Performance Assessment in the Public Sector – A Land of Failed Kings

Patricio Rojas, Ph.D. – projas.ese@uandes.cl

ESE Business School, Universidad de Los Andes

A paper accepted for presentation at the 2015 APPAM Conference
Performance Assessment in the Public Sector - A Land of Failed Kings

ABSTRACT
This exploratory study aims to provide new concepts and contribute with theory extension regarding the dynamics of assessment in the public sector. The paper begins by presenting the paradoxical situation of governments that even though they are accomplishing their electoral promises and have their economic and social indicators at their best, also have their approval poll ratings plummeted to record lows. To address this paradox the paper draws on two case studies to illustrate the enhanced framework that results from combining concepts from the Interpretation literature about the cognitive process by which people attach meaning to information, with the contextual political dynamics described by the Public Sector literature.

The field data collected and analyzed in this study suggest that the lack of positive response of citizens to their governments’ achievements would be partly explained by the “Credibility Threshold”, that is a nonlinear relationship between the performance measures reported by public agencies and how these measures are interpreted by the general public. The “Credibility Threshold” is a side effect of the “confidence gap”, that is the declining confidence shown by citizens in their political, social, and economic institutions. The effects of the “Credibility Threshold” are not limited to politically elected officers, but also affect public agencies carrying out the most diverse types of services and activities.

This paper contributes to the increasing research on the intersection of the public sector and the interpretation literatures, and provides new concepts that enhance our understanding of the dynamics of assessment in political environments.

Keywords:
Public sector; performance assessment; confidence gap; credibility threshold; interpretation asymmetries; politicization dynamics
INTRODUCTION
Imagine that you are the President of a small country. While nations around the world are suffering one of the worst economic crises in history, your country’s economy is growing at a healthy 5 per cent, among the fastest in the OECD, and more than 500 thousand jobs have been created in the year and a half since you have assumed office, with almost 3 percent increase in wages. Furthermore, presidents of neighboring countries say “This country is an example for us, it has the best indicators in the continent not only economic, but also social. It would be great to be like them”. Your fellow country people should love you, shouldn’t they?

The paradox of Mr. Piñera, former president of Chile was that under the same conditions described above he had the dubious honor of being the most unpopular president since Chile’s return to democracy. When asked by the Financial Times whether low poll ratings may compromise his ability to govern and do the things he planned to do, he answered: “Don’t judge us by our intentions, but by our results”. He argued that his government had a very ambitious program, and that they were working very hard to honor their promises. “But people always want more, and are impatient”, he said (Rathbone, Webber, & Financial Times, 2011).

Since the Financial Times’ interview things did not improved much for President Piñera. By September 2012 his popularity has stagnated around a 35 percent approval rating, with the added worrisome sign that the problem was not just limited to the person of the President, but spilled over into all domains of government. When asked about the Chilean case, Colombian President Juan Manuel Santos said that it was a mystery. He mentioned that this paradox was recently a matter of discussion with many South American leaders, and that the Chilean contradiction was a

---

1 President Piñera’s electoral program promised to accomplish a series of clearly quantifiable targets for 39 specific goals in seven main domains. For instance, some promises were achieving 6% GDP annual growth, increasing GDP per capita from to 16,000 to 22,000 by 2018, and creating one million new and good jobs during his term in office.

2 A survey by Adimark asks respondents to rate government job in nine domains: Education, Healthcare, Delinquency, Environment Protection, Corruption, Employment, Public Transport, the Economy, and International Relations. A comparison of Piñera’s first semester in office with his fifth one shows a significant and widespread reduction of approval in all but two dimensions. The average reduction is 16 percentage points, with the largest reduction of being 33 points.
warning signal for the region. “There is something we are ignoring and maybe we will realize it too late. This is a very interesting phenomenon to analyze”, he said (Quezada, 2011).

This case highlights that the dynamics of assessment in the public domain are not well understood, in spite of the relevance of the Public Sector. It is important to recall that public expending represents a large share of the economy of most countries, for instance around 40% of the GDP in the United States, and around 25% in Chile. Furthermore, in addition to its direct impact on the economy, the public domain also affects countries’ economic development, firm’s competitiveness, and the wellbeing of the whole population. To explore the dynamics of assessment in the public sector, this paper uses concepts from two literature threads and combines them with insights coming from field data. Even though each of the two literatures considered – (a) the Interpretation literature, and (b) the Public Sector literature – contribute with key concepts to understand assessment in the public sector, none of them provides a comprehensive view by itself. The review of the literature starts with that on Interpretation, which explains the cognitive processes through with meaning is attached to information and argues that people confronted with the same information may reach fundamentally different conclusions. Parties interested to influence assessment do not need to manipulate information; they can do so through framing and manipulating interpretations, taking advantage of people’s tendency to resort to heuristics and biases when dealing with complex information. The review continues with the literature on the Public Sector that informs on the contextual aspects of the public sector, such as the dynamics and effects of politicization that influence the nature of assessment in that domain. Each literature contributes with a particular emphasis and specific set of concepts, that together result in a more comprehensive framework to understand the dynamics of assessment in the public domain. The paper uses two cases to illustrate the results of

---

3 As a reference, during 2014 public expending in the United States was 41.6% of GDP, while in Chile it was 23.2%.
combining ideas from the two literatures described above. This paper also describes and provides evidence of the “Credibility Threshold”, that is a non-linear relationship between the performance measures reported by public agencies, and how these measures are interpreted by the general public. This phenomenon results from the interaction of (i) the heuristics and biases associated with interpretations, and the (ii) political dynamics that operate in the public sector.

**RESEARCH DESIGN**

This study is part of a larger research effort that explores processes of change and institutional innovation in public sector organizations. The project was structured as a multi-case study that combined data collected from interviews\(^4\) and information coming from other public and private sources. In this paper I chose to use the data of Chilean public sector units, as in that country a new ruling-coalitions took office in 2010. This was the first political transition in 20 years\(^5\) and provided a good vantage point to study the dynamics of assessment of public agencies, as those who formerly run the country had to perform a supervisory role as political opposition.

**LITERATURE REVIEW**

**Literature on Interpretation**

This literature stream is mostly centered on the cognitive processes that influence the way people attach meaning to the information they receive, and the review begins referring to heuristics and points out that they can result in systematic and predictable errors in judgment, known as biases. Next, it describes how interested parties may attempt to take advantage of other people’s biases.

\(^4\) All interview protocols used in this study were designed following the suggestions and guidelines of Seidman (2006), Kvale and Brinkmann (2008), and McCracken (1988).

\(^5\) Even though government changes are quite frequent around the world, this one is special from a research perspective because it is a natural experiment on management style change in the public sector, as several new public officers were recruited from the private sector with the idea of bringing breath of fresh air to government bureaucracy.
Assessment, references and heuristics

A central aspect of assessment is determining whether things are getting better or worse. For instance, Sapolsky (2004) describes parents of children who had a 25 percent chance of dying of cancer but curiously these parents show only moderate levels of stress. How could that be possible? The reason is that these children are all in remission, after a period in which the odds of death had been far higher. Surprisingly, a one in four chance of one’s child dying can be good news. These results are part of a large body of literature indicating that assessment of actual situations and possible outcomes depends on the reference point. Situations and outcomes are seen as losses when they are below the reference level and as gains when they are above. Tversky and Kahneman’s (1979; 1991) extended the referent-dependent ideas addressing the phenomenon of loss aversion. Their model was able to account for the endowment effect and the status quo bias, arguing that both phenomena are explained by losses looming larger than equivalent gains. Recent research by political scientists has extended Tversky and Kahneman’s (1991) reference-dependent model of decision-making to explore the determinants of political behavior, identifying the “keeping up with Joneses” phenomenon. Killian et al. (2008) studied the considerations that affect individuals’ decision of whether to vote in elections. They found that citizens are concerned about whether they are keeping up with the Joneses, that is, with the average citizen. Individuals who perceive that they are falling behind economically relative to the rest of society are spurred to action and thus are more willing to vote during elections. In contrast, those who are getting ahead or keeping up with the Joneses are less likely to vote. Interestingly, these researchers show that rough perceptions of losses or disadvantages are enough to trigger people’s active engagement.

---

6 The endowment effect refers to the observation that the loss of utility of giving up a something is tends to be greater than the utility gain associated with receiving it.
Another relevant aspect of assessment in addition to reference points is how people living in a complex world make sense of the information they get. The literature on judgment and decision-making, in particular the work of Kahneman, Tversky, Slovic and colleagues (Kahneman & Tversky, 1979; Kahneman, 2011; Slovic, 1999; Slovic, Finucane, Peters, & MacGregor, 2007; Tversky & Kahneman, 1974; Tversky & Kahneman, 1991) has identified the preeminent role of heuristics and biases in information processing. Heuristics are simple procedures that help people find adequate, though often imperfect, answers to difficult questions. When faced with a difficult question, people often answer an easier one instead, usually without noticing the substitution. For instance the difficult question “How good this candidate will be in office?” is replaced by the easier one “How much do I like this candidate?” These intuitive mental short-cuts are generally effective and serve us quiet well, but can also lead us to biases, i.e. systematic and predictable errors in judgment. For instance, Slovic et al. (1999; 2007) identified an “affect heuristic” in which people make judgments and decisions by consulting their emotions. They observed that while the advantages and disadvantages from various activities and technologies (e.g., nuclear power, commercial aviation) tend to be positively associated in real life, they are inversely correlated in people’s minds. Hence, if people like a choice, they will perceive its advantages are substantial and its costs more manageable than the costs of alternatives. Research by Wells et al. (2009) provides evidence that this heuristic not only operates in the technical domain that Slovic and colleagues studied, but also in politically charged contexts. They studied voters participating in an initiative election that repealed a workplace safety regulation. Their results show that citizens that hold more partisan views tend to develop a distorted set of beliefs and become systematically wrong over time, coming to believe things that are not true. Partisan voters tend to

---

7 Higher perceived benefit is associated with lower perceived disadvantages; lower perceived benefit is associated with higher perceived disadvantage.
refine their preferences by seeking out and accepting only the messages of like-minded sources, while rejecting or ignoring information that challenges their preconceived notions. In contrast, non-partisan people are not affected by this biasing dynamics. Along similar lines, Nyhan and Reif (2010) find that people with committed partisan views are unwilling to revise their beliefs in the face of corrective information, and attempts to correct those mistaken beliefs only make matters worse, creating a backfire effect that actually increased misperceptions among partisan individuals. Another cognitive mechanism is the “availability heuristic” (Kahneman, 2011; Tversky & Kahneman, 1974). This concept describes the process of judging the frequency and importance of an issue by the ease and fluency with which instances come to mind. People tend to generalize from easily available information, regardless of its validity, and rely on personal experiences, publicly available information, dramatic news, vivid examples, and topics frequently mentioned in conversations to assess the relative importance and likelihood of an issue. Like other heuristics of judgment, this one substitutes one question for another: one wishes to estimate how likely is to be robbed at down-town, but one reports an impression based on a recent workmate’s experience or by salient recent news.

When working together, loss aversion, the affect heuristic, and the availability heuristic can create “availability cascades”, which are self-reinforcing processes of belief formation in which an issue charged with emotional intensity becomes available to an increasing larger number of people (Kuran & Sunstein, 1999; C. R. Sunstein, 2005; C. Sunstein, 2006). An availability cascade can be triggered by a media story about a problem that catches the attention of a segment of the public, which becomes aroused and concerned. This emotional reaction becomes a story in itself, attracting additional coverage in the media, which in turn triggers even greater worry and involvement as the issue is perceived increasingly important. This process may also be enhanced by "availability entrepreneurs" (Kuran & Sunstein, 1999), individuals or interest groups that can
advance their interests by maintaining and increasing the mass media coverage of the issue, calling people's attention to specific aspects, and providing their favored interpretations of the matter. As the problem is constantly reinforced by the media, it becomes politically important because it is on everyone’s mind, and the response of the government and politicians is guided by the intensity of public sentiment. Through this process, availability cascades end up resetting priorities and distorting the allocation of public resources.

**Framing, biases, and its limitations**

As we have seen, the literature on interpretation is full of examples of factors that we human beings are mostly unaware of but have powerful influences on our thinking. Loss aversion, the status quo bias, the affect heuristic, and the availability heuristic are all fundamental aspects of how people perceive and process their experiences. We are able to make sense of partial information in a complex world because our brain is highly efficient in constructing the best possible story using the limited information available. The more coherent the story our brain manages to create, the more confident we are of our judgments. The amount and quality of the data on which the story is based are largely irrelevant, and thus we are likely to jumping to conclusions on the basis of limited evidence. This weakness leave us susceptible to be influence by third parties – as the “availability entrepreneurs” mentioned above – who can influence our decisions without distorting or suppressing information, merely by the framing of outcomes and contingencies. For instance, it has been found that citizens’ attitudes regarding social welfare programs provisioning of benefits to the poor depend on how programs are framed. On the one hand, people decline to share with needy individuals whose needs reflect a lack of effort, but on the other hand they feel morally obliged to share with individuals whose needs reflect bad luck. Lack of effort in alleviating one's own need triggers anger; while effort triggers compassion towards the needy (Petersen, 2011).
These framing effects may be fortuitous, without participants being aware of the impact of the frame in the ultimate decision, but they can also the purposively exploited to manipulate the relative attractiveness of options. Effective manipulation relies on people’s unchallenged biases, so as to make it easy for them to believe what they already are predisposed to believe. Not surprisingly, attempts to frame situations are common both in the business sector and in politics. People trying to make sense of novel situations would respond positively to positions or choices framed in terms of socially valuable ideals such as freedom, choice, rights, fairness or public officers’ accountability, while they would react negatively to those framed in terms of socially disapproved concepts such as discrimination, inequality, or benefits to the privileged few. Parties advocating contentious positions or issues actively use this human tendency to advance their agendas, for instance the concept of “rights” used by the women’s movement (Meyer & Whittier, 1994), or the idea of “fairness” leveraged by those wanting the extension of employees’ healthcare benefits to domestic partners (Briscoe & Safford, 2008).

**Main insights on assessment from the interpretation literature**

So far the discussion can be summarized as follows. When judging situations and making assessments, people follow a reference-dependent decision making approach that focuses attention on differences regarding a baseline, which constitutes the actual expectation. Our cognitive machinery evolved with the primary focus of noticing and avoiding dangers, and thus decisions are mostly guided by the principle of loss aversion, with opportunities for gains being less salient. People use heuristics as shortcuts to deal with complexity. Two are particularly relevant: the affect bias, and the availability heuristic. Parties interested in advancing their positions do not need to manipulate sources of information to do so because biases allow the framing and manipulation of interpretation, which may be just enough.

To conclude, the literature on Interpretation indicates that human cognitive processes lead
individuals to reach fundamentally different conclusions despite having access to the same information. Moreover, the same person may reach different conclusions depending on how the information is framed. Hence, even in the hypothetical case of having perfect information, there will still be problems due to non-intended and unpredictable behaviors induced by the way in which people interpret information.

**Literature on the Public Sector**

The Public Sector literature highlights the preponderant role of the general public as ultimate assessor of politically elected officers. The fact that representatives are voted in office results in some politicians doing whatever is necessary so as to maintain or increase the support of ordinary citizens. The politicization dynamics that result from the struggle of political competition play a relevant role in this literature stream, which describes some politicians’ use of exaggerations and misleading statements, and the resulting erosion of general public trust in politicians and public officers, known as the “confidence gap”. Moreover, politicization episodes impair assessment as judgment regarding an agency’ performance no longer considers the full dimensionality of its activities and outcomes, but just the single aspect politicized.

**The nature of public organizations**

There is near-unanimity among public sector scholars that public organizations are substantially different from their private counterparts, particularly in two aspects: (i) the absence of a “bottom line”, and (ii) the fact that their performance is judged by multiple constituencies. Performance monitoring in the private sector is argued to be simpler and straightforward, as financial returns provide a unifying concept that allows focusing attention into just an unambiguous and one-dimensional measure (e.g. Smith, 1993; Stewart & Walsh, 1994). In contrast, in the public sector the absence of clear performance measures is said to limit public managers’ capability to control and develop their organizations, forcing them to play a more political and expository role so as to
get support for their programs and obtain continued funding (Rainey, 2009). In turn, the second aspect emphasizes that public agencies can be held accountable by numerous external actors with often-conflicting interests; each using their own criteria to judged agencies’ services (e.g. Boyne, 2004; Dixit, 2002; Fernandez & Pitts, 2007; Propper & Wilson, 2003; Sanderson, 2001). In consequence, public agencies are highly exposed to the interventions, interruptions, and involvement of numerous external parties engaged in non-cooperative tactics to influence their actions (Dixit, 2002). Not surprisingly, there is virtual universal agreement among public sector scholars that public organizations’ goals are more ambiguous and complex that those of private firms (Rainey & Bozeman, 2000).

**Democracy and political incentives**

In democratic societies every citizen has the right to participate in decisions concerning the public domain (Deleon & Deleon, 2002). Political parties are key mediators between citizens and the state, because they aggregate individual preferences into political demands to public agencies and also in programs with potential majority appeal to become a viable alternative of government (Alford & Friedland, 1985). Hence, political parties are continuously trying to gain the support of ordinary citizens so as to be better positioned to take office or to be re-elected at next election time, and this political competition results in a permanent state of uncertainty and struggle, with those who currently run the country trying to avoid losing legitimacy and popular support, while those who nowadays perform a supervisory role as political opposition eagerly trying to amass the votes of disaffected constituencies so as to increase the chances of getting into office during next elections. For doing so politicians compete with each other in many different ways: they trade on personal popularity; they attack the integrity and character of their opponents; they

---

8 According to Lipset (1960), democracy is "a social mechanism for resolving the problem of societal decision making among conflicting interest groups with minimum force and maximal consensus".
exploit or are victims of biased coverage in the news media; they play dirty tricks (Laver & Sergenti, 2011; Moe, 1997). In consequence, conflict is structurally installed in this competitive context, as political opposition usually has formal oversight rights (e.g. opposition representatives in the legislative branch) and in addition has the incentive to “help” government fail. Along these lines, one interviewee commented: “the major difference with respect to the private sector is that here, in the public sector, your main competitors – the political opposition - is perusing and controlling your expenses and results”. Problems or failures in public agencies are opportunities susceptible to be exploited politically by opposition, not only by their own relevance as problems but also as part of the larger political game (Brändström & Kuipers, 2003). Hence, issues affecting a specific public agency can become the arena in which political parties and advocacy groups dispute the larger battle for getting public support and increase their influence, power and legitimacy.

**Legitimacy proficiency**

Legitimacy is an important concept to understand organizations and their relation with their social environment (Parsons, 1956; Pfeffer & Salancik, 1978). Organizations are part of the larger social system and depend on society’s support for their continued subsistence. Since organizations consume society’s resources, organizations are continually being assessed on the legitimacy and appropriateness of their activities and the usefulness of their output within a socially constructed system of norms, values, and beliefs (Parsons, 1956; Pfeffer & Salancik, 1978; Suchman, 1995). If objections are raised by some individuals regarding the legitimacy of an organization’s activities, the illegitimacy problem will be a function of how widely the objections are dispersed and whether sufficient public interest is generated to support disapproval. Brändström and Kuipers’ (2003) study of selective politicization in the public sector highlights that failures in public agencies are commonplace, but they can be interpreted as normal
incidents or as political crises. Failures rarely become political battles by themselves. On the contrary, there are actors driving politicization processes as there are gains and losses at stake. Brändström and Kuipers’ evidence illustrates how actions and events become “politicized” when influential actors in the political arena succeed in framing them as blameworthy violations of crucial public values⁹, and thus events become a matter of societal debate, which may result in the delegitimization of key programs, individuals, and organizations.

A noteworthy aspect of politicization processes is that they collapse assessment into a single performance dimension of worth, whatever represents the central issue in dispute (e.g. Dixit, 2002). Instead of taking into account the whole dimensionality of the problem, parties involved in discussing and judging the issue will be focused on the single aspect that apparently violates essential social values. This reductionist assessment is facilitated by bounded rationality, which according to Herbert Simon (1995) also affects political behavior. He argues that attention will be usually focused just on the most salient aspect of decisions, a condition that can be strategically use by political actors. According to Simon, the main purpose of political campaigning and communicating efforts is not to change people's values or beliefs but to turn their attention to those issues on which majority views may favor the campaigner, letting others aspects outside the focus of public attention. Hence, the definition of what is legitimate will be continually shifting and evolving, partially in response to politicization events triggered by political actors.

**From political advocacy to the manipulation of interpretations**

In principle political attempts to influence the general publics’ perception are not detrimental per se. Diverse constituencies may hold different but valid views regarding particular issues, and political debate may enhance the full comprehension of the problem and reinforce the legitimacy of whatever decision is made. But in practice influence attempts can mutate to the manipulation

⁹ For instance, justice, democracy, equality and liberty
of interpretation and cheating that results in the systematic destruction of trust in politicians, government, and public agencies. Along these lines, recent research indicates that false beliefs and misinterpretations are all too easy to create. According to Nyhan (2010), misleading statements by politicians and flawed media coverage contribute to widespread misperception by the general public in matters as varied as public healthcare, social security, and foreign policy, among others. For instance, Jerit and Barabas’ (2006) study shows that one third of Americans got to have wrong beliefs about Social Security policy due to biases political rhetoric. Even more, after controlling for personal-level factors such as education, they found that the likelihood of people getting correct beliefs nearly doubled when not exposed to politically biased assertions. Interestingly, outright lying has been found to be much less common in political rhetoric than one may think. Most so-called political lying are not exactly bald lies, but exaggerations and misleading statements that lead people to make false inferences (Clemmitt, 2011). Thus, citizens may find themselves confused and ill-informed about complex public matters, and fall prey to manipulation at hands of interest groups and political actors (Gastil, Reedy, & Wells, 2007). But although some voters may get mislead in the short run, manipulation attempts by politicians do not go unnoticed. A survey by Ramsay and colleagues (2010) concluded that over 90 percent of American voters believe they encountered false or misleading information during 2010 elections, and 54 percent said the amount of misleading information was greater than any they had heard or seen in past elections. The resulting effect is that political dishonesty can boost public distrust and skepticism, creating a crisis of confidence not limited to politicians but that spreads to other public institutions (Lipset & Schneider, 1983; 1987). Recent evidence shows how this phenomenon can reach unprecedented highs. According to a survey by the Pew Research Center in year 2010, record levels of discontent with politicians poisoned trust in U.S. government. Just 24% of citizens had positive view on Congress and about as many said the same about the federal
government. Favorable ratings for public agencies have also felt significantly, with a decline particularly large for some agencies such as the Department of Education, the Food and Drug Administration, and the Social Security Administration among others. Furthermore, nearly 80 percent of U.S. citizens said they do not trust government is doing what is right, a level that is the highest in a half-century (Pew Research, 2010). My analysis of data from years 2006 to 2012 - not reported - shows a similar pattern in the Chilean public domain. Increased levels of politicization significantly reduce support and augment rejection to political parties, and in addition negatively affect perceptions regarding whether government and public agencies are doing well their job in matters as varied as education, public healthcare, public transport, delinquency prevention, and unemployment among others.

Several commentators (e.g. Lipset & Schneider, 1983; 1987) have pointed out that the effects of this “confidence gap” – i.e. the declining confidence shown by citizens in their political, social, and economic institutions – are augmented by several interested parties that have an incentive to sustain the belief that things are worse than they really are. Their assertion is related with the fact that politicians and the media have to come with really outlandish stuff to get people’s attention (Clemmitt, 2011), and several have found that “bad news” are more effective than “good news” in competing for the attention of the general public (Lipset & Schneider, 1983). This phenomenon is complicated by the fragility of trust. Trust is typically created slowly, but in can be destroyed in a moment. And once trust is lost, it can take a long time to rebuild it (Slovic, 1999). In consequence, for public agencies and their managers, the problem is that the public expect them to fulfill promises, meet increasing demands, and improve service standards while being increasingly skeptical about their claims to have done so (Clarke, 2005).

Main insights on assessment from the public sector literature

The distinctive characteristic of public sector organizations is that they are likely to fail prey to
politicization processes. A first main effect of politicization is that organizational assessment no longer considers the full dimensionality of organizations’ activities and outcomes, but is framed so as attention and judgment are centered in a single dimension of worth, the one associated with the supposed blameworthy violation of accepted social values. A second main effect of politicization is that it fosters the “confidence gap”, a sustained climate of suspicion and mistrust in public organizations and actors: politicians, government, public agencies, and public officers, no-one gets away. Indeed, referring to the U.S., Lipset and Schneider (1987) commented that “there seems to be no end to the series of events that create and sustain the confidence gap”.

TWO CASES

The case of the Housing Reconstruction Program

The Housing Reconstruction Program (HRP) was one of the Chilean government initiatives under the direction of the Housing Ministry - MINVU

10 to address the drama of the 800,000 people affected by the February 2010 earthquake and tsunami that shaken continental and peninsular Chilean territory. The magnitude of the earthquake was huge: it is the fifth major one registered in human history, and caused widespread devastation, destroying patrimony for US 30 billion

11. It affected 900 cities, towns, and villages, and damaged or destroyed more than 370,000 dwellings. 1.25 million children lost their schools, 140 hospitals and medical centers were damaged, and 1,554 kilometers of roadways were destroyed.

A new administration led by president Piñera assumed office in March 11th, just two weeks after the earthquake. One of the first measures taken was to announce a reconstruction program and create a voluntary register of affected citizens. From this process it was estimated that there were

10 MINVU is the acronym for “Ministerio de Vivienda y Urbanismo”.
11 U.S. 30 billion are equivalent to 18 percent of Chilean GDP.
220,000 families eligible for subsidy\textsuperscript{12}, and that the whole housing reconstruction program will require public funds for U.S. 2.5 billion. The response to the housing problem was organized in two phases. The first one addressed the urgent need for shelter of those vulnerable families affected by the earthquake due to the incoming winter, cold and rain. In 100 days 50,000\textsuperscript{13} mediaguas - prefabricated structures that function as temporary shelters - were installed. As a reference, this amount is larger than that of all mediaguas built in the prior 10 year period in Chile. The second phase, the Housing Reconstruction Program, is probably one of the greatest challenges of planning and implementation that Chile has experienced in its 200 year history.

\textit{The technical challenge}

The HRP had two main technical challenges: the (i) widespread damage cause by the earthquake, together with the enormous variety of problems affecting each dwelling and (ii) the fact that nearly 95\% of the new dwellings would need to be constructed in the same terrain families were already living. The usual solution to address housing deficit had been constructing a large number of new dwellings in new terrains, an approach that provides economies of scale and guarantees that there will be no problems with legal property rights of dwellings and land. The problem posed by the earthquake made this solution unfeasible. A “one size fits all” approach would not solve appropriately the different circumstances in which each family lived. For instance, there were affected citizens living scattered, in highly isolated zones. Other lived in their city’s historic centre, characterized by long and narrow plots of land, with narrow street frontages. Some houses were inhabitable and thus full reconstruction was needed, while others could be repaired. Some people were proprietors, while others were not. Furthermore, it was

\textsuperscript{12} Wisdom from prior natural disasters indicates that public assistance can be subject to abuses and exploitation by unscrupulous individuals. For instance, after the 2007 earthquake that affected 3,700 families in northern Chile, in some towns there were more subsidy applications than inhabitants.

\textsuperscript{13} This amount is by itself a record for Chilean standards, but even more relevant when taking into account the large size of the zone affected - 120,000 squared kilometers approximately - and the fact that the road and productive infrastructure of the zone were seriously compromised.
considered important to avoid uprooting people from the neighborhoods in which they had been living for years so as to preserve social cohesion. In sum, a new approach was required.

In addition to the technical construction problem, there was the “subsidy conundrum”: it was essential to guarantee that public funds were used properly. Regulation required each subsidy being assigned to an eligible person having the legal title to the terrain. This requisite, although simple at first sight, proved to be quite challenging. An initial estimation indicated that more than 18,000 beneficiaries had problems with their titles. Hence, the law regulating ownership was amended so as to make easier the process of normalizing and legalizing property titles. Multi-family dwellings such as apartments and condominiums presented additional issues as not necessarily all dwellers were eligible for subsidy. But due to common property of certain areas in the building any solution required considering both eligible and ineligible people.

The target President Piñera set was that all subsidies have to be allocated by December 2011, all constructions and reparations have to be started by December 2012, and that all solutions have to be delivered by February 2014. Addressing this challenge required making processes more flexible and timing requirements set by a variety of regulations. For instance, the prior subsidy allocation process usually took 18 months and required applicants to submit 170 different documents. The process was redesigned and reduced to require just 18 documents, a change that is estimated to have saved paperwork to the State and applicants equivalent to issuing and presenting 3 million documents. Regarding the supply of new dwellings, it was decided to change the signal sent to the construction market. The prior approach was to set minimum quality standards and make bidders compete on price. The new approach made bidders compete on quality instead, fixing the price per dwelling at US$ 17,200. The result was that 60 companies
offered 82 different types of dwellings, largely exceeding prior quality standards\textsuperscript{14}. People were allowed to choose the type of dwelling in participative processes, in which all those beneficiaries living in a neighborhood had the opportunity to see all the alternatives and vote for the one they preferred. The final decision was made based on majority vote, and in most processes the type of dwelling finally chosen received more than 90\% of preferences. Managing the reconstruction process proved to be a tricky and complex matter. Anyway, by December 2010, 121,000 subsidies had been allocated and more than 50,000 constructions initiated. A year later, December 2011, the totality of subsidies had been allocated, 136,000 constructions started, and 72,000 dwelling delivered.

\textit{The communicational and political dimension}

As can be expected, the disaster and devastation caused by the earthquake and tsunami received extensive media coverage. There were plenty of images shown, stories and testimonies told, and there was widespread demand for information not only in Chile but also abroad. Evident signs of recovery and reconstruction coexisted with those of frustration and despair of people that, besides living in poverty conditions, now had their homes damaged. There was much to do and increasing pressures for delivering quickly. Several of my interviewees assert that pressures did not come just from their bosses but from everywhere and everybody: the affected citizens, local authorities, the media, the general public, politicians of all colors, in sum the whole country.

By September 2010, six months after President Pi\ñera assumed office, a process of politicization regarding the reconstruction program became evident and political tension got exacerbated when two opposition congressmen led a request for interpellation to Housing Minister Matte. The reason argued was the slowness of the reconstruction process in the areas affected by the

\textsuperscript{14} For instance, prior social dwellings were 42 sq-meters while new ones averaged 50 sq-meters. The social impact of this change has been relevant. A 42 sq-meters dwelling is typical of the first quintile of poverty, while a 50 sq-meters dwelling is typical of the second quintile. The new bidding criterion provided second quintile houses to first quintile families.
earthquake and tsunami. The interpellation took place by the end of December 2010, few days after the Housing Ministry informed that the target of awarding 100,000 subsidies during 2010 had been surpassed. During the interpellation Minister Magdalena Matte was accused of (i) misinforming the President with false information, (ii) prioritizing municipalities run by government-minded majors, (iii) and a generalized neglect and abandonment of the areas affected by the earthquake. Although the veracity of HRP’s advance reports was later certified by the Comptroller, the concept of the slowness in the reconstruction process got installed in the general public. Physical evidence of construction and delivery of dwellings came increasingly visible by the first term of 2012, but that did not reduce the intensity of the debate. On the one hand government argued that in spite of the immensity of the task, work was on schedule and meeting targets, while opposition focused on the human side of the problem. For instance, one opposition Senator said “**One continues seeing the drama that people is living. Although houses are being delivered, that is not enough given the extension of the damage**”. The same Senator was present during a new houses handover ceremony, in December 2010, ten months after the earthquake. These were among the first houses delivered, and were located in a very isolated and remote region. When a journalist asked her opinion about the new houses, she replied “**I think they are great, but it is a shame Government has taken so long**”. One interviewee who witnessed this episode commented: “**If you take into account the devastation caused by the earthquake and that these houses were delivered just nine months after we took office, this comment is surrealist! The reconstruction program is not like playing with Lego bricks**”. With the concept of slowness installed at the center of the debate, the Reconstruction Program’s approval rating systematically decreased over time, moving from 65% just after the earthquake to below 40% by the “one year anniversary” on February 2011. The second anniversary provided again the stage for partisan discussion about HRP achievement or lack of it, but this time there was a difference. While most
opposition politicians harshly criticized the level of advance, the presidents of two of the main opposition political parties made a more balanced assessment. For instance, Deputy Andrade, president of the Socialist Party, said “I have no problem recognizing the things that have been done well. But as I am from the opposition I have to emphasize those things I think should be done better. Because if we - from the opposition - do not say this, then nobody else will”.

When thinking retrospectively on the way in which the politicization process evolved, some of my interviewees comment that they lost an important battle, the communicational one. As explained by one interviewee, “good news doesn’t sell, and given that the news media will not look for these cases, we have to provide newspapers and TV channels with that information. For instance, they will report the case of Mrs. Juana, who is still waiting for her house, living in a temporary shelter, suffering 41° Celsius summer heat. Our goal is that together with the case of Mrs. Juana, the news media show also the case of Mrs. Rosa, who received her nice new house in record time and feels she now lives in a palace. It is easy to make news reports denouncing problems: in 220,000 individual cases you have all kind of situations. But it is also possible to deliver news reports that provide a more balanced view”.

Another issue was that most of the work done was invisible to the general public. For instance, one interviewee referred to the case of Robinson Crusoe Island\(^\text{15}\): “Before initiating any kind of reconstruction work, it was necessary to remove all the debris that resulted from the tsunami, something not trivial in the island’s mountainous and undulating terrain, 600 kilometers offshore continental Chile. In total, 900 tons of debris was removed and now the terrain is clear, free of debris. The conditions of the terrain made necessary that work was done by hand, with the single aid of a backhoe. Another issue was where to take the 900 tons, which finally were shipped to the continent”. He also mentioned the huge difference between photos taken after the tsunami and

---

\(^{15}\) The tsunami following the earthquake struck the island and destroyed most of the town where 95% of inhabitants live.
the later ones showing the terrain clear of any debris, ready for the reconstruction work. It was evident that only those comparing both photos would notice all the work done. Without that comparison, all that somebody will see will be that after one year, no house has been finished yet.

The case of the CASEN affair

In Chile the magnitude of poverty is measured using the CASEN national household survey that is conducted by the government every two years, and analyzed by the United Nation’s Economic Commission for Latin America and the Caribbean (ECLAC). The proportion of the country's population living in conditions of poverty and indigence - extreme poverty - are determined with respect to reference values known as “poverty lines”. Indigent households are those whose income is not enough to meet members' basic food needs, and poor households are those whose income is not enough to meet the basic food and non-food needs. In order to calculate poverty and indigence it is necessary to update these reference values to reflect changes in the prices of goods and services. The practice traditionally employed by ECLAC was to update the reference values for poverty and indigence taking into account changes in food prices\(^\text{16}\). In 2007 ECLAC introduced a change, by which the indigence line for all countries working together with ECLAC was updated on the basis of variations in the food component of the Consumer Prices Index (CPI), while the poverty line for spending on nonfood items was updated using the variation in the CPI. This change was motivated by the regional surge in food prices in 2007 and 2008, which was not accompanied by a similar rise in the price of other goods and services.

When it took to analyze 2009 survey data in year 2010, Chilean Ministry of Social Development (MSD) decided not to adopt the ECLAC’s modification and instead preserve comparability by updating both reference lines for indigence and poverty on the basis of food CPI variation. This

---

\(^{16}\)This practice ensured that the ratio between the poverty and indigence reference values remained constant, with the family income associated with poverty being two times that of indigence.
decision had a relevant effect because in Chile the CPI for food grew by 32.4%, while the CPI for other goods increased by just 6.4% during the 2006-2009 period.

The result of maintaining prior methodology was that poverty rate increased to 15.1% from the 2006 value of 13.7% (ECLAC, 2010). These 1.4 percentage points were equivalent to 355,000 additional Chileans pushed into poverty, a situation that broke down the trend of poverty reduction that had been sustained for more than 20 years. In contrast, using ECLAC’s new method the resulting poverty rate would be 11.5%, which was 2.2 percentage points lower than the 2006 year figure. Interestingly, the political debate centered not on whether the 15.1% value was correct or not, but on a “blame assignment” game. Government blamed prior administration of mismanaging public policy and, in turn, CASEN results created political rift among political opposition. On the one hand, former Finance Minister Velasco and others blamed the world financial crisis and a steep rise in food prices, while other members of the opposition conglomerate blamed their former colleague Velasco and other technocrats of their own side for not having done enough when in office.

Poverty was measured again in 2011, and results were made public in July 2012. Mr. Lavin, head of the Ministry of Social Development (MDS), proudly announced that poverty had fallen to 14.4 percent from 15.1 percent in 2009, and that extreme poverty was at a record low of 2.8 percent, down from 3.7 in previous measurement\textsuperscript{17}. He said that these results were a relevant milestone in the way of meeting President Piñera’s promise of eliminating indigence during his term in office. In addition these results demonstrated that Chile had recovered the ability to reduce poverty. But few weeks later it was known that ECLAC initial analysis reported a poverty rate of 15%, essentially the same as that of 2009, and that MSD requested including an item about informal income earned by the unemployed, on the basis that it was appropriate for preserving

\textsuperscript{17} Due to space limitations the comparison chart of Chilean poverty data from 2006 to 2011 is available on request.
comparability\textsuperscript{18} (ECLAC, 2012b). This change was questioned and interpreted politically by different parties. Prior government’s Finance Minister Velasco – the one blamed of not having done enough by his political comrades during the discussion of the CASEN 2009 results - accused government of having been massaging poverty figures, while a group of 30 economists called on the government to clarify the survey’s methodology. Although other changes to the survey in the past had been heatedly debated, this occasion was the first time that questions about the quality of Chilean social and economic statistics were raised in major international media. The situation created a negative perception that damaged the image of Chile, as illustrated by the articles published by The Financial Times and The Economist (The Economist, 2012; Webber & Financial Times, 2012), and also triggered the resignation of two officials, one from government’s MSD and the other from ECLAC. Andrés Hernando, the resigned MSD chief official working on CASEN, criticized that the survey was used for politics by both sides: “While government is getting out champagne and celebrating a reduction of poverty that is not significant, the opposition has cast doubts on the credibility of the survey”. Few weeks later ECLAC announced the decision to no longer collaborate on the CASEN survey. ECLAC argued that since this survey began in 1987, Chile has achieved sufficient institutional and technical maturity to carry out the process independently, thereby rendering ECLAC’s continued collaboration unnecessary (ECLAC, 2012a). Although ECLAC explained its decision in positive terms, it is difficult not to connect this resolution with the politicized events that had compromised ECLAC’s prestige. The conclusions of two prestigious Chilean economists - well respected across party lines, one considered center-left and the other center-right – provide a good overview of the CASEN affair. First, the lack of statistical significance made the whole

\textsuperscript{18} In September 1\textsuperscript{st} ECLAC published a press release addressing the controversy and providing detailed information. Regarding the informal income item they say: “Undoubtedly, this item represents a type of labor income actually received by people, so it was finally decided to include it as part of total household income in 2011”.

24
poverty reduction debate a nonsense discussion from the beginning. And second, only one thing is certain, and it is that the clear loser in this affair was the prestige and credibility of Chilean institutions (Engel, 2012; Montes, 2012).

**DISCUSSION**

In this paper I have explored two literatures to better understand the dynamics of assessment in the public domain. The literature on Interpretation highlights that when making judgments, people use heuristics as shortcuts to deal with complexity. Assessment is reference-dependent and is also influenced by the “affect” and the “availability” biases among others. In consequence, people confronted with exactly the same information may reach fundamentally different conclusions, resulting in interpretation asymmetries (Rojas, 2015). In addition, the human tendency to use heuristics and resort to biases facilitates the framing and manipulation of interpretations by third parties. While framing is related with persuasion and does not necessarily have negative connotations, the manipulation of interpretations is purposively deceitful and leverages on the aforementioned weaknesses of human judgment.

In turn, the Public Sector literature highlights the citizens are the “ultimate assessor” in the public domain, and thus political parties are continuously competing for gaining the support of ordinary citizens. This struggle results in a tendency of “issue politicization” in the public domain as any event may become the arena in which the larger political game is disputed. A direct effect of politicization is that assessment is collapsed into a single dimension of worth, hence increasing the likelihood of improper assessment, particularly in complex matters. For instance, public agencies can be labeled as underperformers due to issues affecting a single dimension of assessment, even if they would be considered performing adequately when taking into account the full dimensionality of their activities.
Citizens are not blind to manipulation attempts and the result is the “confidence gap”, which is the erosion of general public’s trust in their political leaders, public servants, and public institutions. In consequence, the ability of public agencies to link their activities to socially accepted values becomes increasingly important for ensuring a positive assessment. In addition, there is the problem of the expertise and information that are necessary for proper assessment. Democracy is inevitably messy, in part because the general public’s opinion matters in spite of its eventual little familiarity and limited knowledge regarding certain topics, but decisions in the public sphere that do not gain citizens’ support produce popular rejection. As a result of the combined effects of the confidence gap and the necessity of considering the standing of non-expert citizens, public agencies and their officers face a relevant handicap. Even if their performance would be considered outstanding taking into account available evidence, they are likely of not being considered as such due to citizens’ skepticism about the agency’s success claims, and citizen’s lack of expertise to properly judge them. Under these circumstances the general public will not respond to claims of improvement until the level of progress is such that it is hard to have any doubts, and thus a “credibility threshold” has been surpassed. Hence, for certain variables one may observe a non-linear relationship between performance measures as reported by public agencies, and how these measures are interpreted by the general public, in particular measures that are claimed to be indicative of good performance

The Credibility Threshold and the literature review

The literature reviewed in this paper provides a series of arguments explaining this non-linear relationship. First, information reported by government comes from large samples that represent the whole population. In contrast, daily life experience provides ordinary citizens a much smaller sample, as they will see only a small bit of the range of situations that are actually out there. For instance, let’s consider the unemployment rate. When ordinary citizens think about
unemployment, data coming from their immediate and smaller context have a larger weight on their judgment, and it is a statistical fact that small samples yield extreme result more often than large samples do. Hence, every time government reports unemployment statistics there will be citizens perceiving that the actual situation is worse than reported. In turn, the literature on Interpretation highlights that (i) perceptions are influenced by the availability heuristic, that (ii) judgment is reference-dependent, and that (iii) behavior is motivated by loss aversion. The “availability” phenomenon provides supplementary support to the idea that judgment is influenced by small sampling. But the concepts of reference-dependence and loss aversion provide a new additional insight. As indicated by Kahneman (2011) and others, people that perceive themselves below the reference point will feel that they are losing and respond strongly to reduce that loss, being very active and determined in doing whatever is at hand. The Housing Reconstruction Program provides a variety of examples that illustrate this reaction. For instance, Major Morales, vice president of the Chilean Association of Municipalities, criticized progress reports issued by government. Commenting a report informing that the Housing Reconstruction Program’ progress was 62%, with nearly 100 thousand houses delivered, he said: “The problem is that these figures create false expectations. In my territory the reality is different, the program progress is much less, and when people listen about government figures the go to the city hall to protest, because they have not received they houses yet” (Scheuch & Radio UChile, 2012).

Moreover, one of HRP’s problems was that a lot of the work done was invisible to the general public, and that which is not communicated does not exist (Lee, 2000). While HRP management was solving bureaucratic problems, designing new solutions, and planning the whole program, thousands of families were waiting for their houses. As an opposition congressman said, “What people need is to see shovels digging the soil, and see their houses under construction”. Given this gap between the work done by the HRP team and general public’s perception about it, it is
not surprising that even though the reconstruction program was progressing according to plan, its approval rating decreased systematically over time. The HRP case illustrates that public agencies are not necessarily assessed by their activities and outcomes but by how the general public may come to interpret and perceive them. The confidence gap, associated with citizens’ skepticism and suspicion regarding success claims by public officers, makes less likely that well-performing public agencies and managers will be assessed as such.

A final contribution explaining the Credibility Threshold comes from the literature on the Public Sector that highlights the politicization dynamics that affect public sector institutions and the problem posed by the “confidence gap”. Firstly, people tend to receive government’s claims of target accomplishment with skepticism. And secondly, political debate that casts doubts on those claims is likely to lead to even worse perceptions, with an increased effect among opposition-minded citizens that go even further complicated when performance measures are imperfect and noisy. The CASEN affair provides a good illustration of these dynamics. From the beginning the main error made by government was to treat CASEN’s poverty rate as if were a hard and unquestionable number, without stopping to think in the consequences of the survey’s 0.7% margin of error. It was needed a scandal and 30 economists demanding clarifications to make government step back. One immediate effect was that government lost credibility. In the month corresponding to the CASEN affair, Minister Lavin’s approval rating fell 5 percentage points, from 61% to 56%. This variation was the largest fall for all the 15 Ministers evaluated during that period. In addition, government’s excess of enthusiasm, lack of prudence, and sloppiness in managing the survey led to an unfortunate chain of events that resulted in ECLAC’s decision of terminating its collaboration with CASEN.

19 A commented earlier, Mr. Lavin was the head of the Ministry of Social Development, the government unit in charge of the CASEN survey.
In sum, as a consequence of the combined effects of all the dynamics described above, accomplishment that should be considered good news by the general public will not necessarily be considered as such until the level of achievement is large and evident enough that it is difficult to dismiss. Furthermore, public officers who assume that indexes that improve at the population level will be appreciated by the general public can shot themselves in the foot. The larger the variance, the more likely than somebody will cast doubts on achievement claims, or even consider them outright lies, triggering politicization dynamics and fostering a climate of disbelief.

For illustration purposes, figure 1 shows Chilean monthly data on the employment rate and on the approval rating to government regarding employment between years 2010 and 2014. Unemployment in Chile has been below 10% during the whole period. The data in Figure 1 shows that there is a dramatic increase in approval as the employment rate goes in excess of 93%, i.e. unemployment goes below 7%. The pattern of the data is the one that would result in the presence of an assessment threshold by which people assign good grades only when superior results are evident, and thus credible. Results using Hansen’s (2000) estimation threshold procedure indicate that the null hypothesis of no threshold is rejected at the one percent significance level in favor of the alternative of one threshold.

---

In order to further explore the combined effects of partisan orientation and the “Credibility Threshold”, I developed a simulation that models interpretation as a logistic function of the performance signal within the range that the threshold operates and illustrates the non-linear relation between performance measures and citizens’ interpretation. The simulation considers the

---

20 Results available on request, not reported due to space limitations.
aggregate effects of three subpopulations: non-partisan citizens, pro-government citizens, and opposition-minded citizens. The results of the simulation show a pattern that is consistent with that of the data on employment and government approval-on-employment shown above.

**Descriptive model**

“Assessment” results from assessors considering noisy measures of performance. Measure “m,” is a proxy of the actual underlying dimension of performance “AP.” When judging performance, individual-j will apply a weight “wij” to each of the pertinent measures according to his or her preferences, and also apply an interpretation bias “ψj” and a political bias “πj,” although these later two with different grades of consciousness. The interpretation literature contributes by pointing out that that several parties may try to influence the way in which performance measures and other signals of performance are interpreted. For doing so they will suggest that certain underweighted dimensions are worthy of larger consideration, and they will frame the situation such that interpretation biases favor their interests. The following expression describes assessment as a weighted interpretation of actual performance measures.

\[
Assessment_j = \sum_i w_{ij} \times I_j(m_i) = \sum_i w_{ij} \times (\beta_i AP_i + \epsilon_i + \psi_j + \pi_j)
\]

This expression can model a variety of situations. For instance, politicization processes would result in large and polarized values for the political bias “π”. To simulate the workings of the “Credibility Threshold”, I used a simple version that considers just a single measure of performance. Political bias is represented by the range over which the “Credibility Threshold” operates. For pro-government population the threshold does not exist, while for the opposition-minded population the bias makes them to dismiss performance claims until results are very hard to dismiss. The interpretation bias is modeled applying a logistic function to the performance signal, which goes from giving little credibility to success claims to fully believing them.
The simulation

For simplicity, the model for simulating and illustrating the workings of the “Credibility Threshold” considers that the total population of citizens is composed of three sub-populations, the (i) non-partisan population, the (ii) pro-government population, and the (iii) opposition-minded population. The essential difference among them is that opposition-minded citizens are more skeptical than non-partisan citizens regarding government claims of success, while pro-government citizens believe whatever government says.

The performance threshold operates in an interval. Performance signals under the lower endpoint are dismissed as evidence of success, while signals above the upper point are considered as full evidence. Within the interval, citizens interpret signals by applying a logistic function\(^{21}\) that results in an “S-shaped” outcome. In turn, the performance signal is a noisy representation of underlying performance. Noise was modeled as a random value that varies within a range. The graph below plots the results of a simulation with the following parameters:

- The lower endpoint for the threshold intervals are 20 and 60 for the non-partisan and opposition-minded populations respectively. In both cases the range of the threshold interval is 30, that is, the upper endpoint for both populations is equal to the lower endpoint plus 30. Pro-government citizens are not affected by the threshold, they just believe the performance signal provided.
- Noise has zero-mean and a range of (+/-) 3.5.
- Non-partisan population equals to 40% of the total population, while both pro-government and opposition-minded populations equal 30% each.

The graph plots actual performance versus perceived performance, and shows four series of data.

---

\(^{21}\) The logistic function was chosen because it is applied in a variety of fields, including artificial neural networks, biology, demography, economics, chemistry, mathematical psychology, probability, sociology, political science, and statistics.
The “dash” series corresponds to the interpretation of pro-government population, the “plus-sign” series corresponds to the interpretation of the non-partisan population, the “triangle” series corresponds to the opposition-minded population, and the “black-circle” series correspond to the weighted average of the three sub-populations, that is, the interpretation of the aggregate population. Regarding this later series, it is interesting to point out that it is influenced by two different effects. First, signals are affected by the performance threshold, thus non-partisan and opposition-minded people need significant improvements to interpret success claims by government as valid. Second, opposition-minded population is more skeptical, and because it requires even larger improvements to believe government, the total result is that the smaller is the pro-government population, the more amazing and heroic success needs to be so as to receive good rating. The figure below suggests that public officers cannot rely on their agencies’ performance being self-evident.

CONCLUSIONS

This paper begins by presenting the paradoxical situation of former President Piñera. Even though he was accomplishing his electoral promises and his country’s economic and social indicators were at their best, Mr. Piñera’s approval poll ratings plummeted to record lows. When interview by the Financial Times, he requested to be judged not by his intentions, but by results (Rathbone et al., 2011). He had the conviction that soon or later the numbers will speak for themselves. But the passage of time showed that “what the numbers say” is of little value if citizens pay scant attention to or dismiss them as not relevant, and instead use other criteria of assessment. This is not surprising as the Public Sector literature indicates that citizens are neither
much aware of performance reports, nor interested in reading them (e.g. OPSR, 2003).

Several commentators have provided a variety of explanations to Piñera’s paradox. It has been said that it has to do with a general discontent not only with Mr. Piñera but with the political class as a whole (Segovia & CEP, 2011). Others argue that the explanation lies in a social rejection to the “Piñera Way” and the installation of the private sector logic in the state (Tironi, 2011).

Finally, there are some that highlight the affective dimension, saying that this paradox happens because people “simply do not like him” (Mosciatti, 2011).

In this paper I advance an alternative proposition, which combines ideas from the Interpretation literature about the cognitive process by which people attach meaning to information, with the contextual political dynamics described by the Public Sector literature. The lack of positive response of Chileans to their government’s achievements would be partly explained by the “Credibility Threshold”, a side effect of the “confidence gap”, that is the declining confidence shown by citizens in their political, social, and economic institutions. The effects of the “Credibility Threshold” are not limited to politically elected officers, but also affect public agencies carrying out the most diverse types of services and activities. As a result, claims of accomplishment by public agencies are received with skepticism by the general public, and doubts start to vanish only when agencies’ level of achievement is such that it is difficult to dismiss.
Implications for research

This paper contributes to the increasing literature on the intersection of the public sector, interpretations and cognitions, highlighting that research on assessment in the public domain should begin to include the effects of interpretation asymmetries. The fact that public agencies are exposed to a much broader social environment than their private counterparts suggests that diverse interpretations to the same data are more likely in the public sector. Bearing in mind this particularity of the public sector seems essential for avoiding confounds in research efforts.

Implications for practice

Political leaders and public officer would do well to take into account the “Credibility Threshold”. Excess of enthusiasm in claiming that electoral promises and government targets have been met is likely to be received by citizens with skepticism instead of with satisfaction. Government triumphalism provides an opportunity to opposition for highlighting the weakest aspects of performance, and can create a backfire effect through which success claims are received with increasing disbelief by people who may even consider them outright lies. Indeed, even in cases where public officers can recount extensive evidence in support of their success claims, they are likely to experience the same fortune as Nuñez, Wells’ (1904) failed “King of the Blind”. In Wells’ story, Nuñez22, the one seeing man in a Country of Blind people begins reciting to himself "In the Country of the Blind the One-Eyed Man is King", assuming he could rule blind

---

22 In Wells’ (1904) story, Nuñez is an explorer that accidentally arrives to the Country of the Blind.
people. But villagers had no concept of sight and did not understand his attempts to explain this fifth sense to them, and finally he found himself at a disadvantage in relation to people whose lives were ordered on the basis of just four senses. In a similar fashion, Public Officers may became “failed Kings”. The average citizen usually has no way of distinguishing between truthful and deceitful claims, and in addition is affected by the cognitive biases described by the Interpretation literature. Hence, even if Public Officers have evidence supporting that targets have been met, what matters is whether people believe or dismiss success claims. Citizens’ skepticisms regarding success claims and the generalized “confidence gap” suggest that the public domain is a land where the one-eyed man is doomed. That is, Public Officers with superior information may be simply not believed.

Finally, this paper also illustrates that solutions to the problem of assessment are unlikely to work if based only in the provision of performance measures, without considering interpretations. Additional measures may work when there is a sophisticated audience that can deal with the increased complexity of information, but this is unlikely to be the case of the average citizen. Whichever the improvement of reporting systems, what makes the difference is assessors’ capacity to discriminate good from bad performance\textsuperscript{23}, or the participation of prestigious and independent third parties that provide legitimacy to the information and interpretation provided by government, lending credibility to public agencies’ claims.

\textsuperscript{23} In this sense, first-hand experience has shown to be more successful that arm’s length reporting in the public sector (e.g. OPSR, 2003).
Future Research

This study was designed to be exploratory and fundamentally oriented to theory extension regarding the dynamics of assessment in the public sector. As such, the findings and conclusions are tentative and require further research. One avenue for further inquiry is to explore the extent to which the “Credibility Threshold” is affected by citizens’ first-hand experience. For instance, a large proportion of citizens have a daily and recurrent personal experience of public transport, and thus they will judge government performance in this domain based on their direct experience. In contrast, in countries with a healthy economy just a reduced proportion of citizens directly experience unemployment. The information they use to judge this domain comes mostly from the news media, and conversations with other people. Future research can determine whether direct experience and mediated experience differ regarding the Credibility Threshold.

Another avenue of research is exploring the extent to which the “Credibility Threshold” is a widespread phenomenon, or alternatively is restricted to the data analyzed in this paper.

Finally, I hope the concepts advanced in this paper will contribute to a better understanding of the dynamics of organizational assessment in political environments, and that they can be of help to future research going deeper into these dynamics.

References


Scheuch, M., & Radio UChile. (2012, June 27th, 2012). Las cifras de reconstrucción son una exageración que no corresponden a la realidad.


Tironi, E. (2011). *¿Por que no me quieren? de piñera way A la rebelion de los estudiantes*. Santiago de Chile, Chile: Uqbar Editores.


Wells, H. G. (1904). *The country of the blind*
“Employment Rate” data comes from the monthly survey carried out by the Chilean National Institute of Statistics (INE), while “Approval Rate” comes from a survey carried out in Chile by Adimark. The series with black diamond markers corresponds to the actual data from INE and Adimark, while the series with white circles corresponds to a threshold model included for comparison.

Figure 2 “Credibility Threshold” simulation – Actual vs Perceived Performance