The tone from above: The effect of communicating a supportive regulatory strategy on reporting quality

ABSTRACT:
As part of their regulatory strategy, authorities may request firms to submit mandatory self-assessments on a periodic basis. The effectiveness of these strategies highly depends on the quality of the information that firms provide. We conduct a field experiment in the financial intermediation setting in the Netherlands, and assess whether firms’ reporting quality is influenced by the extent to which official communications reflect a supportive regulatory strategy (low vs. high support). In collaboration with the Authority for the Financial Markets, we manipulate the content of official letters that instruct firms to submit a mandatory self-assessment. We hypothesize that, as compared to low support letters, high support letters lead to higher reporting quality (as indicated by comprehensiveness and accuracy of reporting, and whether firms respond to non-mandatory questions). We further expect this effect to be weaker for larger firms, and even reversed when firms are short-term oriented.

Please do not cite or distribute.

This paper is written as part of a two-stage editorial process
(also known as a Registered Report Process).
1. Introduction

Across diverse regulatory settings, authorities request proprietary information from firms as part of their monitoring and supervision strategies. This is particularly relevant in contexts in which there is a significant information gap between authorities and firms (Besanko and Sappington [1987], Pautz and Rinfret [2016]). For instance, firms requested to comply with tax, customs, financial, or environmental regulations are more commonly confronted with mandatory self-assessments in which they are requested to report sensitive information regarding their accounting, management, or specific economic activities. Evidently, the effectiveness of such supervision strategies directly depends on the quality of the information provided by firms (Peeters [2006]). Inaccurate, incomplete, or untimely information obstructs monitoring and makes supervision more difficult, time consuming, and costly.

Self-reporting commonly involves a series of actions, which include gathering detailed information, completing extensive questionnaires, and meeting tight deadlines (Ernst & Young [2012]). The quality of the information reported to the authorities is determined by these actions, as they directly influence accuracy and comprehensiveness, as well as the extent to which firms voluntarily report additional information. In this sense, reporting quality can be conceptualized as a behavioral outcome that depends on firms’ motivation and capacity to attain high reporting quality.

Because effective supervision benefits greatly from high reporting quality, a fundamental question is whether authorities can exert influence that leads to improved reporting quality. This is a non-trivial problem, as it may not be sufficiently addressed through deterrence (i.e., by threatening firms with penalties for not reporting information with the highest level of quality). Firms may comply with minimum reporting requirements, and yet provide low reporting quality (e.g., by reporting approximate figures). Establishing whether low reporting quality sufficiently reflects reality – and whether it deserves a punishment – imposes additional challenges for the authorities. From a supervisory standpoint, the core problem is that authorities may explicitly request the highest level of reporting quality, but in practice may have limited resources to monitor and assess the level of reporting quality itself.
Following the literature on responsive regulation (Braithwaite [2003, 2007]), one way in which authorities might be able to influence reporting quality is by signaling a supportive regulatory strategy towards firms. The rationale is that, as the strategy focuses more on guidance and support, firms’ motivations to attain high reporting quality increase. The concept of responsive regulation establishes that there are distinct types of regulatory strategies, which vary in terms of how supportive they are (Braithwaite [2007]). A strategy that focuses less on deterrence and more on education and service-delivery motivates firms to share information. Moreover, an educational approach helps firms to “make sense” of what is requested from them and encourages compliance (Fairman and Yapp [2005]). In this sense, support enhances the perception that the authority is fair and trustworthy, and prompts reciprocity from at least a fraction of supervised firms (Alford and Speed [2006]). Overall, there are distinct reasons to expect a positive effect of a supportive strategy on how firms respond to the authority’s requests, including the request to provide high quality information.

Because not all firms can be visited at the same time, authorities have limited options to present themselves as helpful and exert this type of supportive influence at an industry-wide scale. For a large number of firms, interactions with the authority are constrained to infrequent official communications (in the form of formal letters). Taking this point into consideration, in this study we examine whether reporting quality is influenced by the extent to which official communications reflect a supportive regulatory strategy. Our hypothesis is that, as compared to less supportive communications, more supportive communications lead to higher reporting quality. We acknowledge, however, that the strength of the expected effect may vary depending on firms’ specific characteristics and motivations (Braithwaite [2003, 2007]). Hence, we further hypothesize that the expected effect is weaker for larger firms, which might be less sensitive to the content of official communications, and even reversed for firms with a short-time horizon, as these more likely behave in an opportunistic manner, maximizing short-term gains and minimizing unnecessary efforts.

To test these expectations, we conduct a field experiment in the regulatory setting of financial intermediation in the Netherlands. The Authority for the Financial Markets (AFM), which supervises the financial markets and financial services providers in this country, granted us permission to introduce an experimental treatment in their 2017 communications, and granted us
access to (anonymized) firm data. The AFM employs a supervision instrument called Market Monitor, which is a mandatory, standardized self-assessment instrument set out yearly among approximately 7,000 financial intermediaries. Through this instrument, firms are requested to provide an elaborate amount of sensitive information related to, for example, financial products offered, specific financial intermediation activities, number of employees and customer relationships, as well as detailed accounting and managerial information. Every year, the AFM sends three letters to each firm providing information on how to complete and when to submit the Market Monitor.

Based on prior experience, the AFM recognizes that intermediaries vary in the amount and accuracy of the information that they report. After a series of discussions with the AFM, it was jointly determined that there are valid reasons to expect that the quality of reporting can be influenced by the content of the letters in which firms are instructed to complete and submit the Market Monitor. As a result, we designed a field experiment in which we employ two distinct sets of letters (which contain information about the instructions, username, and password required to complete the Market Monitor), and which reflect a regulatory strategy based on low vs. high support. We identify three elements that relate to support, and which can be manipulated in order to influence firms’ perceptions and reporting behavior. These elements are: accessibility (absence of obstacles or barriers to ask for help or guidance), purpose (goal alignment between the authority and the firm), and tone. High support letters emphasize that the regulator is accessible to provide support on how to complete the Market Monitor, identify shared goals, and have a supportive tone. In contrast, low support letters only indicate where to find information without expressing the willingness of the regulator to provide support, do not identify shared goals, and have a strict tone. Our objective is to compare the effects that each set of letters has on reporting quality, as indicated by comprehensiveness, accuracy, and voluntary provision of information (i.e., answering non-mandatory questions). In addition, we plan to examine how the strength of this effect varies with firm size and time horizon. Our results have the potential to inform regulators about the value of signaling different enforcement strategies through official communications to different types of firms.

The study aims at contributing to the literature in several ways. In the empirical accounting literature, reporting quality is typically examined with respect to financial accounting reports
using publicly available data (e.g., Balakrishnan, Core and Verdi [2012], Biddle, Hilary and Verdi [2009]). In contrast, we examine reporting quality in a regulatory context where firms face the compulsory demand to self-report sensitive information to an authority. Analyses of regulatory reporting are uncommon, likely due to difficulties in obtaining access to sensitive and confidential data. Similar to the argument that financial reporting quality reduces information asymmetry with investors (Balakrishnan et al. [2012]), high reporting quality reduces the information asymmetry with the authority, enabling it to better fulfill its objective of regulatory oversight.

A second contribution relates to how authorities may be able to influence firms’ reporting quality. In a setting with infrequent contact with firms, communicating a particular regulatory strategy may have significant impact on reporting behavior. This provides critical importance to the way an authority presents itself in its official communications. A third contribution relates to the possibility of identifying boundary conditions for the expected effect, as it may only apply to certain types of firms.

From a broader perspective, regulators are generally reluctant to participate in field experiments given the political sensitivity and magnitude of the responsibility associated with using differential treatments between firms. However, there is great value in examining these types of questions in regulatory practice, as there is uncertainty on whether regulation-related questions are suitable for lab experiments (Choo, Fonseca and Myles [2016], Lindeboom, Van der Klaauw and Vriend [2016]). Prior field experiments on regulatory strategies have focused primarily on tax reporting, testing whether the communication of particular strategies affects tax compliance. These experiments mainly include theory-based manipulations related to deterrence and support, such as varying monitoring or punishment levels (Blumenthal, Christian and Slemrod [2001], Wenzel and Taylor [2004], Hasseldine, Hite, James and Toumi [2007], Ariel [2012], Fellner, Sausgruber and Traxler [2013], Hallsworth [2014], Castro and Scartascini [2015]), and normative appeals and social persuasion (Ariel [2012]; Hallsworth, List, Metcalfe and Vlaev [2017]; Torgler [2004]). Despite the pervasiveness of self-reporting in other settings than tax regulation, field experiments in these other settings are rare. This provides much scope for the use of a field experiment to examine the effects of theory-based manipulations on
reporting behavior. In this sense, we respond to the recent call for field experiments that examine regulatory effects in practice (Leuz and Wysocki [2016, p.602]).

2. Theory and Hypotheses

a. Self-reporting

There are several reasons why supervisory authorities may be interested in employing self-reporting instruments. Seminal theory in this area indicates that self-reporting can save costs associated with the effective detection of noncompliance, and reduce uncertainty about punishments for both the authority and regulated agents (Kaplow and Shavell [1994], Heyes [2000]). Subsequent theoretical developments have focused on the contexts of tax, environmental, and financial regulations, which are encompassing and permeate a variety of industries. A common element across the literature, however, is the emphasis on information asymmetry, whereby authorities have limited access to information about firms’ activities and operations (this is also referred to as an information gap; Baron and Myerson [1982], Sappington [1982], Besanko and Sappington [1987], Pautz and Rinfret [2016]). The fundamental problem for authorities is that monitoring is costly, as they have limited capacity and resources to process technical and extensive information in order to effectively assess whether firms are failing to comply with regulations. Self-reporting can help reduce information asymmetries, and in this sense, it can be seen as a complement of a given supervisory strategy. As a result, modern regulatory regimes may transfer control and reporting functions to corporate management (referred to as “first instance monitoring”), while the final responsibility for oversight, inspections, and sanctions remains on the side of the authorities (Parker [2002]).

Self-reporting instruments can differ in terms of whether they are mandatory vs. voluntary, periodic vs. sporadic (submitted only upon request), completed and submitted in paper vs. online, or in terms of how extensive and detailed they are. In most cases, however, self-reporting instruments are based on the request of sensitive information, and for this reason, they tend to be treated as confidential. Self-reporting instruments can also have limitations. These mainly relate to potential low response rates (due to, e.g., unclear instructions, incapacity to reach the targeted population, unmeetable requirements or deadlines), untruthful reporting or intentional
misreporting (Friesen and Gangadharan [2013]), and low reporting quality (i.e., a level of quality that is insufficient in terms of timeliness, accuracy, or completeness). Although untruthful reporting has by definition low reporting quality, the distinction between these two potential limitations is fundamental. This is because self-reporting can be truthful, and yet be limited to the bare minimum that is required by law. For example, low reporting quality can involve rough approximations of numerical or date-related responses, and relate to firms’ unwillingness to answer non-mandatory questions or provide additional clarifications and perspectives on a voluntary basis (in, e.g., open-ended questions). For supervisory authorities, high reporting quality is therefore the most desirable outcome.

b. Reporting quality as a behavioral outcome

Self-reporting comprises a series of actions, which broadly relate to properly registering contact information so that the firm can be reached by the authorities, reading and understanding the instructions of the self-reported instrument (on how and when to complete and submit responses), requesting information internally (e.g., to different employees or departments within the firm), aggregating and processing gathered information (e.g., making calculations or synthetizing evaluations), inputting and reviewing responses (by, e.g., board members or heads of departments), and submitting responses on time.

Reporting quality – in terms of accuracy, comprehensiveness, and the provision of voluntary responses – can be affected by any of the above-mentioned actions, and therefore can be conceptualized as a behavioral outcome. As such, reporting quality can be explained by firms’ capacity and motivations to strive for high reporting quality.

c. The potential influence of a supportive regulatory strategy

This study builds on the theory of responsive regulation (Braithwaite [2003, 2007]), which establishes that regulation is more effective when distinct regulatory strategies are properly aligned with corresponding compliance motivations. Regulatory strategies can vary in terms of how intrusive they are. More intrusive strategies rely more heavily on deterrence (involving audits, inspections, punishments, and direct control), whereas less intrusive strategies rely more heavily on persuasion and education (Healy and Braithwaite [2006]). One key objective for the
authorities, therefore, is to identify which motivations firms have, and respond accordingly with suitable regulatory strategies.

Firms may have economic and non-economic motivations to comply with regulations (Nielsen and Parker [2012]). Non-economic motivations relate to social or normative considerations (i.e., giving value to the perceptions of significant others or being committed to act responsibly – over and above economic cost-benefit analyses). Abundant research suggests that normative motivations are enhanced by the perception that the authorities’ rules and decisions are fair, trustworthy, and legitimate (Malloy [2003], Parker and Nielsen [2011]).

There are several reasons why a supportive strategy fits well into the context of supervision and self-reporting, and may motivate firms to strive for high reporting quality. First, a supportive strategy places particular value on the exchange of information. A regulatory strategy oriented towards education and service-delivery is expected to motivate firms to share information and engage in open dialogue with the authority (Braithwaite [2007]). The focus is not on detecting wrong-doing, but on using information to attain more important goals. Firms are treated not with suspicion or cold detachment, but rather with respect and in a fair manner (i.e., less as distant regulatees and more as clients or business partners; Alford and Speed [2006], Kirchler, Hoelzl and Wahl [2008]).

Second, the authority can use this strategy to explain why it needs extensive or detailed information. This is particularly important, as providing explanations about legal requirements and advice on how to implement them helps firms to “make sense” of the what is requested from them and encourages compliance (Fairman and Yapp [2005]). For example, the authority can indicate that the information will be used to gain a better understanding of firms’ states and needs, and consequently design and implement tools and strategies that better align firms’ interests with the interests of the authority and the industry as a whole. Notably, an effective shift from external to internal monitoring requires convincing explanations of why reliable internal diagnostics are useful for both the authority and the firm (e.g., for tracking and improving performance over time; Sinclair [1997]).

A third reason, which follows from the prior two, is that a supportive strategy enhances perceptions of informational fairness (provision of information in a sufficient, clear, and timely manner), procedural fairness (incorporation of firms’ perspectives and needs on how to respond
to official requests), and interpersonal fairness (treatment of firms with dignity and respect through for example a supportive tone; for a review of fairness dimensions, see Colquitt, Conlon, Wesson, Porter and Yee [2001]). These perceptions enhance the legitimacy of the authority’s requests, and increase the willingness to submit and obey on a voluntary basis (Tyler [2006]). Indeed, these perceptions of fairness may lead to reciprocating behaviors of supervised firms (Smith and Stalans [1991], Smith [1992]).

These elements suggest that a supportive regulatory strategy can be used to explicitly recognize and further promote both economic and normative motivations. Firms are better able to identify the material benefits of high reporting quality when they are provided with straightforward explanations of why it is needed and justified (which relates to the outcome of the request). In addition, a supportive regulatory strategy signals that the authority strives for responsiveness, simplicity, and accessibility, and that it cares about the firm and the relationship they share (which relates to the process of how the request is made and handled; Alford and Speed [2006]). By implication, a supportive regulatory strategy involves the respectful and empathic treatment of firms. All these elements have the potential to enhance fairness perceptions and normative motivations to improve reporting quality and to draw attention on its economic value. In terms of social exchange, a strategy based on support, guidance, and service-delivery can be seen as a “gift” from the authority, and may trigger reciprocation from at least a fraction of firms (Alford and Speed [2006]).

d. Official communications that reflect a supportive regulatory strategy

Official communications are at the core of a supportive regulatory strategy, especially in the context of self-reporting. They reveal the social distance between the authority and the firm (Braithwaite [2003]), and indicate how the firm is perceived and treated by the authority (e.g., as a suspect vs. a client or business partner; with empathy and respect vs. suspicion and detachment). Furthermore, official communications are a principal mechanism through which a service-delivery approach is materialized at an industry-wide scale, as these are needed to properly (i) explain how to obtain materials, tools, and information that firms may require to successfully respond to the authority’s request (accessibility); (ii) describe and justify the purpose of the request (goal alignment), and (iii) display a respectful and open attitude towards the firm (tone). For these reasons, official communications are inherently linked to how firms in general
develop perceptions about the regulatory strategy – not only in terms content, but also in terms form, style and tone.

This logic applies to the specific case of official communications in which the authority explicitly requests high reporting quality, which is the main focus of our analysis. Hence, we state our first hypothesis as follows:

_Hypothesis 1: The extent to which official communication reflects a supportive regulatory strategy is positively associated with reporting quality._

*e. The moderating effects of firm size and firms’ time horizon*

Hypothesis 1 predicts average differences in reporting quality between firms that receive communications of different nature. However, there are also arguments against this expectation. In a firm-level – instead of an individual-level – setting, the content of the communication can be ignored or insufficiently noticed by firms, and may therefore have little or no impact on behavior. In addition, regulations in the setting of our experiment (financial intermediation) have expanded and tightened in recent years, so letters reflecting less intrusiveness and more support could be perceived as “soft”, diluting the authority’s mandate to legally require firms to comply with its requests. One limitation of communicating a supportive regulatory strategy is that it can induce opportunistic behavior, as it relies less on deterrence (Alford and Speed [2006]). Taking these counter-arguments into consideration, we identify two relevant factors that can moderate the expected association between reporting quality and the nature of the communication.

_Firm size._ Firm size is indicative of the availability of resources and organizational specialization and professionalism. Reporting and compliance-related tasks are typically performed by owners or high-level managers in smaller firms, and by specialized departments or personnel in larger firms. As firms become larger, these tasks are more likely systematic and standardized, and interactions with the authority may be more frequent (due to relative importance in the market). For these reasons, larger firms might be less sensitive to the content of official communications.

One could also argue that compliance officers in larger firms may have relatively more resources available for gathering and reporting information, and that communicating a supportive
regulatory strategy could make them walk the extra mile. However, these officers are likely evaluated based on how they avoid issues with – and raise no concerns from – the authorities. Hence, we expect larger firms to be more responsive, independent of the level of support they receive. Our second hypothesis is depicted in Figure 1, and is stated as follows:

_Hypothesis 2: Firm size weakens the positive effect of communicating a supportive regulatory strategy on reporting quality._

This predicted moderation effect of firm size is also practically relevant as it informs regulators about the possible benefits of customizing communication for small firms and individuals as compared to large firms.

_Time horizon._ This refers to the extent to which firms value future payoffs. It plays a crucial role in inter-temporal decisions, in which costs and benefits are spread out over several periods of time (Loewenstein and Thaler [1989]). Reporting information can be analyzed as an inter-temporal decision-making process. High reporting quality is costly, as it demands time and effort, and thus can be seen as a cost in the short-run, but as an investment in the long-run. This is because high reporting quality can reduce future (expected) costs associated with potential audits and inspections, penalties, and legal disputes, and can also reduce time and expedite the process of gathering, storing, and retrieving information in the future. This goes in-line with the idea that long-term oriented agents value cost-reductions derived from lower supervision, and can thus be monitored to a lesser extent (Mendoza and Wielhouwer [2015]).

Short-term oriented firms are expected to act in a more opportunistic manner, focusing predominantly on minimizing costs. They may even blame the authority for the circumstances they are in, and perceive regulation as mainly costly and obstructive. This can trigger uncooperative motivational postures towards the authority and its requests (Braithwaite [2003]). But more importantly, short-term oriented firms may interpret a supportive tone as part of a weak enforcement strategy, which is relatively less focused on deterrence and therefore less likely effective in detecting and punishing deviance or noncompliance. These firms may see a supportive regulatory strategy as an opportunity to reduce costs, and assign less value to the potential benefits of high reporting quality and a cooperative long-term relationship with the authority. Our third hypothesis is depicted in Figure 1, and is stated as follows:
Hypothesis 3: The positive effect of communicating a supportive regulatory strategy on reporting quality is enhanced by long-term orientation and reversed by (becomes negative with) short-term orientation.

Figure 1. The moderating effects of firm size and short-term orientation

3. Data

Research setting: Financial intermediation

To test our hypotheses, we conduct a field experiment in the regulatory setting of financial intermediation in the Netherlands. In general, financial intermediation regulations aim to encourage financial intermediaries to provide sound and transparent advice to potential customers of financial products, such as life insurances or mortgage loans. Because financial intermediaries tend to have more knowledge than customers about financial matters, and charge commissions or fees for the products they sell, they have incentives to provide inaccurate or unsuitable advice to customers in order to increase private gains (a practice referred to as *mis-selling*; Inderst and
Ottaviani [2009]). Hence, the core objective of regulations in this setting is to prevent misselling behavior.

To conduct the field experiment, we entered into a research collaboration agreement with the Authority for the Financial Markets (AFM), which is the supervisor of the financial markets in the Netherlands (AFM [2017]). The AFM supervises approximately 7,000 financial intermediaries, who sell or provide advice on financial products such as mortgage loans, income insurances, and damage insurances.

**Market Monitor**

As part of its supervision strategy, the AFM employs a mandatory, online self-assessment instrument called *Market Monitor* (MM), which gathers detailed information from all financial intermediaries on a yearly basis. This information relates to licenses and registrations, economic activities, financial products the firm specializes in, number and types of customers, number and function of employees, sales per product, and other specifics related to accounting and financial matters (*Appendices B and C* contain the original and translated versions of the 2016 MM, which was used to gather data from 2015). The MM provides the AFM with a comprehensive overview of the accounting, financial, and managerial states of individual firms as well as the industry as a whole. At the same time, the MM helps firms to become aware of internal processes and performance, and to track their progress with regard to recommended practices.

The MM is a central element of the AFM’s regulatory strategy. Although the AFM also receives customer complaints and performs random audits, completing and submitting the MM is the only or main way in which most firms interact with the AFM (unless they contact the AFM for other purposes or specific questions). Because of the size of the market, the AFM needs to heavily rely on signals and risk-based supervision to effectively regulate financial intermediaries. Notably, the MM is the only means through which the AFM can gather detailed information of all market participants simultaneously and on a yearly basis. The key implication is that, for most firms, interactions with the AFM are restricted to communications directly related to the MM, and in this sense, the AFM’s official communications lie at the basis upon which financial intermediaries form perceptions about the AFM’s regulatory strategy.

The instructions to complete and submit the MM are sent out in three batches. The number of firms included in the first, second, and third batches of 2017 are 200, 3,544, and 3,498,
respectively (these numbers differ based on the AFM’s estimates about the time required by firms to complete the MM). The instruction letter was sent to all firms on Jan. 26, 2017. Subsequently, firms receive a (second) letter with their username, and one day later, they receive a (third) letter with the password to log into the system. The first batch received these letters on Jan. 31 and Feb. 1, the second batch on Feb. 7 and 8, and the third batch on Feb. 21 and 22. Firms in the first batch were requested to submit the completed MM between Feb. 6 and Mar. 12, firms in the second batch between Feb. 13 and Mar. 19, and firms in the last batch between Feb. 27 and Apr. 2. Upon receiving the first (instruction) letter, firms can – and are advised to – prepare all the information requested by the AFM. One week before the deadline, firms that have not completed the MM receive a reminder letter. Immediately after the deadline, firms that still have not completed the MM receive an additional reminder letter. The translated and original reminder letters are included in Appendix C and E, respectively. Firms’ responses to the MM are expected to be accessible to the research team before the summer of 2017, after they have been anonymized.

**Experimental manipulation**

We were granted permission to manipulate the content of the three letters that the AFM uses to communicate the instructions, username, and password required to complete and submit the MM. Our objective is to introduce a manipulation that is consistent throughout these three letters, and assess its effect on reporting quality. Firms are randomly assigned to one of the following three experimental manipulations:

1. Letters that do not reflect a supportive regulatory strategy (referred to as low support letters)
2. Letters that reflect a supportive regulatory strategy, and thus signal a cooperative attitude towards the firm (high support letters)
3. Letters that are similar to the letter the authority sent out in the prior year (which the AFM requested to include). The responses of this group of firms will not be part of our analyses.

We identify three elements that a supportive regulatory strategy has, and which can be signaled via official communications. These are referred to as tone (Wenzel [2006]), purpose (Greenberg [1994]; Alford and Speed [2006]), and accessibility (Alford and Speed [2006]). As
explained next, these elements are expected to affect firms’ perceptions about support in the same direction.

*Tone.* High support letters use a supportive tone, while low support letters use a more distant, formal and authoritative tone. Wenzel [2006] showed that an authority’s tone can influence agents’ perceptions of interpersonal fairness (which relates to whether firms feel that they are treated with fairness and respect). As opposed to low support letters, high support letters are expected to enhance firms’ perceptions of interpersonal fairness, and motivations to reciprocate the authority’s cooperative attitude and comply with its requests (Smith and Stalans [1991], Smith [1992]). Hence, support in this context is expected to have a positive impact on reporting quality.

*Purpose.* In the high support letters, the authority provides information about the goals of the MM, describes how these goals are aligned with the goals of the firm, and explains why it needs high quality information from the firm. In the low support letters, the authority neither provides information about the goals of the MM nor provides explicit reasons of why high quality information is valuable. These explanations are crucial, as they educate firms and allow them to “make sense” of what the AFM is requesting from them, and can therefore increase firms’ cooperativeness and willingness to comply. Communicating the purpose of the MM may motivate firms to reciprocate the authority’s open attitude, voluntarily share information, and engage in an open dialogue (Smith and Stalans [1991], Smith [1992]). Providing information about purpose can also enhance perceptions of procedural fairness, which relate to how firms perceive that their perspective is being taken into account (Greenberg [1994]).

*Accessibility.* Accessibility is one of the key elements that can enhance firms’ positive experiences with the authority (Alford and Speed [2006]). We manipulate accessibility by altering the extent to which the AFM signals willingness to help firms complete the MM, but not the options firms have to contact and receive help from the AFM. The tone and formulations of the low support letters signal that there is low willingness to help firms complete the MM, while in the high support letters, the AFM signals a great willingness to help. We expect a signal of high (low) willingness to negatively (positively) influence firms’ perceived barrier to contact the authority. A perceived low barrier to contact the authority may open a dialogue with the authority, and increase information sharing. Accessibility can also positively influence
perceptions of informational and interpersonal fairness. Agents may feel they are treated with respect and provided with all information to contact the authority to successfully respond to the authority’s request, which through reciprocating behavior, is motivating them to increase reporting quality (Smith and Stalans [1991]).

Appendix A contains the translated version of the letters. Highlights are applied to show where the letters are manipulated. We conducted two pretests in different populations to assess the strength of the proposed manipulations in the instruction letter. First, we followed a similar approach as used in Wenzel [2006] by directly assessing group differences in perceptions based on different instruction letters that experimental participants receive. Specifically, we provided participants with similar invitation letters as will be used in the field experiment, but then related to a fictional setting of tax reporting. We solicited participation of 106 bachelor students in an accounting course at a Dutch university to whom we randomly assigned the letters. After reading the letter, students filled out a set of questions about perceptions related to motivational postures towards the authority, informational and interpersonal fairness, and the authority’s regulatory strategy. Results show that the high support letter was perceived as being more informationally fair, and more directed towards support than the low support letter. However, absolute differences between conditions were relatively small. Participants’ motivational postures towards the authority were not altered by the treatment. We discussed these results with the AFM and decided to strengthen the manipulation by increasing the differences between the high and low support letters (e.g., different tone and further adjusting information about goals and support).

The adjusted letters were pre-tested among 165 MSc accounting students most of whom are studying part-time and employed in a reporting function. This test is based on a corporate tax reporting setting. Half of the students indicated to have experience in filing corporate tax returns and existing perceptions about the tax authority, which may be important to estimate the strength of the manipulation (Choo et al. [2016]). Participants were randomly assigned to one of the three instruction letters: high support, low support, or last year’s letter (which we included to consult the AFM about their existing letter). After reading the letter, participants undertook a tax reporting task, and then answered a set of questions capturing motivational postures towards the authority, and perceptions of fairness, the authority, and its regulatory strategy. This allowed us

1 Details on the results of the pretests are available upon request.
to assess perceptions in response to the letter, and to have participants conduct an actual task (i.e.,
distraction) before gathering their perceptions. The results of this second pretest showed that the
high support letter was perceived to be more informationally and interpersonally fair, and aimed
at support, particularly by those students with corporate tax reporting experience. This finding
may well be explained by the fact that these students have a reference framework with which
they elicit the authority’s regulatory strategy. This finding suggests that changes in
communication may be effective when there were earlier interactions, which is applicable to most
firms in our sample.

Dependent and moderating variables

Across different regulatory settings, firms are expected to report information in an accurate
and complete manner (Peeters [2006], Ernst & Young [2012]), and these aspects can be directly
associated with reporting quality. Following this conceptualization, and based on the 2016 MM
data as well as on-site discussions with AFM representatives, we identified the following three
elements of reporting quality: comprehensiveness, accuracy, and voluntary reporting.

Comprehensiveness of reporting covers how comprehensive firms are in their reporting to
the authority. The AFM does not allow intermediaries to skip questions in the online MM (i.e.,
forced answering). However, intermediaries can still avoid responding to questions by
deliberately choosing those responses that avoid or limit follow up questions (e.g., questions that
apply to specific firm characteristics, products, and activities). Our measure of
comprehensiveness of reporting captures the total number of questions the firm responded to.
Although this does not allow us to observe whether or not individual firms deliberately avoided
follow up questions (e.g., by choosing options that avoid follow up questions), it does tell us
whether or not the average number of questions responded to is equal across the different
experimental groups because of the randomization of the treatment. Group differences will be
consistent with response avoiding behavior, and thus reflect differences in comprehensiveness of
reporting in response to the treatment condition. The 2016 MM data that we obtained access to

---

2 For example, firms with part-time employees are asked to respond to subsequent items associated with part-time
employees. The same applies to specific financial products, such as mortgages or life insurance. See for example
item 3.1 and 3.2 in appendix B and D.
show significant variation in the number of questions responded to across all intermediaries (mean = 59.6580, st. dev. = 7.7891, min = 30, max = 94).

Accuracy of reporting covers the precision of firms’ responses to the supervisor. Less precise estimates are more likely to involve greater judgment (in contrast to retrieval of exact information) which increases the likelihood of inaccuracies in the reported information. Various questions in the MM request the firm to provide numeric information, such as sales, and number of customers. For most of these questions, the AFM explicitly allows intermediaries to provide estimates instead of exact numbers.

We use nine questions in the MM to construct a measure of accuracy (among which are reported sales, provisions and commissions, see Table 1). We selected these questions because (i) they are relevant for many firms (in the 2016 MM firms indeed responded well to these questions), (ii) the reported information in these questions is proprietary and requires the intermediary to gather a larger amount of information to be able to provide more precise estimates, and (iii) the responses have multiple digits. We compute accuracy of reporting by one minus the number of zero digits at the end of the reported number divided by the total number of digits reported \(1 - \frac{\text{Number of zero digits at the end of the reported number}}{\text{Total number of digits reported}}\).

We compute the average rounding score across the answered items per intermediary, in order to compare the accuracy of reporting across treatments. For example, if the firm only answered five out of nine questions, we calculate accuracy by taking the average score across the five items. The possible values for this variable range from just above 0 (only the first digit is not equal to zero) to 1 (the firm did not report a zero at the end the number). Higher scores represent more accurate reporting (i.e., less zero digits at the end of the responses). In the 2016 MM data we observe significant variation in the accuracy of reporting (mean = 0.6764, st. dev. = 0.1407, min. = 0.2083, and max. = 1).

Voluntary reporting covers how firms respond to requests to report voluntary narrative information to the supervisor. We use responses to a non-mandatory item in the MM which asks intermediaries to voluntarily provide a description of the added value of their firm as compared to their competitors. Respondents can choose whether or not to provide this information, and if they do, they can use 250 words at maximum for the description. Accordingly, we capture the extent
to which intermediaries voluntarily provide narrative information with an indicator variable whether (1) or not (0) the firm responded to the question. In a robustness test, we use the number of words provided as an alternative measure of voluntary reporting. In the 2016 MM 2,930 (41.20%) intermediaries completed this non-compulsory question.

*Overall reporting quality* is a measure that combines the three different components of reporting quality indicated above. These three components are expected to represent three different dimensions of reporting quality, which need not go in line. Firms may, for instance, decide to spend time to be very comprehensive but less accurate to compensate for this increased effort. The existence of individual dimensions of reporting quality is supported by low correlations between comprehensiveness, accuracy and voluntary reporting in the 2016 MM data. To obtain an overall measure of reporting quality, we standardize comprehensiveness and accuracy, and add these to the standardized dummy measure of voluntary reporting. A higher score on the measure reflects better reporting quality. The 2016 data show that there is much variation in the measure (average = 0.000; st. dev. = 1.742; min. = -5.994 and max. = 6.787).

*Time horizon* is measured by the combination of two questions included in the MM regarding the firm’s expectations about its (1) future number of employees (increasing, decreasing, or the same) and (2) future sales (increasing, decreasing or the same). The intuition behind the measure is that an expected decrease (increase) in employees and/or sales indicates organizational decline (growth) and a shorter (longer) time horizon. If a firm believes sales will grow in the future, this indicates that future sales are relatively more important compared to current sales, and that the net present value (NPV) of the firm depends relatively more on the future. Similarly, if a firm expects to hire new employees, it is likely to expect increasing business, which makes the future relatively more important. Hiring new employees indicates that firms are willing to invest and this will only pay off in the future. This suggests a relatively higher importance of the long run as compared to the short run. We rate the responses to each of both questions from 1 to 3 where 1 indicates an expected decrease, 2 expected stability, and 3 an expected increase in sales or number of employees. We use the average score of the two items as a proxy for time horizon, with a higher score reflecting a longer horizon.

*Firm size* is captured by the number of employees of the firm. The 2016 data show a non-normal distribution, which is why we will use a log transformation to normalize the distribution.
Control variables

We control for (1) the number of different products types that intermediaries sell as well as specific types of products, (2) the experience of firms with completing the MM (which can be inferred from the completion of previous Market Monitors), (3) firms’ connections to branch organization(s), (4) firms’ franchise connections, and (5) the batch to which intermediaries were assigned.

Intermediaries can sell 13 different types of products for which they need a license from the AFM. As a greater product range makes more questions in the MM applicable, we control for the number of product types that intermediaries sell. The maximum value of this variable is 13. We also include dummy variables to control for whether (1) or not (0) firms sell the three most frequently sold product types, which may affect which questions to answer in the MM: damage insurance, mortgages, and income insurance. We further control for the potential effects of experience by including an indicator variable whether (1) or not (0) the firm reported in a MM before 2017. Experience with the questions in the MM and interacting with the supervisor can influence how intermediaries respond to the treatment. Prior experience with the MM may make firms better prepared for the information request. Also, if intermediaries have positive or negative past experiences with the AFM, these may influence their attitudes towards the authority and its requests, potentially reducing the impact of the treatment. On the other hand, our pre-tests signaled that experience can make the treatment effect stronger as the firm has a reference to compare the treatment to. We further control for whether the firm is member of one or more branch organizations as interactions with these connections may also affect intermediaries’ perceptions. Similarly, we control for whether or not the firm is a franchisee which may affect the firm’s information environment and connections. Finally, we include two dummy variables to control for the batch number the firm was part of as the MM is sent out in different batches.

Table 1 provides an overview of the constructs and their operationalization.

---

3 Other products for which intermediaries need an AFM license are: consumer credit, electronic money, capital products, health insurance policies, pension insurance policies, premium pension claims, participation rights in investment companies, investment objects, payment accounts, and savings accounts.
Table 1. Operationalization of variables

<table>
<thead>
<tr>
<th>Construct</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicated supportiveness of regulatory strategy</td>
<td>Experimental treatment: High support letters (1) vs. Low support letters (0) (see appendix A for translated letters).</td>
</tr>
<tr>
<td>Reporting quality</td>
<td>Sum of standardized measures for comprehensiveness of reporting, accuracy of reporting, and voluntary reporting.</td>
</tr>
<tr>
<td>Comprehensiveness of reporting</td>
<td>The number of questions responded to.</td>
</tr>
<tr>
<td>Accuracy of reporting</td>
<td>Precision of numerical information (average of scores on items 2.11, 4.11, 5.1, 5.2, 5.5, 5.10, 5.11, 6.1 and 6.2) measured as 1 minus the number of zero digits at the end of a number divided by the total number of digits in the number.</td>
</tr>
<tr>
<td>Voluntary reporting</td>
<td>Indicator whether (1) or not (0) voluntary narrative information is provided to item 1.18.</td>
</tr>
<tr>
<td>Firm size</td>
<td>Natural logarithm of the number of full-time employees (item 3.2).</td>
</tr>
<tr>
<td>Time horizon</td>
<td>Average of the expected change [increase = 3, no change = 2, decrease = 1] in the number of employees (item 3.5), and in firm sales (item 5.18).</td>
</tr>
<tr>
<td>Type of products</td>
<td>Three indicator variables: mortgage, damage insurance, income insurance (item 1.4)</td>
</tr>
<tr>
<td>Number of product types</td>
<td>Sum of the types of products provided (item 1.4)</td>
</tr>
<tr>
<td>Experience</td>
<td>Previous experience with the Market Monitor yes/no (item 1.2)</td>
</tr>
<tr>
<td>Member branch-organizations</td>
<td>Membership of a branch organization yes/no (item 1.13)</td>
</tr>
<tr>
<td>Franchisee</td>
<td>Franchisee of a franchise organization yes/no (item 1.11)</td>
</tr>
<tr>
<td>Batch</td>
<td>Indicator variable for the batch the firm was in (2 dummy variables: group 1 or 2)</td>
</tr>
</tbody>
</table>

Note: Item numbers refer to questions in the Market Monitor added in appendix B (translated) and D (original).
Random assignment

In order to be able to test our hypotheses, we used a stratified random sampling approach to assign the treatment to firms. Using the 2016 MM data, we first derived the score for each firm on the time horizon measure. This yields five different time horizon groups (average score on the two items 3.5 and 5.18: 1, 1.5, 2, 2.5, and 3). Within these groups we ordered firms by size, and then randomly assigned the treatments to the firms. This way we assured that within each time horizon level, treatments are randomly assigned to intermediaries with different sizes. Figure 2 depicts the random stratification procedure. In each of the treatment conditions, there are approximately 2,400 intermediaries.

Figure 2. Random stratification of intermediaries

We used the 2016 MM data to examine if the two stratification conditions (firm size and time horizon) as well as existing perceptions about the authorities information provision concerning the MM are associated with the treatment conditions. Associations between these variables and the treatment could affect the results of the experiment. Existing perceptions on the authorities information provision concerning the MM are captured by the responses to two statements in the MM, being the extent to which respondents believed that (1) they were timely informed about the MM, and (2) it was clear to them which information they needed to collect to complete it. We find no significant differences between the treatment groups for firm size (t-stat.
-0.3958), time horizon (t-stat. 0.0427), and these perceptions (t-stat. 0.6814 and 0.4403). This provides confidence that our random stratification procedure was successful.

4. Planned analysis and research design

Summary statistics

Table 2 shows the summary statistics of the dependent, moderating and control variables. Table 3 shows the Pearson correlation coefficients between the dependent and moderating variables.

Table 2. Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehensiveness of reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy of reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voluntary reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damage insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of products</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch organization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Franchise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows how the summary statistics (mean, standard deviation (SD), min (minimum value), and max (maximum value)) of the dependent, moderating and control variables will be reported once the data have been collected.
Table 3. Pearson correlation coefficients

<table>
<thead>
<tr>
<th></th>
<th>Reporting quality</th>
<th>Comprehensive-ness of reporting</th>
<th>Accuracy of reporting</th>
<th>Voluntary reporting</th>
<th>Firm size</th>
<th>Time horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting quality</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehensive-ness of reporting</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy of reporting</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voluntary reporting</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Time Horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3 will show the Pearson correlation coefficients between the variables of interest. ***, **, and * means significant at the 0.01, 0.05, and 0.10 level, respectively.
### Statistical tests

Given the expected distribution properties of the dependent and independent variables, we plan to conduct *t*-tests and chi-square tests to assess the hypothesized group differences. Additionally, we will examine whether the treatment conditions predict the overall measure of reporting quality and its constituting elements (comprehensiveness, accuracy of reporting, and voluntary reporting), while controlling for the firm characteristics summarized in Table 1. We use OLS regressions to estimate effects on reporting quality (as an overall indicator), and on comprehensiveness and accuracy of reporting (as individual indicators). We use a logit regression to estimate effects on voluntary reporting. In these tests we include the dummy variable *High support letter* (indicating whether the intermediary received an instruction letter that reflects a high support-based regulatory strategy (1) or not (0))) to assess the effects of the experimental treatment. The following regression model will be used to test the first hypothesis.

\[
(1) \text{Reporting quality} = \alpha + \beta_1 * \text{High support-based letter} + \beta_2 * \text{Time horizon} + \beta_3 * \text{Firm size} + \sum_{i=4..12} \beta_i \text{Control}_i + \epsilon
\]

Based on Hypothesis 1, we expect $\beta_1$ to be positive and significant. To test hypothesis 2, we will include the interaction between firm size and the treatment variable in regression model (1). We expect the interaction term to be negative and significant.

Following Hypothesis 3, we expect the treatment–time horizon interaction to be positive for firms that are long-term oriented, and negative for firms that are short-term oriented. We test Hypothesis 3 in two ways. First, we incorporate a dummy for short-term oriented firms (taking the value 1 if the firm’s time horizon score lies below the midpoint of the 5-point scale, and 0 otherwise), and a dummy for long-term oriented firms (taking the value 1 if the score lies above the midpoint of the scale, and 0 otherwise).\(^4\) The effect of scores lying at the midpoint of the scale is therefore captured by the regression intercept (medium-term oriented firms are treated as the reference category). The signs of the short-term and long-term dummies are expected to be negative and positive, respectively.

Second, we generate a time-horizon (categorical) variable that has three levels: short-term oriented (taking the value 1 if the score lies below the midpoint), medium-term (taking the value 1 if the score lies above the midpoint), and long-term oriented (taking the value 1 if the score lies between the midpoints). If in the 2017 MM the number of observations with a score of 1 or 3 on the time horizon scale will be sufficient, we will also consider an additional robustness test in which we contrast these two groups with the middle category.

\(^4\) Applying this to the 2016 MM data results in 17.8% of firms being classified as short-term oriented and 42.4% as long-term oriented, while 39.8% of firms scores the mid-point of the scale. If in the 2017 MM the number of observations with a score of 1 or 3 on the time horizon scale will be sufficient, we will also consider an additional robustness test in which we contrast these two groups with the middle category.
value 2 if the score is equal to the midpoint), and long-term oriented (taking the value 3 if the score lies above the midpoint). In a separate regression, we incorporate the interaction of the treatment and this categorical variable. As we expect an antagonizing interaction effect (negative for the first level, and positive for the third level), we do not formulate an empirical expectation about the sign or the significance of the interaction term. Instead, our approach consists in estimating the predicted marginal effects of the four relevant combinations (i.e., high vs. low support x long vs. short time horizon), and testing whether these marginal effects differ in line with our expectations. In particular, we assess whether the marginal effect of high support and long-term orientation is larger than for the other combinations, and whether the marginal effect of high support and short-term orientation is smaller than for the other combinations. We expect the difference between these estimates to be significant.

5. Results

Group differences in reporting quality

We assess whether firms which receive the high support letter (i.e., communication that reflects a supportive regulatory strategy) respond with higher reporting quality than those receiving the low support letter. We first use t-tests to examine the hypothesized group differences in the overall measure of reporting quality, comprehensiveness of reporting, and the accuracy of reporting. Table 4 shows the average scores on the dependent variables across the two different experimental groups, the differences in average scores between groups, and the significance of these differences.

Table 4. T-test results hypothesis 1

<table>
<thead>
<tr>
<th></th>
<th>High support letter</th>
<th>Low support letter</th>
<th>Difference in mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehensiveness of reporting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy of reporting</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The left panel of table 4 will show the average scores on the dependent variables across the experimental conditions. The right panel of table 4 will show the differences in means and the corresponding significance.
To test for the hypothesized differences in voluntary reporting, we use chi-square difference tests. Table 5 shows the corresponding contingency table. Results of the tests will be narratively explained.

**Table 5. Contingency table hypothesis 1**

<table>
<thead>
<tr>
<th></th>
<th>High support letter</th>
<th>Low support letter</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntarily reporting [1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No voluntary reporting [0]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cells include the frequency of voluntary reporting across experimental conditions.

**Regression analyses**

We run regression analyses to assess the effects of the experimental treatment on reporting quality, while controlling for other possible effects. We also use these analyses to assess whether time horizon and firm size moderate the influence of the treatment variable. Table 6 shows the results.

Jointly identifying treatment effects for a set of outcomes could lead to biased conclusions (List, Shaikh and Xu [2016]). This multiple hypotheses testing problem refers to the risk of incorrectly rejecting a true null hypothesis. List et al. [2016] developed an asymptotically testing approach that controls for the family-wise error rate to deal with this problem. In addition to the ‘regular’ $p$-values, we use this approach [theorem 3.1 and remark 3.1] of List et al. [2016] to report the corrected $p$-values in the regression analyses.

**Marginal effects**

To test Hypothesis 3, we estimate the predicted marginal effects of the combinations resulting from the interaction between the treatment and the categorical (three-level) time horizon measure (with scores below, equal to, and above the midpoint of the time horizon scale). Table 7 shows the predicted marginal effects, $z$-values, and significance. Results and difference tests will be narratively explained.
Table 6. Regression results

<table>
<thead>
<tr>
<th></th>
<th>Expected signs</th>
<th>Comprehensiveness of reporting</th>
<th>Accuracy of reporting</th>
<th>Voluntary reporting</th>
<th>Reporting quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High support letter (HS)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm size (FS)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time horizon (TH)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS x FS</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS1 x STH</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS1 x LTH</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damage insurance</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income insurance</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of products</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch organization</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Franchise</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch 1</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch 2</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6 will show the regression coefficients of the independent, moderating, and control variables. To test the predicted effects in Hypothesis 3, the variable time horizon is replaced by the two indicator variables (i.e., to indicate a short-time horizon (score is lower than 1.5) or to indicate a long time horizon (score is higher than 2.5)) as described in section 4. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level based on ‘regular’ p-values, respectively. +++, ++, and + indicate significance at the 0.01, 0.05, and 0.10 level for the corrected p-values.
Table 7. Marginal effects of short and long time horizons

<table>
<thead>
<tr>
<th>Marginal effect</th>
<th>Z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS x LTH</td>
<td></td>
</tr>
<tr>
<td>HS x STH</td>
<td></td>
</tr>
<tr>
<td>LS x LTH</td>
<td></td>
</tr>
<tr>
<td>LS x STH</td>
<td></td>
</tr>
</tbody>
</table>

HS and LS denote high and low support-based letters, respectively. LTH and STH denote long and short time horizons (above and below the midpoint of the time horizon scale), respectively. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 level based on p-values.

**Statistical power of analyses**

Our main analyses focus on detecting differences in reporting quality across two treatment conditions. The statistical power of such analyses is determined by a combination of different factors: (1) the alpha (the risk of rejecting a true null hypothesis, default 0.05), (2) the sample size and number of groups, (3) power (type II error which is “the degree to which the null hypothesis is believed to be false” (Cohen [1992, p.156])), and (4) effect size (Cohen [1992]). A combination of three factors is sufficient to determine the value of the fourth. In order to estimate the statistical power of our analyses, we compare the number of subjects in our sample with the number of subjects needed to detect small effects. According to Cohen [1992], a sample size of at least 464 or 322 subjects is needed to detect small effects with an alpha of 0.01 and 0.05, respectively. As our sample size is much larger, we will be able to detect even small differences between groups.

**Additional analyses**

To provide a more comprehensive analysis of the treatment on reporting quality, we will perform several additional analyses.

A. Change relative to the previous Market Monitor. In order to increase the validity of our findings, we examine how the three dimensions of reporting quality changed from the 2016 MM to the 2017 MM and whether these changes relate to the treatment intermediaries received. We use firm identifiers to match data over years to make such analyses possible. We run the same tests as in the main section, but use the changes in dependent variables instead of levels.
B. Testing perceptions of informational fairness. We intend to use the two statements of question 8.1 of the MM to create a measure of informational fairness. Translated, the statements read as follows: (1) “I was informed about the Market Monitor in a timely manner”, and (2) “It was sufficiently clear which information my organization had to gather in order to provide answers to the questions in the Market Monitor”. These items are indicative of perceived informational fairness as they measure whether firms believe that they were informed sufficiently and timely. Firms respond to these items on a five-point Likert scale, which ranges from “completely disagree” to “completely agree”. We test whether the individual and average scores on the items differ across experimental conditions. The 2016 MM data showed that there were no differences between experimental groups before the experiment was conducted (t-stat. 0.6814 and 0.4403; see also Section 3). Furthermore, we examine the correlations between the average scores on these items and the measures of reporting quality. This analysis provides us more insight in how fairness perceptions are associated with reporting quality.

C. Benford’s law. Benford’s law describes the phenomenon that the first four digits of random reported numbers follow specific frequency distributions (Benford [1938], Hill [1995]). Frequency distributions that deviate from Benford’s law may indicate that the reported numbers are fraudulent or inaccurate (examples of applications of Benford’s law are in Carslaw [1998], Durtschi, Hilison and Pacini [2004], Nigrini [1996]). For example, if reporters round numbers, we expect to see increases in the frequency of zeros relative to the base prediction of Benford’s law. As an illustration, using reported sales in the 2016 MM Figure 3 shows for the first four digits the frequency distributions relative to Benford’s law. The first digit (Panel A) follows Benford’s law well (p =0.999), while the distribution of the second digit (Panel B) shows a more frequent observation of zeros and fives although the difference is not significant (p =0.269). For the third and fourth digits, however, differences are highly significant (both p <0.001). As panel C and D show, observed zeros are much more frequent than Benford’s frequency distributions would predict, which is consistent with common use of rounding by intermediaries at these digits.
We estimate and test differences in accuracy between groups, by comparing their frequency distributions of digits for reported sales (question 6.1) with Benford’s frequency distributions using Chi-square difference tests. Hypothesis 1 predicts that the low support group will report with lower accuracy than the high support group. Accordingly, we expect that the low support group will show larger deviations from Benford’s frequency distributions than the high support group, and that deviations may occur at an earlier digit (e.g., at the second instead of third digit).

D. Timeliness in reporting. The AFM requires intermediaries to report the requested information within a particular timeframe. To be relevant for effective supervision, reported information also needs to be timely. Therefore, in this additional test, we examine how our
treatment influences timeliness in reporting; i.e., how long before the deadline intermediaries report to the AFM. On the one hand, high support letters may enhance intermediaries’ normative motivations to report early. On the other hand, these letters may motivate them to be more accurate which increases the time needed to report the information. It is to be expected, however, that when a supportive regulatory strategy enhances normative motivations, firms will report in time. We use an indicator variable to indicate whether intermediaries reported their information in time (0) or not (1). As we are yet unsure how large the number of late responses will be, we may also use a second indicator that includes in the ‘late’ group responses received close to the deadline (i.e., last day before the deadline). We run the same tests as in the main analysis, with timeliness as the dependent variable and running logistic regressions.

E. Effort in responding. More comprehensive and accurate reporting is expected to be more time-consuming. However, efficient preparation and well organized (and accurate) data systems may decrease the time spend on the MM. We examine how our treatment influences the self-reported time for preparing and completing the MM (number of hours) and the time they actually spend online to complete the MM as recorded by the AFM. We run the same tests as in the main analysis, only replacing the dependent variable by these self-reported and recorded time variables.

6. Conclusion

7. References

AFM. Strategische doelstellingen. Consulted on April 27, 2017 Available at: https://www.afm.nl/nl-nl/over-afm/werkzaamheden/strategische-doelstellingen.


