



HAL
open science

Quantitative technologies and reflexivity: The role of tools and their layouts in the case of credit risk management

Céline Baud, Nathalie Lallemand-Stempak

► To cite this version:

Céline Baud, Nathalie Lallemand-Stempak. Quantitative technologies and reflexivity: The role of tools and their layouts in the case of credit risk management. *Accounting, Organizations and Society*, 2024, 112, pp.101533. 10.1016/j.aos.2023.101533 . hal-04419872

HAL Id: hal-04419872

<https://hal.science/hal-04419872>

Submitted on 26 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quantitative technologies and reflexivity: the role of tools and their layouts in the case of credit risk management

Authors

Céline Baud
Université Paris-Dauphine, Université PSL
CNRS, DRM
Place du Maréchal de Lattre de Tassigny
75016 Paris
France
Celine.baud@dauphine.psl.eu
(Corresponding author)

Nathalie Lallemand-Stempak
IAE Paris - Sorbonne Business School
8 bis rue de la Croix Jarry
75013 Paris
France
nathalie.lallemand-stempak@iae.pantheonsorbonne.fr

Quantitative technologies and reflexivity: the role of tools and their layouts in the case of credit risk management

Abstract

The development of quantitative technologies is increasingly challenging professional practices and raises questions about whether and how organizations may foster plural and reflexive practices. In this paper, we outline the role played by tools and their layouts in this process. Tools can sustain the enactment of plural views, logics and evaluative principles. However, it is not clear why, in some cases, designing or using these tools triggers intractable conflicts instead of helping to sustain reflexivity in a “productive” way. To address this issue, we explore the case of a French bank that introduced in its credit management processes a new statistical approach of risk management, which conflicted with the professional approach that prevailed at the time. Relying on Boltanski’s (2011) work on critique, we highlight how “productive” reflexivity emerges, not only from critique, but from a dynamic relationship between *critique*, *confirmation* and *practical action*. This framework allows us to bring a fresh look at the layouts identified in the literature as able to sustain pluralism by exposing their differences regarding whether and how they may contribute to trigger reflexivity. We especially show that, when quantitative technologies are involved, the creation of *compromising accounts* may prompt dynamics of escalating conflict, while *combinations* may help organising a pluralism of modes of evaluation that nurtures reflexivity without inhibiting action. Moreover, our study shows how, in credit risk management, quantitative technologies can be implemented, even in the most operational processes, without bringing about an unreflexive “illusion” of control.

Quantitative technologies and reflexivity: the role of tools and their layouts in the case of credit risk management

1. INTRODUCTION

If the day is coming, and I'm afraid our profession is tending a bit in this direction, when decisions will be based exclusively on ratings [...] then I have to admit that we will possibly have a problem [...]. Part of our job is also deliberating the grounds and reasons for our decisions. (Credit manager 1, Headquarters)

The ongoing expansion of new “quantitative technologies” (Mennicken & Espeland, 2019) is transforming information and the production of knowledge for decision-making. It is increasingly challenging professional practices, in risk management and in accounting in particular (Power, 2009 ; Huber and Scheytt, 2013 ; Bhimani & Willcocks, 2014 ; Quattrone, 2016). As the above quotation shows, this raises questions about whether and how, in a quantifying world, the reflexivity fostered by professional practices - here defined as the ability to pause and think about what oneself is doing in the process of doing it – can be sustained.

Previous research has explored the role played by *calculative cultures* in fostering a more or less enthusiastic relationship to quantification techniques (Mikes, 2009, 2011). Meanwhile, accounting scholars also pointed out that whether actors become reflexive and question the appropriateness of their tools depends, among others, on the design characteristics of the tools, on their consequences for the activity and on the availability of competing indicators (Dambrin & Robson, 2011; Jordan & Messner, 2012). This echoes the general idea, developed by Boltanski (2011), that whether actors reflect on a given course of action, criticize it, or “make do” with it, not only depends on their *culture*, but also on the specificities of the *situation* they are facing. Building on this body of work, we focus in this paper on the role played by tools and their layouts in creating situations that favor reflexivity, and we discuss the conditions under which it may allow organizations to avoid the risks associated with an unreflexive use of quantitative technologies.

Scholars have long outlined how tools are a locus of conflicts and controversies that, at the same time, offer opportunities for mediating between groups of actors with distinct interests and concerns (Briers & Chua, 2001; Mouritsen, Hansen & Hansen, 2009). More attention is now also paid to the forms of this mediation and the variety of arrangements that can sustain the enactment of multiple logics and evaluative principles (Chenhall, Hall & Smith, 2013 ; Carlsson-Wall et al., 2016). However, it is not clear why, in some cases, this can lead to intractable conflicts rather than help sustain reflexivity in a “productive” way (Fisher & Ferlie, 2013). This suggests that, as stressed by Stark (2009), even if the overlapping of multiple evaluative principles, that is, pluralism, is key for reflexivity to emerge in organizations, it is not sufficient. This is the issue that we set out to explore in this paper.

For this purpose, we draw on the conceptual framework developed by Boltanski in *On critique: A sociology of emancipation* (2011). This framework highlights that critique is necessary to reflexivity but shall not be conflated with it because critique, alone, leads to intractable conflicts. From this framework, we bring out that, to support a productive form of reflexivity, tools should allow for the emergence of *critique*, but also help organize moments and situations where institutionalized discourses of justification – *confirmation* - and less reflexive moments

– namely, *practical moments* - would also prevail. Understanding reflexivity as emerging not only from critique but from a dynamic relationship between *critique*, *confirmation* and *practical action* allows us to bring a fresh look at how tools and their specific layouts may favor and sustain it. This led us to examine the variety of the configurations that are able to sustain plural views, logics or evaluative principles within organizations; and to discuss the conditions under which they may favor reflexivity when organizations implement quantitative technologies. To do so, we draw on the literature on pluralism and on the distinction operated by Pache and Santos (2013) between *compromises* and *combinations*. This allows us to highlight why the materialization of compromises through the construction of *compromising accounts* (Chenhall et al., 2013) can be especially challenging when quantitative technologies are involved. Contrasting, we show how combinations in the form of *selective couplings* (Pache and Santos, 2013) and *layering* (Polzer et al., 2016) may help avoiding the creation of intractable conflicts while favoring a reflexive use of quantitative technologies.

Empirically, we use a tool-based methodological approach (Chiapello & Gilbert, 2019) to analyze how a new “statistical” approach has been introduced in the daily credit management practices of a French bank, the Mutual Bank (MB). This happened following the adoption of the Basel rules on capital measurement and capital standards in Europe. The central argument behind this regulatory change was that, contrary to traditional risk management which relies on professional expertise, statistical risk management based on quantitative technologies and actuarial calculations fosters transparency and accountability and allows control at a distance, thereby helping managers and regulators altogether to prevent crises (Baud & Chiapello, 2017; Carruthers, 2022). We examine how this change disrupted the strong “professional” approach prevailing at the time at the MB and compelled the bank to adapt its tools, indicators and practices. We study this “riskwork” (Power, 2016) at the operational level of the bank and consequently focus on the routine activities where professional judgment is most frequently replaced by quantitative technologies (Mikes, 2009, 2011 ; Wahlstrom, 2006, 2009, 2013). As excessive reliance on ratings and their “statistical” approach of risk management induced the decline in underwriting standards that enabled the 2008–2009 financial crisis in the USA (Keys et al., 2010, Rajan et al., 2015), the risks associated with an unreflexive use of quantitative technologies at the operational level of credit risk management are obvious. Accordingly, the literature before and after the financial crisis suggests that the development of highly quantified routines of risk management in organizations may actually result in an impoverished mode of control (Power, 2004, 2007, 2009). However, how a more reflexive use of these quantitative technologies can be structurally supported remains an open question. At the MB, to comply with the new regulatory requirements, senior risk managers at first intended to forge a *compromise*, merging into one tool the “professional” and the new “statistical” modes of evaluation of credit risks. But, they soon dropped their efforts in order to avoid creating intractable conflicts. Instead, they developed a new statistical-based tool, the rating system, and used it in *combination* with their traditional professional-based scoring system. These two separate tools have since persisted. Exploring these tools, their design, their history, but also what actors say about them and how they are used and combined, we outline the role played by tools and their layouts in creating conditions that favor a reflexive use of these quantitative technologies and show how – at least for credit risk management – quantitative technologies can be implemented, even in the most operational processes, without bringing about an “illusion” of control (Power, 2009 ; Huber and Scheytt, 2013).

The rest of the paper proceeds as follows. The next section presents Boltanski’s (2011) framework on critique, reflexivity and pragmatism and its relevance for studies of accounting tools and indicators. We then review the specificities and varieties of arrangements of tools and

practices that may sustain pluralism and discuss how they may impact reflexivity. This is followed by a presentation of the case study and of our “tool-based” research design and method (Chiapello & Gilbert, 2019). Subsequently, we outline the two main approaches – professional and statistical – that pervade the field of credit risk management, present their related tools and expose the difficulties encountered by the MB in its attempts to build compromising accounts. We then show how specific combinations of tools, namely, selective coupling and layering, have been developed and highlight how they trigger and help organizing reflexivity at the MB. A discussion of key findings and a conclusion follow.

2. REFLEXIVITY AND PRAGMATISM IN SITUATIONS OF UNCERTAINTY

Drawing on Knight’s distinction between risk and uncertainty, Boltanski (2011, p. 58) outlines that, in a necessarily uncertain world, “the design of describing the world, in what would be its entirety, is not within anyone’s grasp.” Risk indicators such as ratings or scorings or even systems of such indicators are necessarily incomplete and imperfect. When confronted with the incompleteness of indicators (Jordan & Messner, 2012) or the ambiguities of mixed systems (Power, 2010), actors may adopt a more or less “pragmatic attitude.” Although the notion of “pragmatism” has been introduced in the accounting literature as a specific form of “calculative culture” (Power, 2005, 2007, 2010, Mikes, 2009, 2011), Jordan and Messner (2012), as well as Dambrin and Robson (2011), suggest that being more or less “pragmatic” is not only linked to a culture or a preexisting disposition towards quantification: It also depends on the specifics of a given situation, such as the design features of the overall system, the nature and availability of competing indicators, the consequences of the system for the activity, and its links to preexisting ambiguities. This means that organizational actors often “learn to live” and “make do” with imperfect indicators, whether incomplete or divergent. This is especially the case when imperfections can be compensated for through local repairs or flexible use. It can also be the case when indicators are opaque and, thus, prevent actors from understanding and criticizing their internal functioning; or if, despite their imperfections, they help actors to deal with daily problems or ambiguous situations that were not previously addressed. In the latter case, imperfect indicators may even be regarded as enabling.

The work of Jordan and Messner (2012) and Dambrin and Robson (2011) echoes Boltanski’s claim that actors can be pragmatic in some situations while being critical, or reflexive, in others. This calls for taking “situations” and not only “actors” as objects of analysis, or, as Boltanski puts it, to study “actors *en situation*” (2011, p. 24).¹ To recognize actors’ capacity to be pragmatic in some situations while being critical, or reflexive, in others, Boltanski (2011) proposes distinguishing two main registers of action: the practical and the metapragmatic registers. The practical register² is broadly consistent with a pragmatic attitude. In this register, action is “primarily directed towards something ‘to be done,’ a task to be performed, with the concern of ‘getting by’ [...] most often simply so that it is possible to move on to something else” (Boltanski, 2011, pp. 62-63). It implies that a certain tolerance for differences and contradictions prevails in practical moments. This tolerance triggers a self-limitation of disputes that facilitates coordination.³ One of the distinguishing features of the practical register is its low level of reflexivity. In the practical register, “repairs [...] and adjustments intervene

¹ [in situation], the phrase remains in French in the English version of Boltanski’s book.

² Boltanski also calls it the “pragmatic” register.

³ It “facilitates the pursuit of an action in common that would be threatened by the transformation of differences in ways of doing in divergences over the way in which they *must* be performed” (p. 65).

constantly but possess a local character” (Boltanski, 2011, p. 64): Action is situated. It is not random but organized around habits and routines. What matters in this register is that these habits and routines provide footholds to help coordinate actions. In other words, in the practical register:

“people involved in the course of action act as if they more or less knew (...) what they are in the process of doing—and/or as if the others, or some others, in whom one can have confidence, knew it (and this even if the definition of the task performed in common is rather vague)” (Boltanski, 2011, p.62).

The practical register is predominant in our daily social life. Because it is “unreflexive and non-cumulative, it allows for the maintenance of tacit disagreements that do not extend to disputes.” But this tolerance prevails only “up to a certain threshold” (p. 67): When uncertainty surfaces about how to qualify the situation (e.g., when someone asks “What is it that we’re doing?”), this threshold is crossed, and the situation switches to a different register, which Boltanski calls “metapragmatic.” This idea of a threshold that would trigger a switch from the practical to the metapragmatic register echoes the distinction between “broad” and “narrow” forms of incompleteness elaborated by Jordan & Messner (2012, p. 550). Narrow forms of incompleteness relate to concerns about the representational qualities of indicators that can be addressed by “rather minor repair work.” By contrast, with broad forms of incompleteness, “the general idea behind the indicator is called into question.” This does not necessarily mean that the indicator is intrinsically problematic or flawed but rather that it is so different from what the actor’s own reference point would suggest that reconciliation appears impossible.

For Boltanski, two symmetrical modalities of expression constitute the metapragmatic register: confirmation and critique. Confirmation designates forms of intervention aiming to “establish what is and preserve it as being” (Boltanski, 2011, p. 61). The register of confirmation is closely linked to institutional action. As the author explains:

“To institutions falls the task of saying and confirming what matters. This operation assumes the establishment of types, which must be fixed [...] and often stored in definitions, so as to be available, when the need arises, to qualify, in a situation of uncertainty, states of affairs that are the object of ambiguous or contradictory usages and interpretations” (Boltanski, 2011, p.75).

Amid uncertainty, confirmation is a qualification process that allows confirming that what is done is relevant and coherent. It provides semantic security that dissipates the feeling of unease provoked by these situations. Yet, in extreme scenarios, confirmation leads organizations on a maniacal quest for coherence.⁴ When it comes to issues of quantification, this pursuit can be linked to the “calculative idealism” described by Power (2005, 2007, 2010), where actors have little tolerance for the ambiguities of mixed systems, which may result in an impoverished mode of control.

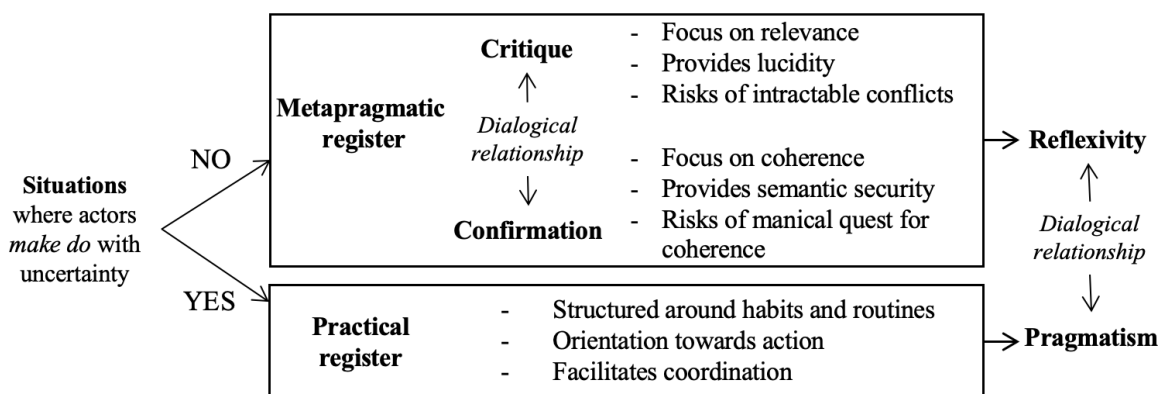
Critique, on the other hand, draws on factors of uncertainty to challenge the reality “on what presents itself as being,” in manifestations of common sense, or in institutions’ official expressions (Boltanski, 2011, p. 62). Therefore, critique poses a constant threat to institutions. But, for Boltanski, confirmation and critique are not antagonistic; rather, they belong to the same movement: A critique “counter-poses its lucidity” (Boltanski, 2011, p. 103) to the

⁴ This shows that Boltanski does not think about incoherences and disagreements as problematic per se. On the contrary, he suggests that “what should prompt identification of a situation as pathological is [...] the maniacal quest for coherence” (p. 118).

institutional systems of confirmation, but the latter ensures the minimal amount of semantic security without which it would not be possible to even express the contours of any critique. As such, the relationship between confirmation and critique “has nothing dialectical about it, in the sense of concluding in a synthesis” (Boltanski, 2011, p. 57). It is rather to be understood as a “dialogical relationship” (Boltanski, 2011, p. 62) that fosters dynamics of change (Boltanski, 2011, pp. 119ff).

Overall, Boltanski’s conceptual framework enables us to better distinguish the issues at stake when organizations and actors face the uncertainty that stems, in particular, from potential incoherence and contradictions between indicators stemming from distinct modes of evaluation. More specifically, as summarized in Figure 1 below, the distinction made by Boltanski between practical and metapragmatic registers helps us better understand how to sustain reflexivity while avoiding intractable conflicts. Stark suggests that it “requires attention to the structure of temporal processes” and “a collective sense of rhythm and timing – of when to make temporary settlements to get the job done” (2009, p.35). Boltanski’s conceptual framework allows to understand what drives this rhythm. As defined by the author, pragmatism and reflexivity are mutually exclusive, alternative, and situation-related moments. In other words, when confronted to situations of uncertainty, actors sometimes *make do* with it, and sometimes do not. Their attitude depends on multiple factors, including the level of uncertainty itself, but also one actor’s own tolerance threshold for uncertainty and the context of the situation, which varies, for instance, according to the level of hierarchical scrutiny or pressure to take rapid decisions exerted at the moment; or the precedence of other similar situations experienced by actors. For Boltanski, the alternance of situations at each actor’s level furthers a dialogical relationship between practical and metapragmatic moments. Reflexivity itself, according to the author, is the articulation of two opposite types of discourse: confirmation and critique, whose dialogical relationship pushes further the idea of a structurally unstable equilibrium.

Figure 1: Boltanski’s framework



The question is then to identify the internal arrangements that organize the use of evaluative principles in an organization integrating new quantitative technologies and to analyze how they contribute to trigger reflexivity while avoiding the risks associated with it, namely, a maniacal quest for coherence, or intractable conflicts.

3. ORGANIZING PLURALISM OF EVALUATIVE PRINCIPLES

Following Stark (2009), pluralism, defined as “the overlapping of multiple evaluative principles” is key, yet not sufficient, for reflexivity to emerge in organizations. In these settings, actors encounter “situations of multiple and potentially conflicting interests, logics and regimes of accountability” (Chenhall et al., 2013).

Accounting scholars have paid significant attention to compromises as a way to organize this pluralism. Despite failures and limitations (Huault & Rainelli-Weiss, 2011; Clune & O’Dwyer, 2020), “compromising accounts” that take the form of tools that crystallize compromises between different “modes of evaluation” have been proved useful to sustain organizational reflexivity (Chenhall et al., 2013), notably because the design phase of these tools requires dialogue and discussion between individuals and groups with differing views. However, once they have reached a material form, compromising accounts leave only a few alternatives to actors: “making do” with them, confirming them, or getting engaged in a new “cycle of critique, dispute, and accord” (Annisette et al., 2017), the latter option being at the risk of not finding an accord and getting “stuck” (Jay, 2013) into dynamics of escalating and intractable conflict between groups demanding adherence to their preferred system (Fischer & Ferlie, 2013).

Extending the work of Chenhall, Hall & Smith (2013), Carlsson-Wall et al. (2016) suggest that instead of crystallizing compromises within “compromising accounts”, compromises can be reached through the building of a system where each metric or mode of evaluation at play corresponds to a distinct tool. These systems allow managers to make compromises that are *situation-specific*. The modes of evaluation are accorded different priorities “depending on the particular situation as represented through the performance measures” (Carlsson-Wall et al., 2016, p. 46). Thus, for each particular situation, managers have to critically assess their tools and question their relevance to choose what to do and later be able to justify their actions. Such systems rely on professionals’ reflexivity and avoid the risk of “stuckness” involved in compromising accounts. But this reflexivity comes at a cost. Indeed, by relying on a plurality of evaluative principles without structurally integrating them, they turn coordination into a problematic issue. The question of how to control actions and limit the scope for discretion becomes especially acute. Moreover, in these systems, reflexivity is not distributed but concentrated “at the top.” These are particularly challenging issues in the context of banking. First, centralizing all decisions “at the top” may be difficult when organizations get large. Furthermore, and more fundamentally, such attempts to maintain pluralism only “from the top” and “at a distance” from the local projects are at risk of downplaying the disciplinary power of standardized risk management systems since these systems precisely disregard what is idiosyncratic – and, thus, difficult to assess at a distance – about a project. For Huber and Scheytt (2013), to reconstruct a more responsive approach to uncertainties, it is necessary to rely more on practitioners’ ethics and responsibility and to defend their degrees of freedom at the micro-level of practices. But banks clearly cannot simply evade their regulatory obligations any more than they can leave risk management entirely up to field practitioners.

Organizational scholars have also paid much attention to the challenges of pluralism, also labelled hybridity (Pache & Santos, 2013), or institutional complexity (Greenwood et al., 2011). A central concept in this stream of literature, which has recently gained prominence in the field of accounting (e.g. Carlsson-Wall et al., 2016; Clune & O’Dwyer, 2020; Anderson-Gough et al., 2022), is the one of “institutional logic,” which refer to “taken-for-granted beliefs and practices” which “prescribe what constitutes legitimate behavior” and “provide conceptions of what goals are appropriate and what means are legitimate to achieve these goals” (Pache & Santos, 2013 p.973; Friedland & Alford, 1991; Thornton, Ocasio, & Lounsbury, 2012). Based on this notion, pluralism is defined in reference to the co-existence of conflicting “institutional

logics” in a given field and the pursuit of multiple objectives (Denis et al. 2007). Drawing from this acceptation, pluralism of institutional logics cannot be assimilated with pluralism of evaluative principles, insofar as distinct modes of evaluation may well serve the same purpose – and actually most often do. However, we suggest that the heterogeneous, and often conflicting, nature of the constituent elements of institutional logics, combined with the fact that logics instantiate in practices and tools, poses similar issues as the coexistence of multiple modes of evaluation in a given field does.

In this respect, this body of literature may prove insightful for better understanding what is at stake in these situations. As a matter of fact, the literature in organization studies also traditionally associates pluralism with compromise (Oliver, 1991; Kraatz & Block, 2008). In this context, compromise involves an attempt to craft an acceptable balance between the conflicting demands stemming from institutional referents, either by conforming to the minimum standards expected, by “crafting a new behavior that brings together elements of the conflicting demands,” or through a bargain with institutional referents to make them alter their demands (Pache & Santos, 2013, p. 975). However, in the recent years, moving past this traditional body of literature, organizational scholars have sought to explore alternatives to compromises as a way to organize pluralism. They suggest that, despite the forces that tear them apart, conflicting logics can coexist without being brought together within scripts of behaviors as in the case of compromises, through the proxy of structures, tools and practices. This arrangement of logics is named: “combinations.”

There are three main forms of combinations. First, combinations can be achieved through “structural differentiation,” where “separate subunits deal with particular logics.” Thus, the “structurally differentiated” organization is compartmentalized “into different mindsets, normative orders, practices, and processes” (Greenwood et al., 2011, p. 354). Different logics are present in the organization but without being combined at the core of organizational units, practices, and tools. This amounts in our view to the situation described by Carlsson-Wall et al. (2016), where compromises that are *situation-specific* may emerge, but where coordination and control become an issue. The two other common forms of combinations discussed in the literature are “selective coupling” and “layering.” As defined by Pache & Santos (2013), “selective coupling” refers to a selection of intact elements of practices drawn from distinct logics. It combines logics at the core of organizational units, practices and processes, but the elements drawn from each logic at stake remain potentially identifiable: Logics are combined but not blended. Relying on a comparative case study of French work integration social enterprises, Pache and Santos (2013) suggest that organizations often rely on such selective couplings, possibly to spare themselves the need to negotiate with institutional referents or to craft new compromising practices. Layering differs from selective coupling in that there is no selection in this form of combination: Several logics are simultaneously present within a given organization, reflecting a process in which “new elements are introduced on top of, or alongside, existing ones, similar to sediment layers in geology” (Polzer et al., 2016, p. 76). In this sense, layering stresses the “legacy effect of past logics” (Raynard, Lounsbury, and Greenwood, 2013, p. 2).

In this section, we looked at the specificities and varieties of the layouts of tools and practices that can sustain pluralism. Building on the literature that analyses pluralism at the institutional level, we suggest that, when it comes to modes of evaluation, compromises and combinations may have different impacts on the expression of critique and, therefore, on how reflexivity is sustained, organized, and/or repressed in a given organization. Therefore, paying attention to how each of them distinguishes itself from the others appears critical to gain a better

understanding of the role tools and their layouts might play in creating conditions favorable to both reflexivity and pragmatism within organizations that implement new quantitative technologies. In the next section, we present the tool-based methodological framework (Chiapello & Gilbert, 2019) we adopted to study the emergence of these layouts at the Mutual Bank (MB).

4. RISKWORK AT THE MUTUAL BANK: A TOOL-BASED METHODOLOGICAL FRAMEWORK

The Mutual Bank

The Mutual Bank (MB) is a French mutual (cooperative) bank backed by one of the major French banking groups, the UMB (Union of Mutual Banks). As about half of the French banks, it has a mutualistic heritage. The MB in particular was created by organizations whose specific characteristics were poorly understood by conventional for-profit banks and who wished to jointly develop access to banking services. Its customers, mostly SMEs, are all members of the bank, and their primary objective is to achieve a high-quality banking service rather than maximum profitability. Accordingly, the banking relationship at the MB demonstrates a commitment to supporting customers in their borrowing needs.

Providing services with a good quality/price ratio to small organizations, whose diversity makes mass processing very difficult, requires developing a detailed knowledge of their business while controlling costs. To this end, the MB was structured as a decentralized organization of “field-based” professionals with some degree of autonomy and initiative. However, since the early 2000s, the MB has also developed all the risk management and audit processes necessary to comply with the regulations imposed on French banks. As a result, the bank introduced highly quantified and standardized risk management processes in most of its operational processes for credit management (Baud & Chiapello, 2015). This triggered intensive riskwork (Power, 2016), but most of the previous tools and processes also remained in use. Thus, the MB constitutes a perfect case to understand how an organization can embrace quantitative innovations by developing a pluralistic approach to credit risk management and to explore whether and how this approach may help foster reflexivity.

Research design

The literature on riskwork (Power, 2016) has already largely demonstrated that risk management, like most other managerial issues, is not only about techniques but rather the provisional outcome of ongoing encounters between institutional demands, material infrastructures, and situated actors.

The tool-based methodological framework developed by Chiapello and Gilbert (2019) builds on the literature that shows that tools relay power relations but that humans maintain some agency,⁵ individually, or, more often, collectively, in using them. In risk management as in other fields, technologies contribute to produce the worlds described by their underlying theories and frameworks (Themsen & Skærbæk, 2018). They play an active part in mediating between the worlds of different actors (Briers & Chua, 2001) and may hybridize by incorporating principles from other kinds of expertise (Miller, Kurunmäki and O’Leary, 2008).

⁵ “[T]ools forge discipline and subjectification, this is not a totalitarian influence as there are failures and resistance. Individuals are neither completely free nor completely constrained.” (Chiapello & Gilbert, 2019, p. 234).

But these processes often produce unexpected results (Christensen and Skærbæk, 2007). Accordingly, Chiapello and Gilbert explain that “tools are malleable,” which means that “there is no absolute technological determinism” and “there is always uncertainty in their use” (p. 234). But, at the same time, with a given tool, “not everything is possible; the game is not completely open” (p. 234). Action is consequently shaped through the complex interactions between humans and artefacts and “[t]here are no simple, linear causalities” (p.233), which suggests to look for “process effects.”⁶

From their extensive review of the literature on tools and technologies, Chiapello and Gilbert (2019) draw an original methodological framework intended to grasp these “process effects” and the complex interactions between the human and non-human elements that constitute them. This involves following tools in all their *states*, from the most *circulating* ones, where they take the form of rules of law, professional or institutional prescriptions, or operating manuals that are “designed to travel far and be adopted (...) in a large number of situations” (p. 236), to the most *inscribed* ones. In the latter states, tools carry “features that have been forged during the production of (their) circulating form, but that are then the object of translation and inscription tasks necessary to make it effective” (p. 238). In this state, the tool “only really exists in its situated uses: [...] it is what it is used for – whether by individuals or groups” (p. 237). This framework is especially relevant in this study since, by putting tools at the center of the analysis, it offers a methodologically consistent thread that makes it possible to reconstitute and analyze the successive stages and events through which tools and their layouts affect the social processes that they contribute to mediate. The study of each of the various *states* of tools “call for analysis and data collection methods that are adapted to them” (see Figure 2 below). In each case, it is “a question of examining the tool and the particular form it takes in its context and linking it to the actors who are pertinent to understand the tool as well as those who designed it” (p. 237).

Data collection and analysis

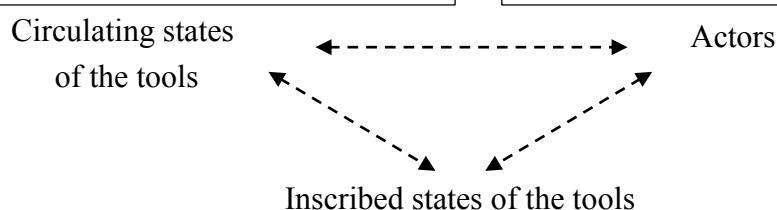
Figure 2 shows how we gathered and confronted three sources of empirical materials in this study: internal documentation (including all procedures, memorandum, and forms related to risk assessments and their integration in work practices), interviews with designers and users of the tools, and direct observation, notably by shadowing the actors.⁷ While the internal documents mostly encapsulate the circulating state of the tools under scrutiny, the shadowing allowed us to describe them in their inscribed state. The interviews with actors eventually revealed the reflexivity involved in their design and their uses.

⁶ “Causality is attributable neither to human actors directly, nor to artefacts alone. The effects [...] emerge, borne by a processual rationality that cannot be reduced to the properties of the human and non-human elements that constitute it” (Chiapello & Gilbert, 2019, p. 233-234).

⁷ We followed each of them for at least half a day by accompanying them in their daily routines, which ranged from solitary activities (e.g., completing credit files or updating client ratings) to attending appointments with clients and discussing cases with their managers, controllers at the headquarters, or colleagues. We consequently observed them not only analyzing their cases and making up their minds but also discussing and justifying their choices to their managers and clients.

Figure 2: A tool-based approach – design and data collection

Documentary resources	Number of documents	Interviews	Number of interviewees	Interviews total duration (hours)
General organization memorandum	4	Loan officers	22	39
Delegation procedures, memorandum and forms	12	Branch managers	5	10
Credit files and risk scoring procedures and memorandum	4	Regional directors	5	11,25
Risk Rating procedures and memorandum	20	Credit managers and their managers (Loans Department)	5	11
Risk Pricing procedures, memorandum and forms	5	Risk controllers and their managers (Risk Management Department)	9	15,75
Total Procedures, memorandum and forms:	45	Other relevant organizational actors	7	10,75
Risk Rating reports and analyses	9	Total :	53	97,75
Over-limit and accrued credits reports	7			
Risk Pricing reports and analyses	2			
Total Internal Reports and Analyses:	18			
Paper credit files and related annexes	23			
Electronic credit files, related annexes and corresponding information available in the information systems	57			
Total Credit files:	80			
TOTAL Number of internal documents:	143			



Observation	Number of working days	Formalised shadowing of practices	Number of actors shadowed	Number of half-day sessions
Headquarters	37	Risk Management Department:	3	5
Branches of the bank	40	Loans Department	4	4
Total:	77	Branch 1: a "big" (> 20 employees) branch in the Paris area	4	12
		Branch 2: a "medium" (10 to 20 employees) branch in the Paris area	5	12
		Branch 3: a "small" (< 10 employees) branch in the Paris area	3	6
		Branch 4: a "medium" (10 to 20 employees) branch outside Paris	2	3
		Branch 5: a "big" (> 20 employees) branch outside Paris	3	5
		Total:	24	47

The empirical evidence comes from a survey that one of the authors conducted at the MB's headquarters and in five of its branches in 2009–2010. Following Chiapello & Gilbert (2019)⁸, we reconstituted the history of the tools in the organization, identified how actors locate them in the organizational context, how they relate to the tools, use them, and talk about them, and identified the different effects of the tools, the situations where these effects are observable, and the transformations of practices that accompany them. To this end, following Pernkopf-Konhäuser (2014), Cloutier & Langley (2013), and Boltanski (2011), we regarded actors as “competent” both to understand and describe a given situation and to assess what is “appropriate” to handle it. Therefore, we viewed actors as able to justify or criticize specific rules and processes, and suggest that following how the risk management tools of the MB are designed, used, envisioned, and criticized, allows us to understand how risk management tools may contribute to creating conditions favorable to both pragmatism and reflexivity when organizations implement new quantitative technologies.

⁸ See “Appendix 2 - Studying the tool in its various states” (Chiapello & Gilbert, 2019, pp. 250-253).

In this process, the immersion of one of the authors in daily routines and concrete work situations helped better understand the viewpoints expressed by actors and the interests at stake. It also proved useful to reassure them of our intentions and to let them go of their defense mechanisms, especially when non-standard (and possibly non-compliant) practices or critical views were involved. Nevertheless, interviewees' accounts were contrasted and compared with each other and with observed practices and formal data available in our documentary resources (data from the information systems, internal reports, procedures, memorandum, etc.). This allowed us to assess the trustworthiness of specific narratives and to distinguish shared issues and central mechanisms from more marginal problems or views⁹. These issues were also continually debated with the other author, whose position in the teaching field makes it possible to maintain an ongoing discussion with dozens of bankers from the main banks operating in France. These interactions helped us envision and discuss alternative interpretations and ensure that the study focuses on issues currently permeating the French banking field. Finally, we drafted an in-depth description of the case to enable readers to evaluate both its trustworthiness and its relevance (Lincoln & Guba, 1985).

The next section presents the two approaches to risk management at work in the bank, which reflect recent evolutions in the bank's institutional environment. We show how these approaches infuse the bank's risk management tools and actors' perception of their work, and we explain how attempts at crystallizing compromises between different modes of evaluation (Chenhall et al. 2013) have failed. In section 6, we show how specific combinations between these approaches and their tools eventually emerged, and how their role in the overall credit management process was defined and stabilized. This leads us to reflect on how this system may have triggered reflexivity and pragmatism. To analyze our data, we make use of Boltanski's (2011) framework as presented in Figure 1 and, in particular, the distinction between the practical and metapragmatic registers of action.

5. TWO APPROACHES TO RISK MANAGEMENT AND THE DIFFICULTIES IN CREATING COMPROMISING ACCOUNTS

In this paper, we distinguish between two main approaches pervading the field of credit risk management: professional and statistical. While the professional approach has traditionally dominated the field of credit risk management in Europe, the statistical approach has been adopted and translated into practices by European banks in the context of the passage of the Basel II Agreement (2004) and its implementation into European law in 2006¹⁰.

In the professional approach, the grounding idea is that credit risk is a social phenomenon that have to be "assessed" (Kalthoff, 2004: 69)¹¹ and bankers performing risk management take for granted the fact that there always will be areas of judgment in their activity. As described by Kalthoff (2005: 70), this judgement operates at different stages of the risk management process: In the fact, for instance, that bankers use their "*own model of economic representation*" of their prospect or client's situation; that they usually proceed to "*direct observations*" of their client or prospect's "*economic reality*" in order "*to gain an impression*" and forge "*sensory*

⁹ Two discussions of the preliminary results – one organized as an open forum during the annual seminar of the MB and the other with the executive committee – also helped us in this process.

¹⁰ This occurred through the adoption of Directives 2006/48/EC2 and 2006/49/EC3, which went into effect on December 31, 2006.

¹¹ As Kalthoff (2004) puts it: "banks granting loans to corporations are faced with the problem of how to assess whether their loan will be repaid or not."

certainty”; and, eventually, that they add “*hand-written comments*” to the internal evaluation that they made of their client.

This mode of evaluation opposes the one of the statistical approach, which is grounded in the idea that the “true” level of credit risk can be – and thus should be – “measured,” which implies that human judgment should not interfere in the process. While the professional approach to risk management instantiates in the confrontation of judgements made by professionals who have built up their expertise throughout their careers, that is, disciplinary objectivity (Porter 1995), the statistical approach to risk management relies heavily on a quantitative technology that applies explicit statistical methods and tests (Dubois, 2006), that is, mechanical objectivity (Porter, 1995), to a set of data. Statistics are specific quantitative technologies in the sense that they require predefined data sets, and cannot perform on-the-fly adaptations, as machine-learning would do (Kitchin & McArdle, 2016). But, beyond these specificities, they are, like all quantitative technologies, meant to process from data to calculations in a mathematically coherent way. They consequently provide a form of impartiality that, following Porter (1995), is the very reason why quantitative technologies have been more and more favored by regulators. These technologies can also be defined by the specificities of the “moral economy of science” (Daston, 1995) that accompany them. They are “concerned with the same question: how do we learn from data?” (Wall, 2018, p. 57). In other words, quantitative technologies’ proponents consider “data-driven” as “the standard of legitimate knowledge” (Porter, 2020, p. viii), and use data both to build and assess quantitative models. This data-driven “new empiricism”, in turn, revitalizes the promises of metrological realism (Desrosières, 2008) and calculative idealism (Power, 2005, 2007; Mikes, 2009), that is, a way to “know everything”, “of anything”, “free from bias”, and “with a high degree of certainty” (Rieder & Simon, 2017, p.91). In this context, increases in knowledge are only hampered by practical issues regarding data availability, quality or method development (Jungherr & Theocharis, 2017).

As Porter explains, what advocates of quantitative technologies “of course mean, but may have neglected to say, is that the data should be reliable and suited to the purpose”, (2020, p.viii). Yet, data are necessarily incomplete (Jungherr & Theocharis, 2017) and conservative (Sætra, 2018). Moreover, not all data meet the criteria of impartiality and impersonality that matter to regulators, fewer even may do so while providing the “rigid continuity” that is required to produce the “robust, comparable longitudinal datasets” that are suitable for statistical analysis (Kitchin & McArdle, 2016, p.8). This challenge is particularly acute for small samples (Coveney et al., 2016). To process from data to calculations, then, the technical requirements and constraints linked to the standards of “proof” (Desrosières, 2008) of the statistical mode of evaluation may become highly structuring.

The statistical approach to risk management as defined in this paper echoes the “IT-based system” studied by Thomsen and Skærbæk (2018), which enacts a specific “reality” of risk management, carrying its own “purpose” as well as its own “assemblage of technologies” expected to provide statistical reliability. It is also evident in the “style” of risk management “by-the-numbers” identified by Mikes (2011) in the banking industry, where “metrics have colonized hitherto uncontrolled areas of uncertainty” and expanded to the domain of human judgement.

The first part of our analysis presents how these two approaches instantiate into the risk management tools at MB. We then analyze why attempts to create compromising accounts by mixing the modes of evaluation derived from each approach have failed at the MB.

The tools conveying the professional approach to risk management at the MB

The MB is one of the French banks with a mutualistic heritage: Its banking relationships demonstrate a commitment to supporting its clients' development. To this end, loan officers strive to have a detailed knowledge of their clients' activities and projects, even though most of their clients are small or very small organizations. They regularly visit their clients and actively participate in the local networks and groups to which they are attached (occupational, trade, professional, and community associations, etc.). Unlike most of the retail banking sector, whose local agents are expected to display mostly commercial skills, the MB – as is typical in a mutual bank – is structured as a decentralized organization of “field-based” professionals with some degree of autonomy and initiative:

Compared with other banking networks [...] our branch offices are closer to their business centers. In the other networks, [...] as soon as the turnover becomes consequential, they send it to the business centers. Expertise regarding the analysis and the processing of applications is in the business centers. In the business centers, the directors manage teams of skilled professionals. We are kind of an intermediary between the two. (Regional Director 1)

Specific tools and processes have been developed to facilitate coordination and transmission of expertise among loan officers but also to control their actions. Loan application files and the scoring system are the most prominent examples of this double purpose.

To be approved, any loan application requires the completion of a loan application file by the loan officer in charge of the account. It is made of at least eight sections, organized as follows:

1. Type, purpose, and terms of loan;
2. Business description and background;
3. Financial analysis;
4. Forecast analysis;
5. Justification for the request;
6. Final risk analysis (estimate of the credit risk borne by the bank);
7. Conclusions (opinion of the loan officer and local/regional/central management if required);
8. Decision form (specifying the approved terms).

Past, current, and projected financial statements are also annexed to the file.

Unlike with consumer and residential loans, business loan applications at the MB are not highly standardized: Each section is completed by the loan officer in charge of the account according to the specificities of the case. This mode of assessment has the advantage of adapting to specific situations, thus allowing “fine-tuned” (Huber and Scheytt, 2013) analyses. However, this approach makes it difficult to compare risks, aggregate them, or even get a rapid idea of what they are. To this end, the MB has established a complementary system called the “credit-scoring system.”

In the “credit-scoring system,” the score goes from A to E: A is “very good,” B “good,” C “should be monitored,” D “excessive,” and E “troubled case.” The score reflects the loan officer's opinion about the quality of a customer from the point of view of risk. In this context, non-quantifiable and qualitative elements are supposed to play a “very important” role in the scoring process.¹²

As one credit officer underscored:

Internal scoring [...] is the scoring we assign to our client, so it incorporates all the information, including the quality of management, future developments, and sectorial specificities, so our scoring encompasses a bunch of judgments that are not financially observable. Therefore, it is

¹² “Principles and description of the scoring system” – MB internal document, June 1999.

more subjective than anything else, but on the other hand, it really adds value because it gives a sense of the internal vision [of the banker], which is interesting. It's usually a good complement to the conclusions of the last loan application; both are correlated. (Banker 3, Branch 1)

However, the score is not supposed to reflect only a personal opinion: Its objectivity is ensured by a systematic review by branch, regional, or central managers. As they generally do not amend scores without prior discussion with the actors in the branch, the score helps coordinate professional judgments about risk assessments within the MB. Indeed, as the Risk Management Director explains, the score represents “a hyper-summarized expression of our opinion on a file,” but “this whole system of risk quality, it took years to bring it about in our culture, and now, it is so entrenched in our daily practices that saying A, B, C, D, E has turned into a way to communicate: ‘How much is it?’ – ‘B’ – ‘Oh, okay, a B.’”

Both tools carry a mode of evaluation of risks where objectivity is based on the shared judgment of experienced professionals who study each project in detail through ad-hoc analyses. However, this approach is not the sole one governing risk management at the MB. During the 2000s, the MB also adopted a statistical approach to risk management in most of its operational activities.

The tools conveying the statistical approach to risk management at the MB

The statistical approach to risk management at the MB takes the form of a dedicated computer application that internal actors call the “risk-rating engine.” For all clients whose turnover is less than EUR 5 million and whose bank commitments are less than EUR 1 million (the so-called “retail” segment), the “engine” automatically assigns a grade that cannot be changed manually. For companies whose commitments exceed EUR 1 million or whose turnover exceeds EUR 5 million (the “corporate” segment), limited rating adjustments may be proposed by loan officers and eventually incorporated by risk managers.

The system distinguishes between 18 broad risk categories: Two categories (D and C) are for customers with late payments or overdraft excess lasting more than 90 or 30 days, and 16 “regular” grades (ranging from 1 for the “less risky” exposures to 16 for the “most risky” exposures) distinguish the other clients. Each of the grades corresponds to a given probability that a default (late payments or overdraft excess lasting more than 90 days) will occur over the next 12 months. As the probabilities of default associated with each loan are used to assess the bank’s capital adequacy for its credit activities, the rating process is a very sensitive issue, and rating criteria are tightly controlled and monitored.

To comply with the standards of proofs of the statistical mode of evaluation, the bank used historical data on prior default experiences to build its rating system and estimate the probabilities of default corresponding to each grade. The rating criteria are statistically identified to differentiate between risks, group the homogenous exposures, and ensure the predictive power of the selected probabilities of default.

The development of such a system presupposes that standardized information is recorded in the bank’s information systems and is sufficient to form statistically representative cohorts, which are necessary for the rating engine to reliably discriminate between borrowers and predict their likelihood of default. It follows that the factors the “engine” considers are based on standardized and formal representations of organizations rather than the specific knowledge acquired by bankers.

For the retail segment, the rating system relies on four sets of data: (1) identity information (country, industry, turnover, amounts outstanding at the MB), (2) external ratings, when available, (3) information from financial statements, and (4) data on banking behavior (current account balance, average maximum lending balance, amount of savings, duration of banking relationship, presence of overruns of overdraft or unpaid bills, etc.). For the corporate segment, the system requires collecting a fifth set of data regarding the environment, identity, and strategy of the firm. In practice, only a few variables are identified as being able to predict the risk of defaulting for each sub-group of clients. For example, the engine uses only eight variables to rate transportation companies: Four of them are financial ratios (working capital and capital turnover ratios; debt-to-income ratio; value-added rate), which determine around 40% of the final grade, and four are linked to recent banking behavior (average of overdraft in excess over a recent period; the number of payments delayed due to insufficient funds over a certain period; total savings in cash or kind; the sum of credit transactions on the account), which makes up around 60% of the final grade.

Another important aspect of the rating system is that – unlike staff expertise – it is based on replicable formulas and is, therefore, vulnerable to intellectual property theft. These formulas and their “ingredient lists,” such as the one described above for transportation companies, are very closely guarded. The system is then almost totally opaque to field bankers, as well as their line managers, which creates issues:

I totally lack a frame of reference for this system. Totally. Sure, if the firm is rated 3 or 4, it's great. To me, it's a good rating. But, then, what is the difference between a firm rated 5 and another that is rated 9? Often, I don't know. What is the most important? Is it the functioning of the accounts? The debts? The financial health of the company? [...] It's very obscure. Very, very, very obscure. (Banker 3, Branch 4)

This opacity is in line with the overall rationale of the system, which is to provide an independent risk assessment. It prevents any potential attempts from those in charge of assigning ratings to develop “gaming” strategies. For the same reasons, “human intervention” in the form of manually adjusting the ratings is strictly limited¹³ and even impossible for clients belonging to the retail segment. Even the fifth set of data mentioned above, which is used only for the corporate segment, refers to a statistical approach in that it relies on standardization and quantification. It stems from a one-size-fits-all multiple-choice questionnaire with about 40 questions that translate qualitative data into variables to be used for quantitative measures.

In sum, the rating engine relies on an extensive use of data but the technical requirements linked to statistical evaluation and the regulatory push to frame and limit the use of professional judgment restrict the perimeter of the analysis to a small number of highly standardized data and its scope to very broad and abstract categories of exposures, that are then treated as perfectly homogeneous. This strongly contrasts with the mode of evaluation of the loan application file and its related credit-scoring system.

The failure of attempts to build compromising accounts

On two occasions, the bank has attempted to craft a compromise between the two modes of evaluation but has failed in this endeavor: they have, for the most part, been kept separate at the MB.

¹³ It requires filling in two additional documents and getting validation from the risk management team.

At first, the risk management team of the bank had thought they would be able to comply with regulatory requirements by building a statistically reliable risk assessment tool based on the preexisting scoring system. But they soon gave up on this idea:

So, when the whole Basel II project began, [...] we thought a bit about it, about how it would turn out for the Mutual Bank. [...] We met the risk management team of the UMB group and presented our stuff. It wasn't that stupid, you know. It was far too simplistic but not stupid. And the first reaction of the Group Risk Management Director was: "It will never be approved." To determine probabilities of default, our idea was to rely on the risk assessments we usually used in the credit files, not on such a big machine with backtests, statistics, and so on. We hadn't grasped the scope of it [...], that it has to be running like a well-oiled machine to get approved. (Chief analyst, Risk Management Department)

This first attempt to merge both modes of evaluation in one tool failed because the bank's institutional referents expected strict compliance to the statistical approach and its data-driven systems of proof. The bank then committed to finding another compromise by adding the professional approach to the statistical approach rather than the other way around. Specifically, it drafted a rating tool in line with the regulatory requirements conveying this approach but complemented this tool with an element that referred to the professional approach. It took the form of the fifth set of data that bankers have to collect to complete the ratings for the corporate segment. As we have seen, it relies on standardization and quantification. Yet, by incorporating non-financial data and mobilizing professional knowledge of the field within which clients operate, the questionnaire also provides visibility to professional evaluative principles, thereby acknowledging their importance.

Following Chenhall et al. (2013), visibility is key to building productive friction between multiple evaluative principles. However, we observed quite the opposite: The questionnaire highlights the latent divergences between both modes of evaluation and fuels contempt and criticism. When bankers are filling out the questionnaire, they regard it as irrelevant at best:

Personally, I think the questionnaire is a bit of a sham because it should have way more details. (Banker 2, Branch 2)

In many cases, when having to answer the questions, they cannot resist making ironic or critical comments:

Are the offered services diverse enough? I don't know, he's a plumber, and we all need plumbers. Let's stop putting more and more constraints on ourselves; we can manage a great deal of these cases quite well without all this, [...] without putting everyone in a box. (Banker 1, Branch 3).

The bankers perceive the questionnaire as being unable to ensure that they can communicate their knowledge about the specificities of their client because it does not respect the idiosyncratic nature of such professional knowledge. More generally, in this situation, the two modes of evaluation are perceived to be incompatible with each other. In particular, what matters most in the professional mode is understanding the specifics of each case, which makes a statistical approach based on broad aggregations and comparisons irrelevant:

[The average representative firm] doesn't exist; it's the same thing when we're talking about industries: Industry brings together thousands of professions, so it's difficult to compare one company [with another]. For example, in my file, I can have truckers, I can have people who do bending, I can have people who do laser cutting, I can have people who do other kinds of cutting, who work with the press, all kinds of employment; we can have thousands of types of jobs, and each has its own particular problems. We just can't compare them. (Banker 4, Branch 4)

Conversely, bankers are aware that their own assessments may be viewed as “subjective” (Banker 3, Branch 1) because they are not based on clearly identifiable and quantifiable elements. They acknowledge that the professional assessment is “not an exact science” (Banker 1, Branch 1) and “does not look rational” (Credit Manager 2, Headquarters). This suggests they do not criticize the statistical approach because they see it as irrelevant per se but because, in this situation, there is a contradiction that they cannot resolve.

While the bank had come up with the new rating tool and the compromise it represented to replace traditional scoring, this is not what happened. Clients applying for a loan continued to be “scored” in addition to being “rated,” which we interpret as a failure of the compromise. Pluralism remained, however, since two distinct indicators sustained by two distinct approaches were kept in operation. This raises questions about how to organize their overlap to support the kind of reflexivity that pluralism may trigger while, at the same time, ensuring coordination, compliance and avoiding intractable conflicts. As the next section shows, this situation has neither driven the MB into an escalation of intractable conflicts nor suffocated the professional approach under coercive requirements to fully adhere to the statistical approach. Instead, we observe creative combinations between tools that allow the MB to comply with the regulatory requirements commanding a statistical approach to risk management while, at the same time, preserving its distinctive professional feature. This process deeply transformed the MB’s practices and sparked criticism and fears. However, it also helped identify and clarify some of the ambiguities that the previous system had not addressed.

6. LAYERING AND SELECTIVE COUPLING IN RISK MANAGEMENT PRACTICES

To understand the configurations of modes of evaluation in the MB risk management practices, it is necessary to get into the details of operational processes. As previously mentioned, at the MB, any loan application requires the completion of a loan application file. This task is conducted by the loan officer at the branch office. The file presents the case and the banker’s conclusions about it. A reminder of past and current risk scores follows these conclusions. The rating from the rating engine is also systematically recorded on the first page. In total, the file contains three distinct assessments of the client’s riskiness: (a) the score attributed by the branch officer in charge of the client, (b) the internal rating provided by the rating engine, and (c) the credit rating assigned to the client by an external institution, the Banque de France (BDF). The latter rating is based on reported trade bill and loan payment incidents at the national level. As we have seen, the MB’s internal score – or scoring – is embedded in a professional approach, while the internal rating follows a statistical approach to risk management.

Three decisions are made based on this file: Should the loan be granted? If yes, at what price? And, who is to decide? This translates into three risk management concepts, namely, (1) risk selection, (2) risk pricing and (3) delegation and control. The three decisions are tightly linked together. If the demand is rejected, there will be no risk pricing. Similarly, the decision to grant the loan is correlated with the risk it poses, and, therefore, the level of delegation required from the banker who will authorize or refuse the loan. Initially, these decisions were made more or less altogether. With the introduction of the statistical approach, however, they end up being considered separately because the way they rely on one approach and/or the other differentiate them: Risk selection is coupled with the professional approach, but risk pricing relies on the statistical approach, and delegation and control involve both. After analyzing the configurations of modes of evaluation in the credit risks management process in each of these three moments,

we identify whether and how reflexivity is involved and, when relevant, what causes the switch from the practical to the metapragmatic register and the emergence of critique.

A risk selection process coupled with the professional approach

Surprisingly, the MB provides no rules or written recommendations regarding how to use, articulate, or prioritize the three risk assessments contained in the file. The conclusions provided by the banker on which decision will be made to grant the loan or not, are not following any specific guidelines. Resulting, in practice, loan assessments are very clearly led by the professional approach to risk management. In the 80 loan application files we examined, even when the rating was good, bankers never referred to it to justify their opinion in the concluding section of the file. This is partly because, as it is opaque, the rating provides little informational content for the risk analysis:

It remains a tool that's difficult to understand [...]. So, when we tell ourselves, "My god, this is really a risky client," it's not because of [the rating], that's for sure. (Banker 3, Branch 4)

But, more fundamentally, if loan officers do not rely on the ratings to assess the riskiness of a file, it is because they believe they are in a better position to fully evaluate the situation. As the same banker goes on to say:

We're the ones who are on the field, so we are seeing risk every day in the form of cash flow gaps, a company's overall shortcomings – sometimes, even risk pertaining to an individual as well. The death of a director, the absence of a key person, things like that. Then, of course, the financial statements; a business that doesn't promise any profitability is certainly risky, right? And behind that, there's the question of guarantees: Is there sufficient capital? Things that are pretty straightforward, but sometimes we do temper them by this relationship, by the fact that we know the company, that we know what kind of machines it has, which technologies it's developing, what plans it has, its ambition, or what commercial possibilities it has, etc. So, we're always tempering; we tend to have more qualitative indicators that allow us to see things differently. (Banker 3, Branch 4)

In this quotation, the banker adheres to the professional approach to risk management, as he insists on his and his colleagues' intimate knowledge of the client, which allows them to reevaluate – or “temper” – the riskiness of a client. As opposed to knowledge embedded in a statistical approach, their knowledge here is individual and non-replicable. The predominance of the professional approach may be explained by the fact that bankers are personally engaged in loan applications. They have to justify their decisions and cannot do so based on a risk assessment that is opaque to them.

Even at the upper levels of the hierarchy, when the intervention of managers is required to authorize a loan, following the rules of delegation (as we will see below), the professional approach dominates. The control exerted on branch offices bankers is seen as a way to further develop rather than contain their professional expertise:

What I do when I review the files, as I have some experience, is I try to see things from the perspective of those who will receive the file, and I ask myself, "They don't know the client, they never saw him, they don't know the file, so will they understand?" Because the young rep tends to go through important points too quickly because it seems obvious to them. [...] Then, I also help those who are less experienced and haven't seen what is important; I help them improve their financial analyses because it is important for credit activities. It's a pedagogical work: drawing their attention to specific points requiring further explanation, refining the financial analysis, and all the commercial work: why we will grant on overdraft facility of EUR 20,000 and not EUR 15,000 or EUR 25,000, which supporting documents we will require, and what will happen if it goes wrong. (Manager, Branch 5)

In sum, even if the ratings are put on the first page of the loan application files, it is the professional approach that drives risk selection. The bank relies on bankers' intimate knowledge of their clients and the market in which they operate and on disciplinary forms of control by managers to authorize a loan. Here, the coupling of the risk selection practices with the professional approach involves a decoupling with the statistical approach: The bankers are not trying to craft practices that would be a compromise between the two modes of evaluation. On the contrary, they are very reluctant to rely on the ratings for risk selection and pleased that there is no explicit incentive to do so.

This is true at the branch level:

In some other banks, reps have objectives linked to the ratings [...] reps have objectives to establish relations with clients who got a good quality rating, and, in some banks, it's even prohibited to establish relations with clients who got a poor rating, which is fortunately not the case today at the MB. (Manager, Branch 2)

It is also true at the headquarters level:

If the day is coming, and I'm afraid our profession is tending a bit in this direction, when decisions will be based exclusively on ratings [...] then I have to admit that we will possibly have a problem [...]. Part of our job is also deliberating the grounds and reasons for our decisions. (Credit manager 1, Headquarters)

However, even though the professional approach dominates risk selection, the statistical approach is not left out from daily practices; in fact, it is now profoundly structuring risk pricing.

A risk pricing process coupled with the statistical approach

At the MB, the decision regarding the rate that will be offered to a client for their loan always comes after the decision has been made to authorize the loan¹⁴. The rate of a loan has three components: (1) the funding cost incurred by the bank; (2) the operating costs of servicing the loan, including direct and indirect costs; and (3) a risk premium to compensate the bank for the degree of default risk associated with the loan. Funding and operating costs are calculated for each kind of product, while the risk premium is linked to the credit rating of the borrower. The adjustment of the rate to risk is significant. For long-term loans, for example, it may increase rates up to 160 points (1.6%). This risk-adjusted rate is a target rate. Once it is defined, the rate can be adjusted up or down by the bankers, but downward adjustments are limited to 20 points (0.2%). Branch managers can grant adjustments of up to 50 points (0.5%), but any further adjustment is very strictly regulated and requires special authorization from the headquarters.

Surprisingly, although field actors ignore the ratings in the risk selection processes and criticize their overall relevance as a risk assessment tool, they do not seem reluctant to rely on the ratings with regard to risk pricing. For instance, the same manager who told us how important experience was in risk selection processes explains:

I think [the pricing system] sensitizes our reps. Since [...] the rates we can offer depend on the ratings, it is smarter... because we knew it, informally, but we weren't really doing it. Roughly speaking, when we wanted to offer a loan, we used to make it cheap, and when we didn't want to do it, we used to make it expensive. But, today [...] we realize that a client of poor quality

¹⁴ Accordingly, it is always the part of their loan applications files that the bankers fill in last. However, this is an unwritten rule as the rate being offered is reported, together with the rating, on the first page of the files.

with poor collateral doesn't deserve the same rate as a client of very good quality with good collateral. (Manager, Branch 5)

Bankers at the branch level who are responsible for the customer relationship also emphasize the advantages of this system:

In the previous system, we were only concerned with competition. With some good-quality clients, we made them pay expensive rates because they wouldn't look around elsewhere, and some poor-quality customers were so multi-banked and so concerned about competition all the time that we were giving them an unbeatable rate. And it's true that we realized this and wondered: "But why are we doing this when, in the end, this customer represents a significant risk but pays less?" We felt this wasn't necessarily logical, especially because some banks already had risk-based pricing systems, and we didn't. (Banker 4, Branch 1)

As this quotation shows, without the rating-based pricing system, bankers had difficulties dealing with price competition to limit the bargaining power of multi-banked clients looking for good prices. By providing reference points to deal with this issue, the rating system now counter-intuitively provides bankers with an alternative to crude exposure to price competition and, thus, becomes useful.

Adapting tools and practices and confronting the problem of price-based competition, even with an imperfect tool, is seen as a better option than leaving bankers to deal with the commercial bargaining issues on their own:

So if they pay 4.5% or 5%, basically that won't fundamentally change anything. But at the same time, it doesn't help ... In the end, the system is made like that, and it has its rationale: Inevitably, if there is more risk, it's a cost to be passed on. I still find the system more logical than one where we only considered the commercial aspect. (Banker 1, Branch 4)

They also more readily endorse the pricing system because pricing is not their main concern. For most bankers at the MB, risk selection is what matters the most, and risk pricing is a secondary issue that is not necessarily worth fighting for:

The reality is that for complicated loan applications, the problem isn't the interest rate. In the end, what's really at stake is knowing whether or not you're going to give the loan; so if I have an application that's problematic in terms of risk but I still feel is worth investing in, I don't worry much about the rate, which may be comparatively higher than what we usually give. My real challenge is getting the loan accepted. In the end, I mean, the rate rarely has a significant impact on whether or not the project survives. (Banker 6, Branch 5)

Overall, this suggests that the statistical mode of evaluation encompassed in the rating engine dominates the risk-pricing practices at the MB because it fills a gap. The professional approach does not provide norms and principles to address competitive pressures and the bargaining power of customers. Besides, the institutional pressures pushing for the adoption of a ratings-based pricing policy are very strong, and, in a context where this could be an issue, this is considered a secondary issue, behind loan acceptance.

A delegation process layering risks assessments

Bankers' autonomy regarding decision making is governed by a specific system that awards bankers levels of delegation, sets their lending limits, and specifies controls and monitoring processes. This delegation system determines who will decide on a given loan: namely, the banker, their branch manager, or managers at the headquarters. It also determines whether control will be exercised ex-post or ex-ante. Thus, it distinguishes between situations where it is required that two different people at distinct hierarchical levels give their opinion on a case before making a decision, from situations where only one banker's opinion is needed. In that,

the delegation system ensures an important function of institutions, which is to sort out “what can only be considered once, in association with a context [...], and [...] what it is appropriate to look at twice” (Boltanski, 2011, p. 75).

At the MB, the system is based on the banker’s position inside the organization. At the branch level, there are five different levels of delegation, depending on seniority and experience: The lower two are reserved for branch officers, and the three highest mostly – albeit not exclusively – for branch managers. Five higher categories concern only regional and central managers. A senior manager at a branch can, for example, authorize loans of up to EUR 230,000. For larger amounts, the loan will have to be pre-authorized by a loans manager at the headquarters. Moreover, the system is risk-adjusted: Depending on the risk category to which the client belongs, bankers may have more or less autonomy. Three main risk categories have been created to this end: “superior rating,” “intermediate rating,” and “exclusion.” For the “superior ratings,” decision-making autonomy is increased. The limit applied to senior managers in branches, for instance, is extended from EUR 230,000 to EUR 300,000. By contrast, “exclusion” leads to a total withdrawal of decision-making power from the branch offices’ level and a switch to ex-ante control at the headquarters.

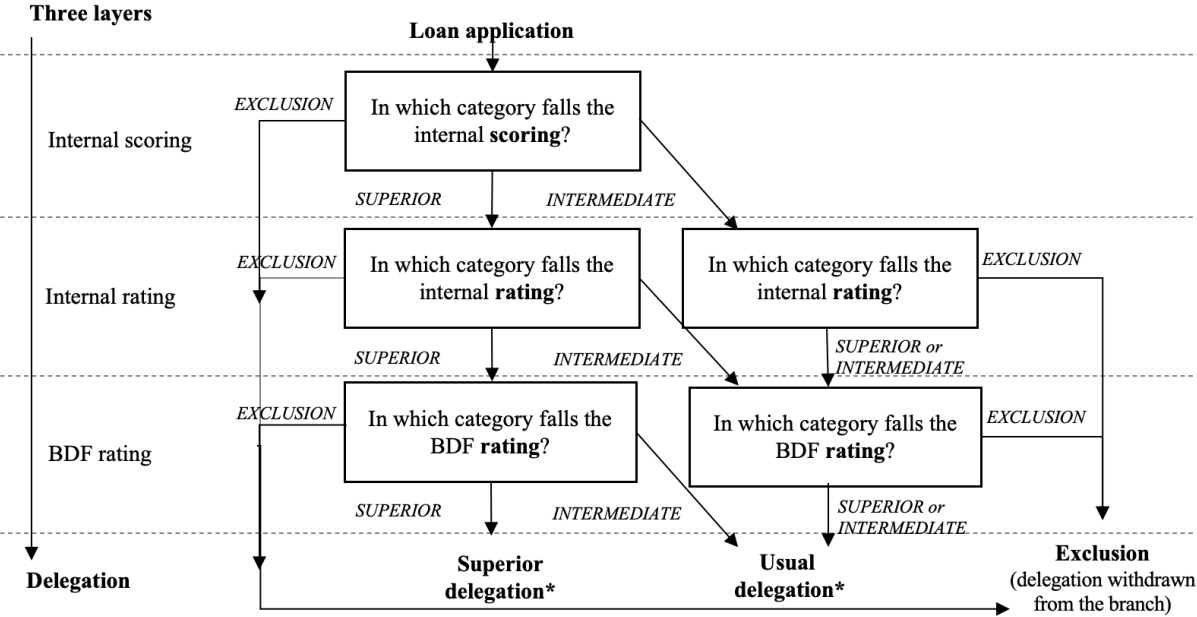
To classify clients among the three main risk categories, the MB uses the three different risk assessments required on the loan application file: (a) the score attributed by the branch officer in charge of the client, (b) the internal rating provided by the rating engine, and (c) the credit rating assigned to the client by the Banque de France (BDF).

As displayed in Figure 3 below, to determine whether a client will have a “superior” or “intermediate” rating or enter the “exclusion” process, these three risk assessments are neither merged nor ranked – they are layered. This “layering” of risks assessments functions as follows: Rules and constraints introduced by each of the three tools providing the assessments are added on top of each other. There is no possibility to compensate for a poor assessment from one tool by relying on one or the two others. Accordingly, any client classified by one of the tools as excessively risky falls into the “exclusion” process, regardless of the classifications obtained from the other tools. By contrast, the layering of constraints implies that in order to be classified as “superior” a client must reach this category by all three tools. In the other cases, clients are classified as “intermediate.”¹⁵

¹⁵ The thresholds defining each of these categories for the three tools are:

- Superior ratings: A or B for the internal score; 1 to 7 for the internal rating; 3++ to 4 for the BDF rating.
- Exclusion process: E for internal score; 15,16, DX, CX for the internal rating; 8, 9, P for the BDF rating.

Figure 3: Delegation process at the MB



* The level of delegation depends on the position of the banker instructing the loan application and the profile of the client applying for a loan (e.g., corporate, retail, etc.).

The categorization obtained through the three kinds of assessments can diverge and bankers may lose some autonomy. Because, with this layering configuration, the delegation system takes into account “the worst” among potentially three distinct assessments, it increases the chances that a given application will be re-examined by another actor, who stands at a distance from the case and at a higher hierarchical level.

From practical to metapragmatic moments: pluralism and the layout of tools

In this section, we analyze, for each of the configurations presented above, the situations that made bankers switch from practical moments to metapragmatic moments and become reflexive, in order to identify the thresholds that triggered this switch and the role tools play in this.

As previously shown, the statistical approach that infuses the rating tool is perceived as irrelevant for assessing the riskiness of a client. When asked their opinion regarding risk selection, some bankers clearly engage in critique, suggesting that statistical rating is intrinsically problematic, while other explain that it is too difficult to understand to be able to make sense of it. In both cases, the rating appears so alien from their own reference point that reconciliation appears impossible, as illustrates the following quotation:

[T]he assessment that we make [...] doesn't always match up with the rating. [...]. We have loan applications, [...] where we find ourselves with companies that are very well rated but are absolutely disgusting and, on the other hand, [we'll encounter] companies that are very badly rated and yet are financially functioning quite well. So why is that? I really don't know what's in the engine. I don't know how it's handled; I don't know how it works. (Credit Manager 1, Headquarters)

However, in practice, most of the time, this situation of “broad incompleteness” is not salient. This is because the rating and the score often converge, which makes bankers conclude that “most of the time [the rating] is reliable” (Banker 1, Branch 5). And if they do not converge,

bankers may try to “make do” with the system by justifying, ignoring, and/or minimizing dissonances, as the following quotation illustrates:

Personally, I often keep the automatic rating because it is not so far off the mark. And anyway, if it suggests 10, it wouldn't change much if I put 8 or 9; so, as I believe it's a system that's been made by intelligent people, let's be optimistic: If it suggests 10, it's for [a reason] [...] and ... roughly speaking, it fits [with our assessment]. (Banker 2, Branch 1)

In the case of risk pricing, which is based on selective coupling with the statistical approach, bankers mostly consider it appropriate to rely on ratings. We suggest that this is because it provides a coherent way to deal with a preexisting and unresolved practical issue. This does not imply, however, that, they agree with this pricing system:

So is that justice? I don't know. To say that a customer who isn't doing well must pay more ... is that right? That's the big question! Anyway, today it's like that: The customers who are doing the worst are those who are paying the most. (Banker 4, Branch 1)

This shows that field bankers can criticize a specific mode of evaluation and, at the same time, when it is deemed appropriate, rely on it and even adhere to its principles – or “confirm” them, in Boltanski's words.

The practical register is dominant in situations of risk selection and risk pricing: actors are focused on performing specific tasks, and, as each of these tasks is “selectively coupled” with the “most appropriate” approach, they usually display a low level of reflexivity. They may switch towards the metapragmatic register of action, but in a rather unstructured way, depending on what appears to be mostly an individual sensibility, as the example above shows, or only in very extreme situations where divergences are especially salient. Selective coupling, in this perspective, appears very efficient to ensure daily coordination and limit controversies to extreme cases:

There are also negative effects linked to the fact that bank management has wanted to link pricing with the rating system. This is totally laudable since there are capital adequacy issues. So, it is absolutely not stupid, except that a client got a poor rating only for one reason – because he had an overdraft excess for 30 days, or whatever. There is no reason why his rating should be totally downgraded and that we won't be able to make him an offer corresponding to his financial statements or, at least, corresponding to our idea of the rate competitors will offer. (Manager, Branch 4)

We see here that one of these “extreme cases” is when clients have a poor rating “only” because they had an overdraft excess for 30 days. It constitutes a “threshold” that make this banker shift towards the metapragmatic register, but, since bankers do not possess any coherent rationale that can be readily compared to the rating to assess what could have been the rate offered to the client otherwise, this threshold is only met in “extreme cases”.

This contrasts with what happens with the delegation process, where the layering of risk assessments provides foothold to compare the rating and the scoring and, thus, to assess their potential discrepancies. Indeed, the delegation system provides two reference points. The first is when the client fall from the “superior” to the “intermediate” category, which happens when the internal score is worse than B, the internal rating worse than 7 or the BDF rating worse than 4. The second one, which triggers an exclusion, corresponds to an internal score of E, an internal rating of 15 or worst, or a BDF rating of 8 or worst. In these two cases, the layering process creates situations in which differences between assessments are turned into divergences. Together with the fact that it has practical consequences for the handling of the file, this constitutes a threshold that make the bankers shift towards the metapragmatic register.

As this banker explains:

For our clients, we've determined two categories, intermediate ratings – the poor ones and the good ones. That means we've classified our clients as good or weak. So, then, the important thing is that when you cross the line [...] what is called "superior ratings", it is from 1 to 7, [...] if you reach 8, it makes a difference. But whether you get 1 or 7 [...] doesn't change anything. So, it's not worth refining... (Banker 2, Branch 4)

The “difference” evoked in this quotation is that a rating below 8, as compared to a rating above 8, cannot be ignored anymore since bankers may lose part of their autonomy. It is even more obvious when the rating triggers an exclusion (ratings below 15):

Nowadays, a client with whom we enter into a relationship, who doesn't have any specific problems, [...] who (with regard to the Banque de France) doesn't have any rejections or risk warnings [...] can be transferred to the headquarters just because [...] the rating is negative. [...] But this client could have scored negatively just because they had an overdraft of EUR 2 [...] because of an account that they thought they had closed but had, in fact, not closed [...]. So there we have it: an arbitrary rating. Of course, it's not the end of the world, but it shows the disadvantages of the system. Because we're the ones who have to convince our higher-ups that the rating is unfair, we have to provide complementary arguments in addition to all that we've done from our risk-analysis perspective, just to say: "this rating is unfair." Even then, they don't always listen to us because they attribute such importance to this risk impact as it can have an enormous impact on our capital, as we well know; therefore, we have to work extra hard to [convince them]. (Banker 4, Branch 4)

Three distinct concerns are articulated in this quotation: (1) the fact that ratings can be arbitrary and not reflect the “true” riskiness of a client; (2) the institutional pressure to follow ratings’ assessments through capital requirements; (3) and the time and energy spent to support a loan application despite a deteriorated rating.

The first and second concerns signal an engagement in a metapragmatic register. They are reflexive in the sense that they do not refer to the practical consequences of the rating system but refer to its relevance. That being said, although they are intertwined, the two have yet to be distinguished. In the first concern, the rating is qualified as “arbitrary” and “unfair.” There is a gap between what it is assumed that the rating *should* display, that is, the “true” riskiness of a client, and what it displays *in reality* (Boltanski, 2011, pp. 67-70). Thus, it is clearly located in the register of critique. In the second concern, institutional pressures are invoked to justify the reluctance of superiors to go against ratings. However, this is not a critique but, rather, a “confirmation” of existing institutional arrangements, as suggested by the saying “as we well know” (Boltanski, 2011, p. 72). This illustrates the dialogical relation between critique and confirmation, or, in other words, the fact that critique is grounded on what is institutionally confirmed, to which it “counter-poses its lucidity” (Boltanski, 2011, p. 103). On the contrary, the third concern is practical as it relates to the task to be performed and the effort it takes. This illustrates the dialogical relationship between the practical and metapragmatic registers. The fact that the banker will have to put extra effort into building a case to defend the loan application before their superiors, which would not have happened had the decision remained within their level of delegation, triggers a shift towards the pragmatic registers. As a result, the layering of risk assessments not only ensures a stricter control on these cases, but also creates a situation that favor the emulation of reflexivity and its circulation from the bottom towards the top within the organization on the one hand, while ensuring its pragmatic limitation on the other hand.

Overall, the selective coupling that characterizes risk selection and risk pricing appears very efficient to ensure daily coordination and limit controversies. On the contrary, the layering structurally creates situations where bankers reflect on their practices, justify them, and question, challenge, or support how risks are managed at the MB. However, as engaging in

these situations raise pragmatic issues, this engagement is limited to specific cases and do not impede daily coordination. These findings show that the tensions did not disappear in the MB system after the failed attempts to build compromises but that they were managed in a way that limited their potential conflicts while stimulating reflexivity and its circulation within the organization.

7. DISCUSSION

Scholars have been worried that the development of highly quantified routines of risk management at the bank's operational level would bring about an unreflexive "illusion of control" (Huber and Scheytt, 2013). Owing to compliance imperatives, most banks developed similar "risk measurements systems" based on quantitative technologies and used them as decision-making tools at the most operational levels of risk silo management (Mikes, 2009, 2011). We examined how such tools were displayed in the credit management process of a French bank, introducing a statistical approach that conflicted with the professional approach that prevailed at the time. This happened in the context of strong institutional pressures exerted on the banking system to conform to this new approach, as described by Wahlstrom (2006, 2009, 2013) and Baud & Chiapello (2017). In this paper, we intended to outline the role tools may play in sustaining pluralism and favoring reflexivity within organizations that implement new quantitative technologies. More specifically, we set out to examine what kind of configurations of modes of evaluation could trigger a dialogical relationship between reflexivity and pragmatism when others would open the way for conflicts.

Overall, our study highlights that the layout of modes of evaluation and their related tools matters when it comes to understanding how pluralism may be sustained in an organization, and reflexivity triggered from this pluralism. In our case, the combinations proved to be better adapted than compromising accounts to prevent intractable conflicts in the bank's risk management processes. Combinations in the form of selective couplings, on the one hand, help to limit the expression of critique and to facilitate daily coordination by giving one approach precedence over the other. On the other hand, combinations in the form of layering create structural opportunities for the expression of critique but in specific cases only, which, in turn, nurtures reflexivity at the organizational level without inhibiting action. As a result, a dialogic relationship between pragmatism and reflexivity may emerge.

This work contributes to the ongoing discussion in the literature about the adoption of quantitative technologies and the maintenance of pluralism at the organizational level in four ways.

The difficulties of materializing compromises involving quantitative technologies

The difficulty of reaching a compromise in our case highlights the question of the limits of the "plasticity" of accounts (Mennicken & Power, 2015). One of these limits, we argue, is linked to the other "strength" of accounts, namely their materiality (Mennicken & Power, 2015). The materiality of accounting, that is, the fact that accounting formalizes operations into tangible objects, is central to compromising processes because it helps stabilize practices and diffuses them within time and space. However, our observations show that the "materialization" of compromises is also challenging.

Many professional practices of risk management rely on quantitative techniques, including at the MB. Accordingly, when they had to integrate the statistical approach to risk management, the bank's executives readily engaged to apply statistical techniques to their preexisting

professional mode of evaluation. However, they soon realized that this compromise could not meet the standards of “proof” (Desrosières, 2008) required under the statistical approach to risk management. As a result, the MB built another compromising account – the rating system – that was based on these standards of “proof” but provided visibility to the professional approach by incorporating a standardized list of professional criteria of expertise. When confronted with this system, bankers acknowledged its advantages, rationality, and objectivity. But when it came to using this “rating engine” as the main proxy for assessing a client’s riskiness or even to recognize it as indicating something about their “riskiness,” they could not help but point out the irrelevance of the questions they had to answer and, more generally, the lack of relevance of a standardized approach to risks. This points toward a failure not of the compromise itself but of its materialization: Bankers are not opposed to the statistical mode of evaluation per se, but, in the compromise as it is finally settled, they do not recognize anymore the distinguishing features of the professional evaluative principles. Indeed, by promoting a standardized approach to each case, the rating engine deeply conflicts with the “moral economy” (Daston, 1995) of the professional approach. Therefore, the mixed system elaborated by the MB cannot be “confirmed” and it rather generates ongoing critiques and prompts bankers not to rely on the internal rating when assessing a loan application.

The literature already suggests that compromising accounts are prone to generate endless debates about technicalities and lead to “stuckness” (Jay, 2013) rather than enabling “productive friction” (Chenhall et al. 2103, p. 263). From a metapragmatic point of view, any compromising device is doomed to be revoked at some point because of its inherent lack of coherence. However, our case suggests that compromising with quantitative technologies can be especially challenging for two reasons.

First, as quantitative technologies process from data to calculations, their legitimacy is especially attached to a strict conformation with the standards of “proofs” established for each step of this process. Strict conformation to these standards is supposed to eliminate the conventional – and thus potentially arbitrary – nature of decision making under uncertainty. On the contrary, loosening these standards would put the legitimacy of the whole system at risk, as our case suggests. It also explains why cultures of “calculative idealism” (Power, 2005, 2007) have very little tolerance for the ambiguities of mixed systems.

Second, it appears that, while it may be especially difficult to compromise with the standards of “proof” required by quantitative technologies, conforming to those standards can also be very limiting, which further restrains the possibility of materializing a compromise acceptable under both approaches. In our case, this was partly linked to how regulators framed the criteria of impartiality and impersonality applicable to the use of internal rating systems, which displays a strong suspicion towards human judgement (Baud & Chiapello, 2015). But it was also directly linked to the material requirements of the statistical technologies of risk management. It is indeed necessary to arrange a large number of observations documented through the same variables to product “sound” probabilities of default. Confronted with the limited availability of data, the junction of these requirements gave birth to a system, the “rating engine”, where very broad and abstract categories of exposures are analyzed based on a small number of highly standardized variables. Moreover, most of the details of this analysis are opaque to users. This show how the materialization of the compromise was constrained by technical requirements and issues regarding data availability, “quality”, or length. This matters to our analysis, since, as we observed, it is the materialization of this compromise – and not its underlying requirements – that contradicts the standards that would prevail under the professional approach and explain why, as soon as actors switch toward the metapragmatic register, they fail to find grounds to confirm the compromise and thus criticize it.

One could argue that the limiting effect of technical requirements could be related in our case to the “immaturity” of the MB data-systems or to the specificities of statistics as compared to

other, more flexible, quantitative technologies. However, independently of their capacity to perform on-the-fly adaptations or integrate heterogeneous sets of data, all quantitative technologies rely on data that are incomplete, of which only a fraction is relevant to the purpose and, this, in a way that is often opaque to users. In this context, our study highlights that the conformation with quantitative technologies' technical requirements in a given empirical setting may create limiting effects that contribute to explain the difficulties of materializing compromises with quantitative technologies and cannot be deduced from cultural and institutional factors alone. This suggests that, to better understand how quantification regimes interact with preexisting systems and modes of evaluation, further investigations regarding these technically driven, yet socially consequential, processes are needed, and that, as suggested by Chiapello and Gilbert (2019), they should be conducted down to the most inscribed states of tools.

Pragmatism and reflexivity: combining modes of evaluation within tools

At the MB, the failure of the second managerial attempt to structurally intertwine the professional and statistical approaches by creating a compromising account is evidenced by the fact that, contrary to what was originally intended, the rating system did not replace the preexisting scoring system. Instead, practices emerged that combined these two distinct tools and their approaches. Thus, our case points toward the crafting of combinations as a possible way to sustain a plurality of modes of evaluation in an organization confronted with institutional pressures to adopt new quantitative technologies. It also suggests that this configuration is able to sustain reflexivity. In our case, it takes the form of recurrent, albeit unsystematic, questionings of risk assessments. Two distinct forms of combinations are displayed in this process. First, risk selection and risk pricing are each “selectively coupled” with the “most appropriate” mode of evaluation: the scoring for the former and the rating for the latter. Different accounts are used as reference points for different tasks, which makes it possible to display differences in accounts of riskiness without turning them into divergences. On top of this, comes the “delegation process,” within which these two risk assessments are “layered”, which means that both the scoring and the rating are used to determine whether the loan application file can be handled at the most operational level or whether the branch manager – or even bankers at the headquarters – have to be involved in the decision. Most of the time, rating and scoring results converge: Bankers remain in a practical mode and move on to the next step in the loan application process. Sometimes, however, the layering system turns the differences between assessments into divergences because it associates the two assessments to different categories of delegation. Such cases have practical consequences since the use of the most restrictive of these categories is prescribed. The question of which mode of evaluation is the most appropriate to assess the riskiness of the case then resurfaces. Bankers enter the metapragmatic register and mobilize confirmation and critique, often dialogically, to justify their opinion before their senior managers. These additional efforts of justification help make their position explicit to those managers who then have to rule on the case by either confirming or rejecting this position. This way, the layering fosters a dialogical relationship between critique and confirmation which enables reflexivity and its circulation within the organization. Overall, this process embodies the idea that some cases should be looked at twice whereas others do not require it. However, as engaging in this process is constraining, engagement in such situations is limited and the process causes recurring disputes, but in a manageable number of cases. As a consequence, reflexivity is triggered but does not impede day-to-day activity: It is itself put in a dialogic relationship with pragmatic issues.

These results transpose to modes of evaluation the idea articulated by Pache and Santos (2013) regarding institutional logics, that, instead of adopting strategies of decoupling, compartmentalizing, or compromising, as the literature usually suggests, organizations often successfully rely on combinations to handle pluralism. Our case also shows that it is from a specific form of combination, namely, layering, that the organization achieves to articulate pragmatism and reflexivity and that this combination seems better suited than compromises to do so.

We suggest that the key difference between the compromises and combinations is located in the ability of the tools instantiating them to provide discrepancies or “dissonances” (Stark, 2009). By definition, a compromising account does not provide any discrepancy despite the incompatibility of the modes of evaluation it may encompass. Because of that, contrasting with what happens with layering in particular, actors lack a reference point that could be mobilized in their articulation of critique as much as of confirmation. This is what, in our view, may lead to “intractable conflicts.”

On the contrary, layering modes of evaluation provides such dissonances. This finding is coherent with Stark’s thesis according to which what he calls “pragmatic reflexivity” emerges from rivalry between evaluative principles, which “is not [...] compartmentalization, in which different principles of worth map to separate departments or units, bounded and buffered from contamination” (Stark, 2009, p. 35). Stark further suggests that the internal organization of a pragmatic form of reflexivity requires the recurrent confrontation of principles, but also “temporary settlements” in order to “get the job done” (Stark, 2009, p.35), which echoes Boltanski’s distinction between the practical and metapragmatic registers of action and their dialogical relationship (2011). As our case shows, selective coupling sustains the practical register of action—which is about getting the job done—because it compartmentalizes approaches, thereby limiting the possibility of a confrontation. In these situations, actors may switch to the metapragmatic register, but this will depend on their internalized reference points and be usually limited to extreme situations, where divergences are especially salient. On the contrary, the layering process provides thresholds beyond which differences are turned into divergences. We observed that actors assimilated these thresholds: when one was crossed, they usually switched to the metapragmatic register. This shows how tools and their material arrangements not only contribute to sustain or inhibit reflexivity but play a role in organizing it in a way that does not impede daily coordination. In our case, we suggest that the fact that the thresholds inscribed in the tools were raising attention only to a manageable number of cases and that they were associated with practical consequences helped this assimilation.

Selective coupling and the management of imperfect tools

Our case also contributes to prior work on imperfect tools, whether incomplete or opaque. Incompleteness stems from discrepancies between competing indicators and the legitimacy granted to an external reference point. To address incompleteness issues, Jordan and Messner (2012) point to local repairs and flexibility. In our case, combinations in the form of selective couplings seemed to facilitate the management of incompleteness issues because they allowed actors to rely, for each task, on the indicator they regarded as the most appropriate. This shows that, when tighter control is required, as was the case in this study, incompleteness can be attenuated by “qualifying” the modes of evaluation at play, as we observe in our case through the collective work done to define and specify the uses of the existing tools. This is particularly important in the context of banking since, contrary to what can be observed at the senior managers’ level, where it is possible to display some flexibility (Mikes, 2009, 2011), due to the

compliance imperative, most banks rely on statistical tools and use them as decision-making devices at the operational level (Mikes, 2009, 2011).¹⁶

Unpacking how selective couplings practices operate this qualification at the MB also helps us better understand the issue of the opacity of indicators, which is especially important regarding quantitative technologies since many of them are at least partly opaque to users. In our case, we show that the same opaque tool (i.e., the rating engine) can be heavily criticized because of its opacity while being simultaneously – although for another task – relied on and used in a way that is perceived as enabling *because* of its opacity. We explain these paradoxical results by relying on the idea that the same exigencies of coherence do not prevail in the practical and metapragmatic registers of action (Boltanski, 2011). In our case, attempts to compromise within tools directly call for metapragmatic judgments. In this register, the rating engine is criticized because its mode of evaluation is incompatible with the professional approach used as a reference point¹⁷. On the contrary, when ratings are used in the course of action, notably to deal with pricing issues, this is done in the practical mode. Since this mode requires low reflexivity, what matters is whether a given tool helps to handle a situation and coordinate actions. In such practical situations, actors demonstrate little concern either for the incompleteness or for the opacity of the indicator. In the latter case, one could even suggest, in line with Dambrin & Robson (2011), that the methodological opacity of the ratings becomes an enabling feature. During the action, this opacity allows bankers to “turn a blind eye” (Boltanski, 2011, p. 63) to the weak link between ratings and what they regard as the “true” riskiness of a project. Thus, opacity enables bankers to rely on this indicator and to deal with pricing issues.

Bankers’ ability to rely on a statistical mode of evaluation to perform the task of loan pricing despite their reluctance to move from their professional approach for risk selection suggests that the process we describe cannot be entirely considered as an attempt from bankers to defend their profession and its autonomy from deskilling processes, or only as a nostalgic claim to defend craftsmanship. On the contrary, our study shows that, even at the most operational level, actors may display calculative pragmatism (Power, 2005, 2007; Mikes, 2009) but that it depends, as Carlsson-Wall et al. (2016) suggest, on the situations encountered.

Towards more resilient credit risk management practices

Our study shows that risk management systems combining fine-tuned and quantified risk management are possible, and that they can be developed in an enabling way even at the most operational levels. This demonstrates the possibility to conform to institutional demands for risk sensitivity and risk-adjusted performances while simultaneously maintaining part of the specificities that characterize professional-based banks. Interestingly, what emerges from the MB arrangements is a system where professional control is applied upstream of statistical risk management. Indeed, risk selection, which is coupled with the professional approach to risk management, always intervenes before pricing, which is coupled with the statistical approach to risk management. In other words, bankers only proceed with the files they have already selected on a professional basis. This means that control rooted in the professional mode of evaluation constitutes the external point of view that makes it possible to criticize or confirm what the statistical mode of evaluation suggests. Thus, the combination of modes of evaluation

¹⁶ As shown by Caldarelli et al. (2016) even a small credit cooperative bank aiming to achieve both economic and social profitability may have difficulties transposing its duality of logics down to the most operational level of risk management tools and practices.

¹⁷ Its opacity then raises a supplementary critique: As ratings are opaque, they cannot be used for learning and enriching professionals’ understanding of a situation.

at work at the MB ensures that the statistical approach will not drive credit offers. However, once selected by professionals and, if required, reexamined at a more distant level, files are priced according to the statistical mode of evaluation, which allows building competitive offers for those credits that have been professionally selected. Regulating the supply of credit through pricing leads to bubbles and crises (Stiglitz & Weiss, 1981). Therefore, this combination, which helps maintain professional control over the selection of files while adjusting the price of these offers to account for competitive and statistical factors, may appear quite relevant to all those concerned with systemic stability.

This study consequently contributes to an appreciation of the diversity of banking and the diversity of the roles played by quantified risk management tools in the credit markets. The role and the history of scorings and ratings in the US credit markets are well-documented (Marron, 2009; Poon, 2009; Carruthers, 2013). But the situation in markets such as France, where scores and ratings are not commoditized and remain internal to the banks, is studied less often (Lazarus, 2012). As a matter of fact, our study shows that this situation allows several ways to qualify and quantify credit risk, and many more ways to manage it. However, the viability of these arrangements in the long run as well as outside the MB remains an open question. As demonstrated by Poon (2009), it was *within* the tradition of “credit control-by-screening characterized by simple but rigid barriers of exclusion designed to sift for acceptable credit quality,” that the shift toward the “regime of control-by-risk” that led to the financial crisis occurred. Whatever the specific tools, their combinations, or their paths, this shows that the capacity of organizations to trigger and maintain pragmatic reflexivity about their use of quantified technologies is and will remain a critical part of their resilience.

8. CONCLUSION

As quantitative technologies gain footholds in many walks of life, our results contribute to better understand their relationship to professional work. However, this paper does not focus on professions *per se*, but rather on what they are supposed to foster in a quantifying world, that is, reflexivity. Our study sheds a new light on reflexivity itself, showing that it is not only attached to the survival of the professional approach to risk management alone, but may instead gain momentum in the confrontation of the two worlds. It emphasizes the potential of combinations and the role played by tools in triggering and maintaining a kind of reflexivity able to give way to pragmatism in a timely manner, but further investigations need to be done to achieve uncovering the conditions of possibility of pragmatic reflexivity, notably in terms of pre-existing cultures. Hopefully, our work constitutes a step toward a better understanding of how quantitative technologies could be put, along with professions, at the service of organizational innovation and resilience, and socio-technical approaches developed as part of both a critical and emancipatory project (Chiapello & Gilbert, 2019).

References

- Annisette, M., Vesty, G., & Amslem, T. (2017), Accounting Values, Controversies, and Compromises in Tests of Worth, *Research in the Sociology of Organizations*, 52, 209-239.
- Baud, C., & Chiapello, E. (2015). Comment les firmes se financiarisent : le rôle de la réglementation et des instruments de gestion. *Revue française de sociologie*, 56(3), 439-468.

- Baud, C., & Chiapello, E. (2017). Understanding the disciplinary aspects of neoliberal regulations: The case of credit-risk regulation under the Basel Accords. *Critical Perspectives on Accounting*, 46, 3–23.
- Bhimani, A., & Willcocks, L. (2014). Digitisation, ‘Big Data’ and the transformation of accounting information. *Accounting and Business Research*, 44(4), 469–490.
- Bishop, S., & Waring, J. (2016). Becoming hybrid: The negotiated order on the front line of public–private partnerships. *Human Relations*, 69(10), 1937–1958.
- Boltanski, L. (2011), *On critique: A sociology of emancipation*, Polity.
- Briers, M., & Chua, W. F. (2001), The role of actor-networks and boundary objects in management accounting change: A field study of an implementation of activity-based costing, *Accounting, Organizations and Society*, 26(3), 237–269.
- Caldarelli, A., Fiondella, C., Maffei, M., & Zagaria, C. (2016), Managing risk in credit cooperative banks: Lessons from a case study, *Management Accounting Research*, 32, 1-15.
- Carlsson-Wall, M., Kraus, K., & Messner, M. (2016), Performance measurement systems and the enactment of different institutional logics: insights from a football organization, *Management Accounting Research*, 32, 45-61.
- Carruthers, B. G. (2013), From uncertainty toward risk: the case of credit ratings, *Socio-Economic Review*, 11(3), 525-551.
- Chenhall, R. H., Hall, M., & Smith, D. (2013), Performance measurement, modes of evaluation and the development of compromising accounts, *Accounting, Organizations and Society*, 38(4), 268-287.
- Chiapello, E., & Gilbert, P. (2019). *Management Tools: A Social Sciences Perspective* (Cambridge University Press).
- Christensen, M. and Skærbæk, P., 2007. Framing and overflow of public sector accountability innovations – a comparative study of reporting practices. *Accounting, Auditing and Accountability Journal*, 20 (1), 101 – 132.
- Cloutier, C., & Langley, A. (2013), The logic of institutional logics: Insights from French pragmatist sociology, *Journal of Management Inquiry*, 22(4), 360-380.
- Clune, C., & O’Dwyer, B. (2020). Organizing dissonance through institutional work: The embedding of social and environmental accountability in an investment field. *Accounting, Organizations and Society*, 85, 101130.
- Coveney, P. V., Dougherty, E. R., & Highfield, R. R. (2016). Big data need big theory too. *Philosophical Transactions of the Royal Society A*, 374, 20160153.
- Dambrin, C., & Robson, K. (2011), Tracing performance in the pharmaceutical industry: Ambivalence, opacity, and the performativity of flawed measures, *Accounting, Organizations and Society*, 36(7), 428–455.
- Daston, L. (1995). The moral economy of science. *Osiris*, 10, 2-24.
- Desrosières, A. 2008. *Pour une sociologie historique de la quantification : L’Argument statistique I*. Paris: Presses des Mines.
- Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis*, 51(1), 47- 69.
- Fischer, M. D., & Ferlie, E. (2013), Resisting hybridisation between modes of clinical risk management: Contradiction, contest, and the production of intractable conflict, *Accounting, Organizations and Society*, 38(1), 30–49.
- Friedland, R., & Alford, R. (1991), Bringing Society Back In: Symbols, Practices and Institutional Contradictions. In W. Powell & P. Dimaggio (Eds.), *The New Institutionalism in Organizational Analysis*. University Of Chicago Press, 232–263.
- Gehman, J., Lounsbury, M., & Greenwood, R. (Eds.). (2016). *How institutions matter!*. Emerald Group Publishing.

- Greenwood, R., Raynard, M., Kodeih, F., Micelotta, E. R., & Lounsbury, M. (2011), Institutional complexity and organizational responses, *Academy of Management annals*, 5(1), 317-371.
- Huault, I., & Rainelli-Weiss, H. (2011), A market for weather risk? Conflicting metrics, attempts at compromise, and limits to commensuration, *Organization Studies*, 32(10), 1395–1419.
- Huber, C., & Scheytt, T. (2013), The dispositif of risk management: Reconstructing risk management after the financial crisis, *Management Accounting Research*, 24(2), 88-99.
- Jay, J. (2013). Navigating Paradox as a Mechanism of Change and Innovation in Hybrid Organizations. *Academy of Management Journal*, 56(1), 137–159.
- Jordan, S., & Messner, M. (2012), Enabling control and the problem of incomplete performance indicators, *Accounting, Organizations and Society*, 37, 544–564.
- Jungherr, A., & Theocharis, Y. (2017). The empiricist's challenge: Asking meaningful questions in political science in the age of big data. *Journal of Information Technology & Politics*, 14(2), 97-109.
- Kalthoff, H. (2005). Practices of calculation: Economic representations and risk management. *Theory, Culture & Society*, 22(2), 69-97.
- Keys, B. J., Mukherjee, T., Seru, A., & Vig, V. (2010), Did securitization lead to lax screening? Evidence from subprime loans, *The Quarterly journal of economics*, 125(1), 307-362.
- Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1), 1-10.
- Lazarus, J. (2012), Prévoir la défaillance de crédit : l'ambition du scoring, *Raisons politiques*, 48, 103-118.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
- Marron, D. (2009), *Consumer Credit in the United States*, Palgrave Macmillan.
- Mennicken, A., & Espeland, W. N. (2019). What's New with Numbers? Sociological Approaches to the Study of Quantification. *Annual Review of Sociology*, 45(1), 223.
- Mennicken, A., & Power, M. (2015), Accounting and the plasticity of valuation. In A. Berthoin Antal, M. Hutter and D. Stark (Eds.), *Moments of valuation: Exploring sites of dissonance*, Oxford University Press, 208-228.
- Mikes, A. (2009), Risk management and calculative cultures, *Management Accounting Research*, 20(1), 18–40.
- Mikes, A. (2011), From counting risk to making risk count: Boundary-work in risk management, *Accounting, Organizations and Society*, 36(4–5), 226–245.
- Mouritsen, J., Hansen, A., & Hansen, C. Ø. (2009). Short and long translations : Management accounting calculations and innovation management. *Accounting, Organizations and Society*, 34(6), 738-754.
- Pache, A. C., & Santos, F. (2013), Inside the hybrid organization: Selective coupling as a response to competing institutional logics, *Academy of Management Journal*, 56(4), 972-1001.
- Pernkopf-Konhäusner, K. (2014), The competent actor: Bridging institutional logics and French pragmatist sociology, *Journal of Management Inquiry*, 23(3), 333-337.
- Polzer, T., Meyer, R. E., Höllerer, M. A., & Seiwald, J. (2016). Institutional Hybridity in Public Sector Reform: Replacement, Blending, or Layering of Administrative Paradigms. In *How Institutions Matter!*. Emerald Group Publishing Limited.
- Poon, M. (2009), From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance, *Accounting, Organisations and Society*, 34, 654-674.

- Porter, T. (1995), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University Press.
- Porter, T. (2020), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University Press (2nd. Ed.).
- Power, M. (2004), *The risk management of everything: rethinking the politics of uncertainty*. Demos.
- Power, M. (2005). The invention of operational risk. *Review of International Political Economy*, 12(4), 577-599.
- Power, M. (2007), *Organized uncertainty designing a world of risk management*. Oxford University Press.
- Power, M. (2009), The risk management of nothing, *Accounting, Organizations and Society*, 34(6), 844–855.
- Power, M. (2010). Fair value accounting, financial economics and the transformation of reliability. *Accounting and business research*, 40(3), 197-210.
- Power, M. (2016), *Riskwork: Essays on the Organizational Life of Risk Management*, Oxford University Press.
- Quattrone, P. (2016). Management accounting goes digital: Will the move make it wiser? *Management Accounting Research*, 31, 118–122.
- Raynard, M., Lounsbury, M., & Greenwood, R. (2013). Legacies of logics: Sources of community variation in CSR implementation in China. In *Institutional Logics in Action*, Part A. Emerald Group Publishing Limited.
- Rajan, U., Seru, A., & Vig, V. (2015), The failure of models that predict failure: Distance, incentives, and defaults, *Journal of Financial Economics*, 115(2), 237-260.
- Rieder G., Simon J. (2017) Big Data: A New Empiricism and its Epistemic and Socio-Political Consequences. In: Pietsch W., Wernecke J., Ott M. (eds) *Berechenbarkeit der Welt?*. Springer VS, Wiesbaden.
- Sætra, H. S. (2018). Science as a vocation in the era of big data: The philosophy of science behind big data and humanity's continued part in science. *Integrative Psychological and Behavioral Science*, 52(4), 508-522.
- Stark, D. (2009), *The sense of dissonance: Accounts of worth in economic life*, Princeton University Press.
- Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American economic review*, 71(3), 393-410.
- Themsen, T. N., & Skærbæk, P. (2018). The performativity of risk management frameworks and technologies: The translation of uncertainties into pure and impure risks. *Accounting, Organizations and Society*, 67, 20-33.
- Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012), *The Institutional Logics Perspective: A New Approach to Culture, Structure, and Process*. Oxford University Press.
- Wahlström, G. (2006), Worrying but accepting new measurements: The case of Swedish bankers and operational risk, *Critical Perspectives on Accounting*, 17(4), 493–522.
- Wahlström, G. (2009), Risk management versus operational action: Basel II in a Swedish context, *Management Accounting Research*, 20(1), 53–68.
- Wahlström, G. (2013), Bank Risk Measurement: A Critical Evaluation at a European Bank, *Accounting and Finance Research*, 2(3), 24-36.
- Wall, L. D. (2018). Some financial regulatory implications of artificial intelligence. *Journal of Economics and Business*, 100, 55-63.