfront capital, which is paid back only if target results are actually met, they usually bear the whole financial risk (Dagher, 2013; Lehner & Nicholls, 2014; Loder *et al.*, 2011; Nicholls & Tomkinson, 2013; Warner, 2013; Arena *et al.*, 2016). In some cases, financial risks have reportedly led projects to “trade” hard to achieve outcomes (e.g. higher-level qualifications) with softer outcomes (e.g. improvements in attitudes and school attendance) (Gustafsson-Wright, 2020b). It should be noted however, that there is a downside associated with a stronger focus on financial and budgetary impacts as it can potentially distract stakeholders from the ultimate goal of maximizing impacts on beneficiary populations (Golden *et al.*, 2017).

Financial and reputational risk, as well as the associated perverse incentives, can be accounted for by establishing contract rules (e.g. financial rewards and default risks) to motivate "consummate" or "win-win" behavior amongst the involved parties (Brown *et al.*, 2016). According to the reputations proposition, exchange relationships that impact the reputations of the buyer and the seller increase incentives for cooperative behavior; which will in turn increase gains in future contracts (Brown & Potoski, 2004). In the case of SIBs, some evidence suggests that positive performance and outcome achievement from providers has resulted in increased funding and the ability to scale the intervention (Gustafsson-Wright *et al.*, 2020). Early collaboration in contract negotiations and project planning can improve mutual understanding, as procedures become more predictable and verifiable (Brown *et al.*, 2016). Throughout the project, information exchanges among partners can increase trust, which will lower the legal and transactional costs in case of disputes (Lee & Lim, 2005).

² The latter will be further explored throughout this dissertation, within the Portuguese context.

Through an aligned focus on outcomes, SIBs can allow stakeholders to better react to information and adapt the intervention in real-time (Gustafsson-Wright *et al.*, 2020; Social Finance, 2015a).

This collaborative approach can be potentiated by the presence of a third party (Cooper, 2003). In the case of SIBs, these third parties are usually positioned as managerial intermediaries and are perceived as crucial in the implementation and payment processes (Bafford, 2012; Haffar, 2014). They usually add the most value by providing a link between the program actors, whilst contributing with specialized skills and solutions to delivery challenges (Callanan & Law, 2012; Gustafsson-Wright *et al.*, 2020).

In conclusion, it is suggested that the attribution of risk between investors and service providers is still highly specific to the context of each intervention, making it difficult to standardize or even adapt (Pauly & Swanson, 2013; Gustafsson-Wright *et al.*, 2020). When developing a financial plan for a specific project, for example, a challenging consideration is the appropriate measurement period. In financial terms, interventions requiring longer measurement periods will increase the risk for investors and providers. On the other hand, projects with shorter time-periods will allow investors to quickly recycle their investments, but can ultimately result in further operational and administrative issues for the project (Social Finance, 2015a). An additional challenge might be restrictions placed on the use of outcome funding. For example, in Portugal, where most of the outcome funding is channeled from the European Social Fund through Portugal Social Innovation, it is currently not possible to make direct returns on outcomes - instead investors receive an indirect return through a tax break (Gustafsson-Wright, 2020a). The Portuguese case will be presented in the following section.

2.4 Portuguese SIBs

The emergence of Social Finance and social investment in Portugal had its major breakthroughs, first, in 2013, with the initiative of the European Union - Portugal Social Innovation (PSI), and in 2014, with the launch of the Portuguese Social investment Taskforce, which included entities such as Maze, Social Finance U.K. and Calouste Gulbenkian Foundation (Social Finance, 2015b).

With a central role in the Portuguese social investment ecosystem, PSI is a public sector Mission Structure financed by the European Social Fund within the scope of the Portugal 2020 Partnership Agreement. A Mission Structure is created by the council of ministers as a public administration unit to accomplish a specific mission. The mission in this case, with an allocated budget of 150 million euros, is to stimulate innovative solutions and intervention models, through four financing tools, including social impact bonds, in order to provide answers for priority social problems and complement existing responses in several social policy areas such as health, education, justice and gender equality (Portugal Social Innovation, n.d.). The other financial tools include two grant mechanisms and an equity investment financial instrument.

As the entity responsible for the outcome payments, to date, PSI has partnered in 12 SIBs, with over 3 million euros of committed outcome payments out of a total outcome fund of 15 million (Gustafsson-Wright, 2020a; Portugal Social Innovation, n.d.). The only SIB to be launched outside PSI scope was the first ever Portuguese SIB - Code Academy Junior, in 2015. In July 2017, the three first SIBs in partnership with PSI were launched, with projects currently entering the final stages in 2020. The main partners included Calouste Gulbenkian Foundation as an investor and Maze (known as Social Investment Laboratory at the time) as the intermediary.

In conclusion, despite being the fourth country worldwide with the most launched SIBs, there is still an apparent lack of academic research on SIB development in Portugal, especially research focused on measurement and underlying incentives (Gustafsson-Wright, 2020a; Gustafsson-Wright & Osborn, 2020b).

3. METHODOLOGY

The methodology section provides an overview of the procedures that were applied to answer the research questions, as well as the broader philosophy and strategy used for the research.

3.1 Research Philosophy

Business and management research philosophies are scattered along a multidimensional set of continua (Niglas, 2010). The first ideological dimension positions research between two opposing extremes - objectivist and subjectivist. Along this continua, this thesis positions further to the subjectivist side as the interpretation of cases embraces social constructionism, which puts forward that reality is constructed through social interaction (Bryman & Bell, 2011). This dissertation follows a constructionist view on reality, which reinforces the need to study a situation in detail, including historical, geographical and socio-cultural contexts. Unlike an objectivist researcher who seeks to uncover universal facts and laws, a subjectivist researcher is interested in different narratives and motives that can help to account for different social realities of different social actors (Saunders *et al.*, 2016). The second ideological dimension also has two opposing poles in "sociology of regulation" and "sociology of radical change" (Burrell & Morgan, 1979). This thesis' philosophy is positioned within the regulation perspective as it seeks to suggest how organizational affairs may be improved within the current framework rather than radically challenging the current position (Saunders *et al.*, 2016). According to the 4 paradigms matrix developed by Burrell & Morgan (1979), a subjectivist research within a regulation perspective represents the interpretive paradigm. One of objectives when working within an interpretive paradigm is to attempt to uncover reasons behind failures or irrationalities, which might not be as apparent to the actors involved (Saunders *et al.*, 2016).

3.2 Research Strategy and Design

Before considering the research approach and design, it is important to classify the research purpose for this dissertation as partially exploratory, explanatory and evaluative (Saunders *et al.*, 2016).

In terms of the methodological approach, this dissertation uses predominantly an inductive approach. As suggested by Saunders *et al.* (2016), with research into a topic that is relatively

recent, an inductive approach may be more appropriate, particularly by identifying themes and patterns and locating them in existing Literature, in order to refine, extend or generate theory (Ridder *et al.*, 2014). A deductive approach, to a lesser extent, was also used in defining the initial direction for the research. Particular existing theory was used to provide a focus for the research and to limit its scope (Tashakkori & Teddlie, 2010).

Through the respective associations to the interpretive research philosophy assumptions and the inductive research approach (Saunders *et al.*, 2016), a qualitative research design was the most coherent method for this thesis. In this case, the qualitative research design used more than one qualitative data collection technique defining it as a multi-method qualitative study, which facilitates a combination of exploratory and evaluative research (Saunders *et al.*, 2016).

In order to analyze the study's focal theme, a case study approach was selected. This approach has the advantage to provide a detailed explanation of behavior, the underlying assumptions and subjective constructs of decisions (Yin, 2014). The selected cases have the objective to provide an in-depth understanding of Portuguese social impact bonds and the surrounding social ecosystem, with a particular focus on measurement and underlying incentives. With this in mind, both semi-structured in-depth interviews and secondary data sources were used to collect qualitative data regarding each case study. The use of secondary data was chosen, not only as it complements the primary data in terms of content, but also because it adds significant validation implications (Saunders *et al.*, 2016).

3.3 Case Studies

In this section a short summary of the cases will be presented in Table 1, as well as the basis for the selection of the six cases.

Name of project	ACJunior	ACB SIB	PF SIB	FFWD SIB	CQC SIB	T2W
Social issue	School performance	Youth unemployment	Children in care	Youth unemployment	Support to informal caregivers	Youth unemployment
Start Date	2015	2017	2017	2017	2019	2019
Intervention model	Computer programming basics to primary school students in a classroom setting	14-week coding bootcamps for unemployed young adults	6-week intensive support to families followed by 12-month ongoing support	10-month mentoring, coaching and job readiness program	7-month training for multidisciplinary teams to create a specialized response to informal caregivers needs	6-month of individualized coaching and group training to increase VET students' transversal soft-skills and promote job readiness
Outcome metric(s)	Improvement in Logical Reasoning (90% weight), Improvement in Maths and Portuguese (10% weight), vs. control group	Entry into employment of 11 students (50% of cohort)	Care placement avoided and confirmed by court order of 12 CYPR (50% of cohort)	Entry into employment of 12 (40%) 6-month job maintenance of 6 (20%)	5% improvement in average score of the overall quality of life of the caregivers	Entry into employment of 65 VET students within 3-months of completing their training (Note that there are two outputs also linked to payment)
Impact Evaluation	Controlled experiment Treatment vs. Control-group	Does not apply (Outcome verified through administrative evidence)	Does not apply (Outcome verified through administrative evidence)	Does not apply (Outcome verified through administrative evidence)	Distance-travelled	Does not apply (Outcome verified through administrative evidence)
Investment size	125,000€	723,500€	433,276€	387,848€	239,660€	<i>Does not apply</i>
Beneficiaries	65 students	180 young adults	216 children at risk	150 young adults	240 caregivers	200 final-year trainees from an Apprenticeship Course
Location	Lisbon	Fundão (centre of Portugal)	Porto	Porto	Porto	Viseu, Amadora and Seixal
Status	Completed	Ongoing	Ongoing	Ongoing	Ongoing	Ongoing

Table 1: Cases summary

The selection of the cases was based on the objective to explore OBC practices, specifically through Social Impact Bonds, within the Portuguese context. With this goal in mind, the cases selected were defined by two factors: first, the access to the organizations and stakeholders involved in the projects; and second, the representativeness of the overall Portuguese SIB ecosystem. Ultimately, six cases were selected.

In the first case - SIB Code Academy Junior (ACJunior) - was selected, as it was the first ever SIB to be launched in Portugal and because, as it stands, it is the only Portuguese SIB where Portugal Social Innovation (PSI) is not the outcome payer. The three following cases - SIB Academia de Código Bootcamps (ACB), SIB Projeto Família (PF) and SIB Faz-te Forward (FFWD) - were selected as they constituted the first wave of SIB launches in partnership with PSI, and are currently in their respective final stages. Then, while still under PSI scope, SIB Cuidar de Quem Cuida (CQC) was selected as it focuses on a recent case. Finally, the case of Training2Work (T2W)³ was selected for the opportunity to directly compare two different outcome-based financing mechanisms, since the project shared a common service provider with SIB Faz-te Forward (FFWD).

³ *Note that Training2Work was the only Portuguese case of non-SIB social service outcomes-based commissioning identified throughout this thesis' research.*

3.4 The Data Collection Process

As mentioned in the previous section, data collection focused on primary and secondary data. Respectively, throughout this section, the two data collection methods will be presented.

3.4.1 Primary Data

From the selected cases, the interviewees were chosen based on theoretical sampling, namely through their ability to illuminate the research questions (Yin, 2011). Additionally, it was important to consider the perspective of different stakeholders within the same case. A figure summarizing the most relevant stakeholders which were contacted from each selected case can be found below (Figure 4). The color note highlights the organizations where it was possible to schedule an interview. The contacts were established by email, based on the identification of relevant actors within each organization.

Program	Policy Field	Service Provider	Outcome Payer	Investor	Third Party
SIB - AC Junior	Education	Academia de Código (ACB) - André Machado	Municipality of Lisbon	Calouste Gulbenkian Foundation - Francisco Palmares	Maze - Margarida Anselmo
SIB - ACB	Youth Employment/Training	Academia de Código (ACB) - André Machado	Portugal Social Innovation (GOV) - Vera Barracho	ASSOP - Ricardo Gonçalves, Gulbenkian - Francisco Palmares	Maze - Margarida Anselmo
SIB - PF	Children in Care	Movimento Defesa da Vida - Carmelita Dinis, Mariana Delgado	Portugal Social Innovation (GOV) - Vera Barracho	Calouste Gulbenkian Foundation - Francisco Palmares	Maze - Margarida Anselmo
SIB - FFWD	Youth Employment/Training (NEETs)	TESE (Association for Development) - Joana Guimarães, Vanessa Mendes	Portugal Social Innovation (GOV) - Vera Barracho	Calouste Gulbenkian Foundation - Francisco Palmares	Maze - Margarida Anselmo
SIB - CQC	Informal Caregivers	CASTIIS	Portugal Social Innovation (GOV) - Vera Barracho	Calouste Gulbenkian Foundation - Francisco Palmares	Maze - Margarida Anselmo
T2W	Youth Employment/Training (NEETs)	TESE (Association for Development) - Joana Guimarães, Vanessa Mendes	IEFP (Institute of Employment and Vocational Training (Public))	N/A	Maze - Margarida Anselmo

Figure 4: Interviewees summary

There were six interviews in total, which were conducted on the basis of the interview guide (Appendix 9.1), with durations from forty minutes to slightly over one hour. The interview guide was slightly adapted from interview to interview based, not only on the specific profile of the interviewee, but also the feedback and content from the previous interviews. There were a total of 8 interviewees representing six different organizations: one public entity; two service providers; two investors; and one specialized third party. The interviews identified

above were performed through real-time electronic interviews in Microsoft teams, mainly due to COVID-19 restrictions and precautions. All the interviews were video recorded and manually transcribed with the permission of interviewees.

The interviewees, the respective organizations and the role of each organization in the projects are identified below, ordered by interview date, from first until the most recent. A summary of each organization represented in the interviews can be seen in *Appendix 9.2*.

3.4.1.1 Interviewees

- Margarida Anselmo (MAZE Impact)
- Francisco Palmares (Calouste Gulbenkian Foundation)
- Ricardo Gonçalves (Municipality of Fundão, in representation of ASSOP)
- Vera Barracho (Portugal Social Innovation)
- Carmelita Dinis and Mariana Mira Delgado (Movimento Defesa da Vida)
- Joana Guimarães and Vanessa Mendes (TESE - Association for Development)

3.4.2 Secondary Data

From a review of existing literature on outcomes-based commissioning and social impact bonds, the key themes were established as the theoretical framework, which shaped the structure of the interview guide and the findings. Definitions and insights from numerous academics and professionals regarding these themes have been scrutinized and contrasted to assure validity and reliability.

Secondly, the most significant sources of secondary data regarding the analyzed case studies consisted of select official reports and websites. For four of the six cases, intermediary reports from Maze were key secondary sources to collect qualitative, but mostly quantitative data regarding each project analyzed (Maze, 2018, 2020a, 2020b, 2020c). For the other two cases, CQC and T2W, a combination of official documents from the European Commission and the official organizational websites were used to collect both quantitative and qualitative information.

3.5 Data Quality and Ethical Considerations

With a total of six semi-structured interviews, it is important to recognize the limitations of the case study approach when providing generalizations regarding the research topics. However, despite the small sample size, the representativeness of the sample provides, to some extent an alternative quality criteria (Saunders *et al.*, 2016). The sample of interviews is quite representative of the majority of stakeholders in each case study (in particular), and the key stakeholders of the Portuguese SIB ecosystem (in general). It should be also noted that, as the primary goal of this dissertation focuses on Portuguese SIBs, 5 out of a total of 13 SIBs launched in Portugal are being considered.

Additionally, two validation techniques were used: triangulation and member validation (Saunders *et al.*, 2016). First, by introducing a multi-method research, two different sources of data were used to confirm the validity, credibility and authenticity of the research data, analysis and interpretation. Second, after the interviews occurred, data provided by the participants was sent back, to allow them to confirm its accuracy, by permitting them to comment or correct. This was especially important since the interviews were performed in Portuguese and the transcript was translated to English.

With regards to ethical considerations, as suggested by Saunders *et al.*, (2016), the interview guide and communications with interviewees included requests to record the interview and transcribe, with its use strictly restricted to the purpose of this dissertation (Appendix 9.1). The provided data was managed with the desired level of integrity and confidentiality. No source, individual or organization was identified without the permission of its creator. Additionally, a copy of the results was provided to all participants.

As a final note, the researcher is naturally influenced by his own values and perceptions, as well as interviewee's values (Bryman & Bell, 2011). The findings are ultimately influenced by the understanding and interpretation of the researcher.

4. CASE STUDY FINDINGS

The purpose of this chapter is to highlight the most significant findings in the research, in an attempt to provide an answer for the respective research questions. The conceptual structure of the chapter was defined based on the overlap of the most relevant themes from the Literature and the most relevant thematic insights from the interviews. In order to provide a more "fluid" reading, quotes from the interviewees were added throughout the sections, where only the organization of the interviewee was identified.

4.1 Measurement

In this section, the results will attempt to answer the following research questions:

- 1) *How are SIBs' outcome-based contracts structured?*
- 2) *Which metrics and monitoring processes do they use?*

4.1.1 Metrics and Targets

"(..) the definition of the metrics during the initial stages of the application process is predominantly investor and provider-led." (MAZE)

In terms of the metrics' definition process, for the cases of the four SIBs under PSI scope (PSI-SIBs⁴), the initial stages of SIB design are very much provider and investor-led. During the application process of each SIB, the investors and providers develop a proposal for the outcome metrics and thresholds and come up with a proposal for the sources of evidence to be used as proof that the targets were met. This initial proposal is then discussed and validated by all additional partners, which must include, according to PSI's commissioning requirements, a public sector entity that represents the theme policy area where the SIB is being structured. For example, in the cases of Academia de Código Bootcamps (ACB) and Faz-te Forward (FFWD), it is the National Institute of employment and training (IEFP), for Projeto Família (PF) the Institute of Social Security, and for Cuidar de Quem Cuida (CQC) the Central Administration of Healthcare (ACSS). Finally, the majority investor submits the final proposal for approval to the outcome payer, in these four cases - Portugal Social

⁴ Throughout the chapter, the use of the term "PSI-SIBs" will refer to the four SIBs in analysis which were launched within PSI scope, that is, following the commissioning practices and requirements of PSI: ACB SIB, FFWD SIB, PF SIB, and CQC SIB.

Innovation (PSI). In the process of approving or not approving the proposal, the outcome payer might also push for some adjustments in any dimension of the metrics and the targets.

"[for ACB] This was an absolute value that was contracted, it was not even percentage, even though the logic was in percentage, (...) which had implications because some bootcamps had less students than was planned." (MAZE)

"(...), as we had some unexpected practicalities and delays in the beginning and after the first outcome, we ended up changing the groups from 24 to 20 children, but the target stayed at 12, which meant an increase in risk, since the threshold moved up to 60%." (MDV - PF)

For the four PSI SIBs, the outcome targets were defined per cohort (e.g. bootcamps, administrative groups), with the total budget of each project being split between the different outcomes. The targets that were defined for each cohort were either in percentage or in absolute value like the case of Academia de Código Bootcamps (ACB) and Projeto Família (PF). For these two cases, even though the logic was percentage, in practice, the contracted outcome targets were in absolute value, which resulted in risk implications as some groups had less students than was initially planned - this resulted in increased risk for the outcomes to be achieved since the threshold effectively moved up in percentage.

"There is a difference between what should happen in theory and what ends up happening in practice. The metrics were defined mostly based on the previous experience from service providers." (Calouste Gulbenkian Foundation)

"The Faz-te Forward project was already being developed by TESE in Lisbon with similar methodologies, which made it easier to define the metrics based on the previous 5 editions." (TESE - FFWD)

The actual definition of the targets and metrics for these four cases were mainly defined based on historical data from the organization's previous projects, as well as guidelines in terms of public policy priorities, adapted to SIB requirements. However, even though the priorities signaled by the public entity and PSI were clear and used as a reference, these constituted more general guidelines as the details were very much defined by the partnership. In the cases of FFWD, PF and CQC, since SIB financing was used to replicate the intervention in different geographies, the historical baseline came from previous successful experiences with similar

methodologies in different cities or regions. In the case of ACB, the employment integration threshold was developed from their historical data in the region of Fundão, where they had previously operated.

"(...) different objectives end up affecting the way the metrics are defined (...)." (Calouste Gulbenkian Foundation)

While in the previous four SIBs the metrics were more useful in terms of validation of the intervention, in the case of Academia de Código Junior (ACJunior), the only Portuguese SIB developed outside PSI scope, the metrics were more informative regarding the potential of the intervention. As there were no previous indicators from the provider to rely on, the outcome metrics were based on literature and similar experiences abroad by other organizations. In fact, one of the biggest takeaways from the project's final report developed by Maze (2018) were the insights regarding the most adequate outcomes indicators and assessment methods for future project designs. For example, the one-year time-period of the intervention was deemed insufficient to accurately assess student's performance improvements over time, with a recommendation of two years for future projects.

Looking at the remaining case, as opposed to the evaluated SIBs, the outcome metrics and targets for Joint Action Plan pilot Training2Work (T2W) were mainly based on data provided by public entity IEFPP. The providers checked the values of the Institute in terms of integration rate, number of participants and tutor-participant ratio and added more ambitious targets on top.

"[thresholds for the 4 PSI SIBs] Never too high." (PSI)

"if we look from today's perspective, specially considering their [ACB] historical success rate, the thresholds aren't capturing the full potential of the intervention." (PSI)

Despite being based on internal historical data, the outcome thresholds defined for the four PSI SIBs also weighted specific characteristics and requirements of the SIB model. When asked whether the defined thresholds were considered too high or too low, the overall response of the parties indicated that the contracted targets were relatively conservative. A

quite illustrative example is the case of ACB where despite the historical employment integration being around 90%, the contracted outcome threshold was slightly above 50%⁵.

"(...) considering the level of rigor of the evidence that is required to prove the outcome was met, the thresholds were not unreasonable, because it's very bounding evidence. " (MAZE)

" Code Academy Bootcamps had a certain percentage of employment integration, and the metric defined for the SIB was slightly more conservative in comparison, to attempt to guarantee that, in this initial intervention, the objectives were met. " (Gulbenkian)

"It was a conscious decision from all the parties involved to set a more conservative threshold." (MDV - PF)

The gap between provider's historical performance and the approved SIB thresholds led to discussions regarding whether or not contracted targets lacked ambition. However, it was ultimately described as a conscious decision as there were different motivations for the parties to lower contracted thresholds. First, lower thresholds naturally controlled for investor risk as they attempt to guarantee all outcomes are met; second, the SIB instrument was seen as an opportunity to re-think or expand the interventions, which added additional challenges and risks for service provision and subsequent outcomes (e.g. MDV - PF); and third, as will be further explained in section 4.2.4. *Expense Reporting and Investor Risk*, considering the level of administrative resources and reporting requirements of demonstrating the outcomes and respective evidences, the contracted thresholds were considered reasonable (e.g. ACB and FFWD).

"So far, six out of nine for Code Academy, and four out of nine for Faz-te Forward have been declared achieved by the outcome payer. But they were all quite challenging to achieve." (MAZE)

Considering the results from the three initial PSI SIBs⁶, which were launched within the PSI's first wave of impact bonds and are currently entering the respective final stages, all the pre-established outcomes have been met and demonstrated, but perhaps not always as comfortably as initially expected by stakeholders.

⁵ the contracted outcome was the entry into employment of 11 students out of cohorts of 18 to 20 students (Maze, 2020a).

⁶ Throughout the chapter, the use of the term "initial PSI-SIBs" refers to the three SIBs in analysis which were launched within the first PSI wave: ACB SIB, FFWD SIB, and PF SIB.

4.1.2 Monitoring

For the four PSI SIBs, the performance management and outcome monitoring was performed predominantly by Maze as the specialized intermediary⁷. Maze's responsibilities included implementing processes to regularly manage and monitor performance, generally through continuous monitoring via AidHound platform⁸, weekly situation reports with training and project management teams, monthly meetings with partners, and variable interactions related with reimbursement requests (Maze, 2020a, 2020b, 2020c).

"(...) other partners on the SIB could also have been the monitoring entity, however, in this case, Maze is the entity with the most ability, experience and interest in the development of SIBs." (Calouste Gulbenkian Foundation)

In terms of the attribution of outcome monitoring and performance management responsibilities, considering the specific contexts of these SIBs, Maze was identified as a suitable choice, not only due to the experience within the Portuguese social innovation ecosystem, especially with SIBs (e.g. performance management in ACJunior SIB), but also due to the reputation and relationships with several national and international social organizations, from social providers and investors to public entities such as mission structure PSI and sectoral public institutes.

"One of the big advantages of this instrument, through a focus on outcomes and results, is that we are continuously monitoring and not only waiting for a final evaluation" (MDV - PF)

"This allowed us to integrate in our methodology and operational model insights from the project." (MDV - PF)

"For us to move in that direction [public integration], it is key the monitorization of the outcomes and achieved results, it is something we give a lot of focus, so that impacts are evidenced and not assumed." (ASSOP)

From the provider side, the focus on monitoring led by Maze was perceived not only as helpful but also as advantageous in strategic terms. The continuous oversight associated with

⁷ Throughout the chapter, the use of the term "intermediary" refers to performance management, not to be confused with financial intermediary.

⁸ AidHound is a data management platform designed for social work. This platform is used to register, organize, and share data between the service providers and MAZE.

the SIB instrument allowed providers and partners to continuously monitor and adjust the intervention without moving away from the essential features and contractual requirements pre-established with all stakeholders. Additionally, the monitoring frequency and rigor also empower provider organizations by providing clarity and a more objective assessment of the intervention design and the actual impacts generated, which allows, not only to refine their impact models, but also to signal their quality to the market, either public or private. In the ACB case, for example, the continuous information loops regarding the bootcamps, involving different stakeholders, from the participants to the recruiting companies, was considered key in assuring and maintaining sustained employability (the ultimate desired outcome).

*"(...) in this case, Maze's main role is to 'agilize' the evidence flow to the [European] Commission."
(TESE - T2W)*

Regarding the Training2Work Joint Action Plan (JAP) pilot, monitoring responsibilities were mainly performed by the provider (TESE). While the participation in the activities was monitored monthly, the indicators for market integration were evaluated periodically since the integration evidence could be presented at any time within a three-month period. Maze also played a role but was mainly focused on verifying the evidence before being sent to the European Commission for final verification and validation, which only takes place in two reporting periods, at an intermediate point (after nine months) and end of the project (after eighteen months).

*"There is also a quite intense monitoring of outputs or activities from the outcome payer side."
(MAZE)*

"In order to apply for the SIB it was necessary to define quantitative outcome targets but we also analyze a lot of other data qualitatively." (MDV - PF)

While the payments are contingent on the achievement and demonstration of target outcomes, it is also necessary to showcase evidence of outputs. Throughout the different stages of the project the outcome payer wants to see some evidence of the intervention's delivery, and, in some situations, for providers and partners to explain why changes were made to the initial features. Additionally, providers highlighted the importance of internally monitoring additional quantitative and qualitative data beyond SIB requirements. In the FFWD project, TESE evaluated several variables such as increase in soft skills and employment

competences, increase in job applications, check whether the market integration is made within the participant's interest areas, and the updated NEET's⁹ classification (regardless of the path). For T2W, the organization monitored similar indicators with the addition of tutor evaluations. In the case of PF, weekly records of the contacts with families via AidHound included several variables like engagement of the families, reasons for risk flagging, initial and final metrics associated with the children participating in the intervention, as well as the analysis of the overall perceived risk of each family.

4.1.3 Measurement and Impact Evaluation

"The presence of an external entity, even though still part of the partner consortium, adds rigor and neutrality to the evaluation." (TESE - FFWD)

"(...) the [outcome] thresholds were set up under the assumption that that would be better than the status quo, which is validated to a certain extent by the approval of the public outcome payer." (MAZE)

"(...) we currently do not have the best proxy for the counterfactual - what would have happened without the intervention." (MAZE)

One of the main goals of the SIB projects in analysis is the ability to inform public policy. A proper impact evaluation is key to legitimize the interventions' ability to generate credible solutions. In the case of the four PSI SIBs, the development of the evaluation framework was a process predominantly led by MAZE as an external party and approved by PSI. However, interviewees reported some limitations. Overall, the main limitation was the quality of the counterfactual proxy. Even though the thresholds were supported by the public sector's representative and approved by PSI as the outcome payer, robust and updated data on what was the baseline with the same profile of beneficiaries that did not participate in the interventions throughout the country was still lacking. The case of PF is quite an illustrative example of such struggles. An independent impact evaluation using a matching approach was planned in collaboration with Nova University. However, stakeholders are facing difficulties launching the impact evaluation due to reluctance from the social security institute in sharing the necessary data, as it is a very sensitive social area (CYPR¹⁰). Project partners are currently

⁹ Not in Education, Employment or Training

¹⁰ Children and Young People at Risk

trying to overcome these issues by establishing a partnership with the National Commission for Children's Rights (CNPDPCCJ). It is worth highlighting that this independent evaluation is being pushed for by the provider, intermediary, and investors and not by the outcome payer.

"The outcomes for these three SIBs [initial PSI SIBs] are binary in a sense." (Calouste Gulbenkian Foundation)

For the initial PSI SIBs, the impact ended up being demonstrated more directly, for example through a working contract in the case of ACB and FFWD, or proof of child preservation within the respective family in the case of PF. However, despite being more straightforward to measure, these indicators might be slightly reductive since they only showcase the impact for those who meet the desired SIB outcome, not those who underwent the intervention but didn't meet the outcome in the pre-defined time period.

"(...) the evidence from this project [ACJunior] is much closer to what we would like to have." (Calouste Gulbenkian Foundation)

" (...) [CQC life-quality pre/post survey] is a questionnaire that is validated scientifically, it has been used before for different studies related to mental well being." (MAZE)

In contrast, JAP pilot T2W, ACJunior SIB, and in some ways the recent experience with CQC SIB, end up having different evaluation models from the other three SIBs. In the ACJunior SIB, there was a more robust evaluation model, namely through the introduction of a randomized control trial, even though there were reported statistical limitations. The results from the exploratory pilot SIB highlighted future measurement suggestions such as a larger sample size, evaluations of treatment and control groups over time (difference-in-differences method), and a more even balance of observable and empirical methods (Maze, 2018). In the case of CQC, there wasn't a control group, but it included rigorous pre and post-evaluations by an independent evaluator (academic institute), evaluating improvements in the outcome metric (quality of life of informal caretakers). According to the interviewees, the surveys have been used in several previous life-quality research pieces and are delivered to caretakers in a controlled environment (e.g. no information regarding financial implications of the answers). It is important to note however, that the contracted outcome of 5% improvement in the survey was considered ambitious by the service provider in partnership with other psychologists, based on historical data (e.g. typically the underlying condition is getting worse) - but still

there was no robust counterfactual. Finally, for T2W, besides evidence monitoring by Maze and a planned internal follow-up, there was an external evaluation by Católica University including a comparison group. The control group in this case was not possible, mainly due to the limited number of participants.

"I would say the preferred evaluation was having control groups, although the structure for the other 3 does not necessarily make their interventions any less valuable." (Calouste Gulbenkian Foundation)

"(...)we wanted to make sure that in the future, it would still be possible to do this evaluation retroactively, and we are actually pushing for that to be done." (MAZE)

Overall, the different projects in analysis showcased different levels of impact evaluation. It should be noted however, that the absence of a control group or more robust evaluation methods do not necessarily imply that the interventions are any less viable. In the case of ACB for example, the ability to provide qualifications and insert unemployed young adults into employment in a short period of time is also an extremely valuable source of insights for the public sector, in this case IEFP. Additionally, despite the absence of more robust evaluations in some projects, through the monitoring of outcomes, but also other non-payment related metrics, it is still possible to perform post-evaluations for any of the six cases. For the three initial PSI SIBs, one of the main reasons for the absence of an independent evaluation during the established time-periods was budget considerations.

4.2 Underlying Incentives

In this section, the findings will attempt to answer the following research questions:

3) *What are the associated underlying incentives?*

4) *Which are the main barriers and drivers that may explain a higher use of such mechanisms in specific contexts?*

4.2.1 Perverse Incentives

"(...) I don't have enough information to provide a definite answer, but I don't have the impression this [creaming and parking] took place." (Calouste Gulbenkian Foundation)

Overall, when asked regarding the possibility of perverse incentives taking place, namely practices such as creaming and parking, SIB's stakeholders reported a lack evidence of such practices taking place. Mostly from insights of the four PSI SIBs, four main drivers for this absence were identified: first, the inherent allocation of risk of the SIB model; second, the conservative level of pre-established thresholds; third, the collaborative relationship amongst stakeholders; and fourth, the objectivity of the participant's selection process.

Regarding the first driver, the SIB model allows providers to transfer most of the financial risk of not achieving the thresholds to investors, which controlled for perverse incentives as their payments, especially in the initial phase, were not directly tied to participant's success. The second driver, conservative thresholds, meant that providers were comfortable in achieving the target metrics, without excessive pressure, a typical pre-cursor of creaming and parking practices. However, it should be noted that the nature of the thresholds, especially in the four PSI SIBs brought additional challenges, which will be covered in following section.

"(...) one thing I thought was really interesting was the proximity between the different parts - who invests, who validates and who executes." (MDV - PF)

"(...) there are also important synergies between us [investor] and Maze [intermediary], which creates trust and further collaboration between the different entities. " (Calouste Gulbenkian Foundation)

The third driver, collaborative relationships, especially reinforced through periodic meetings, allowed SIB stakeholders to align their motivations and build trust. Both previous relationships and the early involvement from the different partners, including the sector-representative public entity, the outcome payer, and especially Maze as the intermediary, were identified as key to align the motivations and incentives of all stakeholders.

"(...) the intervention has to respond to a social problem in Portugal; needs to be innovative; and needs to have a positive impact on the target group. If it didn't have these 3 requirements, for as attractive as it might be, it wouldn't be accepted." (PSI)

"The selection process has several priority criteria (migrant, unemployment years, etc) which typically are tied to more challenges in achieving outcomes - so in a way we choose the other way around." (TESE - FFWD)

The fourth driver that was identified as controlling for perverse incentives in these four SIBs was the objectivity of the project and participant selection processes. First, in the "call-for-applications" phase, PSI as the public outcome payer, but also as a promoter of the social ecosystem in Portugal, sets up the priority areas of public policy. Second, in the application phase, regional representatives assured that the target group was directly related with a clear and observable social problem. This initial screening process by public entities is closely related with the classification of the interventions as a Social Entrepreneurship Innovative Initiative (IIES). Third, and perhaps the most impactful in terms of controlling for perverse incentives, is the actual selection criteria employed by each intervention. In the case of FFWD for example, the participant selection, out of the pool of all the potential participants, is restricted from the beginning by risk factors. In fact, these risk factors are usually aligned with the profile that has the least capacity to achieve outcomes. By setting priority selection criteria (e.g. years unemployed), there is, to a certain extent, already a homogeneity of risk factors within the selected group (not to be confused with overall homogeneity in terms of other variables), making it hard to further discriminate throughout the project through perverse incentives.

4.2.2 Binary Thresholds

"(..) it is a little bit unclear whether the target was too high or too low, what is perhaps even more relevant to consider is the type of incentives that are created when you have these one or zero outcomes." (MAZE)

In the case of the three SIBs currently entering the final stage, as described in section 4.1.3 *Measurement and Impact Evaluation*, the contracted targets were considered binary in the sense that, in payment terms, there were only two possible outcomes: if the thresholds were met, regardless if largely surpassed or not, investors would be reimbursed the full amount¹¹; on the other hand, if the thresholds were not met, even if just barely, the payment would be zero. This "hard" nature of thresholds had a significant effect on impact assessment and stakeholder incentives, especially investors and providers. Taking the case of ACB for example, where the contracted outcome is the employability (and respective evidence) of 11 students per bootcamp: If 11 participants are evidenced as employed, investors get fully reimbursed, if only 10 enter employment, no payment would be received. What this creates is

¹¹ in practice there were exceptions as will be described in section 4.2.4. *Expense Reporting and Investor Risk*

that after the evidence is collected for the 11, the financial and operational incentive to keep collecting evidence is very low.

"(...) if in a certain bootcamp the threshold is missed by just 1 beneficiary, the repayment is 0, while in another bootcamp, if the threshold is largely surpassed, the repayment would be the same as if it was 11 (the contracted threshold)." (Calouste Gulbenkian Foundation)

In the case of the three initial PSI SIBs, the outcomes were divided in cohorts (e.g. for ACB one outcome per bootcamp, for FFWD two outcomes per group), meaning that for each "outcome" there is a separate submission/outcome verification phase and respective reimbursement - this creates an amplification of the aforementioned lack of incentive to surpass contracted targets. Still within the example of the ACB SIB, hypothetically, if out of the first six bootcamps 120 employment evidences were achieved (the maximum amount of 20 per bootcamp), and then in the seventh bootcamp there were only 10 evidences, there would be no payment for that outcome/bootcamp. This can place investors in particularly uncomfortable positions. In contrast, in the case of JAP pilot T2W, this issue is specifically addressed with up to 10% compensation: if in one outcome results are 10% above the threshold, it can compensate an eventual 10% deficit in another outcome.

"(...)the incentive for the service provider to keep collecting contracts drops dramatically after they've reached the contracted thresholds." (MAZE)

The lack of operational incentive to surpass the thresholds is exacerbated as the evidence submission is an extremely bureaucratic process and required evidences are harder to collect than initially expected by the contracting parties. In terms of the effect on impact evaluation, this issue might distort an assessment of the actual number of people that benefitted from the intervention. In the cases of ACB and FFWD, there could be participants that effectively got employed but were not evidenced through a working contract due to reporting costs.

"We could consider in the future if it would be possible that the metric had some kind of global weight to account for the overall outcome (...) or the outcome payments for an employed participant would depend also on the risk factors associated with the participant at the beginning of the project, for example." (Calouste Gulbenkian Foundation)

One way these issues were alleviated was through the adaptation of the contracted thresholds to the difficulty of demonstration for each project. For example, for PF, the targets were higher than ACB, because the outcome was easier to demonstrate (single evidence provided by the courts, not dependent on beneficiary cooperation). Therefore, although there is still no additional incentive for surpassing the threshold, the associated costs of providing additional outcome's evidence are much lower. Furthermore, future solutions to these issues, such as tying payments to progress, the performance of the entire target population or specific risk factors, were proposed by providers, investors and intermediaries alike. Perhaps as a consequence of these considerations, the recent CQC SIB, launched in 2019, had a contracted outcome metric of 5% overall improvement in caregiver's average quality of life (Portugal Social Innovation, n.d.).

4.2.3 Financial, Social and Reputational Incentives

4.2.3.1 Financial Incentives

First, in the case of the four PSI SIBs, before analyzing the most significant financial, incentives for stakeholders, it is important to consider how the payments are anchored in these specific contracts. As Portugal Social Innovation (PSI) is a structure financed by the European Social Fund, within the scope of the Portugal 2020 Partnership Agreement, it must follow very specific guidelines. The financing structure requires that outcome-based payments be made against expenses, which in turn must follow the standard rules established by European norms of social subsidy-financed projects. Therefore, while still contingent on the achievement of outcomes, the payment is ultimately anchored to intervention costs (Maze, 2020a, 2020b, 2020c).

"(...) if they [investors] get fully repaid, and that would give their IRR zero, they can still have a little bit of a profit from the tax benefit. In any case, it's not by any means a competitive return." (MAZE)

In terms of the financial incentives for investors, directly, the potential internal rate of return is capped at zero, indirectly, it can be positive since there is a tax incentive in Portugal for investment in social impact bonds. This incentive allows investors to write down 130% of their investment as a cost, regardless of a potential future reimbursement (Maze, 2020a, 2020b, 2020c). As a hypothetical example, provided all the outcomes were met during the SIB validation process, from an investment of 100 thousand, investors can report it as an expense of 130 thousand and obtain the corresponding tax benefit. It should be noted

however, while generally not the case, some investors might not benefit from this tax incentive if they are legally considered non-profit organizations.

"(...) these four sibs were designed with capital recycling." (MAZE)

As described in section 4.2.2 *Binary Thresholds*, a second important consideration before assessing incentives is the division of outcomes (each corresponding to a cohort) within each SIB. The four PSI SIBs were designed with the idea of capital recycling - investors would make an initial investment, not the full cost of the project, and as the first outcome is met, investors get reimbursed for that specific outcome to then reinvest some of that capital into the subsequent phases, eventually providing funds until the end of the project. This design concept had important implications for investor's incentives, but especially for providers.

"(...) if the first outcome is not met, then there needs to be a decision about what to do." (MAZE)

First, for investors, the recycling model was used to decrease the cash flow risk of providing the entire amount upfront. In this way, the initial upfront amount would be returned if the first outcome was met and then reinvested in subsequent rounds/outcomes. However, some questions emerged in the case of outcomes not being met or not being fully reimbursed - would they still provide the capital for future phases? would the recycling method be adjusted? - overall, in the three PSI SIBs launched in the first wave, there were two different approaches. In the FFWD and PF SIBs, the majority investor, Calouste Gulbenkian Foundation, showed willingness to absorb most capital risk and losses; for the ACB SIB, there were investor constraints that led to a rearrangement of the recycling model that shifted some of the risk to the service provider. The reasons behind the two different approaches are related to the nature and risk preferences of investors as will be further explained in section 4.2.4 *Expense Reporting and Investor risk*.

"(...) [investors] would only provide an upfront investment that covers the first outcomes and then use the recycling model to reinvest in further stages." (Calouste Gulbenkian Foundation)

"All the parties involved in the SIB, from investors to intermediary to the providers, knew that the project could only achieve continuity if the metrics and thresholds were achieved in each bootcamp (...)." (ASSOP)

Second, in terms of the service providers, the obvious implication related to the desire for continuity of the project. One of the reasons for the use of capital recycling was the need to create incentives that would not exist if the entire intervention budget were provided upfront, especially considering the intervention timelines for these SIBs are relatively long (three to four years). In fact, providers reportedly felt, if they didn't perform or deliver outcomes, there was a chance that the project was going to be cut short. Additionally, there was a strong pressure on report timing and quality, as, even if the outcome targets were met, providers still need to report their expenses in the established time-period for payments to be approved.

"In this case [T2W] there is a more serious penalty for the providers, as not achieving results means not getting reimbursed." (TESE)

However, it should be noted that, despite the incentives created by the recycling model, for providers in these four SIBs, perhaps with the exception of ACB, the financial risk for providers was generally small. The experience of TESE with both FFWD and T2W is quite illustrative of different incentives resulting from risk allocation in different financing models. While in T2W the financial risk is kept on the provider side resulting in a strong financial incentive to achieve results, in the FFWD SIB, incentives for the provider are mostly reputational and socially-oriented since the investor is taking on most of the financial risks.

"For Fast Forward, Code Academy, and Projeto Familia, the projects that started earlier, the recycling plan was abandoned or refined." (MAZE)

In the cases of FFWD and PF SIBs that are currently entering their respective conclusions, the recycling plan was eventually abandoned. As will be further explored in section 4.2.4 *Expense Reporting and Investor Risk*, reporting delays and mismatches in expense validation, ultimately led investors to provide the capital necessary until the end of the project, even before receiving payments from previous achieved outcomes, to ensure the projects would not run out of working capital. In the case of ACB, delays and other unexpected financial risks led to a different rearrangement of the financial model. In the two cases where investor payments covered the project's costs until the end, there was no evidence of service providers relaxing or changing the level of effort, perhaps due to the continuous oversight by Maze, the intermediary.

"We believe it [financial returns] is necessary as for us this would be the only way of align the incentives across the project stakeholders, (...)" (Calouste Gulbenkian Foundation)

"(...) a lot of times the issues are thought in terms of punishing organizations that are not meeting the outcomes. But it's also about rewarding the ones that achieve them, and about creating a distinction."

(Maze)

"There is hope that the next community program already includes some financial incentives but it is still unclear. Once again, we are using taxpayer money, which makes it a delicate subject." (PSI)

Overall, providers, investors, intermediaries and PSI to a certain extent, have highlighted the need to consider new metrics and designs in the future that are able to more accurately reward or penalize the level of performance and risk of the interventions. From the experiences with the three initial PSI SIBs, two main barriers for the development of financial incentives were identified. The first is tied to the constraints of PSI and European fund's guidelines. Since the reimbursements are anchored to expense reporting, that is, the outcome payer will only pay investors for service provision costs that were actually incurred and verified, this can create a counter incentive. The provider has no incentive to deliver the maximum number of outcomes for the minimum amount of costs, as neither the provider nor the investors get to capture cost-savings achieved in regards to the original project budget. Additionally, the possibility of adding financial incentives or returns is quite sensitive due to the nature of communitarian funds. The second main barrier relates to administrative complexity and high reporting costs, which will be covered in section 4.2.4 *Expense Reporting and Investor Risk*.

4.2.3.2 Social Incentives

"(...)in a way, outcome payments are anchored to intervention costs and not the value that is being created, which is something that is in our perspective, is not ideal." (MAZE)

"But in FFWD, the lack of financial incentive did exist but there was a social incentive." (TESE - FFWD)

"The 50% or 60% was never the limit for us, it was always the maximum amount of children staying at home in a safe environment. If there is any incentive is the one of social impact." (MDV - PF)

"All these providers are special in their own way, they all want to go to the next level in terms of the quality of their of their services. " (MAZE)

When considering financial incentives throughout the interviews, although there was a clear recognition of its importance, the focus was ultimately on social-impact incentives. The overall response was that what distinguishes social projects in the first place, is the fact that they don't follow financial return as the primary goal. For providers and stakeholders alike, the ultimate goal was to maximize their respective outcomes and develop a more objective assessment of the impact they were actually delivering through the intervention, regardless of bureaucratic requirements attached to payments. SIBs are perceived by providers as opportunities to validate, refine or scale their interventions and methodologies. Additionally, on the investor side, a philanthropic organization like Gulbenkian is also accustomed to invest without expecting or intending returns as the main objective for these projects is to generate social impact and inform public policy.

4.2.3.3 Reputational Incentives

"(...) in reputational terms, of course it would be better to have the maximum rate of success, but in operational terms, it is definitely not a priority." (Calouste Gulbenkian Foundation)

"Within the SIB there is a compromise with investors and obviously the reputation of TESE as competent and efficient in achieving results." (TESE)

"(...) internally, we knew we would have to situate above 90%. (...) It would be hard if we approached companies with an historical success rate of around 50% [contracted outcome]." (ASSOP)

In general, for the four service providers of the four PSI SIBs, there is a lot of reputational risk. The first reason is that, for the most part, the investors within the SIB context were also relevant for providers before and most likely relevant afterwards in different projects. Secondly, there is a strong incentive to showcase competencies and ability to achieve results to all the partners, especially to key stakeholders in each intervention's ecosystem. In the case of ACB, for example, there is a strong focus on the relationship with the recruiting companies by signaling the quality of the program, both through the historical success rate and the quality of work from former participants. Third, the monthly meetings where service

providers and investors are present for monitoring of progress towards each specific outcome promote additional pressure.

4.2.4 Expense Reporting and Investor Risk

"(...) as outcomes payments derive from European Social Fund, and are channeled through the Portuguese government, they come with very specific rules." (MAZE)

"(...)even despite this being an instrument that stimulates social impact, it has an operational dynamic of grants." (PSI)

In the traditional SIB model, as the investor is the one providing upfront capital, they incur most of the financial risk, especially when outcomes are not achieved. In all three PSI-SIBs - Code Academy Bootcamps (ACB), Faz-te Forward (FFWD) and Projeto Família (PF), in order to alleviate this risk, capital recycling was included. However, as these three projects are entering their respective final phases, it was possible to identify that, in practice, three main problems emerged: first, the difference between budgeted expenses and submitted expenses ("sub-execution"); second, the differences between submitted expenses and approved expenses ("cuts"); and finally, the overall administrative complexity of outcome and expense reporting.

"(...)the first outcome is usually worth much more than the following outcomes." (MAZE)

In PF, recycling was initially planned for the first three outcomes, for FFWD in the first two outcomes, and ACB in the first six outcomes (Maze, 2020a, 2020b, 2020c). The reason for a higher focus on the initial outcomes is tied to a higher weight from the first outcome, since it includes a substantial initial investment and higher associated risks (even though the number of beneficiaries and expected difficulty of reaching the outcome target was similar across groups/outcomes). The reason for the different number of recycling rounds per project is determined by the total costs of each project and the amount each investor is willing to provide upfront. In the case of ACB for example, the total cost of the project was comparatively higher.

"(...) our projects work through reimbursement requests. The expenses are only considered eligible in the request." (PSI)

The first main difficulty that ultimately led to the abandonment of the recycling plans in the FFWD and PF SIBs, relates to the mismatch between planned (and pre-approved by PSI) expenses and reported expenses. The projections, in some cases, were too high, which meant investors injected more capital than what was actually necessary for the respective timeframe (Maze, 2020b, 2020c). In some periods, there weren't enough expenses to reach the pre-approved amount for the specific period. Since the payments are based on the actual expenses reported by the provider, this created, if only temporarily, a loss for investors.

"It is very common that the activity has the evidence from 1 month but the payment only occurs in the next month. (...) it was identified as a barrier by the provider." (PSI)

Additionally, the second identified issue is that, for several periods in both projects, submitted expenses were not approved by PO ISE, the expense auditing entity (Maze, 2020a, 2020b, 2020c). This included expenses that were not considered essential for the service provision (i.e. consultancy services), expenses relative to salary increases that were not originally budgeted, expenses that were paid too long after the respective service or product was delivered and therefore fell “outside” of the respective outcome *time interval*, among other cases.

"(...)SIBs were designed within a logic of direct costs. This means that all the reimbursements are evaluated one expense at the time, which incurs a higher risk and a longer administrative process." (PSI)

"In the first outcome delivery there were a lot of delays, with investors only being reimbursed months later, which led to move away from the recycling model and provide anticipated funds to guarantee the sustainability of the project." (MDV)

"(...) this was an unexpected challenge, (...) it was expected that outcomes-based contracting was accompanied by a simplification of the reporting mechanism." (TESE - FFWD)

Finally, the third and perhaps the most significant reason for the decision to abandon capital recycling in these two projects were the repayment delays. The method of capital recycling is based on the assumption that repayments from previous outcomes would be received before the start of the following. In practice, this was not the case for the three PSI SIBs. In the case of FFWD for example, the initial payment was supposed to cover the costs of delivering the

first outcome, including a buffer of three months (Maze, 2020c). What happened was that the submission and approval of the first outcome evidences took much longer than three months, which meant, as the time period for the following outcome was starting, investors were still not repaid for the first outcome. After several months, TESE, the service provider, started to have liquidity pressure and the investors decided to inject new capital to keep the project going. The first repayment was then eventually made, but investors decided to move away from recycling as the second outcome also had projected delays of several months. Between the first and second reimbursement, investors provided the remaining amount to cover expected costs until the end of the project. Very similar issues were reported in the PF SIB, with the decision to abandon recycling happening in parallel.

"(...)what happened was a double complexity of reporting." (TESE - FFWD)

The main reason for the delays was the complexity and rigor of the evidence and expense-reporting requirements within the use of European funds, which resulted in several practical questions and clarifications. This was exacerbated since there were different validating entities, PO ISE for the financial evidence, and PSI for the outcome evidence.

"(...)this mismatch affects the reputation of TESE with Investors. Because ultimately, the responsibility for the expense reporting is from the provider." (TESE - FFWD)

When investors took the decision to abandon the recycling plan and provide the remaining capital until the end of the project, they significantly increased their financial risk and decreased their leverage with the provider and Maze. From that point onwards, providers could not as easily be penalized for not achieving outcomes or submitting expenses under the budget and expenses with a lack of quality that risk disapproval from PO ISE. However, should be noted that reputational and social incentives remained as much, or even more, relevant as before.

"(...) investors, in some cases, want to make sure that the providers reports go well, until the end of the project. (...) to see how much is cut, and why. And then make a decision about what to do." (MAZE)

In the case of ACB, while similar issues were reported, the approach was quite different. In this case, the capital-recycling plan was not only kept, but also extended until the end of the project, which can be attributed to three main reasons.

"(...) investor's reimbursement depends on the quality of the expense reporting from the service provider side." (MAZE)

"It can be difficult when the program defines this (salary updates) as non-eligible, especially when it is going against common sector practices around the country." (ASSOP)

The first reason was the volume of "sub-execution" expenses and the "cut" expenses, which in the ACB SIB were much more significant than the other two SIBs (Maze, 2020a). These cuts derived from different expenses being considered non-eligible. As an example, Code Academy expressed the need for staff salary updates and hiring additional bootcamp teachers, especially considering they are positioned in the competitive market of information technology, however, both expenses are not recognized by PO ISE as eligible.

"(...) on a relationship when there is a lot of trust, the investors find it easier to cover these unexpected expenses." (MAZE)

As a consequence, reporting mismatches could have had a significant impact on the relationship and trust between investors and the providers. The decision to keep the recycling plan, as described in section 4.2.3. *Financial, Social and Reputational Incentives*, is able to create additional incentives and pressure for providers to achieve outcomes and perform quality reporting.

The second implicit reason was the nature of Code Academy as an organization that can better handle cash flow issues, namely through access to credit, especially when compared to the providers from the other two SIBs - TESE and MDV. In fact, until this moment, the ACB provider has borne all the respective expense cuts (Maze, 2020a).

The third one, and perhaps the most relevant, is the fact that ASSOP is the majority investor, not Gulbenkian (Maze, 2020a). Even though it is nearly a "50/50" investment split, this fact might be relevant considering the majority investor is the entity that receives the total

repayment and then distributes to the remaining investors or reinvests it in the providers, which gives it a position of influence. Most likely, ASSOP does not have the same risk tolerance or cash flow flexibility to provide higher amounts upfront as Gulbenkian, due to the different natures and sizes of the organizations.

"We are still waiting for the end of the projects to understand who will absorb the losses." (MAZE)

In the case of ACB, until now, all expense cuts were absorbed by the provider, which shift part of the financial risk from investors. It should be noted however; it is still a possibility, contingent on the performance of the remaining outcomes, that investors will eventually cover these cuts. The parties are still waiting for all the outcomes to be processed to have a more objective perception of the generated impact before attributing expenses.

"While some investors are more worried about getting their capital back, others focus more on staying true to the spirit of social impact bonds, and those investors are more willing to cover [these non-outcome-related financial losses]." (MAZE)

On the other hand, the contrasting cases of PF and FFWD, once again highlight the influence of investor's risk and return preferences on the attribution of risk amongst stakeholder's. For PF and FFWD SIBs, Gulbenkian Foundation, a philanthropic organization, was the majority investor with around 85% of initial investment, which means its influence on the decisions of these projects is quite significant. In both SIBs, the upfront payments were all led by Gulbenkian, who showed the willingness to cover all cuts and "sub-execution" expenses.

"(...) this [whether to absorb expense cuts] is not supposed to be an issue in a social impact bond, its just a characteristic of this particular administrative setting." (MAZE)

"(...) [for T2W] the focus is only on results not on financial reporting as this was already agreed upon and defined. This is indeed a simplification and still a true payment by results." (TESE)

Finally, when assessing the case of Joint Action Plan (JAP) pilot Training2Work (T2W), the issues highlighted in the three PSI-SIBs in question weren't nearly as prevalent, likely due to the use of simplified costs. Under a JAP, reimbursements by the Commission are strictly based on the outputs and/or the results achieved, thus focusing on the impact of an action rather than on tracing every detail of co-financed expenditure (European Commission, n.d.).

All the questions regarding financial execution are discussed and defined at the application time. Each outcome is attributed a unitary cost to be accepted by the Commission, in that moment all the evidence is required to justify these values. This is the only audit trail on which auditors will perform controls (European Commission, n.d.). However, despite the simplification of financial reporting, the significant drawback from this method is in case of outcomes not being achieved, since providers would be the ones incurring the risk and loss, as there are no external investors like in SIBs.

"What we are trying to do, still within European framework is to change this logic to simplified costs. (...) but we still have some doubts." (PSI)

" (...) we need to understand if this will ultimately benefit most entities in the social economy and not just a select few. " (PSI)

In conclusion, from the interview with Vera Barracho from PSI, there was a clear recognition of these problems and the question of direct versus simplified costs is being considered. There are two main reasons why the shift to simplified costs hasn't already taken place. The first is the framework from European Funds. While the European Social Fund (ESF) is starting to consider the adoption of simplified costs in innovation projects to guarantee liquidity, agility and growth, this was still not applicable for the three SIBs at the time of contractual negotiations. The second reason is tied with the nature of some provider organizations, namely IPSSs (Private Social Solidarity Institutions). In a logic of simplified costs, at the time of the application, organizations have to demonstrate that a certain percentage of the budget goes to Human Resources (HR) - HR which contributes **directly** to the intervention in question. However, as an example, these organizations often report HR costs related to management, which most likely would not be considered eligible. Certain technicalities and disadvantages still need to be considered before deciding a definite shift to apparently more agile administrative processes.

4.2.5 Investor and Provider Diversity

"It pushes the private investor to rethink the motivation for doing these investments [SIBs]" (Calouste Gulbenkian Foundation)

"We feel that investors might be the ones that are less enthusiastic, when compared to the initial stages of the project." (MDV)

The decision of whether to attribute the cost cuts to the investor or the provider, also had a significant impact on project dynamics and the motivation of stakeholders to engage with the SIB instrument in the future. In the FFWD and PF SIBs, the investor absorbed all the negative differences, which could have a significant impact on the perceived benefits for investors of partnering in such projects in the future. While in these two cases the SIB investor was a philanthropic organization, which covered all the unexpected expenses and liquidity difficulties, all the parties expressed concerns regarding the motivation for private investors with different profiles to be attracted to the social innovation ecosystem.

"There is usually a grey area when deciding whose responsibility it is to incur those expenses. (...) it is very dependent on the project, investor and overall context." (Calouste Gulbenkian Foundation)

If one of the ultimate objectives for SIBs is to attract more investment to the social innovation area, it is necessary to better account for incentives and risks, namely through more significant financial incentives, and clearer, simpler, payment mechanisms when outcomes are achieved.

"It brought a consumption of time and human resources that wasn't predicted and that diverts the focus from this financing method." (MDV)

"(...) [administrative weight] diverts the attention from the real goals, (...) which doesn't happen in the Training to Work for example." (TESE)

The issues of reporting and administrative complexity, which ultimately led to delays and the re-design of the recycling model in the three PSI-SIBs, significantly diverted stakeholder's attention from the project's primary goal - to generate social impact. The amount of resources spent on answering administrative questions related to expenses, was seen as both unexpected and unjustified by stakeholders. In contrast, in T2W JAP pilot, as mentioned in the previous section, by using a logic of simplified costs where most of these issues are discussed "ex-ante", TESE was able to channel most of the organization's available resources to the impact of the action.

"(...) what we observed and still couldn't really understand, is why some of the projects had to rethink their intervention models. (...) perhaps it has to do with the maturity of the intervention or the provider organization." (PSI)

The challenges illustrated by these three SIBs highlight how the different natures of provider organizations in Portugal can be more or less suitable to deal with the SIB design as it currently stands. While the SIB's underlying principles of innovation and impact are perceived by organizations as highly valuable, when implemented in practice, the model is not as agile or useful as initially planned. This can be detrimental, not only to the SIB project, but also to other projects, outside the SIB, that the organizations are involved with. Being so focused on outcomes and the rigorous delivery of the evidence, some entities expected to step-forward to test the model are currently not engaging.

"(...) if the investors are strictly financing in tranches through a recycling method, this can penalize weaker organizations." (TESE - FFWD)

"(...) it is easier for organizations that have a more organized accounting or a financial team that is used to doing this and can prepare these documents the right way." (MAZE)

The extended length of the projects, from the application until the final reimbursement, can be hard to sustain for organizations that have a more fragile financial and operational structure, incapable of handling the administrative weight and cash flow issues. These limitations could be even more enhanced if the provider lacks experience with communitarian funds or does not have the possibility to receive external managerial support. As evidenced by the three initial SIB cases, MAZE reportedly had a key role in assisting providers throughout the entire duration of the still ongoing projects, especially with financial reporting.

"(...) this is negative for any organization, regardless of the resources." (Gulbenkian)

"(...) sibs are the way they are now, very demanding contracts, (...) also because they are trying to find the truth about what is working and what is not working (...) this can be intimidating for some providers " (MAZE)

On the other hand, within the Portuguese ecosystem, big incumbent social organizations or organizations that are able to generate cash flows from other projects, and who would

theoretically be better equipped to deal with financial and administrative challenges, might also have little incentive to engage in SIBs. Besides the lack of operational incentives, the level of rigor of outcome monitoring and evidencing can put providers in a particularly vulnerable and exposed position.

"I don't think it is a financing instrument for pilot projects. Even though this is inserted within social innovation." (MDV)

"(...) [for Code Academy] the SIB was an opportunity to promote the project and give it a different scale, but we always had very present the idea of going beyond (...)" (ASSOP)

"(...) in strategic terms, we are validating our methodologies to potentially be implemented in other projects." (TESE - FFWD)

Additionally, a distinction should be made between pilot projects and projects by organizations who have an established track record with the methodology of the intervention. Experience and a deeper knowledge of traditional methodologies can be especially helpful for providers for three main reasons: first, as mentioned in section 4.1.1. *Metrics and Targets*, track record is an important factor in the definition of outcome targets; second, past relationships with investors and their familiarity with the methodology can create trust and a higher willingness to take on risk; and third is the confidence of providers when entering the SIB model which enables them to be more strategic. In the case of Code Academy for example, the organization conducted a pilot project, before the ACB SIB, with very positive results. This allowed them to look at the SIB as an opportunity not only to validate their intervention, but also to scale and find a sustainable model beyond the SIB scope, even despite all the administrative complexity, which can act as a detractor from those objectives.

"This is a way in for a smaller organization or a newer organization, or a private organization that has not been operating in the social space to come in and try to do something different." (MAZE)

In conclusion, it is important to affirm that despite the challenges and their potential negative effects on provider and investor diversity, SIBs still offer significant opportunities for organizations looking to enter the social ecosystem and establish a bridge with the public sector. For example, for a provider of an innovative intervention that does not yet have any

relationship with public entities and has struggled to get public financing, a SIB is a way of testing and showcasing their impacts.

"(...) to try and see if it's possible to pay investors for their risk because that is what the return is. But of course for that there needs to be risk. It has to be edgy, in a way." (MAZE)

Nevertheless, for the SIB model to keep its appeal of bringing together a diversity of stakeholders, it is key that the administrative burden of the instrument is reduced and financial incentives, especially for investors, are more significant. These two factors should be assisted by a clear specification in advance of the specific risks and necessary requirements, so that the project design and allocation of resources are able to generate the most amount of impact.

"If the appeal is lost, the viability of SIBs is also lost." (Calouste Gulbenkian Foundation)

4.3 Public OBC

In this section, the results will attempt to answer the following research questions, with a focus on public sector perspective:

3) *What are the associated underlying incentives?*

4) *Which are the main barriers and drivers that may explain a higher use of such mechanisms in specific contexts?*

4.3.1 Public Monitoring

"(...) local activators are mainly focused on guaranteeing that the applicant idea is considered a IIES and doesn't fall short in the secretariat." (PSI)

The monitoring and oversight by sectoral public entities throughout the projects is key for a correct assessment of the intervention's performance. First, for the four PSI SIBs, the involvement of a public entity is mandatory in the initial phase of establishing the outcomes and in periodic meetings throughout the implementation. The contribution of public representatives and local activators is mainly focused on the eligibility of the intervention as a Social Entrepreneurship Innovative Initiative (IIES), which is one of the decisive criteria in the SIB approval.

"Learnings from the different phases of the intervention can be incorporated in public policy, which is why it is extremely important that a public entity is also present in the monitoring throughout."

(Calouste Gulbenkian Foundation)

Second, one of the biggest goals of the SIB mechanism is, not only to invest in innovative interventions, but also to inform public policy. There is no way of doing this without a regular follow-up with the sector-relevant public sector representative (note that the outcome payer is a central, non-specialized public sector structure). In the case of ACB for example, IEFP's involvement was reportedly crucial in the process of collecting the evidences of employment. In other cases, such as PF SIB, the provider organization - MDV, already had past relationships with several public entities from other projects, namely social protection teams and the Institute of Social Security, some of which remained in the context of the SIB. Note that in the case of ACJunior SIB, the outcome payer was the Municipality of Lisbon and that, in Portugal, the primary school system falls under the responsibility of local government.

"(...) our job was more on how to formulate and adequate the global metric of the project as well as the project timeline to OBC within a logic of reimbursement requests." (PSI)

Additionally, as a mission structure for the Portuguese government, PSI provided an important link between the four project SIBs - ACB, FFWD, PF and CQC - and the relevant areas of government. First of all, PSI was in itself a facilitator for the creation of the SIB multi-party contract. During the initial stages of each project, PSI's main objective is to adapt the metrics and project timelines to a logic of outcomes-based contracts and reimbursement requests. When a SIB is approved, a monitoring schedule is approved, including control milestones, where PSI will validate the outcome evidences. In this way, they are able to keep track of the outcome's achievement. On-site visits by PSI teams were also planned but it is unclear how frequently they occurred. However, it should be noted that, in the context of Portugal 2020 and the European Social Fund, designated teams visited the projects for observations, without performing any auditing activities.

"We decided to create an impact evaluation model, encompassing all the intervenients in the instrument, but we decided not to engage right now and wait for 2021." (PSI)

PSI had three main goals as a mission structure: first, to stimulate the social ecosystem in Portugal; second, to create both financing (grant logic) and financial (investment logic)

instruments; and third, to measure the social impact that was created. As explored in section 4.1.3. *Measurement and Impact Evaluation*, for the four PSI-SIBs, a more extensive and robust impact evaluation hasn't yet taken place. However, as PSI's mandate extends until 2023, the plan is to start performing evaluations in 2021 and 2022, at a time where most SIBs have reached the conclusion phase.

4.3.2 Communication

"Even with SIBs being a new instrument, there is the political compromise and identification from the public sector regarding priority sectors where there was a void." (PSI)

The role of PSI as a public entity is to identify critical policy areas and promote open innovation in the ecosystem. In theory it would be desirable for the public entity to define outcomes and the service provider, together with the investors, would attempt to meet the outcomes. However, in the case of these four SIB's, the service providers and the investors, with the help of Maze, have been arguably the most proactive stakeholders.

"There needs to be more ambition when aligning the desired outcomes to public priorities and to the potential savings for the government (...)." (Calouste Gulbenkian Foundation)

In the case of the three ending SIBs, from the social sector organizations' perspective, it is very important that public entities are more proactive in detailing the outcomes and how much they are willing to pay for them. Considering the amount of administrative and operational challenges the projects faced, a clearer willingness-to-pay (WTP) for certain outcomes would have been an important driver in aligning the overall incentives in a single direction. It is argued that the public sector has the ability, likely higher than other entities, to understand how much is being saved and how much is being spent. In the example of ACB, the employment of a young adult, not only means a savings from the alternative public support but will also provide taxes over the future salaries and so on. The exact amounts of these potential benefits could be assessed by public entities and expressed through a structured WTP.

"This is the type of reasoning we are hoping gets adopted in the long term, but we are still some steps behind." (Calouste Gulbenkian Foundation)

In order to better grasp the gains from avoided expenditures of certain interventions, what currently exists is proxy measures. One Value, a platform developed by MAZE and PSI, which resulted from a protocol between Gulbenkian Foundation and the Portuguese government, tries to provide an approximate value for the public cost of a certain social issue. As an example, it could be described how much an unemployed young adult would cost per month, and based on that cost, the public sector would have a better idea on how much they would be willing to pay.

"In the projects, there isn't as much information-sharing regarding predicted and unpredicted outcomes as it would ideally exist. This could either be a managerial or a more communicative issue. (...) but there are struggles from both sides." (PSI)

Besides the need for clearer outcome detailing by social partners, on the other hand, the public sector also reports a lack of communication from social sector organizations. PSI has highlighted the need to establish more formal communicating channels. Through the follow-ups and evidence monitoring procedures, but also through informal communication, it was possible to come across valuable information regarding the interventions, which were not directly provided. In the case of ACB, there were significant positive impacts regarding regional and economic development, however, as these were not directly related to the monitored outcomes of the SIB, this information might not have been integrated in the best possible way.

The role of PSI and public representatives is to create a favorable ecosystem, and provide clear policy guidelines for projects to grow, however, for the interventions to be established as real solutions for public policy, there is still room for improvement in the communication between the social and public sector.

4.3.3 Integration into Public Policy

"(...) one of the problems is the need to structure these opportunities." (PSI)

" (...) [ACB] worked collaboratively with several entities like IEFPP, for this process to be integrated more consistently into public policy." (ASSOP)

" (...) [public entity] was present, was familiar with our method but it never had the formal approach of integration." (MDV)

One of the main objectives for the creation of SIBs within the Portuguese context is to stimulate innovation and allow solutions to be adopted by public administrations. However, throughout the different SIBs, especially the three that are entering the final stages, there was a reported lack of structure for these integrations to potentially occur. Providers positioned in different sectors highlighted some difficulty and resistance by public entities to formally integrate the solutions, at least for the short and medium-term. In the case of ACB, the investors and the provider consider that the intervention has the potential to be integrated or supported by public employment entities. However, although there was a clear focus on the collaboration with IEFP throughout the project, the discussions are being done through the organization's channels and not through a structured process, with the help of PSI for example.

"As a facilitator our role is to help the public sector to see the potential in these solutions (...)." (PSI)

"We are in an experimentation phase. We understand there is interest but there is still a lack of focal points." (PSI)

Portugal Social Innovation is aware of the integration difficulties and is currently considering possible solutions. One of the ideas developed by PSI to account for this issue in the future is the creation of focal points within each of area of public policy. The entities would be responsible for monitoring and evaluating the SIB, and eventually absorbing the service or establishing a partnership. However, this plan is still in progress as there are challenges with the identification and establishment of the appropriate public structures within each ministry.

Additionally, if there was an initial formal compromise with a public entity to integrate the given intervention in their own policies, this pre-established intention, would provide great incentives for the provider organizations, not only to perform under the SIB, but to continuously refine and evolve their solutions. At the moment, PSI is working on a project with the Ministry of Employment, which relies, for the most part, on the results from SIBs to propose a public response. However, once again, this opportunity arose spontaneously.

"(...) we feel there is room for improvement (...). Either within the current financial framework or in the coming one." (PSI)

In conclusion, there are still some challenges and questions with regards to the integration of SIB intervention's insights and methodologies. There is room for improvements in the underlying structure of communication and integration between social sector organizations and the public sector, which would in turn align the incentives and alleviate some of the burden from administrative reporting processes.

5. DISCUSSION

In this chapter, the main findings from the interviews will be located within existing literature, in order to draw comparisons regarding the topics of the research questions.

One of the most recurring scale barriers identified by SIB intervenients worldwide concerns the measurement design and evaluation methodologies (Gustafsson-Wright, 2020a). While the literature often highlights difficulties in the process agreeing on the appropriate measurement targets, time-scales for the outcome monitoring, as well as the specific role of each party in monitoring the outcomes (Tan *et al.*, 2015; Disley *et al.*, 2011), in the analyzed cases, there weren't many reported issues in initial planning stages. As suggested by Brown *et al.* (2016), the early collaboration by the different parties in the five SIB cases, including public sector representation (Charlton *et al.*, 2013; Nicholls & Tomkinson, 2013; Chiodo, 2015), could help to provide an explanation. Additionally, the fact that outcome payments come from a dedicated pool of money, external to ministerial budgets, could result in the public sector being less aggressive in outcome negotiations.

Still, despite the lack of major issues during initial stages, there were still two major reported concerns. On one hand, as predicted by the literature, social sector organizations highlighted the importance and need for more clarity and proactivity from the public commissioner in the specification of outcomes and reporting processes (Chiodo, 2015; Stoesz, 2014; Farr, 2015). Considering the amount of administrative and operational challenges the three initial PSI SIBs faced, for example, a clearer and more structured willingness-to-pay for pre-defined outcomes by the public sector, targeting beneficiaries in key policy sectors, was identified as potentially helpful in aligning incentives and setting clear goals since the beginning (Fraser *et al.*, 2016; Gustafsson-Wright & Osborn, 2020a). On the other hand, the public sector also reports a lack of communication and feedback from social sector organizations, especially through formal communication channels. This goes in line with the literature, which highlights public sector's struggles in measuring performance and accountability (Martikke & Moxham, 2010; Stoesz, 2014).

In terms of the targets defined for the different projects, they were highly specific to the context, occasionally relying on historical data from sectorial public institutes, but most of the time the outcome targets were based on provider's historical success rates, adapted to the

specific requirements of the SIB. There was a clear indication of the lack of a SIB dominant design, which could be replicated and standardized across geographies (Arena *et al.*, 2016; Antadze & Westley, 2012; Gustafsson-Wright, 2020a).

Additionally, the outcome metrics in the three initial PSI SIB's were binary in the sense that reimbursement payments were attached to absolute thresholds, with only two possible results: either full repayment if met or no repayment. As a consequence, there was a reported lack incentives to keep collecting evidence after the threshold was met (Maze, 2020a). Social sector organizations, in line with the literature, highlighted the need to account for these issues by tying payments to progress, the performance of the entire target population, and specific risk factors (Social Finance, 2015a; Gustafsson-Wright, 2020b; Disley *et al.*, 2015).

One way these issues were mitigated in the design stage in the three initial PSI-SIBs however, was through the adaptation of the contracted thresholds to the difficulty of demonstrating evidences for each project.

Furthermore, the lack of operational incentive to surpass the thresholds was especially exacerbated in the Portuguese context, as the outcome's reporting was an extremely costly and bureaucratic process, with administrative and reporting requirements of European funds resulting in a lot of practical cuts and delays. This highlights the challenges with the complexity of reporting and data collection processes promoted by SIB culture (Arena *et al.*, 2016; Gustafsson-Wright *et al.*, 2020; Thomas *et al.*, 2016). On the other hand, in terms of the advantages of the monitoring and evaluation culture typically attached to the traditional SIB model, while the communication was not ideal, there were reported benefits, especially from providers, relating to the discipline and rigor it brought to their operational processes (Social Finance, 2015a; Leventhal, 2012; Jackson, 2013; Clark *et al.*, 2014; Gustafsson-Wright *et al.*, 2020).

Overall, the different projects in analysis showcased different levels of impact evaluation. However, there were common operational and budgetary struggles in PSI-SIBs, which prevented more rigorous counterfactual comparisons and wider evaluations to be performed. This was consistent with the observed lack of longitudinal evaluations in SIBs worldwide (Social Finance, 2015a; Fraser *et al.*, 2016; Gustafsson-Wright *et al.*, 2020; Gustafsson-Wright & Osborn, 2020b). Nevertheless, in the four PSI-SIBs, due to valuable data derived from the continuous monitoring of outcomes and non-payment related metrics (Berlin, 2016),

it is still possible to perform rigorous post-evaluations. As one of PSI's main goals is to measure the impact that was created from these projects, it is planned that impact evaluations will take place in 2021 and 2022, at a time where most SIBs have reached their conclusion. This will be especially important to move beyond hypothetical assumptions and assess if SIB schemes actually translate into savings by public administrations (Tan *et al.*, 2015; Mullins *et al.*, 2011; Gustafsson-Wright *et al.*, 2020).

When analyzing the associated financial risks for stakeholders, in two of the initial PSI-SIBs, FFWD and PF, additional injections of capital and the "sub-execution" and "cut" expenses were all absorbed by the investor. In this way, despite the attempt to decrease the risk by introducing capital recycling, ultimately investors took on most of the project's financial risk (Dagher, 2013; Lehner & Nicholls, 2014; Loder *et al.*, 2011; Nicholls & Tomkinson, 2013; Warner, 2013). The case of ACB however, where these expenses are currently being absorbed by the provider, is consistent with situations where private investors do not borne the entire risk of the intervention (Arena *et al.*, 2016).

The uncertainty regarding attribution of risks (Fraser *et al.*, 2016), which escalated in PSI-SIBs from the unexpected operational setbacks, was recognized as potentially impactful on provider and investor diversity in future SIB contracts. It is important to note that most of the investors involved in these projects, were already present in the third sector and the specific areas of intervention. Having in mind that only a small sample size was analyzed, this was not consistent with Gustafsson-Wright *et al.* (2020) claim that SIBs attract private investors to new sectors and geographies where they may not traditionally work. Additionally, the overwhelming presence of public and philanthropic investors in these SIB contracts might suggest a lack of financial and operational incentives for the private sector to engage, as highlighted by Fraser *et al.* (2016). Still, the case of ACB, where the majority investor had a slightly different profile, showcased how the different natures and risk-preferences of majority investors could have a significant impact on different risk allocations for the different projects (Fraser *et al.*, 2016). On the provider side, the negative effects on organizations with more fragile administrative and operational structures were observed, in line with the adverse effect of conditionality of funding on smaller, voluntary sector organizations (Shutes & Taylor, 2014; McHugh *et al.*, 2013).

The specific challenges of Portuguese SIBs' reimbursement logic might not only disincentive more fragile organizations from adopting the model, but also distracted incumbent SIB providers from their ultimate goal - generating impact. The whole underlying philosophy of the SIB approach is that the resources of service providers should be focused on the impact of the intervention and not on financial and budgetary issues (Golden *et al.*, 2017; Social Finance, 2015a). Additionally, even though stakeholders suggested that SIBs within the Portuguese context might not be appropriate for pilot projects, due to the lack of track record and established methodologies (Gustafsson-Wright *et al.*, 2020), they also highlighted the opportunity for other organizations, such as mature providers with innovative solutions that have been struggling to receive public funding (Fraser *et al.*, 2016; Fox & Albertson, 2011; Mulgan *et al.*, 2011; Warner, 2013; Chiodo, 2015).

When considering the underlying incentives of the mechanism, throughout the four PSI-SIBs, for providers, apart from the decrease in risk by the inherent nature of the traditional SIB model (Social Finance, 2015a), the lack of observed perverse incentives, such as parking and creaming, can be attributed to three main reasons. First, the lower level of thresholds, that accounted for the risks of not achieving outcomes (Gustafsson-Wright & Osborn, 2020b). Second, the collaborative relationships, reinforced through past relationships with partners and periodic meetings throughout, which allowed all SIB intervenients to better align their motivations and build trust (Ng *et al.*, 2009; Brown *et al.*, 2016). Reportedly, this collaborative approach was also highly potentiated by the presence of a third party (MAZE in the four PSI-SIBs), especially in the implementation and reimbursement periods (Bafford, 2012; Haffar, 2014; Gustafsson-Wright & Osborn, 2020b). The third reason was the design of the participant selection mechanism, which had clear objective criteria, in most cases based on risk factors (Cabinet Office, 2013; Social Finance, 2015a).

In conclusion, as a result of the most relevant context-specific findings, and the respective associations to existing theory, three main considerations for a higher adoption of the SIB instrument within the Portuguese context were identified: the need to adopt more sophisticated metrics able to create further incentives for investors and providers; the need to decrease the administrative complexity of outcome reporting processes; the need to formally plan the potential integration of the intervention's methodologies and insights into public policy, supported by more robust evaluations.

6. CONCLUSION

In light of recent economic crisis, Social Impact Bonds (SIBs), aligned with NPM values (Hood, 1991), can provide a solution to the aggravation of complex social problems by aligning social and financial incentives, and promoting collaboration between public, private, non-profit and voluntary sectors (Fraser *et al.*, 2016; Mulgan *et al.* 2011; Cohen, 2011; Borgonovi *et al.*, 2019). Worldwide, the adoption and popularity of the SIB mechanism has been steadily increasing, with the literature highlighting new opportunities and challenges (Chiodo, 2015; Fraser *et al.*, 2016; Gustafsson-Wright & Osborn, 2020a).

When considering the main identified scale barriers by SIB literature, besides the lack of provider competition and access to data, some of the most recurring considerations concern commissioner organizational capacity, measurement design, and evaluation methodologies. As highlighted by Gustafsson-Wright (2020a), a better understanding of the driving forces to SIB development in certain countries could help provide insights into the different paths to create a competitive market for outcomes. Taking this into account, this dissertation focuses on the case of Portugal for two main reasons. The first was the lack of SIB literature on the Portuguese case; and second, the access to potential interviewees from local organizations. Additionally, this dissertation focused on particular characteristics of the SIB model, namely outcome metrics and measurement evaluation, as they were identified as potentially more suitable for replication (Ecorys, 2019).

With this in mind, the study's primary goal was to shed light on the metric and threshold selection in SIB's outcome-based contracting (OBC) mechanisms in Portugal, and to understand the impact of the underlying incentives of these indicators on stakeholder's motivations and the overall success of the instrument. With a predominantly exploratory and explanatory research purpose, six representative OBC case studies were evaluated through semi-structured interviews and secondary data.

From the findings of the six cases in analysis, three main barriers for a higher adoption of the SIB instrument within Portugal were identified: first, the need to develop more sophisticated and significant incentives for all stakeholders; second, to decrease the administrative complexity of outcome reporting processes; and finally, to formally structure a potential integration of the intervention's methodologies and insights into public policy.

First, the limitations associated with payments and financial incentives in Portugal are directly tied to the constraints of Mission Structure Portugal Social Innovation (PSI), the outcome payer for 12 of the 13 implemented SIBs, and European Fund's guidelines. The nature of communitarian funds necessarily anchors investor reimbursements to expense reporting, which not only restrict the possibility of adding more complex financial penalties and rewards, but also decrease the incentive for providers to seek cost-efficiency. From the interview's findings, one of the clear messages was the need to attach significant returns to outcomes, as a way of signaling public interest in developing OBC instruments. In Portugal, to date, there are still legal, moral and operational hurdles that need to be overcome for this to happen.

Additionally, for the SIB model to keep its appeal of bringing together a diversity of stakeholders, it is key that the administrative burden of the instrument in Portugal is alleviated, and assisted by a clearer specification, in advance, of the potential risks and necessary requirements, so that project's designs and allocations of resources are able to generate the most amount of impact - the ultimate goal of the SIB mechanism. As it stands, Portuguese SIBs are not taking full advantage of some of the identified benefits of the traditional SIB model regarding flexibility and alignment towards outcomes. In fact, management resources being overly spent on procedure compliance, as observed throughout the cases, is a typically identified disadvantage of input or output based contracting (Fox & Albertson, 2011; Gustafsson-Wright *et al.*, 2020). While still not addressing the need for more significant financial incentives in Portuguese SIBs, a move from direct costs to a simplified cost logic is currently being considered by PSI as a potential solution to these issues.

Finally, a third consideration when evaluating the main barriers for SIB development in Portugal, concerns the evaluation and integration of the intervention models by the public sector. In line with international SIB literature, there was a reported lack of structured and formal communication between social sector organizations and public sector representatives (Martikke & Moxham, 2010; Chiodo, 2015). Across different policy sectors, providers reported difficulties and resistance by public entities to formally integrate the solutions in the short and medium-term. In order to address these challenges, some of the proposed solutions by PSI include the creation of linking structures within each of area of public policy and the inclusion of a formal pre-established intention in initial contractual negotiations. Furthermore,

from a government perspective, the lack of rigorous and longitudinal impact evaluations of these projects still hinder the attribution of a decrease in public expenditure to the SIB mechanism (Lagarde *et al.*, 2013; Gustafsson-Wright *et al.*, 2020).

In conclusion, as suggested from international SIB Literature, the findings from the cases indicate a clear alignment with the main advantages and struggles faced by practitioners worldwide, with additional challenges resulting from the unique nature of the Portuguese SIB ecosystem. Much like the instrument itself, particular SIB characteristics such as metrics, evaluation and target definition are still in a pre-diffusion stage, with a clear recognition from stakeholders of the need to develop and implement more sophisticated metrics and robust evaluation methods.

While there are clear signs of promise and optimism, the wider adoption of the SIB model by the public sector should always weight in realistic evaluations of the benefits and drawbacks, in order to assess the real potential of the OBC instrument within each context.

7. REFLECTION

In this final section, some suggestions to the main challenges identified throughout this dissertation will be proposed, as well as an identification of the main limitations of the study and potential further research.

7.1 Suggestions

First, it is relevant to note that, in light of recent economic crisis, there might be an opportunity for social impact bonds to provide policy makers and public regulators a partial solution to alleviate the negative impacts on marginalized populations. Some of the areas that are often highly affected by crisis, are also areas where SIBs have had some historical success (Gustafsson-Wright & Osborn, 2020a). This constitutes an opportunity for the public sector, in partnership with social sector organizations, to expand or replicate these interventions that effectively address these challenges.

Within the Portuguese context, besides a potential move to simplified costs, another possible solution for administrative challenges associated with outcome reporting, as hinted by the interviewees in this study, is to attribute a higher share of the reporting responsibilities to partner (non-outcome payer) public entities. This would not only allow the social sector to focus their resources on impact maximization, but also provide the public sector with more realistic data regarding the interventions' performance and potential savings in public expenditure.

Additionally, Portuguese public regulators should consider the possibility of adding further and more sophisticated incentives to current regulation on SIBs and other outcome-based mechanisms. The public sector, both central and locally, should move beyond the use of European funds and allocate resources to outcome-based contracting practices, assisted by a structured willingness-to-pay for certain outcomes. As the SIB model is still in an initial experimentation phase in Portugal, when evaluations by PSI take place in 2021 and 2022, this could be a good opportunity for the government to make an honest assessment of the benefits and challenges of the model in addressing the traditional competitiveness and efficiency deficits of the social sector.

From a wider public policy perspective, a higher investment on outcome-based instruments and payment systems, could develop important synergies and externalities that will positively affect the entire social ecosystem and ultimately attract more external investment for instruments such as SIBs. This investment should be assisted by a higher focus on "big data" or more sophisticated monitoring and performance management tools, which could also drastically improve the reported difficulties in communication and impact evaluation of the projects. Additionally, the creation of technical units inside strategic ministries and pre-established integration commitments could go a long way in improving communication and aligning incentives between the public and social sectors.

7.2 Limitations

As in any study focusing on pre-diffusion instruments and phenomena, the number of Portuguese SIBs is still very small to draw sound conclusions. The level of maturity of the SIB ecosystem in general and the cases in analysis, especially the five still in progress, also limit the interpretation of the findings since more rigorous evaluations are still to come in the medium to long-term.

When considering the theoretical foundation for this research, the academic literature on the topic of SIBs is also limited, particularly when considering the Portuguese case. Additionally, as this dissertation focused on measurement and underlying incentives, most of the available literature was developed by practitioners, public entities and social sector organizations, with a notable lack of scholarly work.

Nevertheless, the findings from the selected cases can still offer relevant insights into common challenges faced by practitioners, especially throughout the development and implementation phases.

7.3 Future Research

In order to support the development and scale of the SIB market nationally and internationally, further research should be conducted on the diffusion of the SIB model over time, incorporating qualitative and quantitative data on a larger scale, with the objective to inform developers and policy makers in the future.

Additionally, for the specific market of Portuguese SIBs, it would be relevant to develop research on the aftermath of the evaluations of the initial SIB waves, planned for 2021 and 2022 by the public commissioner, and keep track of the development of new SIBs, either within or outside Portugal Social Innovation (PSI) scope. Specifically, unlike the SIBs analyzed in this dissertation, there are recent SIBs structured with PSI that do not have an intermediary. It would be interesting, from a research standpoint, to compare and understand if there were additional benefits or challenges in the experiences of stakeholders, particularly in terms of the reporting process, from the absence of an intermediary.

Finally, as mentioned throughout the literature review and the discussion chapters, there is an identified need to perform more rigorous evaluations, both from an impact and cost-saving perspective. Ideally, future research would directly compare social impact bonds with other outcomes-based commissioning models and alternative financing methods, which could be highly relevant in assessing the real impact of this innovative instrument in the provision of social services.

8. REFERENCE LIST

- Allen, P. (2009). Restructuring the NHS again: supply side reform in recent English health care policy. *Financial accountability & management*, 25(4), 373-389.
- Andrews, R. & Van de Walle, S. (2012). New Public Management and Citizens' Perceptions of Local Service Efficiency, Responsiveness, Equity and Effectiveness. *Public Management Review*.
- Antadze, N. & Westley, F. R. (2012). Impact metrics for social innovation: Barriers or bridges to radical change? *Journal of Social Entrepreneurship*, 3, 2: 133–50.
- Arena, M., Bengo, I., Calderini, M. & Chiodo, V. (2016). Social Impact Bonds: Blockbuster or Flash in a Pan?. *International Journal of Public Administration*.
- Audit Commission (2012). Local Payment by Results. London: Audit Commission.^{[1][SEP]}
- Bafford, B. (2012). The feasibility and future of social impact bonds in the United States, *Sanford Journal of Public Policy*, 3, 1: 3–19.
- Banke-Thomas, A. O., Charles, A., Madaj, B., & Van den Broek, N. (2015). Social return on investment (SROI) methodology to account for value for money of public health interventions: A systematic review. *BMC Public Health*, 15(1), 582.
- Barajas, A., Barajas, L., Burt, K., Harper Jr, T., Johnson, P., Larsen, E., Licona, J., Nije, C. O., Parker, W., Reid, K., Sturtevant, L., Tokunaga, M., Vien, Q., & Yeh, C. (2014). *Social Impact Bonds: A new tool for social financing*, New Jersey, NJ: PPIA Program, Princeton University.
- Bengo I., Arena M., Azzone G. & Calderini M. (2016). Indicators and metrics for social business: a review of current approaches, *Journal of Social Entrepreneurship*, 7:1, 1-24.
- Berlin, G. (2016). *Learning from Experience: A Guide to Social Impact Bond Investing*. MDRC. Retrieved from https://www.mdrc.org/sites/default/files/Learning_from_Experience_SIB.pdf
- Bolton, E., & Savell, L. (2010). *Towards a new social economy. Blended value creation through Social Impact Bonds*. London, UK: Social Finance Ltd.
- Borgonovi, E., Bianchi, C. & Rivenbark, W.C. (2019). Pursuing Community Resilience through Outcome-Based Public Policies: Challenges and Opportunities for the Design of Performance Management Systems. *Public Organization Review* 19, 153–158.

- Bridges Fund Management (2019). *Better outcomes, better value: the evolution of social impact bonds in the UK*. UK: <https://www.bridges-fundmanagement.com/wp-content/uploads/2017/08/Bridges-Better-Outcomes-Better-Value-2017-print.pdf>
- Brown, T., & Potoski, M. (2003a). Managing Contract Performance: A Transaction Costs Approach. *Journal of Policy Analysis and Management*, 22, 275-297.
- Brown, T., & Potoski, M. (2003b). Contract-Management Capacity in Municipal and County Governments. *Public Administration Review*, 63(2), 153-164.
- Brown, T., & Potoski, M. (2004). Assessing the management costs of delivering services under alternative institutional arrangements. *Journal of Public Procurement* 4, 375–96.
- Brown, T., Potoski, M., & Slyke, D.V. (2016). Managing Complex Contracts: A Theoretical Approach. *Journal of Public Administration Research and Theory*, 26, 294-308.
- Bryman, A. & Bell, E. (2011). *Business research methods*, New York: Oxford University Press, USA.^{[1][SEP]}
- Bryson, J. M., Crosby, B. C., & Bloomberg, L. (2014). Public value governance: Moving beyond traditional public administration and the new public management. *Public Administration Review*, 74(4), 445–456.
- Bugg-Levine, A., & Emerson, J. (2011). Impact Investing: Transforming How We Make Money While Making a Difference. *Innovations: Technology, Governance, Globalization*, 6, 9-18.
- Burrell, G. & Morgan, G. (1979). *Sociological Paradigms and Organisational Analysis*. London: Heinemann.
- Cabinet Office (2013). Guidance on the template contract for social impact bonds and payment by results. UK: Author. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/645183/20170223_FULL_GUIDANCE_SIB_TEMPLATE.pdf.
- Callanan, L. and Law, J. (2012). *Will Social Impact Bonds Work in the United States?* Washington, DC: McKinsey on Society.
- Charlton, K., Douglas, E., Flatau, P., & Gill, I. (2013). *Understanding Social Impact Partnership*. Perth, AU: UW A, Ecotrust Australia.
- Chiodo, V. (2015). Social Impact Bonds: New Finance or New Procurement? *ACRN Oxford Journal of Finance and Risk Perspectives*, 4(4). Retrieved from <http://www.acrn-journals.eu/resources/jfrp0404j.pdf>

- Clark, M., David Jr, C. C., Hwang, J., Moses, N., Nelson, S. and Torres, N. D. (2014). Next-generation nonprofits. In T. Hansen-Turton and N. D. Torres (eds), *Social Innovation and Impact in Nonprofit Leadership*, New York, NY: Springer, 63–81.
- Cohen, R. (2011). Harnessing social entrepreneurship and investment to bridge the social divide, Paper given at the EU conference on the social economy, Brussels, Belgium. Retrieved from <http://ec.europa.eu/DocsRoom/documents/14506/attachments/24/translations/en/renditions/native>
- Cooper, P. (2003). *Governing by Contract: Challenges and Opportunities for Public Managers*. Washington, DC: Congressional Quarterly Press. *Journal of Public Administration Research and Theory*, 14, 438–440.
- Dagher, P. G. (2013). Social Impact Bonds and the Private Benefit Doctrine: Will Participation Jeopardize a Nonprofit's Tax-Exempt Status? *Fordham Law Review*, 81 (6), 3479-3519.
- DCLG (2014). *Qualitative Evaluation of the London Homelessness Social Impact Bond: First Interim Report*. London: Department for Communities and Local Government, Crown Copyright.^[1]_[SEP]
- DCLG (2015). *Qualitative Evaluation of the London Homelessness Social Impact Bond: Second Interim Report*. London: Department for Communities and Local Government, Crown Copyright.
- Deering, J. (2014). A future for probation? *The Howard Journal of Criminal Justice*, 53,1: 1–15.
- Disley, E., Giacomantonio, C., Kruithof, K., & Sim, M. (2015). *The payment by results Social Impact Bond pilot at HMP Peterborough: final process evaluation report*. RAND Europe for Ministry of Justice (UK). Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486512/social-impact-bond-pilot-peterborough-report.pdf
- Disley, E., Rubin, J., Scraggs, E., Burrowes, N., & Culley, D. (2011). *Lessons learned from the planning and early implementation of the Social Impact Bond at HMP Peterborough*. London, UK: RAND Europe.
- Disley, E., & Rubin, J. (2014). *Phase 2 report from the payment by results Social Impact Bond pilot at HMP Peterborough (UK Ministry of Justice Analytical Series)*. London, UK: RAND Europe.

- Dodd, J. A., & Moody, R. (2011). Outcomes, not process: Towards a new model for European funding in an age of austerity, *Journal of Contemporary European Research*, 7, 1:120–8.
- Domberger, S., & Jensen, P. H. (1997). Contracting out by the public sector: theory, evidence, prospects. *Oxford Review of Economic Policy* 13(4): 67–78.
- Ecorys. (2019). *A study into the challenges and benefits of the Social Impact Bond commissioning process in the UK - Final Report*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/844190/A_study_into_the_challenges_and_benefits_of_the_SIB_commissioning_process_Final_Report.pdf
- European Commission (n.d.). Pilot Joint Action Plans. Author. Retrieved from <https://ec.europa.eu/search/?queryText=joint+action+plans+pilot+training2work>
- European Social Fund (n.d.). Simplified Cost Options and Joint Action Plans. Author. Retrieved October 16, 2020, from <https://ec.europa.eu/esf/main.jsp?catId=1434>
- Farr, M. (2015). Co-Production and Value Co-Creation in Outcome-Based Contracting in Public Services. *Public Management Review*. 18. 1-19.
- Fox, C., & Albertson, K. (2011). Payment by results and social impact bonds in the criminal justice sector: New challenges for the concept of evidence-based policy? *Criminology and Criminal Justice*, 11(5), 395-413.
- Fox, C., & Albertson, K. (2012). Is payment by results the most efficient way to address the challenges faced by the criminal justice sector? *Probation Journal*, 59 (4), 355– 373.
- Fraser, A., Tan, S., Lagarde, M. & Mays, N. (2016). Narratives of promise, narratives of caution: A review of the literature on Social Impact Bonds. *Social Policy & Administration*.
- Golden, M., Kohli, J., & Mignonette, S. (2017). A Focus on Cost Savings May Undermine the Influence of Outcomes-Based Funding Mechanisms. In *What Matters: Investing in Results to Build Strong, Vibrant Communities*. Retrieved from https://investinresults.org/sites/default/files/book-chapter/WM_51_Golden-Kohli-Mignonette_0.pdf
- Griffiths, A., & Meinecke, C. (2014). *Introduction to Social Impact Bonds and Early Intervention*, London: Early Intervention Foundation. Retrieved from <https://www.eif.org.uk/resource/social-impact-bonds-and-early-intervention>

- Gustafsson-Wright, E. (2020a). What is the size and scope of the impact bonds market? UK: Brookings. Retrieved from <https://www.brookings.edu/research/what-is-the-size-and-scope-of-the-impact-bonds-market/>
- Gustafsson-Wright E (2020b). Are impact bonds delivering outcomes and paying out returns? UK: Brookings. Retrieved from <https://www.brookings.edu/research/are-impact-bonds-delivering-outcomes-and-paying-out-returns/>
- Gustafsson-Wright E., & Osborn S. (2020a). Do the benefits outweigh the costs of impact bonds? UK: Brookings. Retrieved from <https://www.brookings.edu/research/do-the-benefits-outweigh-the-costs-of-impact-bonds/>
- Gustafsson-Wright E., & Osborn S. (2020b). Are Impact Bonds Reaching the intended populations? UK: Brookings. Retrieved from <https://www.brookings.edu/research/are-impact-bonds-reaching-the-intended-population/>
- Gustafsson-Wright E., Gardiner, S., & Putcha, V. (2015). *The potential and limitations of impact bonds: Lessons from the first five years of experience worldwide*. UK: Brookings. <https://www.brookings.edu/research/the-potential-and-limitations-of-impact-bonds-lessons-from-the-first-five-years-of-experience-worldwide/>
- Gustafsson-Wright E., Osborn S., & Massey M. (2020). Do impact bonds affect the ecosystem of social services delivery and financing? UK: Brookings. Retrieved from <https://www.brookings.edu/research/do-impact-bonds-affect-the-ecosystem-of-social-services-delivery-and-financing/>
- G8 Social Impact Investment Taskforce. (2014). IMPACT INVESTMENT: The invisible heart of markets. Retrieved from www.socialimpactinvestment.org
- Haffar, T. (2014). Communicating Value in Social Impact Bonds: The Role of the Intermediary, MA thesis, Cornell University.
- Haynes, P. (2003). *Managing Complexity in the Public Services*, Maidenhead: Open University Press. [SEP]
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization* 7(Special): 24–52. [SEP]
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69, 1: 3–19.

- Hudson, M., Phillips J., Ray, K., Vegeris, S., & Davidson, R. (2010). The Influence of Outcome Based Contracting in Provider-Led Pathways to Work. DWP Research Report No 638. London: Department for Work and Pensions.^[1]
- Hyndman, N., & McMahon, D. (2011). The Hand of Government in Shaping Accounting and Reporting in the UK Charity Sector. *Public Money & Management* 31 (3): 167_174.^[2]
- ICF Consulting (2019). *Evaluation of the Fair Chance Fund: Final Report*. Ministry of Housing, Communities and Local Government. UK: Author.https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793810/Fair_Chance_Fund_final_report.pdf
- Jackson, E. T. (2013). Evaluating social impact bonds: questions, challenges, innovations, and possibilities in measuring outcomes in impact investing. *Community Development*, 44(5), 608-616.
- Jensen, P., & Stonecash, R. (2005). Incentives and the Efficiency of Public Sector Outsourcing Contracts. *Journal of Economic Surveys*, 19, 767-787.
- Lagarde, M., Wright, M., Nossiter, J., & Mays, N. (2013), *Challenges of Payment-for-Performance in Health Care and Other Public Services – Design, Implementation and Evaluation*. London, UK: PIRU Publications.
- Lamothe, M., & Lamothe, S. (2012). To trust or not to trust? What matters in local government-vendor relationships. *Journal of Public Administration Research and Theory* 22, 867–892.
- Lee, S. J., & Lim, G. G. (2005). The impact of partnership attributes on EDI implementation success. *Information and Management* 42, 503–516.
- Lehner, O., M., & Nicholls, A. (2014). Social finance and crowdfunding for social enterprises: a public–private case study providing legitimacy and leverage. *Venture Capital: An International Journal of Entrepreneurial Finance*, 16(3), 271-286.
- Leventhal, R. (2012). Effecting progress: Using social impact bonds to finance social services, *NYUJL & Business*, 9, 2: 511–34.
- Liebman, J. B. (2011). *Social Impact Bonds: A promising new financing model to accelerate social innovation and improve government performance*, Washington, DC: Center for American Progress.
- Loder, J., Mulgan, G., Reeder, N., Aylott, M., & Shelupanov, A. (2011). *Financing social value: implementing Social Impact Bonds*. London, UK: The Young Foundation.

- Lovan, W. R., Murray, M., & Shaffer, R. (2017). *Participatory governance: Planning, conflict mediation and public decision-making in civil society*. London: Routledge.
- Martikke, S. & Moxham, C. (2010). Public Sector Commissioning: Experiences of Voluntary Organizations Delivering Health and Social Services, *International Journal of Public Administration*, 33, 790-799.
- Martini, M., Cavenago, D. & Mariani, L. (2017). When the Outcome is Employability: Leading Indicators for the Governance of Labour Market Services. *Public Organization Review*. 18. 1-18.
- Maze (2018). *Learnings and Outcomes of the pilot project "Social Impact Bond - Junior Code Academy": Executive Summary (April 2018)*. Portugal: Author.
- Maze (2020a). *Social Impact Bond - Academia de Código Bootcamps (ACB): Intermediate Report (January 2020)*. Portugal: Author.
- Maze (2020b). *Social Impact Bond - Projeto Família: Intermediate Report (January 2020)*. Portugal: Author.
- Maze (2020c). *Social Impact Bond - Faz-te Forward: Intermediate Report (January 2020)*. Portugal: Author.
- McHugh, N., Sinclair, S., Roy, M., Huckfield, L., & Donaldson, C. (2013). Social impact bonds: A wolf in sheep's clothing? *Journal of Poverty and Social Justice*, 21, 3: 247–57.
- McPherson, A.H., & McDonald, S.M. (2010). Measuring the outcomes and impacts of innovation interventions: assessing the role of additionality. *Int. J. Technology, Policy and Management* 10, Nos. 1/2, pp.137–156.
- Millar, R. (2012). Social Enterprise in Health Organisation and Management: Hybridity or Homogeneity?. *Journal of Health Organization and Management*, 26(2), 143- 148.
- Miller, G. (1992). *Managerial dilemmas: The political economy of hierarchy*. New York: Cambridge University Press.
- Miller, R. (2013). "Third Sector Organisations: Unique or Simply Other Qualified Providers?" *Journal of Public Mental Health* 12 (2), 103–113.
- Moore, M.H. (1997). *Creating public value: Strategic Management in Government*. Cambridge: Harvard University Press.

- Moore, M. L., Westley, F. R., & Nicholls, A. (2012). The social finance and social innovation nexus, *Journal of Social Entrepreneurship*, 3, 2: 115–32.
- Moynagh, M. (2010). Investing for public good. Creating social value through social impact bonds. Norfolk, UK: The Tomorrow Project.
- Mulgan, G., Reeder, N., Aylott, M., & Bo'sher, L. (2011). *Social Impact Investment: The Challenge and Opportunity of Social Impact Bonds*, London: Young Foundation.
- Mullins, D., Rees, J., & Meek, R. (2011). Open public services and the third sector: What's the evidence? Centre for Market and Public Organisation, *Research in Public Policy*, no. 13: 15–17.
- NCVO (2014). *Payment by Results and the Voluntary Sector*. London: National Council for Voluntary Organisations. London, UK: Author.
- Ng, I., Maull, R., & Yip N. (2009). "Outcome-Based Contracts as a Driver for Systems Thinking and Service- Dominant Logic in Service Science: Evidence from the Defence Industry." *European Management Journal* 27 (6): 377–387.
- Nicholls, A., & Tomkinson, E. (2013). *The Peterborough Pilot Social Impact Bond*. Oxford, UK: Saïd Business School, Oxford University.
- Niglas, K. (2010). 'The multidimensional model of research methodology: An integrated set of continua', in A. Tashakkori and C. Teddlie (eds) *The Sage Handbook of Mixed Methods in Social and Behavioural Research*. Thousand Oaks, CA: Sage, pp. 215–36.
- O'Flynn, J. (2007). From new public management to public value: Paradigmatic change and managerial implications. *The Australian Journal of Public Administration*, 66(3), 353–366.
- Osborne, D., & Gaebler, T. (1992). *Reinventing Government: How the Entrepreneurial Spirit is Transforming Government*, Reading, MA: Addison Wesley Public Comp.
- Palandijan, T., & Hughes, J. (2014). A strong Field Framework for SIBs. *Stanford Social Innovation Review*. Retrieved from http://www.ssireview.org/blog/entry/a_strong_field_framework_for_sibs
- Pauly, M. & Swanson, A. (2013), *Social Impact Bonds in Nonprofit Health Care: New Product or New Package?* NBER Working Paper No. 18991.
- Poister, T. H. (2003). *Measuring performance in public and nonprofit organisations*. Chichester: Wiley.

- Polonsky, M. J., & Grau, S. L. (2008). Evaluating the social value of charitable organisations: A conceptual foundation. *Journal of Macromarketing*, 28(2), 130–140.
- Polonsky, M., & Grau, S. L. (2011). Assessing the social impact of charitable organisations – four alternative approaches. *International Journal of Nonprofit and Voluntary Sector Marketing*, 16(2), 195–211.
- Porter, M. E. & Kramer, M. R. (2006), The link between competitive advantage and corporate social responsibility, *Harvard Business Review*, 84, 12: 78–92.
- Portugal Social Innovation. (n.d.). Cuidar de Quem Cuida. Portugal: Author. Retrieved from <https://inovacaosocial.portugal2020.pt/project/cuidar-de-quem-cuida/>
- Power, M. (1999). *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.
- Ridder, H-G., Hoon, C. & McCandless, A. (2014). ‘Entering a dialogue: Positioning case study findings towards theory’, *British Journal of Management*, Vol. 25, No. 2, pp. 373–87.
- Rudd, T., Nicoletti, E., Misner, K. & Bonsu, J. (2013), *Financing Promising Evidence- Based Programs: Early Lessons from the New York City Social Impact Bond*, New York, NY: MDRC, <http://www.mdrc.org/publication/financing-promising-evidence-based-programs/file-full>
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods For Business Students* (7th ed.). Harlow, UK: Pearson Education Limited.^{[1][2]}
- Segal, G. (2002). Performance measures improve city governments. *PA Times*, 25(5), pp. 1,9.
- Shutes, I., & Taylor, R. (2014). “Conditionality and the Financing of Employment Services – Implications for the Social Divisions of Work and Welfare.” *Social Policy & Administration* 48 (2): 204–220.
- Sinclair, S., McHugh, N., Huckfield, L., Roy, M. & Donaldson, C. (2014), Social impact bonds: Shifting the boundaries of citizenship. In K. Farnsworth, Z. Irving and M. Fenger (eds), *Social Policy Review 26: Analysis and Debate in Social Policy*, pp. 119–36.
- Social Finance (2015a). *Technical Guide: Designing outcome metrics*. London, UK: Author. Retrieved from <https://www.socialfinance.org.uk/resources/publications/technical-guide-designing-outcome-metrics>
- Social Finance (2015b). *Portuguese Social Investment Taskforce: A blueprint for Portugal’s emerging social investment market*. UK: Author. Retrieved from

<https://www.socialfinance.org.uk/resources/publications/portuguese-social-investment-taskforce-blueprint-portugal%E2%80%99s-emerging-social>

- Sorci, P. (2019) Governing Local Area Development through Dynamic Performance Management. *Public Organiz Rev* 19, 159–177.
- Stoesz, D. (2014), Evidence-based policy reorganizing social services through accountable care organizations and social impact bonds, *Research on Social Work Practice*, 24, 2: 181–5.
- Sturla, K., Buddy Shah, N., & McManus, J. (2018). The Great DIB-ate: Measurement for Development Impact Bonds. *Stanford Social Innovation Review*.
- Tan, S., Fraser, A., Giacomantonio, C., Kruthof, K., Sim, M., Lagarde, M., Disley, E., Rubin, J., & Mays, N. (2015). *An Evaluation of Social Impact Bonds in Health and Social Care*, London: PIRU, London School of Hygiene & Tropical Medicine and RAND Europe.
- Tashakkori, A. & Teddlie, C. (2010) *The Sage Handbook of Mixed Methods in Social and Behavioural Research* (2nd edn). Thousand Oaks, CA: Sage.
- Warner, M. (2012), Profiting from Public Value? The Case of Social Impact Bonds, Draft Paper prepared for Creating Public Value Conference, 20–22 September, University of Minnesota, USA.
- Warner, M. (2013). Private finance for public goods: social impact bonds. *Journal of Economic Policy Reform*, 16(4), 303-319.
- Whitford, A. B., & Miller, G. J. (2007). The principal’s moral hazard: Constraints on the use of incentives in hierarchy. *Journal of Public Administration Research and Theory* 17:213–33.
- Williams, J. (2019). From visions of promise to signs of struggle: social impact bonds in Canada, the US, and UK. Government Outcomes Lab, Blavatnik School of Government. <https://golab.bsg.ox.ac.uk/community/blogs/visions-promise-signs-struggle-social-impact-bonds-canada-us-and-uk/>
- Yin, R.K. (2014). *Case Study Research: Design and Method* (5th edn). London: Sage. Useful for Case Study strategy.

9. APPENDIX

9.1 Interview Guide

Hello,

First of all I would like to thank you for agreeing to be part this interview. As I explained briefly in the email, the focus of my thesis will be on outcome-based contracts, namely in terms of metrics, targets and respective incentives.

This interview is part of a series of case studies on OBC (mostly SIBs) in Portugal, in the context of my Master's thesis with Catholic University of Lisbon and Bocconi University in Milan.

I am fully aware of the need to treat the data provided with the highest integrity and the desired confidentiality. No source, individual or organization, will be identified without the originator's permission. A copy of the transcript of the interview in question, and a copy of the final report of the thesis results will be sent to all study participants, if it is of interest.

I would also like to request your authorization to record the interview, which is used exclusively for the purposes of this study.

1. Outcome-Based Contracting

(1) What are the contracting out contracts/agreements?

2. Metrics, Targets and Indicators

(2.1) Which metrics/indicators are linked to the payment? Do they have outcome targets?

(2.2) How are the metrics and thresholds defined? Which parties are involved in defining them?

(2.3) How was the outcome payment anchored, was it value creation or intervention costs?

(2.4) Did any of the stakeholders push for repayments to also be tied with outputs?

(2.5) How to demonstrate project's social impact? Is there a plan for an impact evaluation? (2.5.1) Perhaps involving a control group?

(2.6) Are there other additional outcomes (not related to payments) that are controlled for?

3. Monitoring

(3.1) How are these indicators and performance in general verified?

(3.2) Is it a continuous or periodic process (or both)?

(3.3) Who is responsible for the monitoring? And why?

4. Incentives tied to Outcome targets

(4.1) What are the incentives for the service providers to achieve outcome targets? Performance incentives, penalties?

(4.2) Was there any evidence of thresholds being "too high" or "too low"?

(4.2.1) (For too low) Was there any incentive to surpass the threshold?

(4.2.2) (For too high) Was there a higher focus on "easier to help" beneficiaries?

(4.2.3) (For too high) Were there any incentives to work with "harder to help" beneficiaries?

(4.3) Was there room for revisiting/adjusting the targets during the project?

(4.3.0) (if yes) Was there room for adjustments a result of Covid-19?

(4.3.1) Was it discussed the possibility of payments being tied to progress towards the outcomes (soft outcomes), either in negotiations or afterwards?

(4.3.1.1) Was there any evidence of trading hard outcomes (qualifications) with softer ones (improvements)?

(4.4) Was there a mismatch between provider's reported expenses and the amount recognized by the financial auditor?

(4.4.1) Would you consider the providers were aloud a competitive margin on this project?

(4.5) Were there any particular benefits from the involvement of a public entity as a commissioner?

5. Outcome-Based Contracting Drivers

(5.1) Were there any unexpected challenges in the use of this outcome based mechanism?

(5.2) Would you be interested in using OBC in the future? And why?

(5.3) What are, in your opinion, the main barriers for scale of this specific OBC mechanism?

6. Outcome-Based Contracting Effects

(6.1) Do you feel that such mechanisms disproportionately benefit the incumbent providers and/or bigger providers in comparison to smaller organizations?

9.2 Organizations Description

9.2.1 *Portugal Social Innovation (PSI)*

As the entity responsible for the outcome-based payments in most of the SIBs launched in Portugal, Portugal Social Innovation (PSI) is a mission structure created as a Resolution of the Council of Ministers, financed by the European Social Fund within the scope of the Portugal 2020 Partnership Agreement. A Mission Structure is created by the council of ministers as a unit in the public administration of Portugal that is created to accomplish a specific mission. The mission in this case is to stimulate innovative solutions and intervention models, by establishing instruments, such as social impact bonds, in order to provide answers for priority social problems and complement existing responses in several social policy areas such as health, education, justice and gender equality.

9.2.2 *Maze Impact*

MAZE is an impact investment company created by the Calouste Gulbenkian Foundation in 2013. MAZE invests in businesses that promote effective solutions for solving social and environmental problems, supports entrepreneurs to expand their solutions to the European market and works with the sector to improve the impact of public service delivery. MAZE works in close partnership with international partners such as the European Venture Philanthropy Association, Social Finance UK and MaRS Center for Impact Investing (Canada).

9.2.3 *Calouste Gulbenkian Foundation*

The Calouste Gulbenkian Foundation is a Portuguese institution of private law and public utility, whose statutory purposes are Art, Charity, Science and Education.

With more than 50 years of existence, the Calouste Gulbenkian Foundation is one of the most important European foundations, developing a vast activity in Portugal and abroad through its

own projects, or in partnership with other entities, and through the granting of grants and scholarships. The Foundation has delegations in Paris and London.

9.2.4 TESE (Association for Development)

TESE is a Non Governmental Development Organisation (NGO) that uses Social Innovation concept as an anchor of its operations in Portugal and in developing countries. The main objectives of the organization are to investigate, create and implement, building up from a cycle of socially innovative and sustainable solutions in response to traditional and emerging needs.

9.2.5 ASSOP

ASSOP - Association Shared Services & Outsourcing is an association that was created for the single purpose of investing in the SIB – Code Academy Bootcamps. Its members include the Municipality of Fundão, where the project is implemented, as well as several undisclosed IT companies operating in the area.

9.2.6 Movimento Defesa da Vida (MDV)

Movimento Defesa da Vida ("Movement for the Defense of Life") is a Private Institution of Social Solidarity (IPSS), both denominational and apolitical, founded in 1977. It acts in the fight for Human Rights, in the Promotion of Women, in the affirmation of Fundamental Ethical Values, in the search for more quality of and in promoting the conjugal and family relationship. Their scope of action covers the entire Portuguese territory. MDV has several projects with the objective of supporting the families with whom it works in an increasingly personalized way, of which we can highlight the Family Project. This is a pioneering project in Portugal, launched by MDV in 1996, in the area of intensive intervention with families with children at risk.