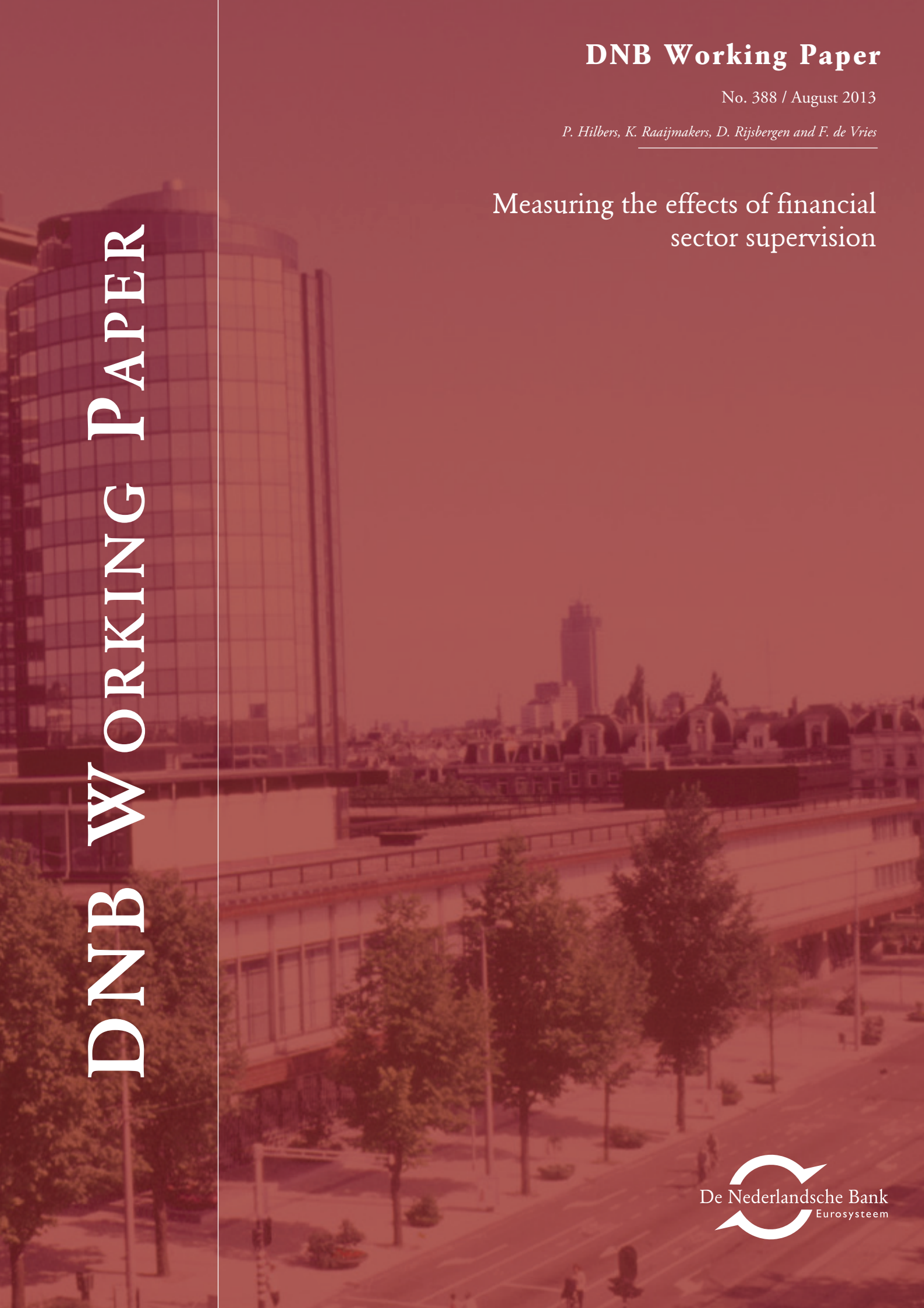
DNB Working Paper

DNB Working Paper

No. 388 / August 2013

P. Hilbers, K. Raaijmakers, D. Rijsbergen and F. de Vries

Measuring the effects of financial sector supervision



De Nederlandsche Bank
Eurosysteem

Measuring the effects of financial sector supervision

Paul Hilbers, Karina Raaijmakers, David Rijsbergen and Femke de Vries *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 388

August 2013

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Measuring the effects of financial sector supervision

Paul Hilbers[*], *Karina Raaijmakers*[†], *David Rijsbergen*[‡] *and Femke de Vries*[§]

August 19, 2012

Abstract

Financial supervisors are increasingly expected to be able to demonstrate the effectiveness of their actions. In practice, however, this proves challenging as it is difficult to prove the causality between supervisory actions and observed effects. In this paper we describe four lessons that help financial supervisors measure the effects of their actions. We also provide suggestions for the development of specific performance indicators to measure the effectiveness of financial supervision.

Keywords: financial supervision, effectiveness, performance indicators

[*] Nyenrode Business University and De Nederlandsche Bank (DNB).
[†] Clear Conduct, a specialized consulting firm that takes on strategic issues regarding regulatory practice, risk management and conduct.
[‡] De Nederlandsche Bank (DNB), Supervisory Policy Division, Strategy Department, PO Box 98, 1000 AB Amsterdam, the Netherlands, Tel: 0031-20-5242320, E-Mail: d.r.rijsbergen@dnb.nl.
[§] De Nederlandsche Bank (DNB), Supervision Expert Centers Division.
The authors would like to thank Malcolm Sparrow (Harvard Kennedy School), Ivo Bruijn (DNB), Paul Cavelaars (DNB) and Edmond Vrancken (DNB) for valuable comments and suggestions. This working paper builds on an earlier Dutch article (see Hilbers et al, 2012). The views expressed in this article are our own and do not necessarily reflect those of DNB.

1. Introduction

The global financial crisis has raised questions concerning the performance of financial supervisors. Did they undertake all the necessary steps to prevent the crisis or could financial supervisors have done more? As their work is being scrutinised, financial supervisors need to become better at demonstrating that their efforts and actions lead to results. Measuring the effects of supervision must consequently become an integral part of the supervisory process. Not only will this promote external accountability, but also – and equally important – it will show whether supervisory actions have contributed to the desired results. This information is important in order to improve the supervisory process and to ensure that the correct priorities are set.

Not surprisingly, performance measurement is becoming increasingly important for financial supervisors such as the Netherlands Authority for the Financial Markets ('AFM') and the Dutch Central Bank ('DNB'). In 2009, a Dutch parliamentary committee examining the causes of the financial crisis called on both the AFM and DNB to display greater openness and transparency to the outside world on their performance of supervisory activities.[1] For reasons of public interest, therefore, financial supervisors must not only be sufficiently independent, but also more accountable for how they perform their tasks (DNB, 2010). That demands a carefully designed accountability regime, not only in reports compiled by the supervisors themselves, but also through external assessments such as the Financial Sector Assessment Program (FSAP) published by the International Monetary Fund (IMF) and the Netherlands Court of Audit's Supervision of Markets Program (Hilbers, 2011).[2]

In financial supervisory practice, however, measuring effectiveness is not straightforward. This is mainly due to the difficulty of proving causality. A change in a financial institution's risk profile, for instance, might have little or nothing to do with any supervisory intervention, as it could simply be the result of a change in economic conditions or some other exogenous factor. Moreover, financial supervisors typically face the legal question of whether they are allowed to report on their interventions. In practice, most financial supervisors face a statutory duty of confidentially and are thus not allowed to report or publish all their actions. As a result of these challenges, the development of performance measurement in financial supervision has long been in its early stages.

---

[1] The Dutch parliamentary committee was chaired by Jan de Wit and published a report on the causes of the financial crisis with the title: 'Lost Credit'. Parliamentary Papers II 2009/10, 31 980, Nos. 3-4, p. 22.
[2] See for instance, Algemene Rekenkamer, *Toezicht van DNB op de stabiliteit van banken*, The Hague, 2011.

This paper aims to fill this gap by analyzing performance measurement in the context of financial supervision. For that, we not only examine the challenges, but also investigate the opportunities for measuring the effects of financial supervision. As such, the study contributes to the development of performance indicators for financial supervision.

The remainder of this paper is structured as follows. Sections 2 describes the different levels at which the effects of supervision can be measured and explores the possible research designs that supervisors can use for performance measurement. Section 3 explains the challenges to measuring the effectiveness in the context of financial supervision. Next, Section 4 examines the opportunities by describing four lessons that supervisors can use to measure their performance.. Section 5 presents different types of performance indicators for financial supervision. Our conclusions are set out in the final section.

2. Measuring effects: goals, levels and methodology

Effectiveness in the context of supervision can be defined as the degree to which supervisory practice contributes to the realization of supervisory objectives. These objectives can be defined in two categories, namely societal objectives and compliance objectives. Societal objectives express what a supervisor would like to contribute to society as a whole. Societal objectives are typically defined at a policy level where the legislator defines rules and standards in order to promote these objectives. Consequently, it is up to the supervisor to make sure that supervised institutions actually comply with these rules and standards (i.e. compliance objectives).

The primary (societal) objective for prudential supervisors such as the De Nederlandsche Bank (DNB) is to safeguard a stable financial system with sound institutions. In order to achieve this goal, DNB not only checks whether institutions meet the supervisory requirements in terms of solvency and liquidity, but also evaluates the business model, strategy and corporate governance within these institutions (DNB, 2009). DNB, for instance, monitors the degree to which financial institutions comply with remuneration policy standards as this contributes, through an enhanced focus on risk-mitigation, to the soundness of those institutions and thus to the societal objective of safeguarding financial stability. As a result, supervisors do not focus on compliance per se, but focus on compliance in order to ensure that supervised institutions contribute to the societal objectives.

2.1 Goals of supervisory performance measurement

The central objective of performance measurement is to provide reliable and valid information on performance (Behn, 2003). This information can be used for different purposes. Hence, financial supervisors may have different goals in mind when measuring performance. These goals can be related to being accountable, determining whether a specific market problem has been successfully mitigated, finding out which interventions are effective in which circumstances, or creating organizational performance incentives.

First, performance measurement at the strategic level enables a supervisor to show his added value to society and other relevant stakeholders: for instance, what are the supervisor's strategic objectives and to which degree are these objectives actually attained? This implies that performance measurement results in an overview of aggregated measures, such as the level of confidence in financial markets, the transparency of financial markets and the degree of financial stability.

Second, performance measurement enables a financial supervisor to determine whether a specific market problem actually becomes smaller due to its interventions. We call this performance measurement at a tactical level. In a problem-based approach, a supervisor tries to identify the most important market problems (e.g. an organizational culture that stimulates excessive risk-taking) and focuses on designing a mix of interventions that will effectively mitigate the problem. Measuring the effects of these interventions is necessary in order to determine whether the problem has been sufficiently mitigated to relocate supervisory resources to another supervisory issue. Performance measurement at a tactical level, consequently, supports supervisors when making decisions about the prioritization of market problems and allocation of resources.

Third, performance management may also be useful in order to enhance knowledge about which interventions are effective under which circumstances. This is generally considered as performance measurement at an operational level. For example, a performance measurement may indicate that a certain supervisory strategy aimed at increasing investor awareness appears to stimulate investment fund managers to be more transparent on the costs of their funds. Consequently, it might be effective for the supervisor to apply that similar strategy in other fields in an attempt to increase transparency there as well. Lastly, supervisors can also use performance measurement to create organizational performance incentives to motivate individual supervisors. By setting targets like reducing the net inflow of capital in a specific risky and non-transparent investment fund or making sure that there will be a solution to lower the

risk profile of a specific financial institution, the performance of individual supervisors is better assessable.

2.2 Optimal methodology for performance measurement

In the context of financial supervision, the key question of performance measurement is whether observed economic or financial outcomes can be attributed to a supervisor's efforts. For that, financial supervisors need to be able to prove a causal relationship between their activities (cause) and the outcome (effect). Proving the internal validity of a causal relation requires a suitable research design. Research designs range from descriptive case studies to experimental designs (Shadish, Cook and Campbell, 2002). In theory, experimental research designs tend to provide the most valid and reliable results. This means that this type of design provides the highest level of certainty with which changes in the outcome (effect) can be attributed to the independent variable (cause). A well-known example of an experimental research design is the randomized controlled trial (RCT), which is widely used in the field of medicine for clinical trials.[3]

In general, an experimental research design (such as a RCT) is characterized by three key components. First, a pre-post test routine is required in order to collect the same data before and after the intervention took place. Second, the design consists of both an experimental group (i.e. a group that experiences the intervention) and a control group in order to control for the possibility that factors not related to the intervention are responsible for the difference between the pre-test and post-test results.[4] And finally, both the experimental group and the control group are randomly assigned. This ensures that both groups are statistically similar (i.e. subjects have an equal chance of being assigned to either group) and reduces the chance of a selection bias. In addition to the traditional experimental design, there are also types of quasi-experimental research designs. The latter generally differs from the traditional design with respect to the selection of the experimental and control groups. In a quasi-experimental design, both groups are typically not assigned randomly.

To evaluate the methodological quality of research studies, Sherman et al. (1998) developed a five point scale named the Maryland Scientific Methods Scale (MSMS). Since inception, the scale is widely used to categorize research and evaluations according to their methodological strengths and weaknesses. The MSMS ranks studies from 1 (weakest) to 5 (strongest) on overall internal validity. A study which lacks a pre-post test design, a control group and a randomized

[3] See for instance: Stephens et al. (1996).
[4] In this context, it is also important that both the treatment group and the control group are of adequate size.

approach is typically classified at the lowest level (1) of the MSMS, as such a study is only able to measure the correlation between a supervisory intervention (cause) and an economic or financial outcome (effect). A measure of causality, on the other hand, requires a study with a higher classification. The highest score (5) on the MSMS is reserved for the traditional experimental research design. In financial supervisory practice, however, research studies that classify high on the Maryland Scientific Methods Scale are rare as it is challenging to form (matching) control groups and randomly assign supervised institutions. We will describe these challenges in more detail in the next section.

3. Measuring effects: the challenges

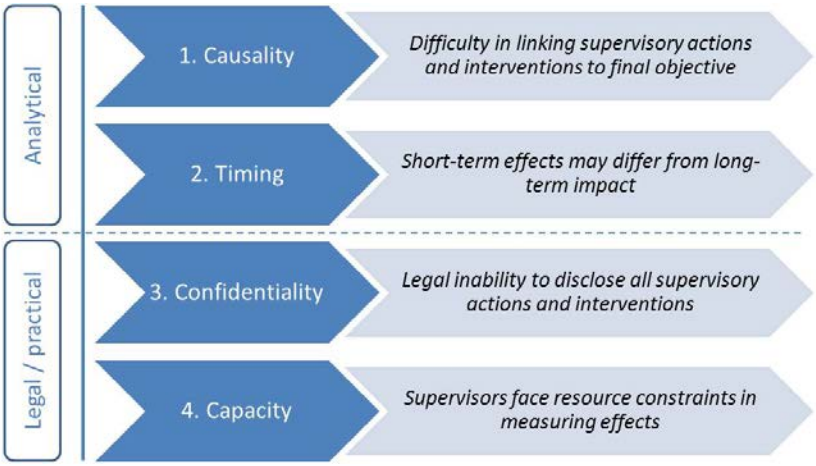
In Section 2 we described the conditions to measure the effectiveness of supervisory actions and interventions. In this paragraph we discuss the four most important challenges in more detail (see figure 1). The first is methodological and relates to the question whether it is possible to prove causality. One way of finding a plausible answer to the causality question is to consider the counterfactual.[5] In other words, to consider what would have happened if there had been *no* supervision, or if the supervisor had *not* intervened (Willemsen, 2009)? As described in Section 2, the ideal approach in these circumstances is to use a traditional experimental research design (Wouters, Derriks and Van der Loop, 2009). However, setting up an experimental research design is challenging in the financial supervisory practice. For one, it is not always possible to effectively separate the control group from the experimental group. Take, for instance, a case in which the supervisor wants to improve the quality level of financial services by making intermediaries accountable for sound and controlled operations. Companies or individuals in the control group are likely to hear of the supervisor's intervention through the media, the internet or their network and are thus hard to isolate in practice. Selecting supervised institutions for inclusion in a control group presents a further problem: the selection process can give rise to questions because it may result in a conscious choice to ignore certain undesirable behaviour demonstrated by institutions within the control group. This means there is no level playing field, which may violate the principle of equality.[6]

---

[5] Gezondheidsraad, *Op weg naar evidence based toezicht. Het onderzoek naar effecten van toezicht door de Inspectie voor de Gezondheidszorg*, The Hague, 2011.

[6] The medical sector, for example, has proved able to compile satisfactory control groups for measuring effects as it uses only patients who have consented to being included in the control group. In many other sectors, however, 'informed consent' is not readily available. See Sparrow (2008).

Figure 1 – Challenges in measuring effects of financial supervision



Without an experimental research design, it is difficult to rule out the influence of exogenous events on the observed economic outcomes, which makes it challenging for financial supervisors to answer the causal question convincingly. This is even more so when performance measurement is performed at the strategic level, where a wide range of external effects (e.g. global economic developments) may influence the desired supervisory outcome. The complexity of the financial sector also contributes to the challenge of performance measurement, as the population of supervised institutions is diverse, while innovation and competition mean that financial markets are also continually changing. As a result, financial supervisors tend to focus on the level of compliance when they are trying to show a causal link between their efforts and a change in market structure or behaviour of market participants. Subsequently, the degree to which societal objectives are attained (i.e. the strategic level) is often neglected. The focus on identifying compliance is confirmed by Winter and De Ridder (2010), who examine 69 international studies of effectiveness. They find that more than half of these studies focus on research into compliance with standards, without consideration of the societal effects.

The second challenge in measuring the effects of supervision relates to the different short and long-term effects that supervisory interventions may have, and which in some cases may even be diametrically opposed to each other. This can best be illustrated by an example. A financial supervisor may instruct a financial institution to sell off previously acquired complex investments because it lacks the appropriate level of risk management. In the short term, such a supervisory intervention may result in additional (transaction) costs and consequently may have an adverse impact on the institution's financial position. In the longer term, however, the financial position may strengthen as the institution is no longer exposed to (investment) risks that it cannot adequately control. If the financial position is taken as the indicator, the initial

effect of the supervisory intervention may be negative. In due course, however, providing the institution can manage its risks properly, the effect may well be positive.[7]

A third challenge relates to how effects of supervision can be demonstrated to public stakeholders. In many cases, the duty of confidentiality that financial supervisors are required to observe makes it difficult to communicate on interventions aimed at individual institutions. DNB and the AFM, for example, are only allowed to publish information on individual institutions to the extent necessary for the performance of their statutory tasks.[8] In general, these tasks do not include measuring the effectiveness of supervisory interventions. In some cases, therefore, supervisors are forced to report *generically* on effects achieved, whereas they would prefer to report *specifically* so as to demonstrate what has been achieved. Depending on a supervisor's objectives, reporting on effects may also conceivably be undesirable. This may be the case in the event of so-called 'near misses'. If, say, matters almost went wrong, but the supervisor managed to avert a disaster, the intervention can be assumed to have been effective. Let us consider the perfect example of a near miss: a situation in which intervention by the supervisor averts a financial institution's impending insolvency. This creates an interesting paradox. By intervening, the supervisor will reinforce the solidity of the financial institution, as well as boost public confidence in the specific institution and probably also in the financial sector as a whole. Communicating this to the outside world, however, could counter that effect if it causes the public to lose confidence in the sector to which the institution that experienced the near miss belongs. The public may believe that if one company is facing difficulties, so too, might others. In that way, communicating about 'near misses' could have an adverse impact on public confidence in the financial sector, thereby countering the very objective of the intervention.

The final challenge relates to the fact that measuring effects demands allocation of time and capacity (Ottow, 2011). If measurements are to be reliable, a proper research method needs to be devised and the relevant data need to be collected. Supervisors also have to ensure the availability of sufficient in-house or external expertise (via contacts in the academic world, for example). Since resources are scarce, supervisors constantly have to set priorities; investing resources in measuring the effects of supervisory interventions implies that certain other activities cannot be performed. Therefore, the task of measuring effects needs to be explicitly included in supervisors' priorities.

[7] The latter claim, however, is not certain as supervisory intervention may serve solely to reduce the risk of a deteriorating financial position.
[8] Section 1:89 Financial Supervision Act (*Wet op het financieel toezicht*).

4. Measuring effects: what is possible?

Due to the discussed limits of measurement, financial supervisors in many cases need to accept some uncertainty when examining the relationship between their activities and economic outcomes. That, however, does not mean to say that it is impossible for supervisors to measure their performance. Measurement can be used to acquire more insight and develop more comfort that the supervisory activities are having a significant impact on the observed economic outcomes (Mayne, 1999). In other words, it can help to show that the supervisor is making a difference. For that, we have set out four lessons in this section that can help establish a practical way of measuring the effects of supervision. These lessons can also be transposed to a supervisory policy environment (see box 2 at the end of this section).

4.1 Define a specific objective and measure at a micro level

As mentioned in Section 2.1, performance measurement is related to the supervisory objectives. These objectives can be broadly categorized as societal and compliance objectives. In practice, however, these objectives are not always defined SMART (i.e. specific, measurable, attainable, relevant and time-bound). To illustrate, consider the typical objectives for financial supervisors such as 'safeguard financial stability' or 'enhance financial market transparency'. Translating these objectives into SMART goals will help supervisors select the most appropriate research design and indicators with which to measure their effectiveness (Winter, 2012). The more clearly a problem is defined, the more precisely the objective (i.e. the question being asked) can be formulated, and the more effectively suitable indicators can be identified.[9]

In this respect, Sparrow (2008) argues that supervisors can increase the plausibility of a causal relationship by reducing the level of abstraction at which effects are measured. He states that defining societal aims such as 'financial stability' or 'solid financial institutions' in concrete terms in the form of relevant problems or risks at a micro or project level makes a causal relationship more plausible (Sparrow, 2008). For instance, it is easier for financial supervisors to demonstrate that additional on-site visits have strengthened the governance structure of a specific institution than to show that those actions improved the financial stability as a whole. Hence, supervisors can demonstrate their effectiveness through a series of successful results at that micro or project level, or as Sparrow (2008) states it, through a "compelling account of harms controlled." Examples of financial supervisors that apply this micro approach in practice

[9] In addition, it also helps to define a clear ambition level or target for most indicators. For instance, to evaluate the performance of a financial supervisor, its managers have to compare that performance with some standard. Such a standard can come from past performance, from the performance of peers, from a professional or industry standard such as the IMF FSAP or from political expectations (Behn, 2003).

are the AFM and DNB. In addition to their regular supervisory work, both supervisors annually select a number of supervisory themes, which will be the focus areas of supervision within that specific year. In 2013, for instance, DNB focuses on themes such as the sustainability of business models and the quality of risk management within financial institutions (DNB, 2013). For each theme, a brief outline of the problem is provided, together with details of the most significant developments. This way, measuring effects essentially becomes a collection of results at a micro or project level. Box 1 gives an example of an indicator that can be used in such measurements. Although such a thematic approach does not directly demonstrate a causal relationship between supervisory interventions and the societal objective, a long series of micro results will show whether the supervisor's actions are result-oriented, analytical and effective. This in turn will help improve the supervisory process as well as the external accountability.

Box 1 - Indicators for measuring effects at a micro or project level: an example

An example of a useful indicator for measuring effects at a micro or project level is the amount of investments in specific complex financial product. A supervisor may, for example, consider a specific product to involve such a high degree of risk that the product is not suitable for private investors. The aim of the supervisor's intervention will then be to prevent private individuals from investing in this product. An obvious indicator for measuring the effects in this situation would be the inflow of capital from private investors. If these amounts decrease, the supervisory intervention could be claimed to be effective. However, it remains important to establish whether there is a causal relationship. For instance, the decline in capital inflow from private investors may also simply be the result of deteriorated market conditions, which resulted in a general decrease in private investment. It is important, therefore, to determine by how much the capital inflow would have decreased if the supervisor had not intervened (the counterfactual). One way to approximate this is by determining the change in the inflow of capital to products with a similar risk profile.

4.2 Develop a plausible theory of change and contribution story

Supervisors can further increase the plausibility of the causal relationship between their activities and the observed outcomes by developing a logical theory of change. A theory of change explains how a supervisory policy or intervention is expected to result in the desired outcome (Weiss, 1997). The logical reasoning behind such a theory may be well-established on past supervisory experience and can help the supervisor build a credible 'contribution story'. Supervisors can thus provide reasonable evidence about their contribution to an observed outcome by verifying the theory of change on which the supervisory intervention or policy is based. This approach is generally known as 'contribution analysis' and provides an alternative

for the assessment of cause and effect when experimental research designs are not practical or not feasible (Mayne, 2008).

Key tools in building a credible contribution story are a results chain or logic chart. A results chain or logic chart displays how the supervisory intervention or policy is supposed to work. That is, how the various outputs of an intervention or policy are believed to produce a number of results that will lead to the intended outcome (Mayne, 1999). Results chains of logic charts typically come in many different forms and have a number of advantages for supervisors. Mayne (1999) names several benefits such as developing consensus on what the intervention or policy is trying to accomplish, developing an understanding of how it is believed to be working by identifying the cause-effect links implicit in the supervisor's activities and identifying the key measures of performance.

Figure 2 shows an example of a logic chart that focuses on the indicators of performance measurement and is a useful tool when selecting suitable indicators. It shows that, in general, a distinction can be made between effort indicators and effect indicators. Effort indicators like input and throughput do not show supervisory effectiveness, but may certainly be useful as a measure of supervisory efficiency (i.e. how much time and resources does it take to realize supervisory objectives?). Effect indicators measure supervisory performance and are typically divided in three categories: output, intermediate outcome and final outcome. Output can be defined as the direct consequences of supervisory intervention. Output indicators that are typically used by supervisors are the number of fines imposed or the number of revoked licenses. Although these numbers give an impression of the concrete output resulting from the supervisory efforts, they do not show whether financial institutions at stake are actually improving their behaviour in order to comply with rules and standards and to produce a socially desirable outcome. Therefore, in order to measure impact on societal objectives, supervisors need to use outcome indicators as well.

In contrast to output, outcome is not easy to measure. Figure 2 shows two reasons for this. First, outcome might be affected by different external events, which may or may not be controlled by the supervisor. As described in Section 2, this leads to the challenge of demonstrating causality: to what extent did supervision contribute to the outcome? Second, supervisory interventions may cause unanticipated, and sometimes even undesirable side effects, which may not always be visible in the short term and require study as part of an impact analysis. For example, from 2013 onwards, regulation prohibits financial intermediaries in the Netherlands to receive fees from financial products providers such as banks and insurers. The rationale behind this new

regulation is that financial intermediaries that are being paid by financial product providers may have an incentive not to act in the best interest of their clients, but to advise financial products with a high fee. The prohibition to receive fees forces intermediaries to develop new earning models. In general, these models will be in the interest of their clients, but that may not automatically hold true. A careful impact analysis is therefore required in order to limit any negative side effects.

Figure 2 – Effects chain: an example of the supervision of remuneration policies

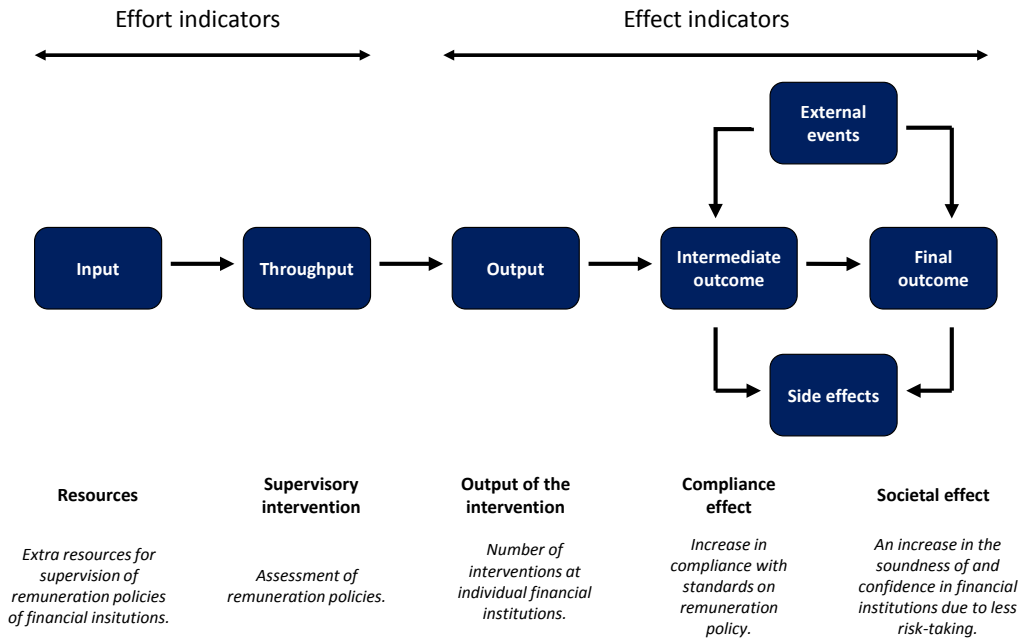
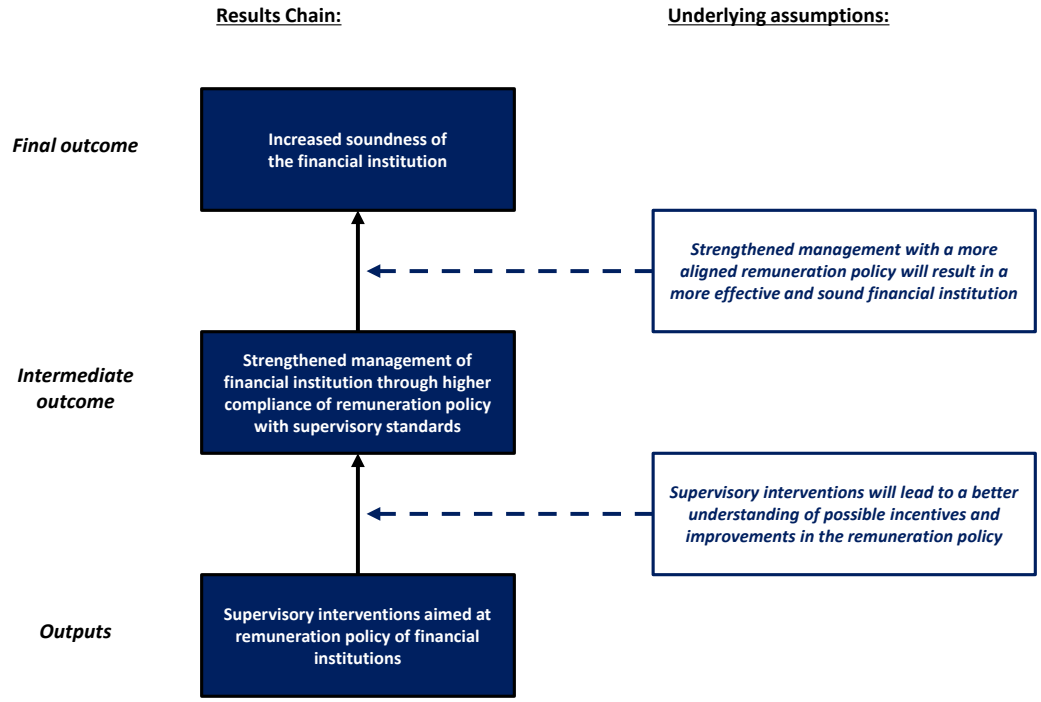


Figure 3 presents a different example of a tool that can help supervisors build a credible contribution story, namely a results chain. In addition to identifying the steps in the theory of change (from output to intermediate outcome to final outcome), the results chain typically also illustrates the assumptions behind the theory. For instance, a financial supervisor may intervene in the remuneration policy of a financial institution (output) under the assumption that this will give the financial institution a better understanding of the possible risks from an unbalanced remuneration policy, and an incentive to improve its remuneration policy. This will subsequently result in a remuneration policy that is more compliant with the supervisory standards and that will strengthen the management quality within the financial institution (intermediate outcome). Finally, the supervisor assumes in its theory of change that a strengthening of management will increase the soundness of that institution and thus contribute to financial stability (final outcome). So through presenting and discussing the reasoning behind the intervention or policy, the supervisor can exactly lay out what the major assumptions are concerning the contribution of its activities. Moreover, Mayne (2008) argues that a supervisor

can gather evidence on the occurrence of these key results to validate the theory of change and the underlying assumptions (i.e. did the intervention actually result in a changed remuneration policy and a strengthening of management?). This provides the supervisor with an opportunity to test which links (or assumptions) in the results chain are strong (e.g. strong logic, sufficient evidence available) and which are not. All this adds up to a stronger and more credible contribution story with which the supervisor can present reasonable evidence of its effectiveness in contributing to the desired outcomes.

Figure 3 – Example of a results chain: theory of change for supervision of remuneration policies



4.3 Develop a portfolio of performance indicators

The third lesson relates to the importance of constructing a coherent set of performance indicators. Monitoring a portfolio of indicators instead of a single parameter allows supervisors to better evaluate their effectiveness, as well as their efficiency, for at least two reasons.

First, a portfolio of indicators is better able to incorporate different perspectives than a single performance indicator. In this respect, Mayne (1999) advocates the "multiple lines of evidence" argument. While the power of evidence stemming from a single indicator is often not convincing, a larger set of different and complementary evidence can become much more compelling. For that, supervisors may learn from the private sector, where Kaplan and Norton (1992) developed the so-called Balanced Scorecard to provide businesses with a varied set of performance measures. The Balanced Scorecard not only evaluates a firm's most recent financial numbers (i.e.

the financial perspective), but also incorporates the customer perspective, the operational perspective and the learning and growth perspective. The central thesis underlying the Balanced Scorecard, namely that non-financial measures are also important in measuring business performance, is also relevant for supervisors as non-financial measures may have a significant impact on the soundness of financial institutions under supervision (Moore, 2003).

A second advantage of applying multiple metrics is that a portfolio of indicators is less sensitive to outliers than a single parameter. Take in mind the situation where a supervisor uses the number of complaints received from the clients of a specific insurer to determine whether the company is improving its level of fair treatment of customers. When it is part of the supervisory strategy to increase the customer awareness of the quality of advice given by insurers in general it might be possible that the number of complaints regarding a specific insurer rises, at least in the beginning. This may give the impression that the insurer is not treating its customers fairly while this does not have to be the case. Customer awareness in general may have simply increased. This example does not imply that the number of complaints is a wrong indicator (in fact, it can be a very informative indicator) but it shows that it is important to develop a set of indicators and to monitor these indicators over time. In this context, benchmarking might also be a useful tool to determine whether a specific pattern is unusual.

When developing a set of performance indicators, it is important for financial supervisors to include different types of indicators in their portfolio. The choice of indicators is strongly related to the objective of the performance measurement, as the latter will determine what type of indicator is most appropriate. If the objective is to evaluate the effectiveness of the supervisor, then outcome indicators are generally most suited. Only by applying outcome indicators can supervisors answer the effectiveness question: did the supervisor achieve the results it set out to produce (Behn, 2003)? However, using performance measurement for budgeting or efficiency purposes requires more than outcome indicators alone. This becomes increasingly important as more attention is being paid to the cost/benefit analysis of supervision. Financial supervisors are expected to decrease market failure. But when the costs of supervisory resources are higher than the benefits in terms of compliance and societal effects, the solution to market failure may be worse than the failure itself (supervisory failure).[10] For efficiency purposes, Moore and Braga (2003) argue that supervisors need a mix of outcome, output and input measures. To illustrate, consider a financial supervisor who not only wants to determine his effectiveness, but also wants to analyse his efficiency in achieving the desired objectives. For that, the supervisor needs

[10] See New Perspectives on Regulation (2009) and Stiglitz (2009).

to monitor the soundness of the financial institutions under supervision (outcome), but also needs to incorporate the number of on-site inspections performed (output) and the level of supervisory resources deployed (input). Hence, using a combination of input, output and outcome measures is a useful approach to measure supervisory efficiency.

Moreover, it is also important to include indicators that focus on the different levels of performance measurement. When it is carried out with the purpose of accountability, performance measurement should not only be focused at the strategic level (where the causal relation between supervisory interventions and objectives is not always easy to establish), but also at the tactical and operational level. A portfolio of indicators at all three levels will give insight into the actual performance of the supervisor without only focusing on the objectives that it can influence directly. A focus on the latter would not do justice do the expectations of the public and other stakeholders that financial supervisors take also responsibility for achieving strategic objectives. It is in the combination of indicators that a convincing picture of the actual performance of the supervisor can be made visible.

4.4. Find alternative research designs to increase the plausibility of the causal relationship

Although demonstrating the causality between efforts and effects in financial supervision is challenging, supervisors can increase the plausibility of whether a specific effect is attributable to its efforts by selecting a research framework that can exclude the impact of exogenous factors. Next to experimental research designs such as a randomised controlled experiment, there are also other research methods that can be used to measure the (strength of the) causal relation between variables. An example is the so-called General Elimination Methodology or GEM approach which aims to systematically exclude alternative causes for an observed effect (Scriven, 2008). Another example is provided by Klomp and De Haan (2011) who apply factor analysis and a quantile regression model to examine the relationship between bank supervision and banking risk for the years 2002 to 2008. The authors use several indicators to model banking risk (including indicators relating to capital adequacy, asset quality and managerial qualities) and for the supervision of banks (including capital requirements and supervisors' authorities). Klomp and De Haan (2011) conclude that not all banks in their sample are affected by regulation and supervision to the same degree. Banks with a high risk profile appear to be significantly more affected by regulation and supervision than banks with a lower risk profile. Moreover, the authors find that not all regulatory and supervisory standards have a significant influence on bank risk. Hence, by applying an appropriate research framework, the authors are able to examine the causal relationship between financial supervision and the risk behaviour of banks.

Box 2 – Measuring the performance of supervisory policy

The four lessons presented in this paper are also relevant for a supervisory policy environment. In this box, we discuss how supervisors can measure the performance related to their policy initiatives.

1. Define a specific objective and measure at a micro level

For most financial (prudential) supervisors, their core mission is to safeguard financial stability and contribute to the safety and soundness of financial institutions under their supervision. A supervisor's policy initiatives should contribute to achieving its mission. However, at the strategic level this may prove challenging due to the influence of exogenous factors. Supervisors, however, can increase the plausibility of the causal relationship between their policy initiatives and the outcome by reducing the level of abstraction at which they measure the effects.

For that, supervisors can translate their mission into specific policy priorities or themes. For instance, a supervisor's mission to safeguard financial stability (*final outcome*) can be transposed into a policy theme that "banks must build adequate financial buffers to be able to absorb unexpected adversities" (*intermediate outcome*). Such a theme is more easily translated into a specific policy-target: "The policy initiatives in international and national fora will significantly contribute to the build-up of adequate capital in the banking sector in 2013".

2. Develop a plausible theory of change and contribution story

Next, it is important for supervisors to develop a plausible theory of change to provide reasonable evidence on the contribution of their policy initiatives to the desired outcomes. This not only requires supervisors to present the policy steps and initiatives (*outputs*) they have undertaken, but also explain how these steps are believed to have contributed to the outcome. In other words, which policy steps and policy initiatives did the supervisor undertake this year and how are they believed to have contributed to the build-up of adequate capital in the banking sector?

3. Construct a portfolio of performance indicators

The performance of supervisory policy can be illustrated more convincingly by following the "multiple lines of evidence" argument and constructing a portfolio of performance indicators (Mayne, 1999). Supervisors can develop different types of outcome, output and input indicators that focus on the policy environment and can further improve their policy effectiveness and efficiency:

- Number of policy initiatives (revisions) successfully adopted in (inter)national fora (*outcome*).
- Number of revisions in national regulation following a supervisor's initiative (*outcome*).
- Number of international fora actively participated in (*output*).
- Number of seminars organized to influence key stakeholders in the policy environment (*output*).
- Number of resources (fte) deployed on a specific supervisory theme (*input*).
- Number of resources (fte) participating in a (inter)national forum (*input*).

4. Find alternative research designs to increase the plausibility of the causal relationship

Qualitative research designs, such as the case study format, add value when measuring policy performance. They can strengthen the plausibility of the theory of change and provide more insight into the context of supervisory policy initiatives. Policy-makers frequently need to reach a compromise, obtaining valuable new insights in the process. In addition, supervisors can 'back test' their policies (i.e. simulate what would have happened if the particular policy initiative had been in place). Back-testing after a crisis or incident, for instance, provides supervisors with a tool to evaluate whether the supervisory (and regulatory) framework was appropriate in retrospect. Back testing can also provide valuable insight into the unintended consequences of policy initiatives and can thus help to improve the quality of the supervisory process.

5. Measuring effects: indicators

As outlined earlier, it is important to select a suitable portfolio of indicators that fit the supervisory objectives. In this section we provide various examples of indicators for measuring the effects of financial supervision. For that, we distinguish the performance indicators into two types: 'hard' indicators based on quantitative date, and 'soft' indicators derived from qualitative data (also see Sijbrand and Rijsbergen, 2013). Table 1, at the end of this section, provides an overview of the presented indicators.

5.1 Hard indicators

Various 'hard' indicators are available to financial supervisors for measuring the effects of their supervision. Firstly, they can use indicators based on market data, such as credit ratings, stock prices or the level of credit default swap (CDS) spreads for a financial institution.[11] These indicators reflect the risk profile of the relevant institution(s), as perceived by market parties, and are reported by supervisors such as the UK Financial Services Authority (FSA) in its annual report (FSA, 2011). Other examples of indicators based on market data are the levels of private investments in high-risk products, the nature of financial products launched and financial institutions' revenue models or the number of criminal investigations resulting from unusual transaction reports of financial institutions to the Financial Intelligence Unit.

Secondly, supervisors can opt for indicators that reflect specific supervisory requirements, such as solvency and liquidity ratios for banks, insurers and other financial institutions. Supervisors, for example, can provide information on various solvency ratios, including the total BIS ratio[12] and Tier 1 ratio[13] for banks. The solvency ratios as well as the indicators based on market data should, however, be adjusted for the economic cycle as they are influenced by the general economic climate. Solvency ratios, for example, are usually higher when the economy is doing well than during economic downturns. Adjusting for the economic cycle is a way to reduce the influence of external factors (i.e. the economic climate) and in turn makes it easier for financial supervisors to demonstrate a causal relationship between improvements in solvency ratios and supervisory intervention.

A third category of 'hard' indicators is related to the number of bankruptcies among supervised institutions and the amount of losses accompanied by these defaults. The Australian prudential

[11] Investors can buy a credit default swap contract as insurance against a counterparty's credit risk. Higher premiums on such contracts imply a higher credit risk.
[12] The BIS ratio represents the ratio of shareholders' equity to outstanding loans.
[13] The Tier 1 ratio shows banks' core equity capital as a percentage of risk-weighted assets.

supervisor APRA, annually reports two 'hard' indicators linked to financial failures (APRA, 2011). The first is the *Performance Entity Ratio (PER),* which reflects the number of supervised entities able to meet their liabilities as a percentage of the total number of entities supervised. The second indicator is the *Money Protection Ratio (MPR)*, which represents the dollar value of liabilities to beneficiaries that remained safe in a specific year, divided by the total amount of liabilities. Furthermore, the Federal Reserve applies a somewhat similar indicator that measures the losses from state member banks to the *Deposit Insurance Fund (DIF)*. The Federal Reserve annually reports the outcome and strives to prevent losses from becoming greater than premiums paid into the DIF by state member banks (Federal Reserve, 2011).

Quantification of supervisory intervention in terms of its economic benefits for consumers is a fourth example of a useful 'hard' indicator. The Netherlands Competition Authority ('NMa'), for example, publishes the directly quantifiable benefits of formal supervisory actions, such as its decisions to impose sanctions on cartels, tariff regulations in the energy market and withdrawals of licence applications (Kemp et al., 2011). The benefits are calculated as the three-year moving average of the initial-year effects of formal actions in a specific year.[14] The economic benefit calculated for consumers came out at € 265 million in 2011. At NMa's request, the underlying calculations were verified by the Netherlands Bureau for Economic Policy Analysis (CPB, 2012).

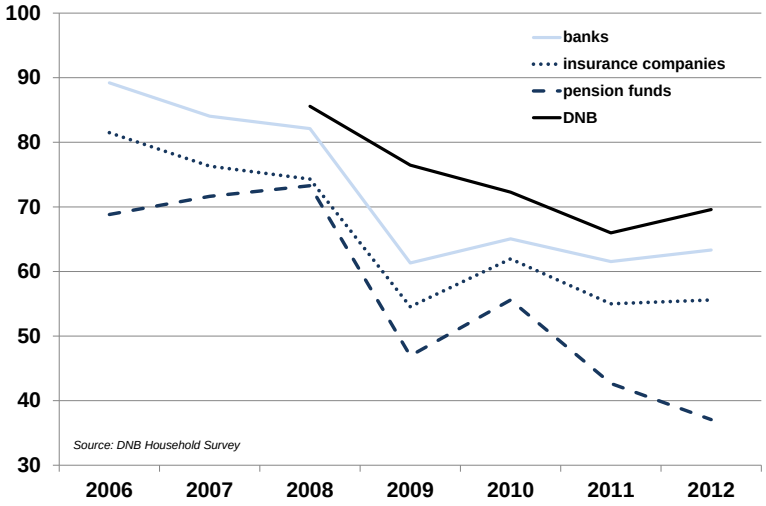
Lastly financial supervisors can use the throughput time for supervisory activities such as procedures, applications, assessments, providing answers to questions and tests as an indicator. The English FSA is one of the parties that periodically report such indicators (FSA, 2011). Another example is the Canadian supervisor OSFI, which measures whether its processing applications for regulatory approval are conducted within the established time frames (OSFI, 2011). Furthermore, the Federal Reserve annually monitors the number of reports of its supervisory examinations that are completed within the established deadlines. These types of indicators can primarily be used to determine the efficiency of a supervisor, which is particularly helpful when measuring effects at an operational level. However, it is important to realise that throughput time in itself has limited meaning to the actual quality of the supervisory activities. A critical assessment of a financial institution's application for market access, for instance, will take considerably more time than just checking the box, but can be essential for achieving the supervisory objective. Therefore, only focussing on throughput time in performance measurement can cause an adverse effect.

[14] The NMa determines these consequences by applying rules of thumb to determine the future effects of decisions taken in a calendar year. These rules are based on findings in economic literature.

5.2 Soft indicators

In addition to 'hard' indicators, the effects of financial supervision can also be measured using 'soft' indicators that are derived from more qualitative information. An example of a 'soft' performance indicator is public confidence in the financial sector or in financial supervisors. This can be measured periodically through random surveys of representative groups of citizens. DNB, for instance, annually measures the public confidence by means of a representative household survey. Figure 3 displays the results of these household surveys between 2006 and 2012. A different approach to measuring public confidence is applied by the FSA. In the annual FSA Consumer Awareness Survey, respondents are asked to indicate whether they are confident that they are treated fairly by financial institutions (FSA, 2011).

Figure 3 – Public confidence in Dutch financial institutions and supervisor



Indicators showing the migration between pre-defined supervisory regimes or risk scores can also be useful. Supervisors usually divide the institutions they supervise into different categories within a risk-based framework in order to optimise their capacity allocation. These categories determine the intensity of supervision required. Alternatively, risk scores may be used to indicate how supervisors should assess risks in a specific part of the business.[15] A financial institution with a poor financial position will, for example, be assigned to an 'increased supervision' regime. Changes in the number of institutions subjected to a specific supervisory regime can then be used to indicate the effect of supervision. The Australian APRA, for example, shows the change in the number of supervised institutions within its own Supervisory Oversight and Response System (APRA, 2011). More specifically, APRA publishes annual information showing the migration of institutions between the four risk categories in the risk assessment

[15] These risk scores can be calculated for individual institutions and also for whole sectors.

system over the previous three and seven years.[16] It should be noted, however, that the assignment to supervisory categories or risk scores is partly determined by external factors and the supervisors themselves. Bafin, the German financial supervisor, has sought to deal with the causality problem by not only reporting institutions' risk classifications over a period of several years, but also reporting the number of supervisory interventions (as a percentage) made in each risk category (Bafin, 2011). Measuring the period during which an institution is subjected to a specific supervisory regime can be useful as an effect indicator. DNB, for example, monitors which institutions are in what specific supervisory regime (especially for 'urgent' and 'high risk' regimes). As a target, DNB has defined that depending on the systemic importance of the financial institution, it is not supposed to stay longer than a certain period of time in a specific supervisory regime. If this period is exceeded, DNB will explicitly assess the effectiveness of its supervisory interventions.

'Soft' indicators can also measure the level of compliance with assessment criteria relating to qualitative aspects of supervision such as business models, strategy and governance within financial institutions. This can involve measuring effects on the basis of standards set by the supervisors. Supervisors regularly draw up standards by way of elaboration of principles laid down in legislation. The extent of compliance with these standards indicates can be an indication of supervisory effectiveness. The DNB standards for information security are an example of this. These standards specify the level at which financial institutions must be in control of their risks in the area of information technology.[17] DNB conducted a zero measurement in 2010 to establish the extent of information security risk control at financial institutions. Those institutions deemed to be insufficiently in control of these risks were required to compile and execute an action plan to bring their risk control into line with the standards. A 'year 1 measurement' will be made at the same institutions in 2013. Being able to demonstrate that more institutions have increased their compliance with the required minimum standards through time can be seen as an effect of the supervision.

Fourth, 'soft' performance indicators can also be based on the outcome of external or peer reviews that measure the level of compliance towards (inter)national supervisory standards. An example of such a study is performed by Cihák and Tieman (2008), who analyze the quality of financial sector supervision by using data from the IMF and World Bank assessments of compliance with international standards and codes. They find significant differences in the

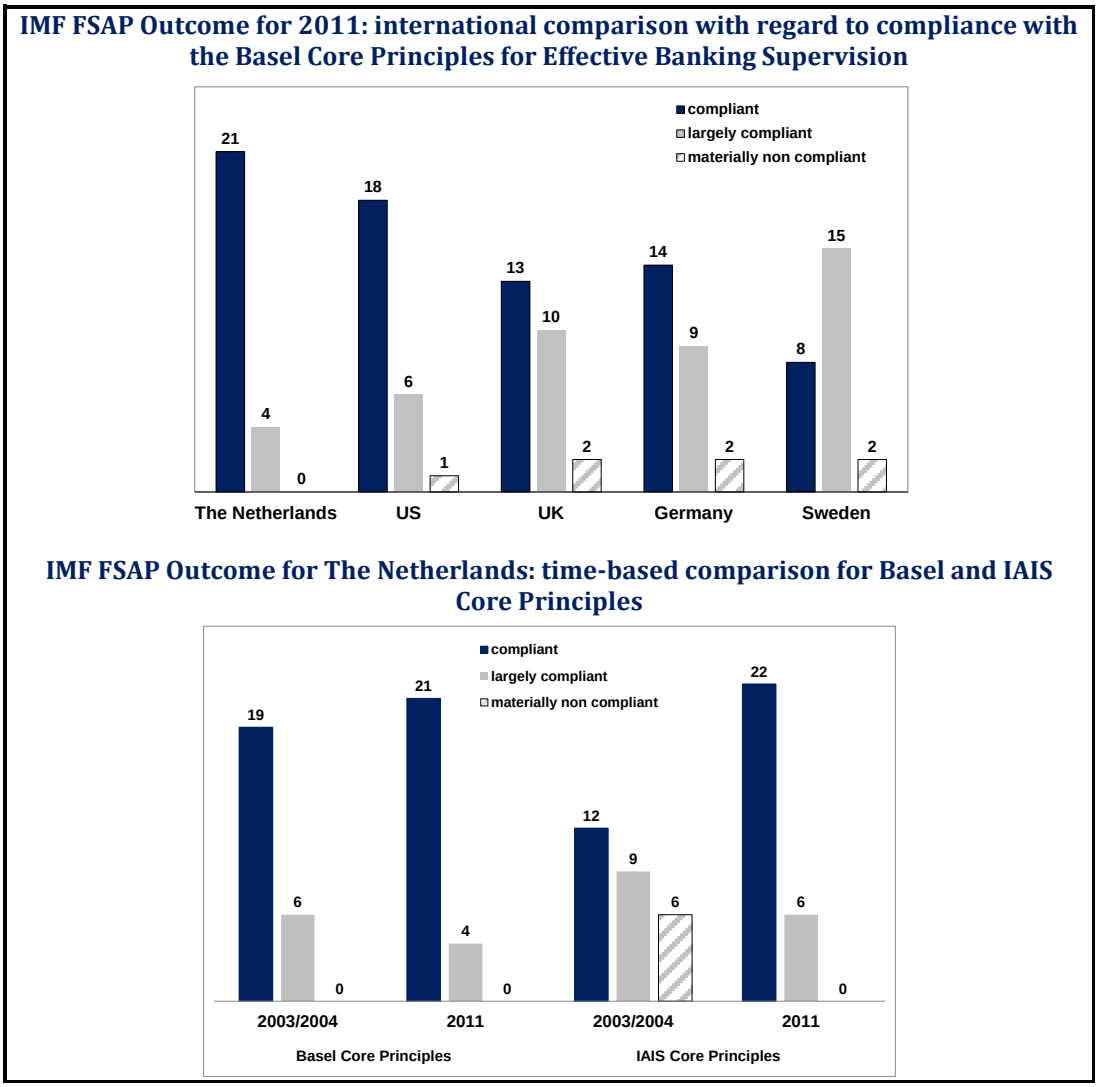
[16] The SOARS method distinguishes between the following four categories (ranked in order of increasing intensity of supervision): normal, oversight, mandated improvement and restructure.
[17] The Assessment Framework for DNB Information Security Examination can be found in DNB's Open Book for Supervision on www.toezicht.dnb.nl/3/50-203304.jsp.

quality of supervisory frameworks across countries, with per capita income being a major factor. In the Financial Sector Assessment Program (FSAP) analyses and in its ROSCs (Reports on the Observance of Standards and Codes), the IMF examines the quality of member states' supervision, based on international standards that represent minimum requirements for the supervision of banks, insurers and securities brokers.[18] These standards include the Basel Core Principles for Effective Supervision and the IAIS Core Principles for Insurance Supervision (IMF, 2005). A country can essentially score compliant, largely compliant, materially non-compliant or non-compliant for each individual standard. Although the individual tests are of a qualitative nature, the results can be aggregated to produce a more quantitative picture of the quality of financial sector supervision. In 2012, the Basel Committee on Banking Supervision (BCBS) started a similar assessment method to determine the degree of compliance with the requirements of the Basel 3 Accord. This program – the Regulatory Consistency Assessment Program (RCAP) – looks at the extent to which domestic Basel 3 regulations are aligned with the requirements in the Accord's rules text. It thereby helps to identify gaps and allows supervisors and regulators to initiate corrective measures, when needed. The RCAPs use the same four scores that are used for assessing compliance with the Basel Core Principles (see figure 4).

Figure 4 presents two graphs based on IMF FSAP results that can be used as a basis for measuring the effectiveness of supervision. The top graph shows an international comparison of the 2011 FSAP results for the Basel Core Principles for Effective Banking Supervision. The bottom graph, on the other hand, shows a comparison of outcomes through time, by presenting the FSAP outcomes in 2003/2004 and 2011 for the Netherlands.[19] It should be noted that the elements examined here are closer to 'output' than 'outcome' as the tests relate to the quality, or improvements in the quality, of supervision rather than to the ultimate results of the supervision in terms of financial health or stability of the sector.

[18] The Basel Core Principles for Effective Banking Supervision for banks, the Insurance Core Principles for insurers and the IOSCO Objectives and Principles for securities brokers are the best known standards. These standards were compiled after extensive international consultations and are regularly updated.
[19] Note that over time the standards will obviously become stricter.

Figure 4 – IMF Financial Sector Assessment Program Outcomes 2011 (IMF, 2011)



Additionally, several financial supervisors employ stakeholder surveys to measure the perceived effectiveness of their supervision. The Canadian OSFI, for instance, conducts anonymous surveys of knowledgeable industry observers to help assess its performance and effectiveness (OSFI, 2011). The Australian APRA conducts stakeholder surveys and consequently publishes the outcomes of these surveys. Stakeholders can score APRA's performance (using a five-point scale) in a wide range of areas ranging from the staff's demonstration of integrity or professionalism to the effectiveness of APRA's external communication (APRA, 2011). A somewhat similar approach is applied by the English FSA. By using a so-called 'firm feedback survey', the FSA annually monitors the firm-supervisory relationship with regards to the quality of risk assessment and risk mitigation. Moreover, the FSA monitors the views of supervised firms about the performance of the FSA in achieving its statutory objectives, such as maintaining confidence in the financial system and securing the

appropriate degree of protection for consumers. The scores (and migration in scores) are published in the annual report (FSA, 2011).

A final example of a "soft" indicator for determining the quality of supervision is 'back testing' supervisory methods. Some financial supervisors already use incidents in the financial sector to 'back test' their supervisory methods, with the intention to analyze whether their reporting mechanisms or the common assessment of risks in their supervisory method would have revealed the specific situation and would have led to supervisory action. A regular practice of back testing supervisory methods can give useful insight into the quality of the supervisory process and its effectiveness.

5.3 Overview of indicators

In this section we have discussed performance indicators, which can broadly be classified in two types: 'hard' and 'soft' indicators. Table 1 provides an overview of the indicators presented in this section. The advantage of 'hard' indicators is the higher degree of objectivity and verifiability associated with them. Many of them are relatively straightforward to understand, easy to monitor over time and can be aggregated for use in a way that complies with the requirement of confidentiality applying to financial supervision. The indicators are usually aggregated at a relatively high level. 'Soft' indicators also make it possible to show the supervisory efforts undertaken in respect of qualitative aspects. This is important because financial supervision is increasingly focused on forward-looking aspects such as governance, conduct and culture. These indicators, too, can be aggregated for use in a way complying with the duty of confidentiality. On the other hand, 'soft' indicators are usually less objectively observable and verifiable than many 'hard' indicators.

Table 1 – Performance indicators at different levels

Suitable for: / Type of indicator:	Strategic level:	Tactical level:	Operational level:
Hard indicators			
1. Market data: • *CDS spreads, credit ratings and equity prices* • *Levels of private investments in high-risk products, nature of financial products launched.*		X	
2. Supervisory ratios: • *BIS and Tier-1 and Tier-2 capital ratios, leverage ratio and liquidity ratios (LCR, NSFR) for banks.* • *Solvency ratios (insurers) and funding ratios (pension funds)*		X	
3. Failures and losses: • *Performance Entity Ratio (PER)* • *Money Protection Ratio (MPR)* • *Pay-outs to Deposit Insurance Fund (DIF)* • *Number of bankruptcies (PD)* • *Losses due to failures (average LGD)*	X	X	
4. Economic benefits: • *Annual economic benefit for consumers* • *Costs of financial crises*	X	X	
5. Efficiency indicators: • *Throughput time for procedures, applications* • *Number of supervisory reports completed on time* • *Number of supervisory staff per unit of currency protected*			X
Soft indicators			
1. Public confidence: • *Household surveys* • *Consumer awareness surveys*	X		
2. Supervisory regimes: • *Total number of institutions migrated between supervisory regimes* • *Total number of institutions in the highest (riskiest) supervisory regime* • *Average length of stay of institutions in supervisory regimes*		X	
3. Compliance with soft assessment criteria: • *Non-compliance with criteria regarding business models* • *Non-compliance with criteria regarding governance*		X	
4. International standards assessments: • *IMF FSAP / ROSCs* • *BCBS / RCAP* • *EBA Surveys*		X	X
5. Stakeholder surveys: • *Firm feedback survey* • *Views of industry observers* • *Surveys among peer supervisors*		X	X

6. Closing comments

Financial supervisors are increasingly expected to be able to demonstrate the effectiveness of their actions. In practice, however, this is not an easy task as many different factors influence the attainment of financial supervisors' societal objectives. Challenges may occur in terms of causality, timing, confidentiality and capacity. The fact that it is difficult to measure effects does not mean, however, that financial supervisors should not attempt to do so as measuring the effects of their work allows them to account for their performance to the outside world, while also improving the supervisory process.

In this paper we have provided an outline for use in measuring the effects of financial supervision. Supervisors will have to decide the level – strategic, tactical or operational – at which they wish to conduct measurements, depending on what they are seeking to achieve. The indicators selected to measure the chosen effects will also need to be 'fit for purpose'. Four lessons may contribute to the success of performance measurement: 1) define a specific objective and measure at a micro level, 2) develop a plausible theory of change and contribution, 3) develop a portfolio of performance indicators, and 4) find alternative research designs to increase the plausibility of the causal relationship. As the examples in this article show, supervisors can select from a range of 'hard' and 'soft' indicators. The opportunities for measuring effects are particularly good at a tactical level, where supervisors measure whether a specific intervention has resulted in any change in behaviour. If interventions are properly chosen, they will also contribute to achieving 'higher' societal objectives. This link is more difficult to establish and measure. However, as Albert Einstein once said, and this may be some consolation to financial supervisors, *'Not everything that counts can be counted. Not everything that can be counted counts'.*

References

Algemene Rekenkamer (2011), *Toezicht van DNB op de stabiliteit van banken*, The Hague, 2011.

Australian Prudential Regulation Authority (2011), *APRA Annual Report 2011*.

Behn, R.D. (2003), 'Why Measure Performance? Different Purposes Require Different Measures', *Public Administration Review 63*, 586-606.

Bundesanstalt fur Finanzdienstleistungsaufsicht (2011), *Bafin Annual report 2011*, Bonn 2011.

Centraal Plan Bureau, *Plausibiliteitstoets Outcome-berekeningen NMa*, 24 January 2012.

Cihák, M. and A. Tieman (2008), 'Quality of Financial Sector Regulation and Supervision Around the World', *IMF Working Paper 08/190*.

De Nederlandsche Bank (2009), *DNB Supervisory Strategy 2010-2014*, DNB, Amsterdam.

De Nederlandsche Bank (2010), *From Analysis to Action: Action plan for a change in the conduct of supervision*, Amsterdam.

De Nederlandsche Bank (2013), *DNB Supervisory Themes 2013*, Amsterdam.

Federal Reserve System (2011), *Government Performance and Results Act – Annual Performance Report*, Washington 2011.

Financial Services Authority, Annual Report 2010/11, London 2011.

Gezondheidsraad (2011), *Op weg naar evidence based toezicht. Het onderzoek naar effecten van toezicht door de Inspectie voor de Gezondheidszorg*, The Hague, 2011.

Hilbers, P.L.C., Rijsbergen, D.R., Raaijmakers, K. and F. de Vries (2012), 'Effectmeting in het toezicht op de financiële sector', *Tijdschrift voor Toezicht (3),* 38-49.

Hilbers, P.L.C. (2011), 'Toezicht 2.0: Nieuwe spelregels voor een sterke financiële sector', *inaugural address* Nyenrode Business University, 21 October 2011.

International Monetary Fund (2011), *Financial Sector Stability Assessment*, FSAP report, Washington DC.

International Monetary Fund (2005), Financial Sector Assessment, A Handbook, Washington DC.

Kaplan, R.S. and D.P. Norton (1992), The Balanced Scorecard – Measures that Drive Performance, *Harvard Business Review 70*, 71-91.

Kemp, R., van Sinderen, J., Mulder, M., Borkent, K., Cordeweners, J. and D. Leliefeld (2011), *Outcome van NMa-optreden; Resultaten voor 2011*, NMa Working Papers, No. 7.

Klomp, J. and J. de Haan (2011), Banking risk and regulation: does one size fit all?, *DNB Working Paper No. 323*.

Mayne, J. (1999), 'Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly', *Discussion Paper, Officer of the Auditor General of Canada*.

Mayne, J. (2008), 'Contribution Analysis: An approach to exploring cause and effect', *International Learning and Change Initiative (ILAC) Brief No. 16.*

Moore, M.H. (2003), 'The Public Value Scorecard: A Rejoinder and an Alternative to "Strategic Performance Measurement and Management in Non-Profit Organizations" by Robert Kaplan', *John F. Kennedy School of Government Working Paper No. 18.*

Moore, M.H. and A.A. Braga (2003), 'Measuring and improving police performance: the lessons of Compstat and its progeny', *Policing: An International Journal of Police Strategies & Management 26*, 439-453.

Moss, D. and J. Cisternino, (ed.) (2009), *New Perspectives on Regulation,* The Tobin Project, Cambridge MA.

Office of the Superintendent of Financial Institutions Canada (2011), *Annual Report 2011*, Ottawa.

Ottow, A.T. (2011), 'De toezicht Bermudadriehoek', *Tijdschrift voor Toezicht*, vol. 4, 48-50.

Scriven, M. (2008), A summarative evaluation of RCT methodology & an alternative approach to casual research, *Journal of MultiDisciplinary Evaluation*, 11-24.

Shadish, W.R., Cook, T.D. and T.D. Campbell (2002), *Experimental and quasi-experimental designs for generalized causal inference*, Boston/New York, Houghon Mifflin Company.

Sherman, L.W., Gottfredson, D.C., MacKenzie, D.L., Eck, J., Reuter, P. and S.D. Bushway (1998), 'Preventing Crime: What Works, What Doesn't, What's Promising', *National Institute of Justice Research in Brief*, Washington DC.

Sijbrand, J. and D.R. Rijsbergen (2013), Managing the quality of financial supervision. In: J. Kellermann, F. de Vries and J. de Haan (eds), Financial Supervision in the 21th Century, Heidelberg, Springer.

Sparrow, M.K. (2008), *The Character of Harms*, Cambridge University Press, Cambridge.

Stephens, N.G., Parsons, A., Schofield, P.M., Kelly, F., Cheeseman, K., Mitchinson, M.J. and M.J. Brown (1996), 'Randomised controlled trial of Vitamen E in patients with coronary disease: Cambridge Heart Antioxidant Study (Chaos)', *The Lancet*, 781-786.

Stiglitz, J. (2009), Government Failure vs. Market Failure: principles of regulation, *Government and Markets, towards a new theory of regulation,* The Tobin Project, Cambridge MA.

Weiss, C.H. (1997), 'Theory-based evaluation: past, present, and future. *New Directions for Evaluation 76*, 41-55.

Willemsen, F. (2009), *Programmeringstudie: Deel II, Designs effectmetingen handhaving en toezicht, Startnotitie WODC-onderzoek*, The Hague.

Winter, H.B. and J. de Ridder (2010), *Toezicht en effectmeting. Het kan! Researchsynthese van onderzoek naar de effectiviteit van toezichtsinterventies*, The Hague, Ministry of Justice.

Winter, H.B. (2012), 'Meten van de effecten van toezicht, *Tijdschrift voor Toezicht (3),* 63-80.

Wouters, P., Derriks, H. and H. van der Loop (2009), *Toezicht tellen*, The Hague: Ministry of Transport, Public Works and Water Management.

Previous DNB Working Papers in 2013

No. 367 **Andrea Colciago and Lorenza Rossi**, Firm Entry, Endogenous Markups and the Dynamics of the Labor Share of Income
No. 368 **Dirk Broeders, Paul Hilbers and David Rijsbergen**, What drives pension indexation in turbulent times? An empirical examination of Dutch pension funds
No. 369 **Luca Arciero, Ronald Heijmans, Richard Heuver, Marco Massarenti, Cristina Picillo and Francesco Vacirca**, How to measure the unsecured money market? The Eurosystem's implementation and validation using TARGET2 data
No. 370 **Stijn Claessens and Neeltje van Horen**, Impact of Foreign Banks
No. 371 **Gabriele Galati, Federica Teppa and Rob Alessie**, Heterogeneity in house price dynamics
No. 372 **Jan Willem van den End**, A macroprudential approach to address liquidity risk with the Loan-to-Deposit ratio
No. 373 **Jon Frost and Ayako Saiki**, Early warning for currency crises: what is the role of financial openness?
No. 374 **Ronald Heijmans, Lola Hernández and Richard Heuver**, Determinants of the rate of the Dutch unsecured overnight money market
No. 375 **Anneke Kosse and Robert Vermeulen**, Migrants' Choice of Remittance Channel: Do General Payment Habits Play a Role?
No. 376 **Jacob Bikker**, Is there an optimal pension fund size? A scale-economy analysis of administrative and investment costs
No. 377 **Beata Bierut**, Global liquidity as an early warning indicator of asset price booms: G5 versus broader measures
No. 378 **Xiao Qin and Chen Zhou**, Systemic Risk Allocation for Systems with A Small Number of Banks
No. 379 **Niels Gilbert, Jeroen Hessel and Silvie Verkaart**, Towards a Stable Monetary Union: What Role for Eurobonds?
No. 380 **Mauro Mastrogiacomo**, Reform of the mortgage interest tax relief system, policy uncertainty and precautionary savings in the Netherlands
No. 381 **Christiaan Pattipeilohy, Jan Willem van den End, Mostafa Tabbae, Jon Frost and Jakob de Haan**, Unconventional monetary policy of the ECB during the financial crisis: An assessment and new evidence
No. 382 **Alexander Popov and Neeltje Van Horen**, The impact of sovereign debt exposure on bank lending: Evidence from the European debt crisis
No. 383 **Andrea Colciago**, Imperfect Competition and Optimal Taxation
No. 384 **Richhild Moessner**, Effects of explicit FOMC policy rate guidance on market interest rates
No. 385 **Steven Ongena, Jose Luis Peydro and Neeltje van Horen,** Shocks Abroad, Pain at Home? Bank-Firm Level Evidence on the International Transmission of Financial Shocks
No. 386 **Jakob de Haan and Razvan Vlahu**, Corporate governance of banks: A survey
No. 387 **Saskia van Ewijk and Ivo Arnold**, How bank business models drive interest margins: Evidence from U.S. bank-level data