Inspections and enforcement as instruments for enhancing environmental behavior

Final report
Inspections and enforcement as instruments for enhancing environmental behavior

Final report

Mathias Herzing and Adam Jacobsson (eds.)
Preface

There is a need for scientific knowledge to demonstrate how inspections and enforcement affect the actions done by the operators. There is also a great need for in-depth knowledge that can help to develop the conditions for inspections and enforcement, and which can help to strengthen the profession.

In 2014 the Swedish Environmental Protection Agency, together with the Swedish Agency for Marine and Water Management, initiated three research projects to increase the knowledge about how different governing instruments function and can be designed for implementation and evaluation in practical environmental work.

This report contains the results from the research project “Inspections and enforcement as instruments for enhancing environmental behavior”. The overall aim of the project has been to investigate how inspections and enforcement can serve as a governing instrument for work towards the environmental goals. Mathias Herzing and Adam Jacobsson have managed the research project, which also included Lars Forsberg, Håkan Källmén and Hans Wickström. The outcomes have complemented and deepened the research carried out in the previous research program “Efficient environmental inspections and enforcement” (Swedish EPA, Report 6558).

The researchers are solely responsible for the content of the report.

Contact persons at the Swedish EPA have been Katariina Parker and Martin Gustafsson.

The projects were financed by the Swedish EPAs environmental research fund.

Stockholm in November 2017

Lena Callermo
Head of Sustainability department
Innehåll

FÖRORD 3

1. SAMMANFATTNING 8

2. SUMMARY 10

3. INTRODUCTION 12

3.1. Aim and scope of the research program 12

3.2. Topic A: Inspectee behavior – applying Motivational Interviewing to inspections and enforcement 13

3.2.1. Motivational Interviewing 13

3.2.2. Measuring MI skills 14

3.2.3. The MI training programs 16

3.2.4. Assessments of inspections and inspection conversations 17

3.2.5. Evaluating inspection effects 17

3.2.6. Aim and scope of studies A1-A3 17

3.2.7. Main results from studies A1-A3 18

3.3. Topic B: Inspections and enforcement activities – explaining and measuring the efficiency and effectiveness of inspections 22

3.3.1. Background 22

3.3.2. Summary of study B1 24

3.3.3. Summary of study B2 25

3.3.4. Summary of study B3 25

3.4. Concluding remarks 26

3.5. References 28

4. APPLYING MOTIVATIONAL INTERVIEWING TO INDUCE COMPLIANCE WITH RADON GAS RADIATION LEGISLATION – A FEASIBILITY STUDY 31

4.1. Introduction 31

4.2. Method 33

4.2.1. Research design 33

4.2.2. The MI training program 35

4.2.3. Inspectors’ assessments 37

4.2.4. Compliance of property owners 37

4.2.5. Statistical analysis 37
4.3. Results

4.3.1. Inspectors’ MI skills

4.3.2. Inspectors’ perceptions of the MI training program

4.3.3. Inspectors’ perceptions of inspection calls

4.3.4. Compliance of property owners

4.4. Discussion

4.5. References

5. Enhancing Compliance with Waste Sorting Regulations through Inspections and Motivational Interviewing

5.1. Introduction

5.2. Methodology

5.2.1. Research design

5.2.2. The MI training program

5.2.3. Assessments by inspectors and inspectees

5.2.4. Statistical analysis

5.3. Results

5.3.1. MI skills

5.3.2. Training day questionnaires

5.3.3. Inspectors’ inspection questionnaires

5.3.4. Restaurants’ perception of inspections

5.3.5. Compliance with waste sorting regulations

5.4. Conclusion

5.5. References

6. Applying Motivational Interviewing in the Context of Cattle Inspections

6.1. Introduction

6.2. Methodology

6.2.1. Research design

6.2.2. The MI training program

6.2.3. Assessments of inspectors’ MI skills

6.2.4. Inspectors’ assessments of MI training and inspections

6.2.5. Cattle keepers’ perceptions of inspections

6.2.6. Compliance of cattle keepers

6.2.7. Statistical analysis
9.7. Appendix

9.7.1. Appendix 1: Feedback letter to pre-schools in Österåker

1. Sammanfattning

Forskningsprogrammet "Tillsynen som styrmedel för förbättrat miljöbeteende" har strukturerats kring de två huvudområdena "Verksamhetsutövares beteende" och "Tillsynsaktiviteter".


Den andra studien, "Enhancing compliance with waste sorting regulation through inspections and Motivational Interviewing", undersöker effekten av inspektioner på restaurangers lagstadgade avfallssortering och analyserar huruvida MI-utbildning av inspektörer i denna specifika kontext ökar restaurangernas benägenhet att avfallssortera. Förbättringarna i inspektörernas MI-kompetens var lika starka som i vår första studie. Dessutom fann vi också positiva effekter av inspektioner på restaurangernas lagefterlevnad. Utöver detta visade vår studie att MI-utbildning av inspektörer förstärkte den positiva effekten av inspektioner på lagefterlevnaden.

I den tredje studien, "Applying Motivational Interviewing in the context of cattle inspections", undersöker vi hur MI-utbildning av inspektörer påverkar inspektioner av djurhållare. Syftet är att utforska effekten av MI-utbildning på inspektörernas kommunikativa färddigheter och på djurhållarnas beteende. Preliminära resultat tyder på att inspektörernas MI-färddigheter inte förbättrades på en aggregerad nivå, i motsats till de andra MI-studierna. Detta resultat belyser viken av att analysera användbarheten och mätbarheten av MI i specifika inspektionssammanhang. I skrivande stund väntar vi fortfarande på de data som vi behöver för att kunna analysera eventuella förändringar i djurhållarnas beteende.

Huvudområde B (Tillsynsaktiviteter) består av en teoretisk, en empirisk och en experimentell studie. Den första studien, "The effectiveness of environmental inspections in oligopolistic markets", undersöker teoretiskt konsekvenserna av att uppnå miljölagefterlevnad genom inspektioner på oligopolistiska marknader. Lagefterlevnad är förknippad med miljövinster, men också med anpassningskostnader för verksamhetsutövarna, vilket minskar...
marknadsöverskotten. Marknadsförhållandena påverkar såväl samhällsvinsten av att uppnå lagefterlevnad som de inspektionskostnader som är nödvändiga för att avskräcka från överträdelser av lagstiftningen. På så sätt kan effekterna av olika marknadskarakteristik på kostnadseffektiviteten av inspektioner utvärderas och därigenom ge vägledning för tillsynsmyndigheter som måste prioritera mellan olika sektorer givet en fast budget. För en tillsynsmyndighet som huvudsakligen bryr sig om miljökonsekvenserna av verksamhetsutövarnas aktiviteter finns det två huvudsakliga slutsatser. För det första är kostnadseffektiviteten av inspektioner högre på marknader med lågt konkurrenstryck (med få företag eller differentierade produkter). För det andra är inspektioner kostnadseffektivare på marknader med svag efterfrågan och höga anpassningskostnader för verksamhetsutövarna när konkurrensten inte är alltför hård; under starkt konkurrenstryck är det omvända sant.

Den andra studien, ”An empirical analysis on the effects of inspections on the environmental behavior of dry cleaners in Stockholm”, använder unik paneldata och utnyttjar exogen variation i inspektionsfrekvenser för att undersöka effekten av inspektioner på kemtvättarnas perkloretylentusläpp. Vårt huvudsakliga resultat är att sannolikheten att utsläppen ligger över riktvärdet under ett givet år är 6,7 procentenheter lägre bland kemtvättar som inspekterades året innan än bland kemtvättar vars senaste inspektion ägde rum för två år sedan. Vår analys indikerar således att inspektionseffekter avtar över tid.

Den sista studien, ”Measuring the effects of feedback from inspections on cleanliness in Swedish pre-schools – an experimental approach”, genomfördes i samarbete med tre svenska kommuner. Våren 2016 mättes renligheten på toaletthandtag i förskolor i alla kommuner. Några månader senare fick endast förskolorna i en av kommunerna ett återkopplingsbrev med information om sina mätvärden, medelvärdet och medianvärdena för förskolorna i den kommunen, deras percentil (vilken visar hur många procent av förskolorna i kommunen som hade ett bättre resultat) och betydelsen av renlighet. Sex veckor senare genomfördes uppföljningsmätningar i alla kommuner. Även om vi inte kunde hitta någon effekt av feedbackbrevet på aggregerad nivå, tyder våra resultat på att det finns en positiv effekt av att ge information till förskolor med relativt dåliga renlighetsresultat (och en negativ effekt på förskolor med ett relativt bra resultat). Vi fann också att information om ett dåligt resultat ökar sannolikheten att vidta åtgärder för att förbättra renligheten, och att vidtagande av åtgärder förbättrar renligheten. Dessa resultat bekräftar konstaterandet från föregående studie att normer kan påverka miljöbeteendet.

2. Summary

The research program Inspections and Enforcement as an Instrument for Enhancing Environmental Behavior (“Tillsynen som styrmedel för förbättrat miljöbeteende”) has been structured around the two topics inspectee behavior and inspections and enforcement activities.

Topic A (inspectee behavior) consists of three studies that all focus on adapting the communicative method "Motivational Interviewing" (MI) to specific inspection contexts. The principal aim of the first study, “Applying Motivational Interviewing to induce compliance with radon gas radiation legislation – a feasibility study” is to analyze the effect of MI training on health safety inspectors' communicative skills. It is shown that the present training program led to improvements in MI skills that clearly exceeded the ones achieved by a previous MI training program, which involved inspectors working in different contexts. The results confirmed our hypothesis that applying MI to a specified inspection setting would generate stronger skills improvements.

The second study, “Enhancing compliance with waste sorting regulation through inspections and Motivational Interviewing”, investigates the effect of inspections of restaurants’ waste sorting and explores whether MI training of inspectors in this specific setting enhances the propensity of restaurants to be compliant with regulations. The improvements in inspectors’ MI skills were as strong as in our first study. Moreover, we also found positive effects of inspections on inspectees’ compliance. In addition, our study showed that MI training of inspectors reinforced the positive impact of inspections on the compliance rate among inspectees.

In the third study, “Applying Motivational Interviewing in the context of cattle inspections”, we examine how MI training of inspectors impacts on inspections of animal keepers. The aim is to explore the effect of MI training on inspectors’ communicative skills and on the behavior of the inspectees. Preliminary results indicate that inspectors’ MI skills did not improve at an aggregate level, in contrast to the other MI studies. This result highlights the importance of analyzing the applicability and measurability of MI in specific inspection settings. At the time of writing we are still awaiting the necessary data that we need in order to analyze any changes in the environmental behavior of the inspectees.

Topic B (inspections and enforcement activities) consists of a theoretical, an empirical and an experimental study. The first study, “The effectiveness of environmental inspections in oligopolistic markets”, theoretically explores the consequences of inducing compliance with environmental legislation through inspections in oligopolistic markets. Adherence to the law is associated with environmental gains, but also with abatement costs that reduce market surpluses. Market conditions affect the net social benefit of achieving compliance as well as
the inspection costs that are necessary to deter breaches of legislation. Thus, the impact of various market characteristics on the effectiveness of inspections can be assessed, thereby providing guidance for environmental inspection agencies that have to prioritize among sectors given a fixed budget. For an inspection agency mainly concerned about the environmental consequences of inspectees’ activities there are two main conclusions. First, the effectiveness of inspections is higher in markets with low competitive pressure (with few firms or differentiated products). Second, inspections are more effective in markets with weak demand and high abatement costs when competition is not too fierce; under strong competitive pressure the reverse is true.

The second study, “An empirical analysis on the effects of inspections on the environmental behavior of dry cleaners in Stockholm”, uses a unique panel data set and exploits exogenous variation in inspection frequencies to investigate the effect of inspections on dry cleaners’ perchlorethylene emissions. Our main finding is that the probability of emitting above the target limit in a given year is 6.7 percentage points lower among dry cleaners that were inspected the year before than among dry cleaners whose last inspection took place two years ago. Hence, our data suggest that inspection effects decay with time.

The final study, “Measuring the effects of feedback from inspections on cleanliness in Swedish pre-schools – an experimental approach”, was carried out in collaboration with three Swedish municipalities. In spring 2016 the cleanliness of toilet handles was measured in pre-schools in all municipalities. A few months later only the pre-schools in one of the municipalities received a feedback letter with information on their measurement values, the mean and median measurement values in that municipality, their percentile score (indicating how many per cent of other pre-schools in the municipality had a better result) and the importance of cleanliness. Six weeks later follow-up measurements were undertaken in all municipalities. Although we could not find any aggregate impact of the feedback letter, our results indicate that there is a positive effect of providing information to pre-schools with relatively bad cleanliness results (and a negative effect on pre-schools with relatively good results). We also found that information about a bad result increases the probability of taking action to improve cleanliness, and that having taken action does improve cleanliness. These results echo the finding from the previous study that norms can affect environmental behavior.
3. Introduction

3.1. Aim and scope of the research program

In June 2013 we submitted the research funding application for this program with the aim of continuing our work on the impact of environmental inspections and enforcement (EIE) on compliance and the environment. Our research team had previously participated in the research program Efficient Environmental Inspections and Enforcement (“Effektiv miljötillsyn” – EMT), which was initiated and financed by the Swedish Environmental Protection Agency (SEPA). EMT started in September 2009 and presented its final report to the SEPA in February 2013. This report (Naturvårdsverket, 2013; see also the English version, Herzing and Jacobsson, 2016) identifies a number of problems in current EIE structures and practices and suggests some concrete remedies for this such as the communicative technique ”Motivational interviewing” (MI), a more differentiated perspective on inspectees and a national prototype of an information system, intended as a supporting tool for inspectors while at the same time facilitating evaluation and analysis of EIE.

It is ultimately the impact of EIE on environmental behavior and actions that determines whether we will reach the environmental quality objectives or not. Therefore, this project has focused on the environmental incentives of firms and other inspectees that are subject to the legislation enshrined in the Environmental Code. It is only by first understanding the driving forces behind environmental behavior that we can evaluate the efficiency of different administrative instruments.

Building on the experiences from the EMT research program we have now further evaluated and refined MI. Besides analyzing inspectee incentives and behavior the focus has also been on EIE activities. In order to achieve and maintain a high degree of efficiency of EIE there is a need for evaluation and analysis. Hitherto, the lack of EIE data has made this difficult. However, EMT begun collecting such data and the current research program has now expanded the EMT database and conducted empirical and theoretical analysis with the ambition of identifying relevant measures of EIE outcomes.

The research program consists of six projects, structured around the two topics inspectee behavior and EIE activities.

Topic A (inspectee behavior) included the following studies:

- Applying motivational interviewing to induce compliance with radon gas radiation legislation – a feasibility study.
- Enhancing compliance with waste sorting regulations through inspections and motivational interviewing.
Topic B (inspections and enforcement activities) consisted of three studies:

- The effectiveness of environmental inspections in oligopolistic markets.
- Measuring the effects of feedback from inspections on cleanliness in Swedish pre-schools – an experimental approach.

In what follows these two topics and the studies that have been carried out are briefly described.

3.2. Topic A: Inspectee behavior – applying Motivational Interviewing to inspections and enforcement

We designed three studies that focus on the adaptation of MI to inspections. Two studies (A1 and A2) were carried out in cooperation with the environmental departments of the five municipalities of the regional collaboration Miljösamverkan Östra Skaraborg (MÖS), and one study (A3) was conducted in collaboration with the Swedish Board of Agriculture and four county administrative boards (see chapters 4-6 for details). In comparison to the EMT MI training program, the current studies targeted more specific inspection settings.

Our aim was both to analyze the impact of the training program on inspectors’ MI skills and to evaluate how inspectees’ behavior is affected by inspections and MI training of inspectors. We expected that inspectors would improve their MI skills to a higher degree than under the EMT training program which was devised for a broad range of inspectors; by focusing MI training on specific types of inspections a higher level of MI skills among participants should be attained. Moreover, we intended to investigate whether MI training of inspectors contributes to changes in inspectees’ behavior. To our knowledge, no study measuring the impact of MI on outcomes has hitherto been carried out in the context of inspections and enforcement.

3.2.1. Motivational Interviewing

MI can briefly be defined as ‘a collaborative conversation style for strengthening a person’s own motivation and commitment to change’ (Miller & Rollnick, 2013). MI is based on the assumption that people prefer to take their own decisions regarding matters that affect them and that they may take offence when their choices are questioned. MI counsellors are trained to interact in an empathic and collaborative manner with clients. Information should be integrated in and adapted to the dialogue with the client to make it more likely to be accepted and understood. MI also requires counsellors to evoke a client’s change talk, i.e.
expressing beliefs in the ability to change undesirable behaviors. Several studies
(Gaume et al., 2013; Lindqvist et al., 2017) have shown correlations between
clients’ change talk (e.g. expressing reasons in favor of or against change) during
conversations and the realization of actual behavior. Other studies, e.g. the meta-
analysis by Apodaca & Longabaugh (2009), have demonstrated that talk about
being unable to change undesirable behaviors is correlated to maintaining these
behaviors.

MI has been widely used for treating health behavior problems and has a solid
research base with more than 200 randomized controlled studies, which have
mainly shown significant low to moderate effects with respect to, e.g. reducing or
stopping problem drinking (Lundahl et al., 2010), stopping the use of illegal drugs
and tobacco (Lundahl et al., 2013), and completing a treatment program (Hette
et al., 2005). The positive effects of MI have contributed to its dissemination within
health care, but rarely beyond this context, with a few exceptions. For example,
Thevos et al. (2000) demonstrate how MI can enhance the adoption of clean
drinking water practices. Moreover, Klonek & Kauffeld (2012) show that MI can
be used in conversations with people about their environmental behavior and that
MI increases pro-environmental verbal behaviors compared to controls.

In a previous research program, Efficient Environmental Inspections and
Enforcement (“Effektiv miljötillsyn” – EMT), our research team carried out a MI
training program aimed at environmental, food safety and health safety inspectors
(Forsberg et al., 2014 and 2016). To our knowledge, the EMT research program
was the first to adapt MI to an inspection setting.

In the present research studies we have further developed the MI training program
to explore the applicability of MI in specific inspection contexts. We expected MI
to be useful in evoking reasons in favor of responsible behavior, e.g. a desire to
preserve the environment for one’s children and grandchildren, which might
counteract the incentives to minimize short-term consequences (e.g. economic
costs). Our hypothesis was that applying MI to more specific settings would
generate stronger improvements in MI skills than the EMT MI training program
which focused on inspections in a more general context.

3.2.2. Measuring MI skills
In our MI-studies (studies A1-A3) each inspector recorded a number of inspection
conversations both before and after the training program with the informed consent
from the inspectees. For the recording a dictaphone was used. The MI skills of the
inspectors were assessed through coding of the recorded inspection calls.
We employed a pre-post design based on the assessments of all inspectors’ MI
skills before and after the training program.

The coding of conversations was conducted by MIC Lab at the Karolinska
Institutet in Stockholm in accordance with the Swedish translation of the
Motivational Interviewing Treatment Integrity Code (MITI) 3.1 manual (Moyers et al., 2009). The conversations were encrypted during the uploading to the MIC Lab homepage and registered in a database at a protected server. When uploaded the interviews were anonymized and also received a serial number before being archived for ten years. The coders at MIC Lab did not know the identities of neither the inspectors nor the inspectees. MITI is commonly used to measure MI skills and has been found to be reliable (Madson & Campbell, 2006; Forsberg et al., 2008). To sustain their competence the coders participate in a quality assurance program twice a year; if necessary they receive extra training or are excluded.

<table>
<thead>
<tr>
<th>GLOBAL ASSESSMENTS (SCALE 1-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction towards change goals – <em>direct the conversation so that it is about the target behavior</em></td>
</tr>
<tr>
<td>Empathy – <em>show an active interest in trying to understand another person and actively communicate this understanding</em></td>
</tr>
<tr>
<td>Collaboration – <em>seek collaboration (dialogue) rather than the advisor solely presenting something (monologue)</em></td>
</tr>
<tr>
<td>Autonomy – <em>assume the capacity of people to decide about their own affairs</em></td>
</tr>
<tr>
<td>Evocation – <em>elicit the individual’s own reasons for change and confidence in being able to implement a change</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FREQUENCY COUNTS OF BEHAVIOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information comments – <em>give information, teach, give feedback, give personal information, express an opinion</em></td>
</tr>
<tr>
<td>Open questions – <em>ask questions that allow a range of possible responses</em></td>
</tr>
<tr>
<td>Closed questions – <em>ask questions that can be answered with “yes” or “no”</em></td>
</tr>
<tr>
<td>Simple reflections – <em>reflect what the individual has said without adding any new meaning</em></td>
</tr>
<tr>
<td>Complex reflections – <em>reflect what the individual is assumed to know, feel, think or experience</em></td>
</tr>
<tr>
<td>MI-adherent utterances – <em>e.g., ask for permission to advise, affirm, emphasize the individual’s control, support.</em></td>
</tr>
<tr>
<td>MI-non-adherent utterances – <em>e.g., give advice without permission, confront, direct, warn.</em></td>
</tr>
</tbody>
</table>

Table 3.1. The MITI 3.1 variables used in the assessment of MI skills

The MITI 3.1 coding manual contains frequency counts of seven verbal behaviors, as well as assessments of five global variables on a Likert scale ranging from 1 (“low”) to 5 (“high”) based on 20 minutes of the conversation (see table 3.1). The verbal behaviors that are counted are: giving information, MI-adherent and MI-non-adherent utterances, questions (closed and open), and reflections (simple and complex). Global assessments of the following variables are made: direction, empathy, and MI spirit (comprising the three sub-variables evocation, collaboration and autonomy support) (see table 5 for more detailed information). The coders may also provide written comments explaining the numbers in the protocol in order to...
facilitate the understanding of feedback and to support the training. Such comments were used for recordings made during the training period. The feedback focused on earlier and current themes.

In our studies inspectors’ MI skills have been defined by four measures derived from the ten MITI 3.1 variables and with hypothesized relationships to future behavior change. The four selected measures are empathy, evocation, MI-non-adherent utterances and the ratio of reflections to questions. Client behavior change has been positively related to empathy (Elliott et al., 2011, and Moyers & Miller, 2013), to evocation (Magill et al., 2014) and to the ratio of reflections to question (Miller & Rose, 2009), while being negatively influenced by the frequency of MI-non-adherent utterances (Magill et al., 2014; Apodaca & Longabaugh, 2009).

3.2.3. The MI training programs

In the MI-studies the inspectors received 30-36 hours of training in MI during five or six days spread out over a 3-6 months period. Every training day focused on a specific theme in accordance with common MI practice (Miller and Moyers, 2007; Miller and Rollnick, 2013) and building on our experiences from the EMT MI training program (Forsberg et al., 2014 and 2016). In addition, the training program also included a brief introduction to fundamental reinforcement principles within psychological learning theory (Sundel and Sundel, 2005).

The training days consisted of short theoretical introductions mixed with experiential exercises of MI skills. From the second to the last training days three hours in the afternoon were devoted to supervision in group format based on recorded professional conversations. The training program was implemented in collaboration with the coding laboratory MIC Lab (see section 3.2.2). Before some of the training days each inspector recorded an inspection that was uploaded to MIC Lab’s homepage to be coded for use as feedback and not to be included in the study. The coders provided written comments explaining the numbers in the protocol to facilitate the understanding of the feedback, which focused on earlier and current themes. In two of the studies each inspector made one verbatim transcription of a recorded inspection call that was used as a pedagogical tool with the purpose of exemplifying MI concepts and MI skill applications. Sections 4.2.2, 5.2.2 and 6.2.2 describe the design of the three training programs in detail.

To evaluate the inspectors’ assessments of the MI training program and its relevance for their practical work the same questionnaire as in the EMT study was used. After every training day the inspectors were asked to rate the usefulness of feedback on recorded inspection calls, of theory reviews and exercises, and of the training program as a whole, as well as the applicability of new MI skills in inspection work. For this purpose a six-graded scale ranging from 1 (“not useful at all”) to 6 (“very useful”) was employed. The questionnaire also contained four open questions regarding what was good and what was bad about the training day,
what might be useful in the inspector’s work, and what advice the inspector would give for future MI training programs. However, answers to the open questions will not be reported for any of our studies, because they do not provide any quantitative measures and were primarily intended as feedback to the MI instructor.

3.2.4. Assessments of inspections and inspection conversations

In all three studies inspectors were asked to assess their conversations with inspectees in a questionnaire immediately after every inspection, for which a permission for recording had been given and the duration was at least ten minutes (the minimal required duration for reliable coding). It consisted of seven statements regarding the inspectee’s attitude and the inspector’s performance, which were rated on a five-graded scale from 1 (“absolutely disagree”) to 5 (“absolutely agree”). The questionnaire had previously also been used during the EMT MI training program, but was adapted to the specific context of each study.

We also obtained assessments regarding inspections from the inspectees in studies A2 and A3 (see sections 5.2.3 and 6.2.3). Inspectees were asked to answer a questionnaire containing questions on the inspector’s performance on a four-graded scale from 1 (“absolutely disagree”) to 4 (“absolutely agree”). The questionnaire also included a question concerning the inspectee’s intentions to comply with the demands of the EIE agency.

3.2.5. Evaluating inspection effects

A further aim beyond the EMT MI study was to actually test if and how inspection conversations have any effects on inspectee behavior. For this purpose, we needed measures of environmental outcomes that could be monitored before and after experimental interventions. To be able to capture effects of inspections, these need to be similar (i.e. focusing on the same type of industry/activity), recurrent (to generate time series data) and objectively measurable (e.g. compliant/not compliant).

The evaluation of the impact of inspections and MI-training differed across the MI studies. Study A1 focused on property owners’ rate of compliance with a request for reporting results from radon gas radiation measurements (see section 4). In study A2 changes in restaurants’ environmental behavior were recorded at subsequent inspections (see section 5). Finally, study A3 has been designed to measure the extent to which cattle keepers after an inspection undertake necessary measures to achieve compliance (see section 6).

3.2.6. Aim and scope of studies A1-A3

In our first topic-A-study, “Applying motivational interviewing to induce compliance with radon gas radiation legislation – a feasibility study”, we report on how MI is adapted to the context of health safety inspections. A MI training program was devised for four inspectors who were involved in a MÖS project
intended to induce property owners to measure radon gas radiation and to submit measurement results. Our principal aim was to analyze the effect of MI training on inspectors' communicative skills and to discuss the outcome in relation to the EMT study that involved training of inspectors in a broader context. A further aim was to explore whether it is possible to establish an effect of MI-trained inspectors on property owners' compliance with the request for reporting radon gas radiation measurement results.

The second study, "Enhancing compliance with waste sorting regulations through inspections and Motivational interviewing", involved seven MÖS inspectors who at routine food safety inspections of restaurants also started checking waste sorting, which had never been done before. Our aim was to investigate the effect of inspections on restaurants’ waste sorting and to explore if and how MI training of inspectors in this specific setting is useful and enhances the propensity of restaurants to comply with regulations.

The third study, “Applying Motivational Interviewing in the context of cattle inspections”, was planned and executed in cooperation with the Swedish Board of Agriculture and four county administrative boards. The aim was to explore the effect of MI training on inspectors’ communicative skills and on the behavior of the cattle keepers (i.e. whether they have undertaken measures that they have been notified about during an inspection).

In all three studies inspectors have answered questionnaires to assess the MI training program as well as questionnaires regarding their perceptions of inspection conversations. Studies A2 and A3 also include results from questionnaires concerning inspectees’ assessments of inspection conversations.

3.2.7. Main results from studies A1-A3

Inspectors' MI skills

In studies A1 (radon gas radiation measurements by property owners) and A2 (restaurants’ waste sorting) the MI training programs resulted in improvements that clearly exceeded the ones achieved by the EMT training program with regard to the four MI skills summary measures empathy, evocation, MI-non-adherent utterances and the ratio of reflections to questions. The results confirm our hypothesis that applying MI to a specified inspection setting would generate stronger skills improvements. However, we could not find any improvements of the inspectors’ MI skills on an aggregate level in study A3 (cattle keepers). This result highlights the importance of analyzing the specific inspection setting and the potential for MI to be a helpful tool. Moreover, as inspections in study A3 took place in a more complex context with a broader range of target behaviors, coding was more difficult than in the other studies, i.e. there are also methodological issues that need to be addressed when designing a MI training program for a specific setting (see section 6.4 for a discussion).
The table below summarizes the MI skills results for our three studies and the EMT study at the aggregate level, before and after the MI training programs. Significant skills improvements are in bold face.

<table>
<thead>
<tr>
<th>MI SKILLS</th>
<th>EMT</th>
<th>A1 (Radon)</th>
<th>A2 (Waste sorting)</th>
<th>A3 (Cattle)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
</tr>
<tr>
<td>Number of observations</td>
<td>29</td>
<td>19</td>
<td>46</td>
<td>40</td>
</tr>
<tr>
<td>Empathy (1=&quot;low&quot; – 5=&quot;high&quot;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>1.36</td>
<td>1.79</td>
<td>1.78</td>
<td>3.35</td>
<td>2.17</td>
</tr>
<tr>
<td>(0.43)</td>
<td>(0.64)</td>
<td>(0.53)</td>
<td>(0.66)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>p-values</td>
<td>0.026</td>
<td>0.000</td>
<td>0.000</td>
<td>0.819</td>
</tr>
<tr>
<td>Evocation (1=&quot;low&quot; – 5=&quot;high&quot;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>1.28</td>
<td>1.83</td>
<td>1.65</td>
<td>3.03</td>
<td>2.14</td>
</tr>
<tr>
<td>(0.44)</td>
<td>(0.62)</td>
<td>(0.53)</td>
<td>(0.65)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>p-values</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.676</td>
</tr>
<tr>
<td>MI-non-adh. utterances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1=&quot;low&quot; – 5=&quot;high&quot;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>3.78</td>
<td>2.01</td>
<td>7.28</td>
<td>0.59</td>
<td>1.95</td>
</tr>
<tr>
<td>(3.40)</td>
<td>(1.40)</td>
<td>(3.86)</td>
<td>(0.99)</td>
<td>(2.56)</td>
</tr>
<tr>
<td>p-values</td>
<td>0.023</td>
<td>0.000</td>
<td>0.007</td>
<td>0.555</td>
</tr>
<tr>
<td>Ratio of refl. to questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1=&quot;low&quot; – 5=&quot;high&quot;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>0.49</td>
<td>0.41</td>
<td>0.73</td>
<td>1.14</td>
<td>0.76</td>
</tr>
<tr>
<td>(0.27)</td>
<td>(0.23)</td>
<td>(0.52)</td>
<td>(0.65)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>p-values</td>
<td>0.268</td>
<td>0.002</td>
<td>0.002</td>
<td>0.259</td>
</tr>
</tbody>
</table>

Table 3.2. Aggregate MI skill measures, means (standard deviations) and p-values, before (pre) and after (post) MI training

Inspectors’ assessments of the MI training program

In all studies the inspectors’ ratings of the MI training days and the applicability of MI in inspections work were higher than the ones of the EMT training program. However, the assessments by inspectors in study A3 were not as high as in studies A1 and A2. See table 3.3 below.

<table>
<thead>
<tr>
<th>MI TRAINING DAY QUESTIONNAIRES</th>
<th>EMT</th>
<th>A1 (Radon)</th>
<th>A2 (Waste sort.)</th>
<th>A3 (Cattle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of training day inspectors</td>
<td>230</td>
<td>19</td>
<td>32</td>
<td>56</td>
</tr>
<tr>
<td>Number of completed questionnaires</td>
<td>194</td>
<td>16</td>
<td>32</td>
<td>56</td>
</tr>
<tr>
<td>1. How would you rate the usefulness of feedback on recorded conversations (afternoon)?</td>
<td>4.86</td>
<td>5.79</td>
<td>5.76</td>
<td>5.32</td>
</tr>
<tr>
<td>2. How would you rate the usefulness of theory reviews and exercises (morning)?</td>
<td>4.80</td>
<td>5.53</td>
<td>5.85</td>
<td>5.22</td>
</tr>
<tr>
<td>3. How would you rate the MI training as a whole?</td>
<td>4.94</td>
<td>5.53</td>
<td>5.89</td>
<td>5.24</td>
</tr>
<tr>
<td>4. Will you be able to apply your new MI-skills in your inspection work?</td>
<td>4.74</td>
<td>5.02</td>
<td>5.02</td>
<td>4.85</td>
</tr>
</tbody>
</table>

Table 3.3. Inspectors’ responses to questions about the MI training days, rated on a six-graded scale (1="not useful at all" – 6="very useful")
Inspectors’ assessments of inspection conversations

Table 3.4 below provides a summary for our three studies, where inspectee refers to property owners, restaurant representatives and cattle keepers, respectively. Significant differences in assessments before (pre) and after (post) the MI training program are in bold face.

<table>
<thead>
<tr>
<th>INSPECTORS’ PERCEPTIONS OF INSPECTION CONVERSATIONS</th>
<th>A1 (Radon)</th>
<th>A2 (Waste sort.)</th>
<th>A3 (Cattle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of submitted questionnaires</td>
<td>49</td>
<td>40</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>203</td>
<td>259</td>
</tr>
<tr>
<td></td>
<td>209</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. My perception at the start of the inspection (call) was that the inspectee had a positive attitude to discussing the measurement of radon gas radiation / discussing waste sorting / the inspection.</td>
<td>4.24</td>
<td>3.83</td>
<td>4.04</td>
</tr>
<tr>
<td></td>
<td>4.33</td>
<td>3.55</td>
<td>3.41</td>
</tr>
<tr>
<td></td>
<td>p = 0.266</td>
<td>p = 0.057</td>
<td>p = 0.090</td>
</tr>
<tr>
<td>2. My perception at the end of the inspection call was that the inspectee had sufficient knowledge of the effects of radon gas radiation / the effects of waste sorting on people’s health / the inspection’s importance for promoting good animal husbandry.</td>
<td>3.16</td>
<td>3.65</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td>4.10</td>
<td>4.08</td>
<td>4.08</td>
</tr>
<tr>
<td></td>
<td>p = 0.005</td>
<td>p = 0.010</td>
<td>p = 0.997</td>
</tr>
<tr>
<td>3. My perception is that the inspectee during the inspection (call) clearly showed that he/she understood the information I wanted to convey.</td>
<td>4.00</td>
<td>3.90</td>
<td>4.04</td>
</tr>
<tr>
<td></td>
<td>4.47</td>
<td>4.07</td>
<td>4.01</td>
</tr>
<tr>
<td></td>
<td>p = 0.070</td>
<td>p = 0.002</td>
<td>p = 0.462</td>
</tr>
<tr>
<td>4. My perception is that the inspectee will undertake a measurement of radon gas radiation before May 31st, 2015 / have a properly functioning waste sorting system at the next inspection / undertake measures to achieve compliance.</td>
<td>4.53</td>
<td>4.40</td>
<td>4.30</td>
</tr>
<tr>
<td></td>
<td>4.04</td>
<td>4.20</td>
<td>4.16</td>
</tr>
<tr>
<td></td>
<td>p = 0.080</td>
<td>p = 0.006</td>
<td>p = 0.781</td>
</tr>
<tr>
<td>5. My perception is that the inspectee expressed his/her own reasons and motivation to conduct a measurement of radon gas radiation / sort waste / undertake measures to achieve compliance.</td>
<td>3.16</td>
<td>3.45</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td>4.20</td>
<td>3.42</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>p = 0.902</td>
<td>p = 0.000</td>
<td>p = 0.152</td>
</tr>
<tr>
<td>6. I am satisfied with my own performance during the inspection (call).</td>
<td>4.04</td>
<td>3.64</td>
<td>3.69</td>
</tr>
<tr>
<td></td>
<td>3.93</td>
<td>3.78</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td>p = 0.011</td>
<td>p = 0.128</td>
<td>p = 0.894</td>
</tr>
<tr>
<td>7. My perception is that the inspectee at the end of the inspection (call) had a positive attitude to radon gas radiation measurement / waste sorting / the inspection.</td>
<td>4.47</td>
<td>4.05</td>
<td>4.25</td>
</tr>
<tr>
<td></td>
<td>4.60</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>p = 0.406</td>
<td>p = 0.027</td>
<td>p = 1.000</td>
</tr>
</tbody>
</table>

Table 3.4. Inspectors’ responses to questions about inspection conversations, rated on a five-graded scale (1=“absolutely disagree” – 5=“absolutely agree”)

Note that the number of answers to questions 4 and 5 are lower, because these questions were not relevant at all inspections in study A3 (the number of answers is reported in brackets).
MI training did not lead to any changes with regard to how inspectors assessed their own performance during inspection conversations in studies A1 and A3. In study A2, however, our results indicate an increase in inspectors’ rating on all seven measures (all but one significant).

**Inspectees’ assessments of inspection conversations**

Here we only have results from studies A2 and A3 (the response rate was far too low in study A1 to allow us to draw any conclusions). In both studies ratings by inspectees are very high and stable over time. Hence, MI training of inspectors did not lead to any changes in how inspectees perceived inspections. Table 3.5 provides summary results for questions that were used in both studies before (pre) and after (post) the MI training program; the questionnaires differed slightly and contained some questions that were not asked in the other study (see sections 5.3.4 and 6.3.4 for complete results).

<table>
<thead>
<tr>
<th>INSPECTEES’ PERCEPTIONS OF INSPECTIONS</th>
<th>A2 (Waste sort.)</th>
<th>A3 (Cattle)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre</td>
<td>post</td>
</tr>
<tr>
<td>Number of submitted questionnaires</td>
<td>39</td>
<td>142</td>
</tr>
<tr>
<td>How satisfied are you with the inspector’s attitude?</td>
<td>3.90</td>
<td>3.86</td>
</tr>
<tr>
<td>How satisfied are you with the inspector’s competence?</td>
<td>3.80</td>
<td>3.86</td>
</tr>
<tr>
<td>How satisfied are you with the inspector’s efficiency?</td>
<td>3.82</td>
<td>3.78</td>
</tr>
<tr>
<td>How satisfied are you with the information that you received?</td>
<td>3.90</td>
<td>3.88</td>
</tr>
</tbody>
</table>

Table 3.5. Inspectees’ responses to questions about inspection conversations, rated on a four-graded scale (1=“not satisfied” – 4=“satisfied”)

**Inspectees’ compliance rate**

We have results regarding the effects of inspections and inspectors’ MI skills on inspectees’ behavior for studies A1 and A2 only, as we are still awaiting the necessary data to analyze changes in compliance among cattle keepers in study A3.

In study A1 we do not find any effects of MI training on the inspectees’ compliance rate. Among property owners contacted before inspectors participated in the MI training program 75 per cent complied with the request to report radon gas radiation measurements. The corresponding rate among property owners contacted after the inspectors had been MI-trained was 69 per cent, i.e. it was actually lower, although insignificantly. In section 4.4 we thoroughly discuss possible explanations for this result.

In study A2, however, we observe strong and significant effects of inspections on restaurants’ rate of compliance with waste sorting regulations. Moreover, it is also shown that MI training of inspectors reinforces the positive impact of inspections on the compliance rate among restaurants (see figure 3.1 below).
3.3. Topic B: Inspections and enforcement activities – explaining and measuring the efficiency and effectiveness of inspections

We have carried out three studies that focus on the efficiency and the effectiveness of inspections. To that end we have used three different – a theoretical, an empirical and an experimental – approaches. Theory is essential to provide a framework within which complex issues can be analyzed and understood. The empirical evaluation of data is necessary for measuring outcomes and establishing correlations and causalities. Finally, experiments enable us to test different EIE methods in a scientifically robust way. We believe that it is important to employ all three perspectives in order to obtain a comprehensive understanding of how EIE can be organized and evaluated.

3.3.1. Background

The efficiency of inspections and the effective use of available resources has received increasing attention by several international organizations in recent years. For example, the Organization for Economic Development (OECD) has established guidelines for regulatory policy practices that emphasize the importance both of "responsive regulation" of specific businesses (OECD, 2014, pp. 33-35) and of
"concentrating resources and efforts where they can deliver the most results" (OECD, 2014, p.38). These recommendations are based on resources for inspection and enforcement usually being limited and not sufficient for achieving full compliance among all inspectees. The OECD therefore advises that "in the absence of comprehensive data and/or fully reliable data, regulatory enforcement agencies should rely on interpreting what data exists (at least to establish which sectors appear to generate the most damage)” (OECD, 2014, p.28).

On the basis of these recommendations our theoretical study (B1) uses a novel approach that accounts for the fact that market conditions vary for different types of inspectees. For example, the degree of competition may vary across markets, such that non-compliance does not only lower a firm's costs, but also impacts on the strategic interaction with its competitors. Given that resources are insufficient for deterring non-compliance in all industries, the EIE agency has to determine which type of activities to monitor. The aim of this study is to draw conclusions regarding the effectiveness of inspections across markets with different characteristics.

Turning to the question on how to measure the effects of EIE on environmental behaviour it is unfortunately difficult to get access to or to produce relevant data for analysis. A crucial ingredient for successful analysis of EIE effects is the ability to link EIE activity data to relevant environmental outcome data, something that was discussed in the final report of the EMT research program (Herzing and Jacobsson, 2016). Another challenge is to establish causality. For example, inspections are often targeted on high-risk inspectees with bad environmental track records, thus generating a causality running from bad environmental outcomes to EIE activities (Gray & Shimshack, 2011). Hence, even if research successfully links EIE activity data to environmental outcome data, using these secondary data requires a careful analysis that ensures proper identification.

In the current program we have used two main approaches to overcome these difficulties. On the one hand, we have carefully investigated a municipal environmental department’s inspection records (study B2). On the other hand, we have conducted experiments with treatment and control groups (study B3 and also studies A1-A3).

An advantage of the experimental approach is the high degree of control of the data generating and identification process, that is, we can be relatively sure that what we want to measure is actually captured. However, these advantages come at a cost. In order to be successfully implemented experiments usually require the full cooperation of EIE agencies. Since experiments can be difficult to integrate with day to day EIE work, the challenge is to find a design that does not burden inspectors too much.
Use of secondary data is not as demanding for EIE agencies, but usually requires a lot of investigative work by researchers to compensate for the loss of control over the data generating processes. These difficulties could be overcome if EIE agencies adopted administrative systems that are better suited for data analysis than they are today. The latter point was discussed in the EMT report where a prototype of such an information system was presented.

3.3.2. Summary of study B1

The first topic-B-study, “The effectiveness of environmental inspections in oligopolistic markets”, employs a theoretical model to explore the consequences of inducing compliance with environmental legislation through inspections in oligopolistic markets. Adherence to the law is associated with improvements in environmental quality, but also with abatement costs that reduce market surpluses. The relative weighting of environmental gains and market surplus losses gives us a measure for the net social benefit of deterring breaches of legislation, which can be related to the inspection costs for achieving compliance. These inspection costs in turn depend on the gains from violating legislation; the larger the increase in profits from being non-compliant, the more frequent inspections need to be to deter breaches of the law.

Market conditions affect the net social benefit of achieving compliance as well as the inspection costs that are necessary to deter violations of legislation and hence, impact on the effectiveness of inspections, measured as the ratio of net social benefits to costs of inspections. Thus, the impact of various market characteristics on the effectiveness of inspections can be assessed, thereby providing guidance for EIE agencies that have to prioritize among sectors given a fixed budget.

The model generates three results. First, the effectiveness of inspections is largest in sectors with low competitive pressure (high degrees of market concentration and product differentiation) when little weight is attributed to market surplus losses. Second, the effectiveness of inspections increases in abatement costs and decreases in demand, except when competition is fierce and the relative weight of market surplus losses is small (in which case the reverse is true). Finally, the inspections are most effective at increasingly higher degrees of competitive pressure as more weight is attached to market surplus losses.

Hence, for an EIE agency mainly concerned with the environmental consequences of inspectees’ activities there are two main conclusions. The effectiveness of inspections is higher in markets with low competitive pressure, i.e. with few firms or differentiated products. And inspections are more effective in markets with weak demand and high abatement costs when competition is not too fierce; under strong competitive pressure the effectiveness increases in demand and decreases in abatement costs.
3.3.3. Summary of study B2

The second study, “An empirical analysis on the effects of inspections on the environmental behavior of dry cleaners in Stockholm”, uses a unique panel data set covering all dry cleaners in Stockholm municipality in the years 2000-2013. Every year these have to submit a report containing information on the amounts of dry cleaned laundry and the toxic substance perchloroethylene, which is used for dry cleaning in most machines. Dry cleaners are by law required to keep their emissions of perchloroethylene at less than 2 per cent of dry cleaned laundry weight; in Stockholm, the environmental department lowered this threshold to 1.5 per cent in 2010. By using data on self-reported perchloroethylene emissions and inspection dates we were able to investigate how inspections affect emission levels. Our analysis focused on three outcomes. First, we measured the impact of inspections at the extensive margin, i.e. whether emissions exceeded the threshold level. Second, we also analyzed effects of inspections at the intensive margin, that is, absolute emission levels. Finally, we examined whether reporting on time – before January 31st – was influenced by inspections.

Our main finding is that the probability of emitting above the target limit in a given year is 6.7 percentage points lower among dry cleaners that were inspected the year before than among dry cleaners whose last inspection took place two years ago. Hence, we established an extensive margin effect suggesting that the impact of inspections decay over time.

However, we did not observe any effect at the intensive margin. Thus, emissions were not reduced following an inspection the year before. This could be explained by the fact that, on average, 84 per cent of dry cleaners emitted less than the prescribed threshold levels. Our results seem to suggest that inspections only have an effect on dry cleaners that are non-compliant, whereas adherent dry cleaners have weak incentives to lower emissions further.

On average, 54 per cent of dry cleaners handed in the annual reports too late. Here, we could also not establish any effect of inspections. Since failing to submit reports on time does not lead to any consequences beyond being sent a reminding letter, the incentives for complying with the deadline are weak; an inspection in the previous year does not affect that.

3.3.4. Summary of study B3

The final study, “Measuring the effects of feedback from inspections on cleanliness in Swedish pre-schools – an experimental approach”, was carried out in collaboration with three Swedish municipalities. By swabbing the handles on the inside of the pre-school toilets and then using an ATP meter to obtain a RLU (relative light units, a measure of bioluminescence) number, a measure of cleanliness was obtained during a first inspection of pre-schools in all three municipalities in spring 2016. Some months later only the pre-schools in one of the municipalities received a feedback letter containing their RLU values, the mean
and median RLU values in that municipality, their percentile score (indicating how many per cent of other pre-schools in the municipality had a better result) and information regarding the importance of cleanliness. Six weeks later follow-up RLU measurements were undertaken in all municipalities.

Employing this experimental approach enabled us to evaluate whether the provision of feedback regarding cleanliness in pre-schools of one municipality led to changes. For reasons of consistency we had to exclude one of the municipalities where no feedback letters were sent. Hence, our results are based on comparisons of pre-schools where feedback was provided after the first inspection (i.e. in the treatment municipality) and of pre-schools in one of the municipalities where no feedback was given (i.e. the control municipality).

At the aggregate level we could find no effect of the feedback letter; in fact, on average, cleanliness actually deteriorated among pre-schools that were informed about their results after the first inspection. However, by comparing individual pre-schools’ RLU values at both inspections different patterns emerge in the two municipalities. In the control municipality we observe persistence in measurements at inspections 1 and 2, whereas a slight ‘boomerang effect’ can be detected among pre-schools in the treatment municipality. Hence, our results indicate that there is a positive effect of providing information to pre-schools with relatively bad cleanliness results (and a negative effect on pre-schools with relatively good results).

By conducting follow-up interviews with pre-schools in the treatment municipality, we found that receiving a high percentile score (i.e. a bad result) increases the probability of taking action to improve cleanliness. Our data reveals that having taken action, on average, leads to improvements in cleanliness. These results echo the finding from the previous study that norms can affect environmental behavior.

3.4. Concluding remarks

What have we learned so far?
The current research program has continued the work of EMT in important ways. One direction has been to further evaluate if and how MI can contribute to a more efficient EIE. Projects A1 and A2 have both shown that inspectors can improve MI skills more if the training program is focused on a specific EIE context as opposed to a more general one, as was the case in the EMT MI study. Moreover, compared to the EMT MI training program, inspectors in all three MI studies, particularly in studies A1 and A2, rated the MI training days and the relevance of MI skills for inspection work higher. These findings have clear implications for how MI training programs for inspectors should be designed.
In our studies A2, B2 and B3 we can find significant effects from inspections on the environmental behavior of inspectees, which underscores the importance of this instrument in enforcing the environmental code. Further, in study A2 we find that, on average, inspectees that were inspected by an MI-trained inspector improved their environmental behavior more than inspectees that were inspected by inspectors who were not MI-trained. This result is a novelty in the scientific literature.

In studies A2 and A3 we also measured how inspectees perceived inspections throughout the duration of the projects. In project A2 the inspectees were equally happy with their inspections before and after inspectors had received MI training. Still, their environmental behavior improved to a greater extent following inspections after compared to before the MI training program. This indicates that measuring inspectees’ satisfaction with inspections is a poor indicator of the efficacy of EIE. In fact, choosing relevant and precise indicators of the conduct and impact of EIE is crucial for evaluating its usefulness. For example, when coding inspection dialogues with the purpose of evaluating MI behavior, it is essential to have a clear definition of the targeted environmental behavior(s).

Complexity is thus an integral part of EIE. Our study B1 proposes a theoretical tool that is helpful in allocating scarce EIE resources where they can be most effective based on the characteristics of the markets where inspectees operate. Yet another reminder of EIE’s complexity is study B3, which shows that inspection feedback can be either good or bad (from an environmental perspective) depending on the specific characteristics of inspectees. It is our hope that our research program has provided a few more pieces of the large EIE puzzle.

**Future research**

The TSFM project has given us new knowledge of practical use and also many ideas for future research. Our work has so far taught us that EIE methods need to be chosen carefully with attention to the specific context. For example, in which settings is MI likely to be an efficient method to use? How should we measure to what extent MI has been applied? What skill level is sufficient for affecting inspectees’ environmental behaviors? Are there any specific components in MI that are especially efficient in promoting pro-environmental behavior? How permanent are the effects of different EIE methods, and how does MI affect the longevity of pro-environmental behavioral changes? To answer these questions, more data needs to be collected in real-life settings in cooperation with EIE agencies. This, in turn, raises questions regarding how studies should be designed for maximum effect given limited resources and the need to incorporate them into the day to day activities of these agencies. We hope to be able to continue our work on answering these (and other) questions in the future.
3.5. References


4. Applying motivational interviewing to induce compliance with radon gas radiation legislation – a feasibility study

Hans Wickström2, Mathias Herzing3, Lars Forsberg4, Adam Jacobsson5 and Håkan Källmén6

4.1. Introduction

In this study, we explore how the communicative method “Motivational Interviewing” (MI) can be adapted to the context of health safety inspections. The study was carried out by the research program Inspections and Enforcement as an Instrument for Enhancing Environmental Behavior (“Tillsynen som styrmedel för ett förbättrat miljöbeteende” – TSFM) as part of an ongoing project in five Swedish municipalities, which aims at inducing property owners to report results from radon gas radiation measurements as required by law.

Radon gas radiation constitutes a health risk and is the second most common cause of lung cancer (after smoking) in Sweden (Strålsäkerhetsmyndigheten, 2016). Therefore, guidelines of the Public Health Agency of Sweden prescribe radiation from radon gas in apartments to be below 200 Bq/m³ no later than 2020 (Folkhälsomyndigheten, 2014). As a consequence, collaboration between the Swedish municipalities Falköping, Hjo, Skövde and Tibro (“Miljösamverkan Östra Skaraborg” – MÖS), initiated a project in 2011 with the aim of establishing the levels of radon gas radiation in rented apartments. All landlords within the region were requested by mail to measure and to report the level of radon gas radiation in their apartments until May 31st, 2012. Among the almost 450 landlords that were contacted and together own more than 1600 blocks of rented apartments, only a third had submitted measurement results by September 30th, 2012 (MÖS, 2012). When also including landlords which had submitted results by May 31st, 2013, 36% of blocks reported radiation levels above 200 Bq/m³ (MÖS, 2013).

The project was extended in 2014 to include co-operative apartments in all municipalities and also to co-operative and rented apartments in Karlsborg (which

2 Meetme Psykologkonsult AB, Gothenburg, email: hans.wickstrom@kbtanalys.se
3 Department of Economics, Stockholm University, email: Mathias.Herzing@ne.su.se
4 MicLab AB, Stockholm, email: lars.forsberg@miclab.se
5 Department of Economics, Stockholm University, email: aja@ne.su.se
6 Karolinska Institutet, Department of Clinical Neuroscience, Center for Psychiatry Research, email: hakan.kallmen@ki.se
joined MÖS in 2013). In contrast to the first part of the project, the health safety inspectors from MÖS now followed up the mail requests with inspection calls to the property owners (landlords in the case of rented apartments or representatives of the housing associations in the case of co-operative apartments7) with the aim of improving the rate of submitted measurement results. Naturally this modified set-up of the project made the communicative skills of the inspectors an important factor for its success. To examine how these skills could be improved, our research team became involved and a research study was designed. The four health safety inspectors involved in this MÖS project were trained in MI, an evidence-based method to affect behavior by means of communication.

The principal aim of this study is to investigate the feasibility of using MI in health safety inspection calls to property owners who are required by law to report results from radon gas radiation measurements. A further aim is to explore whether it is possible to establish an effect of MI-trained inspectors on property owners’ compliance with the request for reporting radon gas radiation measurement results. To our knowledge, no study measuring the impact of MI on outcomes has hitherto been carried out in the context of inspections and enforcement.

Based on our results, we also discuss this study’s specialized MI training of health safety inspectors in light of the EMT training program, which covered a wide range of inspection settings. Since this program was devised for heterogeneous groups of inspectors working in different contexts, there was a need for generalizations that made it difficult to convey the applicability of MI. Nevertheless, results from the EMT study showed that after having completed an MI training program, inspectors significantly improved their skills in conversations with inspectees regarding two of the most important MI variables, showing empathy and avoiding MI-non-adherent utterances (Forsberg et al., 2014 and 2016). The inspectors also experienced MI to be of benefit for inspections and became more satisfied with their own performance at inspections. In contrast to the EMT MI study, the present training program was adapted to a clearly defined setting, allowing the participants to be focused on the communicative requirements in this specific context. Hence, we expect inspectors to attain a higher increase in MI competence than in the previous training program.

The study was approved by the Regional Ethical Review Board in Stockholm (reg nr 2014/1417-32). Results regarding property owners’ compliance are presented on an aggregate level, while results of individual inspectors have been de-identified. The presentation of this study is organized as follows. Section 4.2 presents the methodological approach. Section 4.3 contains the results of the study, and in section 4.4 a discussion of results and possible future research are presented.

7 All co-operative apartments belong to one housing association whose responsibility it is to conduct their members’ radon gas radiation measurements.
4.2. Method

4.2.1. Research design

The present radon gas radiation measurement project involved four health safety inspectors making inspection calls to a total number of 158 property owners, consisting of 118 housing associations and 40 landlords. Almost all landlords were located in Karlsborg, which had not participated in the previous MÖS project. The present study was carried out according to the schedule (see table 4.1).

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>27–28 March 2014</td>
<td>Letter sent to 158 property owners, requesting radon gas radiation measurements to be undertaken</td>
</tr>
<tr>
<td>1–28 April 2014</td>
<td>Before MI training: 79 randomized property owners (the control group) receive inspection call by the four MÖS inspectors</td>
</tr>
<tr>
<td>29 Apr. – 2 Sept. 2014</td>
<td>MI-training of the four MÖS inspectors</td>
</tr>
<tr>
<td>3–28 September 2014</td>
<td>After MI training: remaining 79 property owners (the intervention group) receive inspection call by the four MÖS inspectors</td>
</tr>
<tr>
<td>1 Oct. 2014 – 30 Apr. 2015</td>
<td>Radon gas radiation measurement period</td>
</tr>
<tr>
<td>31 May 2015</td>
<td>Last day for reporting radon gas radiation measurement results</td>
</tr>
<tr>
<td>17 June 2015</td>
<td>Reminding letters sent</td>
</tr>
<tr>
<td>31 August 2015</td>
<td>Last date for MÖS to check whether new or complementary measurements have been reported after May 31st 2015</td>
</tr>
</tbody>
</table>

Table 4.1. Schedule of the study

On March, 27th and 28th, 2014, a letter was sent to all property owners, containing a request for measuring the level of radon gas radiation in their apartments and submitting the measurement results no later than May 31st, 2015. The letter provided information for how these measurements should be carried out and reported. In particular, it stated that measuring should take place for at least two months between October 1st and April 30th, because during this period windows are normally kept closed. Hence, the measurement environment is as free from drought as possible. The letter also emphasized that reported levels above 200 Bq/m³ would require the submission of a plan for undertaking efforts to achieve compliance, and that failing to report measurement results would lead to a formal request and possibly also a fine.

The letter was followed up by an inspection call from one of the four inspectors with the aim of inducing property owners to comply with their obligation to report radon gas radiation measurement levels. Since not all property owners submitted radon gas radiation measurement results by May 31st, 2015, a reminder to report
the results or to carry out the radon gas radiation measurement during the next period (October 2015 – April 2016) was sent in June 2015. On August 31st, 2015, MÖS checked whether further measurement results had been reported, thereby also including property owners who had complied with the request for carrying out radon gas radiation measurements, but had failed to report before the May 31st deadline (e.g. because they had not received the measurements from the laboratory).  

A cross-sectional design was implemented to make it possible to explore the effect of letting the four health safety inspectors acquire MI skills through a, specifically for this project, devised training program (section 2.2 provides a detailed account of the MI training program). Thus, the property owners were randomly divided into two groups of equal size. The first group was contacted by phone in April 2014 before the inspectors had received any MI training, while the second group received an inspection call in September 2014 after the inspectors had completed the MI training program. Thereby we obtained a control group consisting of the 79 property owners contacted in April 2014, and an intervention group with the remaining 79 property owners who were called after the inspectors had acquired MI skills. All inspection calls were made by the same inspectors – the control group was contacted when the four inspectors were MI-untrained, whereas the intervention group was contacted when the inspectors were MI-trained. The order in which the property owners were contacted and the assignment of inspectors to property owners was random in both the control and the intervention group. In total, 46 property owners were excluded due to the following criteria: no contact information available or no contact established; other organization responsible for measuring radon gas radiation (applies to housing associations); no blocks of apartments, no tenants, or no permanent residence (e.g. holiday homes); radon gas radiation measurement already reported and approved by MÖS; property sold or belonging to a deceased person; and additional contact of property with MÖS beyond the letter sent and the inspection call (e.g. regarding some other issue, while at the same time discussing the measurement of radon gas radiation, before the study).

The randomization of property owners took place before the exclusion criteria were applied, because exclusion was not possible before inspection calls had been made. Among the property owners in the control group 19 (24%) were excluded. The corresponding number for the intervention group was 27 (34%). Thus, in terms of evaluating the rate of compliance with the request for submitting radon gas radiation measurement results, the control and intervention groups consisted of 60 and 52 property owners, respectively. Table 2 shows the distribution of the two types of property owners across municipalities. The exclusion rate was much

---

8 After August 31st, 2015, MÖS has continued to remind property owners of their obligation to submit radon gas radiation measurement results and also imposed conditional fines on those failing to do so.
higher among landlords. More than half of housing associations included in our study were located in Skövde, the largest municipality participating in MÖS.

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Contacted before MI training program</th>
<th>Contacted after MI training program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Housing associations</td>
<td>Landlords</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>Included in study</td>
</tr>
<tr>
<td>Falköping</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Hjo</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Karlslöv</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Skövde</td>
<td>32</td>
<td>26</td>
</tr>
<tr>
<td>Tibro</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>TOTAL</td>
<td>59</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 4.2. Distribution of property owners with respect to type and municipality

For the purpose of our study we measure the changes in inspectors’ MI skills before and after having received MI training. Moreover, results from questionnaires regarding the inspectors’ perceptions of the MI training days as well as their perceptions of inspection calls to property owners are also used to evaluate the inspectors’ experience of MI practice. Finally, we determine property owners’ rate of compliance with the request for submitting results from radon gas radiation measurements.

4.2.2. The MI training program

Our study included three female and one male health safety inspectors between the ages of 25 and 34 years. They all had university education in natural sciences, but none had one in social sciences. The inspectors had between two and ten years of inspection and enforcement work experience, and three of them had worked with inspection and enforcement during the last twelve months. Three inspectors had received training in communication skills, but no one had any previous knowledge of MI.

The inspectors received 30 hours of training in MI during five days spread out over a 16 week period (see table 3 below for more details). The first day was devoted to theoretical orientation and practical applications. The other training days consisted of three hours of theory and practical applications and three hours of feedback on recorded conversations. The training program was implemented in collaboration with the coding laboratory MIC Lab (see section 3.2.2). Before each training day, except the first one, each inspector recorded an inspection call that was uploaded to MIC Lab’s homepage to be coded for use as feedback and not to be included in the study (one inspector provided no recorded inspection call before training day 4). These inspection calls were about radon gas radiation measuring and directed at
property owners not participating in the current study, except one that was about a private sewer system. The coders provided written comments explaining the numbers in the protocol to facilitate the understanding of the feedback, which focused on earlier and current themes. Each inspector made one verbatim transcription of a recorded inspection call that was used as a pedagogic tool with the purpose of exemplifying MI concepts and MI skill applications.

The choices of MI themes for each training day were made according to common MI practice (Miller and Moyers, 2007; Miller and Rollnick, 2013) and built on our experiences from the EMT MI training program (Forsberg et al., 2014 and 2016). In addition, the training program also included a brief introduction to fundamental reinforcement principles within psychological learning theory (Sundel and Sundel, 2005).

<table>
<thead>
<tr>
<th>TRAINING DAY</th>
<th>MI THEME</th>
<th>MI THEME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>To listen and to communicate collaboration</td>
<td>– To understand the meaning of MI, four processes</td>
</tr>
<tr>
<td></td>
<td>– To direct towards a change goal</td>
<td>– To convey collaboration and equality</td>
</tr>
<tr>
<td></td>
<td>– To emphasize autonomy</td>
<td>– To ask questions and summarize</td>
</tr>
<tr>
<td>Day 2</td>
<td>To strengthen reinforces toward positive behaviors</td>
<td>– To recognize, elicit and strengthen change talk</td>
</tr>
<tr>
<td></td>
<td>– To use the coding results as feedback</td>
<td></td>
</tr>
<tr>
<td>Day 3</td>
<td>To make efforts to understand the perspectives of the other</td>
<td>– To use empathic listening</td>
</tr>
<tr>
<td></td>
<td>– To exchange information in a dialogue</td>
<td>– To understand and implement positive and negative reinforcement</td>
</tr>
<tr>
<td>Day 4</td>
<td>To meet ambivalence and resistance</td>
<td>– To explore readiness to change and ambivalence</td>
</tr>
<tr>
<td></td>
<td>– To meet and roll with resistance/dissonance</td>
<td>– To avoid MI-non-adherent utterances</td>
</tr>
<tr>
<td>Day 5</td>
<td>To use MI in daily inspection practice</td>
<td>– To summarize the MI training program</td>
</tr>
<tr>
<td></td>
<td>– To form a personal plan for upholding MI proficiency</td>
<td>– To use skills in roleplaying</td>
</tr>
</tbody>
</table>

Table 4.3. MI themes for each of the five days of MI training

Each inspector recorded 8 - 12 live inspection calls both before and after the training program with the informed consent from the property owners. The MI skills of the four inspectors were assessed through coding of the recorded inspection calls, as outlined in section 3.2.2.
Among the 60 property owners in the control group and the 52 property owners in the intervention group, some had to be excluded in the analysis of MI skills, either because the property owner declined to be audio recorded or because the inspection calls were too short for reliable assessment (less than 10 minutes). In total there were 86 recorded conversations, 46 before the program started and 40 after it had been completed, that were used to evaluate MI skills. We employed a pre-post design based on the assessments of all four inspectors’ MI skills before and after the training program. The property owners were not identified when recorded inspection calls were coded. Hence, it is not possible to link an inspector’s MI skills in an inspection call to property owners’ compliance. For this reason we are not able to provide information regarding the distribution of coded conversations across types and municipality of property owners.

4.2.3. Inspectors’ assessments
After every training day the inspectors were asked to answer a questionnaire, as described in section 3.2.3. In total 16 questionnaires were submitted during the course of the training program.

Inspectors also assessed conversations with property owners in a questionnaire immediately after inspection calls, for which recording had been permitted and the duration was at least ten minutes (see section 3.2.4 for more details). In total, 89 questionnaires were submitted, 49 before and 40 after the MI training program.9

4.2.4. Compliance of property owners
The compliance rates of housing associations and landlords that were included in the study were determined for both May 31st and August 31st, 2015. The former date was the deadline mentioned in the letter requesting the submission of radon gas radiation measurement results. The latter date was chosen as it also includes property owners who conducted radon gas radiation measurements between October 2014 and April 2015 as requested, but who failed to submit the results before May 31st. Since what matters is that measurements are actually carried out and given that delays in reporting results can be caused by the firm in charge of the measuring, we believe that it is more accurate to also take into account results that were sent in later than May 31st.

4.2.5. Statistical analysis
The statistical analysis was conducted in the SPSS software environment. All comparisons between means were made by using T-tests for independent groups. Although the inspectors’ MI skills were measured before and after training, we consider the assessments as independent because these were made on different inspection calls.

---

9 Three of the questionnaires submitted before the MI training program concern inspection calls that were not recorded and hence, should not have been submitted. Unfortunately we are not able to identify these and they are therefore included in our analysis.
4.3. Results

4.3.1. Inspectors’ MI skills

Table 4.4 shows the means and standard deviations for the four summary measures of MI skills, before and after the training program, as well as the p-values for the differences between the means for every inspector and totally. At the aggregate level, significant improvements were achieved on all measures. Individually, all inspectors improved their MI skills (with the exception of inspector 1 experiencing a slight decline in the ratio of reflections to questions). In particular, all inspectors became significantly better in terms of empathy, evocation and reducing MI-non-adherent utterances. With regard to the ratio of reflections to questions a significant increase was recorded only for inspector 4.

<table>
<thead>
<tr>
<th></th>
<th>Inspector 1</th>
<th>Inspector 2</th>
<th>Inspector 3</th>
<th>Inspector 4</th>
<th>All inspectors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre post</td>
<td>pre post</td>
<td>pre post</td>
<td>pre post</td>
<td>pre post</td>
</tr>
<tr>
<td>Number of coded conversations</td>
<td>13 13</td>
<td>11 8</td>
<td>11 8</td>
<td>11 11</td>
<td>46 40</td>
</tr>
<tr>
<td>Empathy</td>
<td>1.77 3.46</td>
<td>1.75 3.38</td>
<td>1.70 3.00</td>
<td>1.91 3.45</td>
<td>1.78 3.35</td>
</tr>
<tr>
<td></td>
<td>(0.44 0.52)</td>
<td>(0.45 0.52)</td>
<td>(0.48 0.53)</td>
<td>(0.30 0.52)</td>
<td>(0.42 0.53)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Evocation</td>
<td>1.54 3.23</td>
<td>1.58 3.13</td>
<td>1.70 2.63</td>
<td>1.82 3.00</td>
<td>1.65 3.03</td>
</tr>
<tr>
<td></td>
<td>(0.52 0.60)</td>
<td>(0.51 0.35)</td>
<td>(0.48 0.52)</td>
<td>(0.40 0.45)</td>
<td>(0.48 0.53)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MI-non-adherent utterances</td>
<td>8.51 0.69</td>
<td>7.93 0.00</td>
<td>5.84 1.18</td>
<td>6.44 0.46</td>
<td>7.28 0.59</td>
</tr>
<tr>
<td></td>
<td>(4.38 1.03)</td>
<td>(4.30 0.00)</td>
<td>(3.29 1.46)</td>
<td>(2.94 0.69)</td>
<td>(3.86 0.99)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Ratio reflections to questions</td>
<td>0.95 0.80</td>
<td>0.88 1.05</td>
<td>0.61 0.97</td>
<td>0.46 1.73</td>
<td>0.73 1.14</td>
</tr>
<tr>
<td></td>
<td>(0.81 0.27)</td>
<td>(0.31 0.46)</td>
<td>(0.35 0.43)</td>
<td>(0.25 0.85)</td>
<td>(0.52 0.65)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.561</td>
<td>0.339</td>
<td>0.066</td>
<td>0.000</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Table 4.4. Values of the four MI skills summary measures, before and after the training program, individually and totally, means (standard deviations) and p-values

4.3.2. Inspectors’ perceptions of the MI training program

Table 4.5 below summarizes the responses to the questionnaire that the inspectors were asked to hand in after each training day. MI training as a whole was on average rated at 5.5/6, with a slightly decreasing tendency towards the end of the program; there were no significant changes over time and between inspectors. The perceived usefulness of theory reviews and practical exercises were on average assessed at 5.5/6 and changed little over time. The usefulness of feedback on recorded inspections scored even higher, 5.8/6 on average, and was stable over time. Regarding the applicability of MI skills in inspection work the answers varied between 4 and 6 and reached 5.5/6 after the final training day. However, the
statistical base is weak due to the low number of inspectors and some missing values.

<table>
<thead>
<tr>
<th>MI TRAINING DAY QUESTIONNAIRE</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of training day participants</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>Number of completed questionnaires</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>1. How would you rate the usefulness of feedback on recorded conversations (afternoon)?</td>
<td>-</td>
<td>5.67</td>
<td>6.00</td>
<td>6.00</td>
<td>5.50</td>
<td>5.79</td>
</tr>
<tr>
<td>2. How would you rate the usefulness of theory reviews and exercises (morning)?</td>
<td>5.50</td>
<td>6.00</td>
<td>5.67</td>
<td>5.00</td>
<td>5.50</td>
<td>5.53</td>
</tr>
<tr>
<td>3. How would you rate the MI training as a whole?</td>
<td>6.00</td>
<td>6.00</td>
<td>5.67</td>
<td>5.00</td>
<td>5.00</td>
<td>5.53</td>
</tr>
<tr>
<td>4. Will you be able to apply your new MI skills in your inspection work?</td>
<td>4.25</td>
<td>5.00</td>
<td>5.67</td>
<td>4.67</td>
<td>5.50</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Table 4.5. Responses to questions about the MI training days on a six-graded scale (1 = “not at all useful” – 6 = “very useful”), per day and totally

4.3.3. Inspectors’ perceptions of inspection calls

Table 4.6 contains the results from the inspection call questionnaires at the aggregate level. All but two of the inspection call questionnaires were answered completely (question 3 is missing in one questionnaire before the MI training program, and question 6 is missing in one questionnaire after the MI training program).

The mean ratings of statement 2 (“My perception at the end of the inspection call was that the property owner had sufficient knowledge of the effects of radon gas radiation on people’s health.”) increased significantly from 3.16/5 before to 3.65/5 after the MI training program, while mean ratings of statement 6 (“I am satisfied with my own performance during the inspection call.”) actually decreased significantly from 4.04/5 to 3.64/5.11

Property owners talking to MI-trained inspectors were assessed to have a more negative attitude to radon gas radiation measurements at the start of the inspection calls (statement 1) compared to property owners in the control group. For statement 7 (“My perception is that the property owner at the end of the inspection call had a positive attitude to radon gas radiation measurement.”) the mean rating was 0.23 higher than for statement 1 before the MI training program and 0.22 higher after the MI training program, but these differences are insignificant. Hence, although on average property owners became slightly more positive as a result of the inspection

10 One inspector was unable to participate on the third training day – this person was given a 2.5 hour compensatory review between the third and fourth training days.

11 At an individual level, the increase in the answers to question 2 was significant for inspectors 3 and 4, while the decrease in the answers to question 6 was significant for all inspectors except inspector 4.
call, it seems that MI training has not improved inspectors’ self-perceived ability to influence property owners’ attitudes in inspection calls.

<table>
<thead>
<tr>
<th>INSPECTORS’ PERCEPTIONS OF INSPECTION CALLS</th>
<th>Before MI training</th>
<th>After MI training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of submitted questionnaires</td>
<td>49</td>
<td>40</td>
</tr>
<tr>
<td>1. My perception at the start of the inspection call was that the property owner had a positive attitude to discussing the measurement of radon gas radiation.</td>
<td>4.24</td>
<td>3.83</td>
</tr>
<tr>
<td>2. My perception at the end of the inspection call was that the property owner had sufficient knowledge of the effects of radon gas radiation on people’s health.</td>
<td>3.16</td>
<td>3.65</td>
</tr>
<tr>
<td>3. My perception is that the property owner during the inspection call clearly showed that he/she understood the information I wanted to convey.</td>
<td>4.00</td>
<td>3.90</td>
</tr>
<tr>
<td>4. My perception is that the property owner will undertake a measurement of radon gas radiation before May 31st, 2015.</td>
<td>4.53</td>
<td>4.40</td>
</tr>
<tr>
<td>5. My perception is that the property owner expressed his/her own reasons and motivation to conduct a measurement of radon gas radiation.</td>
<td>3.16</td>
<td>3.45</td>
</tr>
<tr>
<td>6. I am satisfied with my own performance during the inspection call.</td>
<td>4.04</td>
<td>3.64</td>
</tr>
<tr>
<td>7. My perception is that the property owner at the end of the inspection call had a positive attitude to radon gas radiation measurement.</td>
<td>4.47</td>
<td>4.05</td>
</tr>
</tbody>
</table>

Table 4.6. Inspectors’ assessments of inspection calls on a five-graded scale (1=“absolutely disagree” – 5=“absolutely agree”), averages and p-values

However, the average difference between ratings of statement 4 (“My perception is that the property owner will undertake a measurement of radon gas radiation before May 31st, 2015.”) and statement 1 was 0.29 in the control group, while being 0.57 in the intervention group. One interpretation of these differences is that inspection calls to property owners increases their propensity to undertake radon gas radiation measurements and that this effect is reinforced if the inspector has received MI training.

The answers to the inspection call questionnaires show that the property owners contacted by a MI-trained inspector had a more negative attitude to radon gas radiation measurements to begin with (statement 1), that this attitude became more positive during the inspection calls, especially in the intervention group, but that the attitude was still more negative in this group at the end of the inspection calls (statement 4). Thus, one would suspect that the compliance rate would be lower among property owners contacted by a MI-trained inspector.
4.3.4. Compliance of property owners

Table 4.7 presents the compliance rates of housing associations and landlords that were included in the study, across municipalities and totally. The propensity to carry out radon gas radiation measurements as requested was actually lower among property owners that were called by a MI-trained inspector. However, the difference in compliance rates between the control and the intervention groups – 9 per cent units in May and 6 per cent units in August – is not statistically significant.

<table>
<thead>
<tr>
<th></th>
<th>Contacted before MI training</th>
<th>Contacted after MI training</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Housing associations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falköping</td>
<td>10</td>
<td>0.60</td>
</tr>
<tr>
<td>Hjo</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Karlsborg</td>
<td>2</td>
<td>0.50</td>
</tr>
<tr>
<td>Skövde</td>
<td>26</td>
<td>0.73</td>
</tr>
<tr>
<td>Tibro</td>
<td>7</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>48</td>
<td>0.60</td>
</tr>
<tr>
<td><strong>Landlords</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (all in Karlsborg)</td>
<td>12</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>ALL PROPERTY OWNERS</strong></td>
<td>52</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Table 4.7. Compliance rates (CR) of housing associations and landlords in the control and intervention groups, May 31st and August 31st, 2015.

About a fifth of property owners in both groups complied with the request for carrying out radon gas radiation measurements, but missed the May 31st deadline for reporting the results. The compliance rates are well above the one achieved in the previous MÖS project where landlords were contacted by mail only. Although the present study mainly involved housing associations the effect is also present among landlords in our sample, but less pronounced. A possible explanation for the higher compliance rate in the present study could be that information about the previous project directed at landlords had alerted property owners. However, we note that higher compliance rates were achieved also in Karlsborg which had not participated in the previous project.

4.4. Discussion

The principal aim of the study was to evaluate the feasibility of MI in inspection calls to property owners in order to induce the reporting of results from radon gas radiation measurements as required by law. At the aggregate level the four health safety inspectors significantly improved on all four summary measures of MI competences. Individually all inspectors significantly increased their empathy and evocation scores, and significantly reduced the frequency of MI-non-adherent
utterances. Concerning the ratio of reflections to questions three inspectors recorded an increase; however, the increase was significant for only one of them.

It is of interest to compare the present training program with the EMT training program of environmental, health safety and food safety inspectors (Forsberg et al., 2014 and 2016). Under the EMT training program inspectors improved their MI skills, but to a much lesser extent. At the aggregate level the measures for empathy and evocation increased significantly (from 1.36 to 1.79 and from 1.3 to 1.83, respectively) whereas the decrease in MI-non-adherent utterances was insignificant (from 3.8 to 2.0) and the ratio of reflections to questions actually dropped slightly (from 0.46 to 0.43, insignificantly). The four figures below illustrate inspectors’ MI skills before (pre) and after (post) the MI training program in the present study in comparison to the EMT study.

Figures 4.1-4.4. Changes in MI skills in the present study (Radon) and the EMT study, where empathy and evocation are rated from 1 (“low”) to 5 (“high”)
We can conclude that the present training program resulted in improvements that clearly exceeded the ones achieved by the EMT training program with regard to the four MI skills summary measures. The results confirm our hypothesis that applying MI to a more specified inspection setting would generate stronger skills improvements.

Inspectors’ assessments of the present MI training program were very positive and stable; both over time and between inspectors there were no significant differences. Compared to the EMT training program, the average values were higher in the present study. In particular, the perceived usefulness of feedback on recorded inspections was graded almost one unit higher (5.79/6 vis-à-vis 4.86/6), while the usefulness of theory reviews and exercises was graded 0.73 units higher (5.53/6 vis-à-vis 4.80/6) under the present training program. MI training as a whole was rated more than half a grading unit higher than under the EMT training program (5.53/6 vis-à-vis 4.94/6). The applicability of MI skills in inspection work was also rated higher, but to a lesser extent (5.02/6 vis-à-vis 4.74/6), by the inspectors of the present study. As hypothesized, MI training focused on a specific inspection setting seems to increase the perceived value of the program.

Regarding inspectors’ perceptions of inspection calls to property owners two significant changes between responses before and after the MI training program were identified. On the one hand, MI-trained inspectors perceived a better knowledge among property owners about the impact of radon gas radiation on health at the end of inspection calls. On the other hand, MI-trained inspectors were less satisfied concerning their own performance in inspection calls. A possible explanation for this lower degree of satisfaction could be that MI-trained inspectors became more conscious and self-critical after having completed the MI training program. Another explanation could be related to the fact that the MI-trained inspectors rated the attitude of property owners to discuss radon gas radiation measurement lower than MI-untrained inspectors.

The property owners’ compliance rates regarding the request for reporting radon gas radiation measurements were 75 per cent in the control group and 69 per cent in the intervention group. Thus, compliance was lower among property owners contacted by MI-trained inspectors, but not statistically significantly. The result should be interpreted with caution due to the limitations of the present study, which was primarily designed for the evaluation of MI training on inspectors’ communicative skills. The fact that the letters were followed up within a month in the control group, but five months later in the intervention group, suggests that inspection calls have a smaller positive effect if the time span between sending the letter and the call is larger. It is thus impossible to separate the impact of MI skills from the effect of follow-up inspection calls having taken place much later in the intervention group. Results from a parallel study with a stronger research design and focusing on waste sorting inspections of restaurants show a significant positive effect of MI training on compliance (see section 5.4).
Moreover, the compliance rates in both the control and the intervention group were higher than in a previous MÖS project where letters requesting the measurement of radon gas radiation were not followed up with an inspection call. Thus, a following-up inspection call seems to enhance compliance, which is an important result. Finally, the lower compliance rate in the intervention group could also be explained by sampling bias. As already mentioned in section 3.3, the inspectors’ assessments of inspection calls with property owners suggest a more negative attitude to radon gas radiation measurements in the intervention group.

A general limitation of this study is the low number of inspectors (4), which naturally threatens the validity of our results. Another limitation is that we are not able to associate MI skills in specific inspection calls to outcomes in terms of reported radon gas radiation measurements, which is due to the fact that property owners were not identified in the recordings of conversations. In our parallel study (Herzing et al., 2017) these limitations have been addressed by increasing the number of inspectors receiving MI training and identifying inspectees at recorded inspections.

In our research program an attempt has been made to investigate the impact of MI-trained inspectors on behavioral changes among inspectees; to our knowledge, this has not been done so far in any published study. Building on our results we intend to design and carry out further studies to evaluate how improved communicative skills among inspectors can improve compliance with legislation. Our aim is also to further investigate the challenges of training inspectors in MI. It is worth emphasizing that when applying MI to inspections the interaction between the inspector and the inspectee does not take place voluntarily. Since inspectors are representatives of an authority, responsible for monitoring inspectees and enforcing legislation, the nature of potential motivations to change might differ compared to a clinical context. In an inspection setting external factors (i.e. legislation) are likely to carry a relatively stronger weight than internal motivations (e.g. long-term consequences on the environment and hence, children’s health).

4.5. References


5. Enhancing compliance with waste sorting regulations through inspections and motivational interviewing

Mathias Herzing¹², Hans Wickström¹³, Lars Forsberg¹⁴, Adam Jacobsson¹⁵ and Håkan Källmén¹⁶

5.1. Introduction

In this study we explore how the communicative method “Motivational Interviewing” (MI) can be adapted to the context of food safety inspections and how inspections and improved MI skills can improve compliance with legislation. The study was carried out as part of the interdisciplinary research program Inspections and Enforcement as Instruments for Enhancing Environmental Behavior (“Tillsynen som styrmedel för ett förbättrat miljöbeteende” – TSFM) as part of an ongoing project in five Swedish municipalities, which aims at inducing restaurants to sort their waste in accordance with the law.

These five municipalities – Falköping, Hjo, Karlsborg, Skövde and Tibro – coordinate inspections and enforcement of environmental legislation within the regional collaboration Miljösamverkan Östra Skaraborg (MÖS). The project involved seven food safety inspectors who regularly carry out inspections of restaurants to check compliance with food safety legislation. These inspectors had noted that many restaurants do not sort their waste, which is regulated by the Environmental Code’s Ordinance on waste management. Since restaurants are not inspected by environmental inspectors, no active efforts had previously been undertaken to monitor and improve restaurants’ waste sorting. MÖS therefore decided to add waste management to the food safety inspectors’ checklist for restaurants and to actively motivate restaurants to sort waste as required by legislation. The fact that restaurants’ waste sorting had previously not been monitored by inspectors provided a unique opportunity to design a research study analyzing how inspections per se and improvements in inspectors’ MI skills affect

¹² Department of Economics, Stockholm University, email: Mathias.Herzing@ne.su.se
¹³ Meetme Psykologkonsult AB, Gothenburg, email: hans.wickstrom@kbtanalys.se
¹⁴ MicLab AB, Stockholm, email: lars.forsberg@miclab.se
¹⁵ Department of Economics, Stockholm University, email: aja@ne.su.se
¹⁶ Karolinska Institutet, Department of Clinical Neuroscience, Center for Psychiatry Research, email: hakan.kallmen@ki.se
compliance with legislation. To that end the seven MÖS inspectors were trained in MI, an evidence-based method to change behavior through communication.

The aim of this study is to investigate the effectiveness of inspections on compliance with waste sorting legislation and to explore the effects of MI training of inspectors in this specific setting. More specifically, the intention is both to analyze the impact of the training program on inspectors’ MI skills and to evaluate how restaurants’ compliance with regulations is affected by previous inspections (an inspection effect) and the MI training program (a combined inspection and MI effect). Moreover, results from questionnaires regarding the inspectors’ perceptions of the MI training as well as their assessments of conversations with restaurant staff are also used to further evaluate the training program. In addition, we have evaluated responses to questionnaires concerning restaurant staff’s perceptions of conversations with the inspectors.

Results from the earlier EMT study showed that after having completed a MI training program inspectors significantly improved their skills in conversations with inspectees regarding two of the most important MI variables, showing empathy and avoiding MI-non-adherent utterances (Forsberg et al., 2014 and 2016). The inspectors also experienced MI to be of benefit during inspections and became more satisfied with their own performance at inspections. However, the EMT MI training program covered a wide range of inspection settings and was devised for heterogeneous groups of inspectors working in different contexts, which made it difficult to convey the applicability of MI. In contrast to the EMT research program, the present study has been adapted to a specific setting, allowing the participants to be focused on the communicative requirements in their context. Hence, we expected inspectors to attain a higher MI competence than in the previous EMT training program. Results from the parallel study (see section 4) covering inspectors making phone calls to property owners to induce the measuring and reporting of radon gas radiation confirm this expectation.

The study was approved by the Regional Ethical Review Board in Stockholm (reg nr 2014/1417-32). Results concerning restaurants’ compliance are presented on an aggregate level, while results regarding individual inspectors have been de-identified. The presentation of this study is organized as follows. Section 5.2 presents our methodological approach, while section 5.3 contains the results of the study. In section 5.4 we discuss our results and possible future research.

5.2.  Methodology

5.2.1.  Research design

Our research study was designed to enable an evaluation both of how the MI training program influenced the participating inspectors’ MI skills and of how
inspections and MI training of inspectors affected the restaurants’ rate of compliance with waste sorting regulations.

There are seven food safety inspectors at MÖS who usually carry out so-called sprints of about one month’s duration. In each sprint 3-4 inspectors conduct inspections that focus on specific targets (e.g. to check the cleanliness of restaurants), while the remaining inspectors work with all other tasks (e.g. handling complaints and following-up on previous cases). After every sprint new inspector teams are formed and tasks are switched. The goal is that every inspector works with all colleagues and that inspectees are inspected by different inspectors over time. For the purpose of this research study, more inspectors than usually conducted inspections in most of the sprints (see table 5.1 below).

In 2014 MÖS determined that regular food safety inspections of restaurants should also involve controlling compliance with waste sorting regulations. Our study covers the first four sprints where restaurants’ waste sorting was also checked. These four sprints took place between the fall of 2014 and the fall of 2015. Between the first and the second sprints all seven food safety inspectors participated in a MI training program, which consisted of five training days spread out over a three month period. Between the third and fourth sprints the inspectors received two booster MI training half-days. The study covered 181 registered restaurants. A total number of 359 inspections were carried out in the four sprints. One of the inspectors dropped out of the study during the MI training program for health reasons. The table below shows the schedule of the study and provides information on the number of restaurants being inspected and the number of inspectors participating in each sprint.

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>29 September – 24 October 2014</td>
<td>Sprint 1: 62 inspections, 7 inspectors</td>
</tr>
<tr>
<td>November 2014 – January 2015</td>
<td>MI TRAINING PROGRAM</td>
</tr>
<tr>
<td>10 February – 6 March 2015</td>
<td>Sprint 2: 116 inspections, 5 inspectors</td>
</tr>
<tr>
<td>20 April – 7 August 2015</td>
<td>Sprint 3: 46 inspections, 3 inspectors</td>
</tr>
<tr>
<td>August 2015</td>
<td>Extra MI training</td>
</tr>
<tr>
<td>28 September – 12 November 2015</td>
<td>Sprint 4: 135 inspections, 6 inspectors</td>
</tr>
</tbody>
</table>

Table 5.1. Schedule of the study

Restaurants were randomly assigned to inspectors as well as the order in which restaurants were inspected. At inspections inspectors made an ocular assessment to determine whether restaurants comply with waste sorting regulations such that we obtained a binary outcome variable (to keep things simple the degree of non-

---

17 Not all inspections could be carried out within one month and hence, had to be postponed to the summer when one of the inspectors inspected the remaining restaurants.
compliance was not accounted for). In normal routine, non-compliance leads to an injunction prescribing the inspectee to undertake measures to achieve adherence. However, in this research project violations of waste sorting regulations were not followed by an injunction – the purpose was to avoid distortions in the analysis of the impact of inspections and MI on compliance.

To evaluate the effect of the training program on inspectors’ MI skills we employed a pre-post design. For the assessment of MI skills recordings of inspection conversations were used, allowing us to compare the skill levels of all seven inspectors before (sprint 1) and after (sprints 2-4) the training program. For inclusion in the MI skills analysis inspections had to be at least ten minutes long and permission for audio recording been granted. For more details on the measurement of MI skills, we refer to section 3.2.2.

<table>
<thead>
<tr>
<th>Inspector 1</th>
<th>MI-measured inspections</th>
<th>Sprint 1</th>
<th>Sprint 2</th>
<th>Sprint 3</th>
<th>Sprint 4</th>
<th>All sprints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspector 2</td>
<td>MI-measured inspections</td>
<td>14</td>
<td>13</td>
<td>26</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>Inspector 3</td>
<td>MI-measured inspections</td>
<td>18</td>
<td>26</td>
<td>18</td>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>Inspector 4</td>
<td>MI-measured inspections</td>
<td>8</td>
<td>21</td>
<td>22</td>
<td>10</td>
<td>51</td>
</tr>
<tr>
<td>Inspector 5</td>
<td>MI-measured inspections</td>
<td>4</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Inspector 6</td>
<td>MI-measured inspections</td>
<td>5</td>
<td>21</td>
<td>14</td>
<td>36</td>
<td>76</td>
</tr>
<tr>
<td>Inspector 7</td>
<td>MI-measured inspections</td>
<td>7</td>
<td>17</td>
<td>23</td>
<td>45</td>
<td>45</td>
</tr>
</tbody>
</table>

Table 5.2. Number of inspections per inspector and sprint, total and MI-measured

Table 5.2 shows the distribution of inspections per inspector and sprint; it contains both the total number of inspections that were carried out and the number of inspections for which coding results (MI measurements) were available. Originally the aim was to let all seven food safety inspectors participate in the MI training program, but one dropped out after the second training day due to health problems. For this person three conversations were coded in sprint 1; naturally, these were excluded from the MI skills evaluation.
For the analysis of inspectors’ MI skills we obtained complete coding results for 174 out of all 359 inspections. In some cases either no permission for recording was given or recordings were too short (less than 10 minutes) to be reliably coded. Since inspector 7 dropped out of the study after the first sprint, after having conducted seven inspections of which three were coded, the observations for the MI skills analysis were reduced to 171.

To evaluate how inspections and the MI training program affected compliance we had to observe changes in compliance over time and hence, 49 restaurants that were inspected only once had to be excluded. Thus, the number of observations is reduced from 359 to 310 for the analysis of compliance, corresponding to a decrease in the number of included restaurants from 181 to 132. There were 40 restaurants that were inspected more than twice, of which six were inspected four times (i.e. in every sprint).

For every restaurant that was inspected more than once we were able to record changes over time and thus, to establish the effect of the previous inspection. Hence, if an inspector at an inspection succeeded in motivating the restaurant to comply with waste sorting regulations this became apparent at the following inspection of that restaurant.

By employing a cross-sectional design we were able to obtain measures for both a pure inspection effect and a combined inspection and MI effect. For one group of restaurants that was inspected at least twice the first inspection took place during the first sprint before inspectors had received any MI training, i.e. they were inspected by a MI-untrained inspector. By comparing the outcomes at the two first inspections among these restaurants we could identify an inspection effect; the increase in compliance in this group could be attributed to waste sorting having been checked for the first time. A second group of restaurants was inspected for the first time after inspectors had completed the training program and then was inspected at least once more. Since MI-trained inspectors carried out the first inspection, the difference in outcomes between the first two inspections could be attributed to an inspection and MI interaction effect among these restaurants. For the restaurants that were inspected in sprint 1 and then inspected at least two more times we could establish both an inspection effect (after the first inspection) followed by a second inspection and MI interaction effect.

Food safety inspections are by law supposed to be conducted without prior notification. However, sometimes inspections are notified in advance for practical reasons, e.g. if the restaurant has irregular opening hours or if it has been closed at several unannounced inspector visits. In our sample there is a total number of 24

---

18 Note that the second inspection of restaurants in the control group took place in either sprint 2 or 3 or 4.
inspections where restaurants were notified in advance; of these 19 were included in the analysis of MI skills and 15 were used to evaluate changes in compliance.

5.2.2. The MI training program
Six food safety inspectors participated in the study, five women and one man between the ages of 23 and 57 years. They all had university education in natural sciences. The inspectors had 2-13 years of inspection and enforcement work experience, and all of them had conducted food safety inspections during the last twelve months. None had previous knowledge of MI, but all had received some training in communication skills.

<table>
<thead>
<tr>
<th>TRAINING DAY</th>
<th>MI THEME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>To listen and to communicate collaboration</td>
</tr>
<tr>
<td></td>
<td>– To understand the meaning of MI, four processes</td>
</tr>
<tr>
<td></td>
<td>– To direct towards a change goal</td>
</tr>
<tr>
<td></td>
<td>– To convey collaboration and equality</td>
</tr>
<tr>
<td></td>
<td>– To emphasize autonomy</td>
</tr>
<tr>
<td></td>
<td>– To ask questions and summarize</td>
</tr>
<tr>
<td>Day 2</td>
<td>To reinforce positive behaviors</td>
</tr>
<tr>
<td></td>
<td>– To recognize, elicit and strengthen change talk</td>
</tr>
<tr>
<td></td>
<td>– To use the coding results as feedback</td>
</tr>
<tr>
<td>Day 3</td>
<td>To understand the perspectives of the other</td>
</tr>
<tr>
<td></td>
<td>– To use empathic listening</td>
</tr>
<tr>
<td></td>
<td>– To exchange information in a dialogue</td>
</tr>
<tr>
<td></td>
<td>– To understand and implement positive and negative reinforcement</td>
</tr>
<tr>
<td>Day 4</td>
<td>To meet ambivalence and resistance</td>
</tr>
<tr>
<td></td>
<td>– To explore readiness to change and ambivalence</td>
</tr>
<tr>
<td></td>
<td>– To meet and roll with resistance/dissonance</td>
</tr>
<tr>
<td></td>
<td>– To avoid MI-non-adherent utterances</td>
</tr>
<tr>
<td>Day 5</td>
<td>To use MI in daily inspection practice</td>
</tr>
<tr>
<td></td>
<td>– To summarize the MI training program</td>
</tr>
<tr>
<td></td>
<td>– To form a personal plan for upholding MI-proficiency</td>
</tr>
<tr>
<td></td>
<td>– To use skills in roleplaying</td>
</tr>
</tbody>
</table>

Table 5.3. MI themes for each of the five days of MI training

The inspectors received 30 hours of training during five days spread out over a 16 week period. Each training day consisted of short theoretical introductions mixed with experiential exercises of MI skills. During the last three hours of all training days (except the first one) the inspectors gave each other feedback to recorded conversations under the supervision of the instructor. On training days 2 and 4 feedback was given to non-professional conversations that were not coded. Before training days 3 and 5, each participant recorded an inspection conversation about
waste sorting, which was uploaded to MIC Lab’s homepage to be professionally coded for use as feedback by the MI instructor.

Table 5.3 above summarizes the topics covered by the MI training program. The choices of MI themes for each training day were made according to common MI practice (Miller & Moyers, 2007; Miller & Rollnick, 2013) and built on our experiences from the MI training programs of EMT (Forsberg et al., 2014 and 2016) as well as our recent study on the application of MI to induce radon gas radiation measurements by property owners (see section 4). In addition, the training program also included a brief introduction to fundamental reinforcement principles within psychological learning theory (Sundel & Sundel, 2005).

As MÖS had decided to let all its 40 inspectors undergo MI training, the inspectors in our study were offered further training to become MI coaches for their colleagues. Four of the six inspectors agreed to this and were trained for two half-days by the MI instructor after the third sprint. In August 2015 all MÖS inspectors had two days of MI training, led by the MI instructor who was assisted by the four new MI coaches. In order to facilitate the implementation of MI these two days of training were followed by MI group supervision, where each group had monthly meetings led by a MI coach. Between the third and fourth sprint these groups met twice. The two inspectors who declined to become MI coaches participated in the training days and in the supervised group meetings.

5.2.3. Assessments by inspectors and inspectees
At the end of every training day the inspectors were asked to answer a questionnaire, as described in section 3.2.3. In total 32 questionnaires were submitted during the course of the training program.

Inspectors also answered a questionnaire immediately after inspections, for which recording had been permitted and the duration was at least ten minutes (see section 3.2.4 for more details). In total, 250 questionnaires were submitted, 47 before and 203 after the MI training program.

After the inspection the inspector handed over a questionnaire from the research team (including a missive and a stamped envelope). The missive contained brief information regarding the questionnaire and the research project. Restaurants were asked to answer questions about the inspection and to send the completed questionnaire to Karolinska institutet in Stockholm. By letting restaurants respond to a questionnaire, we were able to compare restaurants’ perceptions regarding the inspections and the inspectors before and after the MI training program.

5.2.4. Statistical analysis
The statistical analysis was conducted in the SPSS software environment. All comparisons between means were made by using T-tests for independent groups. Although the inspectors’ MI skills were assessed before and after the training
program, we consider them as independent since the assessments were made on different inspection conversations.

5.3. Results

5.3.1. MI skills

Figures 5.1-4 show how the means for the four MI skills measures have evolved across sprints, individually for every inspector and aggregately.\(^{19}\)

\(^{19}\) Note that only inspectors 2 and 4 participated in every sprint.
The whole group of inspectors achieved strong improvements for the four MI skill measures. There are some variations at the individual level, but all inspectors attained better results in the last sprint compared to the first sprint on all measures. At the aggregate level the improvements in skills were continuous over time, except that the ratio of reflections to questions decreased slightly between sprints 2 and 3. For empathy, evocation and MI-non-adherent utterances the biggest improvement occurred after the first sprint, as expected; the ratio of reflections to questions increased most before the last sprint.

The following table presents the means and standard deviations of these four variables in every sprint for the whole group of inspectors and also contains the p-values between sprints.

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S2/3/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observ.</td>
<td>42</td>
<td>54</td>
<td>23</td>
<td>52</td>
<td>129</td>
</tr>
<tr>
<td>Empathy</td>
<td>2.17 (0.66)</td>
<td>2.87 (0.73)</td>
<td>3.04 (0.64)</td>
<td>3.33 (0.65)</td>
<td>3.09 (0.71)</td>
</tr>
<tr>
<td>p-values</td>
<td>S1-S2: &lt;0.001</td>
<td>S1-S3: &lt;0.001</td>
<td>S2-S3: 0.302</td>
<td>S2-S4: &lt;0.001</td>
<td>S1-S2/3/4: &lt;0.001</td>
</tr>
<tr>
<td>Evocation</td>
<td>2.14 (0.65)</td>
<td>2.76 (0.61)</td>
<td>3.09 (0.60)</td>
<td>3.23 (0.47)</td>
<td>3.01 (0.59)</td>
</tr>
<tr>
<td>p-values</td>
<td>S1-S2: &lt;0.001</td>
<td>S1-S3: &lt;0.001</td>
<td>S2-S3: 0.034</td>
<td>S2-S4: &lt;0.001</td>
<td>S1-S2/3/4: &lt;0.001</td>
</tr>
<tr>
<td>MI-non-adherent utterances</td>
<td>1.95 (2.56)</td>
<td>1.23 (2.04)</td>
<td>0.47 (1.17)</td>
<td>0.43 (1.05)</td>
<td>0.77 (1.60)</td>
</tr>
<tr>
<td>p-values</td>
<td>S1-S2: 0.138</td>
<td>S1-S3: 0.002</td>
<td>S2-S3: 0.042</td>
<td>S2-S4: 0.012</td>
<td>S1-S2/3/4: 0.007</td>
</tr>
<tr>
<td>Ratio of reflections to questions</td>
<td>0.76 (0.54)</td>
<td>0.98 (0.87)</td>
<td>0.83 (0.52)</td>
<td>1.58 (1.61)</td>
<td>1.20 (1.22)</td>
</tr>
<tr>
<td>p-values</td>
<td>S1-S2: 0.127</td>
<td>S1-S3: 0.598</td>
<td>S2-S3: 0.354</td>
<td>S2-S4: 0.001</td>
<td>S1-S2/3/4: 0.002</td>
</tr>
</tbody>
</table>

Table 5.4. MI skill measures per sprint, means (standard deviations), numbers of observations (N) and p-values between sprints for the whole group of inspectors

Comparing the MI skills of inspectors in the first and the last sprints, we can conclude that the improvements at the aggregate level were significant for all four measures. We see that empathy and evocation improved significantly between the first two sprints and also between the second and the last sprints. The frequency of MI-non-adherent utterances decreased significantly between the two first sprints, 20 Individual results should be treated with caution as the number of observations is very small for some inspectors in some sprints (see table 5.2).
while the ratio of reflections to questions increased significantly between the last two sprints.

5.3.2. Training day questionnaires

Table 5.5 summarizes the results of the questionnaire that the inspectors were asked to hand in after each training session.

<table>
<thead>
<tr>
<th>MI TRAINING DAY QUESTIONNAIRE</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of training day participants</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>Number of completed questionnaires</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>1. How would you rate the usefulness of feedback on recorded conversations (afternoon)?</td>
<td>5.71</td>
<td>6.00</td>
<td>5.50</td>
<td>5.83</td>
<td>5.76</td>
<td></td>
</tr>
<tr>
<td>2. How would you rate the usefulness of theory reviews and exercises (morning)?</td>
<td>5.71</td>
<td>5.71</td>
<td>6.00</td>
<td>5.83</td>
<td>6.00</td>
<td>5.85</td>
</tr>
<tr>
<td>3. How would you rate the MI-training as a whole?</td>
<td>5.57</td>
<td>5.86</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td>5.89</td>
</tr>
<tr>
<td>4. Will you be able to apply your new MI-skills in your inspection work?</td>
<td>4.43</td>
<td>4.86</td>
<td>5.17</td>
<td>5.33</td>
<td>5.33</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Table 5.5. Inspectors’ responses to questions about the MI training days, rated on a six-graded scale (1=“not at all useful” – 6=“very useful”), per day and totally

The MI training as a whole was on average rated at 5.89; for the last three training days all inspectors’ ratings were at the highest level. The short theory introductions and the experimental exercises as well as the feedback on recorded conversations were also rated at a consistently high level, 5.85 and 5.76 overall, respectively. The usefulness of MI for inspection work was rated at 5.02 overall, with an increase over time, from 4.43 the first day to 5.33 the last two days.

5.3.3. Inspectors’ inspection questionnaires

In total, the inspectors submitted 250 questionnaires regarding their perceived experiences of the inspection conversations with restaurant staff. The assessments by each inspector concerning his/her performance as well as his/her perception of the restaurant representatives’ attitude are summarized in table 5.6 below.

Ratings were higher for all seven questions in sprints 2-4 (after inspectors had completed the MI-training program) than in sprint 1. The strongest and most significant improvement concerned the assessment of restaurant representatives expressing reasons and motivation to sort waste (question 5), which could be explained by improvements in the inspectors’ evocation skills. The fact that the ratings regarding the attitude of restaurants to waste sorting both at the start (question 1) and at the end (question 7) of the inspection were significantly higher after compared to before the MI training program could be attributed to an
increasing number of previously inspected restaurants over time; having discussed waste sorting before might make restaurants more positive.

<table>
<thead>
<tr>
<th>INSPECTORS’ PERCEPTIONS OF INSPECTIONS</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S2/3/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of submitted questionnaires</td>
<td>47</td>
<td>94</td>
<td>36</td>
<td>73</td>
<td>203</td>
</tr>
<tr>
<td>1. My perception at the start of the inspection was that the restaurant representative had a positive attitude to discussing waste sorting.</td>
<td>4.04 (1.03)</td>
<td>4.22 (1.06)</td>
<td>4.33 (0.76)</td>
<td>4.49 (0.87)</td>
<td>4.33 (0.95)</td>
</tr>
<tr>
<td>2. My perception at the end of the inspection was that the restaurant representative had sufficient knowledge of the effects of waste sorting on people’s health.</td>
<td>3.71 (0.97)</td>
<td>3.86 (1.05)</td>
<td>4.22 (0.76)</td>
<td>4.34 (0.77)</td>
<td>4.10 (0.93)</td>
</tr>
<tr>
<td>3. My perception is that the restaurant representative during the inspection clearly showed that he/she understood the information I wanted to convey.</td>
<td>4.04 (0.92)</td>
<td>4.43 (0.89)</td>
<td>4.50 (0.70)</td>
<td>4.52 (0.82)</td>
<td>4.47 (0.83)</td>
</tr>
<tr>
<td>4. My perception is that the restaurant representative will have a properly functioning waste sorting system at the next inspection.</td>
<td>3.50 (1.32)</td>
<td>3.95 (1.28)</td>
<td>3.78 (1.24)</td>
<td>4.30 (1.01)</td>
<td>4.04 (1.20)</td>
</tr>
<tr>
<td>5. My perception is that the restaurant representative expressed his/her own reasons and motivation to sort waste.</td>
<td>3.27 (1.32)</td>
<td>4.02 (1.03)</td>
<td>4.31 (0.71)</td>
<td>4.37 (0.87)</td>
<td>4.20 (0.94)</td>
</tr>
<tr>
<td>6. I am satisfied with my own performance during the inspection.</td>
<td>3.69 (0.90)</td>
<td>4.01 (1.05)</td>
<td>3.69 (0.98)</td>
<td>3.93 (0.92)</td>
<td>3.93 (0.99)</td>
</tr>
<tr>
<td>7. My perception is that the restaurant representative at the end of the inspection had a positive attitude to waste sorting.</td>
<td>4.25 (0.99)</td>
<td>4.57 (0.71)</td>
<td>4.50 (0.69)</td>
<td>4.68 (0.55)</td>
<td>4.60 (0.65)</td>
</tr>
</tbody>
</table>

Table 5.6. Inspectors’ assessments of inspections, rated on a five-grade scale (1=“absolutely disagree” – 5=“absolutely agree”), averages (standard deviations) and p-values for sprint 1 and sprints 2-4

5.3.4. Restaurants’ perception of inspections
A total number of 181 restaurants submitted questionnaires regarding their perceptions of inspections. The results are summarized in table 5.7. Assessments by restaurant representatives were high and stable over time; any changes were insignificant.
Table 5.7. Restaurant representatives’ ratings of the inspections on a four-graded scale (1=“absolutely disagree” – 4=“absolutely agree”), averages (standard deviations) and p-values for sprint 1 and sprints 2-4.

5.3.5. Compliance with waste sorting regulations

Figure 5.5 below illustrates the distribution of inspected restaurants across sprints and how the compliance rate has changed over time for different groups of restaurants. The green boxes indicate the number of restaurants that are inspected for the first time, the blue boxes those inspected for the second time, the red boxes those inspected for the third time and the yellow box those inspected for the fourth time. The arrows show how the compliance rate changes between inspections; the color of the arrow indicates the number of inspections, e.g. a red arrow contains the compliance rates of restaurants across three inspections which have taken place in a certain sequence of sprints. A total number of 49 restaurants were only inspected once and hence, were irrelevant for our compliance rate analysis; these are represented by the grey arrows which contain the compliance rate among these restaurants.

For example, among the 62 restaurants that were inspected in sprint 1, 10 were not inspected again, 36 had a second inspection in sprint 2, 12 were inspected for the second time in sprint 3 and 4 had a second inspection in sprint 4. There are 28 restaurants that were inspected in sprints 1, 2 and 4; their compliance rate increased from 0.29 in the first sprint to 0.46 in the second sprint to 0.68 in the last sprint.
### Figure 5.5

Distribution of restaurants and inspections across sprints with changes in compliance rates indicated inside the arrows

Table 5.8 shows the rates of compliance with waste sorting regulation among restaurants across sprints based on the number (N) of inspections for each calculated compliance rate.
If all inspections are included, the compliance rate is 37% in sprint 1, then remains stable during the following two sprints before jumping to 62% in sprint 4. Excluding restaurants that were only inspected once (i.e. those for which we were unable to measure changes in compliance) we obtain similar compliance rates in all sprints. The slight decrease in the compliance rate in sprint 2 is due to a large number of restaurants being inspected for the first time; among restaurants that were inspected for the second time in sprint 2 the compliance rate is 44%. However, there is hardly any difference in the compliance rate between restaurants that have been previously inspected and those that are inspected for the first time in sprint 3.

Table 5.9 provides detailed information regarding the compliance rates among restaurants that were inspected at least twice. In contrast to table 10, compliance rates are presented with respect to the inspection order, which allows us to detect changes over time. For example, the third row contains restaurants whose two first inspections occurred in sprints 1 and 2.

There is an increase in compliance between the first and the second inspection from 36% to 55%, which represents an increase of 55 per cent. However, this improvement in the compliance rate is stronger among restaurants that were inspected for the first time by inspectors who had obtained MI training (i.e. in sprint 2 or 3). Compliance rates among the 52 restaurants inspected in sprint 1, before the MI training program, increased from 0.385 to 0.462 (a 20 per cent increase). The improvement in compliance of 0.077 can be seen as a pure inspection effect. Having been inspected induces more compliant behavior, as should be expected. Interestingly, the inspection effect is only present among the 36 restaurants that received a follow-up inspection in sprint 2; compliance among restaurants where follow-up inspections took place in sprint 3 or 4 actually deteriorated. This suggests that the inspection effect is not enduring (although we
should not draw to strong conclusions due to the relatively low number of observations).

<table>
<thead>
<tr>
<th>N</th>
<th>Inspection 1</th>
<th>Inspection 2</th>
<th>Inspection 3</th>
<th>Inspection 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL, INSPECTED AT LEAST TWICE</td>
<td>132</td>
<td>0.356</td>
<td>0.553</td>
<td></td>
</tr>
<tr>
<td>First inspection in sprint 1</td>
<td>52</td>
<td>0.385</td>
<td>0.462</td>
<td></td>
</tr>
<tr>
<td>Second inspection in sprint 2</td>
<td>36</td>
<td>0.278</td>
<td>0.444</td>
<td></td>
</tr>
<tr>
<td>Second inspection in sprint 3</td>
<td>12</td>
<td>0.500</td>
<td>0.417</td>
<td></td>
</tr>
<tr>
<td>Second inspection in sprint 4</td>
<td>4</td>
<td>1.000</td>
<td>0.750</td>
<td></td>
</tr>
<tr>
<td>First inspection in sprint 2 or 3</td>
<td>80</td>
<td>0.338</td>
<td>0.613</td>
<td></td>
</tr>
<tr>
<td>First insp. in S2, second insp. in S3</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>First insp. in S2, second insp. in S4</td>
<td>76</td>
<td>0.342</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>First insp. in S3, second insp. in S4</td>
<td>3</td>
<td>0.333</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td>TOTAL, INSPECTED AT LEAST THRICE</td>
<td>40</td>
<td>0.275</td>
<td>0.400</td>
<td>0.600</td>
</tr>
<tr>
<td>First inspection in sprint 1</td>
<td>39</td>
<td>0.282</td>
<td>0.410</td>
<td>0.615</td>
</tr>
<tr>
<td>Inspected in S1, S2 and S3</td>
<td>7</td>
<td>0.143</td>
<td>0.286</td>
<td>0.429</td>
</tr>
<tr>
<td>Inspected in S1, S2 and S4</td>
<td>28</td>
<td>0.286</td>
<td>0.464</td>
<td>0.679</td>
</tr>
<tr>
<td>Inspected in S1, S3 and S4</td>
<td>4</td>
<td>0.500</td>
<td>0.250</td>
<td>0.500</td>
</tr>
<tr>
<td>First inspection in sprint 2</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>TOTAL, INSPECTED FOUR TIMES</td>
<td>6</td>
<td>0.167</td>
<td>0.333</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Table 5.9. Number of restaurants and compliance rates with respect to inspection

Among the 80 restaurants inspected for the first time in sprint 2 or 3 the compliance rate rose from 0.338 to 0.613 (a 85 per cent increase). In this group the improvement in compliance can be attributed to having been inspected by a MI-trained inspector at the first inspection. Hence, besides the inspection effect a MI effect is also present. The combined effect of a first inspection and a MI-trained inspector amounts to 0.275.

There is a total number of 40 restaurants that were inspected at least thrice. All but one of these was inspected for the first time before the MI training program. Among those 39 restaurants the compliance rate rose by 0.128 after the first inspection in sprint 1 (a 45 per cent increase) and by 0.205 after the second inspection in sprint 2 or 3 (a 50 per cent increase). The first increase represents the pure inspection effect, while the second increase stems from the combined effect of having been inspected twice and having been visited by an MI-trained inspector at the second occasion.

The figure below illustrates the changes in the compliance rate between the first two inspections, among restaurants visited by a MI-untrained inspector at the first inspection (blue line) and among restaurants for which the first inspection was conducted by a MI-trained inspector (red line). Moreover, it also contains the restaurants that were inspected at least thrice where the first inspection was carried
out by a MI-untrained inspector and the second inspection by a MI-trained inspector (green line).

![Graph illustrating inspection and MI effects on the compliance rate](image)

**Figure 5.6. Illustration of inspection and MI effects on the compliance rate (number of restaurants)**

Our results presented in table 5.9 and illustrated in figure 5.6 suggest strong effects from MI training on restaurants’ compliance. By running regressions we can confirm that the MI effect is significant at the 10% level.

### 5.4. Conclusion

The aim of this study was to investigate the effectiveness of inspections on restaurants’ compliance with waste sorting legislation and to explore the effects of MI training of inspectors in this specific setting. Our main conclusions are the following.

*Inspections and MI training of inspectors enhance compliance with waste sorting regulations*

Our results show that the compliance rate of the 132 restaurants that were inspected at least twice rose from 36% at the first inspection to 55% at the second inspection. However, since first inspections were carried out both before and after the MI
training program, this increase in compliance is also affected by inspectors being MI-trained or MI-untrained at the first inspection.

The compliance increased from 38% to 46% between the first two inspections among the 52 restaurants where the first inspection took place before the inspectors had received any MI training. Hence, we observe a pure (first) inspection effect of 8%, i.e. inspections by MI-untrained inspectors led to 20 per cent more restaurants being compliant at the second inspection.

The 80 restaurants whose first two inspections took place after the MI training program increased their compliance rate from 34% at the first inspection to 61% at the second inspection. Thus, there is a combined effect of a (first) inspection and the inspector being MI-trained at the first inspection of 27%, implying that the share of compliant restaurants was boosted by 85 per cent after the first inspection. There were 39 restaurants first inspected before the MI training program and inspected at least two more times. In this group the compliance rate rose from 28% to 41% to 62%. The first increase can be attributed to the pure inspection effect; inspections by MI-untrained inspectors led to a 45 per cent higher compliance rate. The second increase stems from a combined effect of a second inspection and the inspector having completed the MI training program before the second inspection; the share of compliant restaurants rose by 50 per cent between the second and third inspection.

**MI training improves inspectors’ communicative skills**

Our results show that the inspectors have strongly improved their MI skills over time, both at the aggregate and the individual level and for all four quality measures of MI competence. On average, both empathy and evocation increased significantly, from 2.18/5 and 2.11/5 in sprint 1 to 3.36/5 and 3.27/5 in sprint 4. The frequency of MI-non-adherent utterances decreased significantly from 2.20 to 0.37, while the ratio reflections to questions increased significantly from 0.73 to 1.54.

The improvements in MI skills of the present training program are similar to those achieved in our parallel study where the focus was on health safety inspection conversations aiming at inducing property owners to measure and report radon gas radiation results (Wickström et al., 2017). In comparison to the results of the EMT MI training program that targeted a broader range of inspections (Forsberg et al., 2014) improvements in MI skills are stronger.

**Inspectors value the MI training program highly**

Also with regard to inspectors’ perceptions of the MI training program we obtained results similar to the ones in our parallel study (Wickström et al., 2017). Compared to the ratings in the previous EMT study (Forsberg et al., 2016), the perceived usefulness of supervised feedback on recorded conversations and the MI training as a whole were rated almost one unit higher, while the short theory introductions and
the experiential exercises were rated more than one unit higher in the present training program. These ratings were very high and stable across training days, ranging between 5.5/6 and 6/6. As in our parallel study, the usefulness of MI skills for inspection work was graded somewhat lower, but with an increase over time, from 4.43/6 after the first training day to 5.33/6 after the last two training days. Our results confirm that MI training focused on a specific inspection context improves the effectiveness of the training program.

Assessments of inspections improved among inspectors and remained high among restaurants
Regarding the restaurant representatives’ intentions and attitude, the inspectors’ ratings generally increased across sprints, suggesting a positive effect of the MI training program. The strongest and most significant improvement concerned the assessment of restaurant representatives expressing reasons and motivation to sort waste, which could be explained by inspectors’ improved evocation skills. However, the inspectors’ satisfaction with their own performance during the inspection only improved slightly and insignificantly.

Assessments by restaurant representatives were high and stable over time, ranging between 3.4/4 and 3.92/4. Hence, although inspectors perceived a change in attitude among restaurant representatives over time, this does not seem to have influenced the ratings of restaurants’ satisfaction with inspections.

5.5. References


6. Applying motivational interviewing in the context of cattle inspections

Hans Wickström\textsuperscript{21}, Mathias Herzing\textsuperscript{22}, Lars Forsberg\textsuperscript{23}, Adam Jacobsson\textsuperscript{24} and Håkan Källmén\textsuperscript{25}

6.1. Introduction

To ensure adherence to animal welfare legislation Swedish county administrative boards (CABs) conduct inspections of animal husbandry. It is therefore of interest to investigate how efficient these inspections are and how a higher efficiency could be achieved. The purpose of this study is to examine how training of inspectors in the communicative technique \textit{Motivational Interviewing} (MI) impacts on inspections (so called "standard controls") of cattle keepers.

This study has been carried out by the research program \textit{Inspections and Enforcement as an Instrument for Enhancing Environmental Behavior} (“Tillsynen som styrmedel för ett förbättrat miljöbeteende” – TSFM) in collaboration with the Swedish Board of Agriculture (Jordbruksverket) and four CABs (Västra Götaland, Halland, Södermanland and Kronoberg). The impact of MI training has been examined in two further TSFM studies, in the context of inspections regarding radon gas radiation measurements (see section 4) and inspections of restaurants’ waste sorting (see section 5). Previously our research team had tested MI training of inspectors as part of the research program \textit{Efficient Environmental Inspections and Enforcement} (“Effektiv miljötillsyn” – EMT). The EMT study was the first applying MI to an inspection and enforcement context (Forsberg et al., 2016). Since the results were positive the TSFM research program has further developed the MI training program for inspectors and the measuring of its effects, both on inspectors and inspectees (see section 3.2.1).

The MI training program was carried out in the summer of 2015. To enable an assessment of the effect of MI training we collected data from inspections before (2014-15) and after (2015-16) the inspectors participated in the MI training

\textsuperscript{21} Meetme Psykologkonsult AB, Gothenburg, email: hans.wickstrom@kbtanalys.se
\textsuperscript{22} Department of Economics, Stockholm University, email: Mathias.Herzing@ne.su.se
\textsuperscript{23} MicLab AB, Stockholm, email: lars.forsberg@miclab.se
\textsuperscript{24} Department of Economics, Stockholm University, email: aja@ne.su.se
\textsuperscript{25} Karolinska Institutet, Department of Clinical Neuroscience, Center for Psychiatry Research, email: hakan.kallmen@ki.se
program. Moreover, both the inspectors’ and the animal keepers’ perceptions of inspections were assessed through questionnaires. In addition, we also measured the behavior of animal keepers by analyzing inspection reports to assess the extent of non-compliance as well as reports of follow-up inspections that were carried out within 9-14 months after an inspection to check whether measures had been undertaken to achieve compliance.

The study was approved by the Regional Ethical Review Board in Stockholm (reg nr 2014/1417-32). All results regarding animal keepers are presented on an aggregate level, while results of individual inspectors have been de-identified.

The presentation of this study is organized as follows. Section 6.2 describes our methodological approach, while section 6.3 contains the results of the study that have been generated so far. In section 6.4 we provide a brief summary. As we are still awaiting the necessary data to analyze changes in compliance among cattle keepers, this report is preliminary.

6.2. Methodology

6.2.1. Research design

Five CABs were asked about their interest in participating in this study. These were Västra Götaland, Halland, Jönköping, Kronoberg and Södermanland, which were chosen mainly for logistical reasons. All wanted to participate, but Jönköping dropped out at the start of the project due to lack of resources. The study was designed in cooperation with the Swedish Board of Agriculture and the participating CABs. Our intention was to focus on inspections that (i) are sufficiently frequent in all counties; (ii) include sufficiently many inspectors; (iii) are comparable with respect to content and outcome; (iv) generate outcome measures that can be followed up easily and within reasonable time. We agreed on inspections of cattle, mainly because these are standardized to a high degree and are carried out annually and the participating CABs have sufficiently many cattle. However, we were aware that compliance among cattle keepers has been high in the past and hence, a high number of inspection conversations would not be ideal for applying MI. Nevertheless, we chose this context as it provided good opportunities to measure changes in behavior in a consistent way.

At each CAB except Södermanland meetings took place where the MI instructor (Hans Wickström) provided information regarding the study and the training program. In the four participating counties there are about 50 inspectors working with animal welfare. To be included in the study we required the inspectors to have mainly focused on animal welfare inspections for at least one year and most likely would continue to do so in the following two years, as well as to be willing to participate in the MI training program voluntarily. A total number of 33 inspectors satisfied these inclusion criteria (see table 6.1 below). Among these inspectors two
groups were selected for the assessment of MI skills. One group of inspectors participated in the MI training program (the intervention group), while inspectors in the second group did not (the control group), thereby enabling an evaluation of the effects of MI training.

Originally, we intended to let twelve inspectors participate in the MI training program, but this number was reduced to ten as Jönköping withdrew from the project. It was determined that two inspectors in each county should be included in the intervention group, except for Västra Götaland, which as the by far largest county was allowed to let four inspectors participate in the training program. The intention was to have an equal number of inspectors from each county in the intervention and control groups, but as there were only two inspectors in Södermanland that met the inclusion criteria these were automatically assigned to the intervention group and an extra inspector from Västra Götaland was included in the control group. In all counties except Södermanland inspectors were randomly assigned to the intervention and control groups.

<table>
<thead>
<tr>
<th>County</th>
<th>Inspectors satisfying inclusion criteria</th>
<th>Inspectors in intervention group</th>
<th>Inspectors in control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Västra Götaland</td>
<td>21</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Halland</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kronoberg</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Södermanland</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>33</strong></td>
<td><strong>10</strong></td>
<td><strong>9</strong></td>
</tr>
</tbody>
</table>

*Table 6.1. Distribution of inspectors included in our study across CABs*

Animal welfare inspections are conducted annually during the so-called stable period when animals normally are kept indoors (between November 1st and March 31st in this study). Inspections concerning animal husbandry involve huge farms with hundreds of cattle as well as small units with only one animal. For each species there exist specific control points that have been identified by the Swedish Board of Agriculture (SBA, 2012). Regarding inspections of cattle there are 56 control points that have been recommended by the Swedish Board of Agriculture in accordance with the “Regulation (EC) No. 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules” (http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:84005). To facilitate comparability across inspectors and counties we determined in cooperation with the Swedish Board of Agriculture and the participating CABs that 46 of these control points should be checked at all inspections. These included the 38 base control points and further 8 control points for which the non-compliance rate is above 10 per cent nationally or in at least two of the participating CABs (SBA, 2014).
Our study was carried out between November 2014 and December 2016 according to table 6.2 below. During the first stable period, between November 1st, 2014, and March 31st, 2015, all animal keepers were inspected by MI-untrained inspectors. Between April and November 2015 the MI training program was carried out. Hence, the animal keepers were inspected by either MI-trained or MI-untrained inspectors during the second stable period (between November 1st, 2015, and March 31st, 2016). After both stable periods the inspection reports were sent by the CABs to our research team. The CABs also provided information concerning following-up that had been reported by the end of 2015 and by the end of 2016.

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Nov. 2014 – 31 Mar. 2015</td>
<td>Stable period 1: Standard controls of cattle by both the intervention and the control group</td>
</tr>
<tr>
<td>14 Apr. – 2 Nov. 2015</td>
<td>MI training of the inspectors in the intervention group</td>
</tr>
<tr>
<td>31 Dec. 2015</td>
<td>Last date for following-up on inspections in stable period 1</td>
</tr>
<tr>
<td>1 Nov. 2015 – 31 Mars 2016</td>
<td>Stable period 2: Standard controls of cattle by both the intervention and the control group</td>
</tr>
<tr>
<td>31 Dec. 2016</td>
<td>Last date for following-up on inspections in stable period 2</td>
</tr>
</tbody>
</table>

Table 6.2. Schedule of the study

Every year the Swedish Board of Agriculture sends a list with cattle keepers that should be inspected during the next stable period. Cattle keepers are not inspected annually. Hence, none of the cattle keepers in our study was inspected in both stable periods.

Given the list of cattle keepers selected for standard controls, our research team randomly determined which inspector should carry out the inspection. However, an exception was made for inspectors that had to be recused due to previous relations. Moreover, geographical considerations had to be made in Västra Götaland due to its size. Prior to inspection the cattle keepers were notified at most 24 hours in advance. At the beginning of the inspections the cattle keepers were given written information regarding this study and contact details. Inspectors in the intervention group asked for permission to record the conversations and provided information regarding the handling of recordings. Cattle keepers were also told that they would be contacted via phone within one week to answer some questions regarding their perceptions of the inspection. Those who accepted being recorded were asked to provide their written consent.

For the purpose of our study we used the recorded conversations to measure inspectors’ MI skills before and after the MI training program (see section 6.2.3).
Moreover, results from questionnaires regarding the inspectors’ assessments of the MI training days as well as their assessments of inspection conversations are also used to evaluate the inspectors’ experience of MI practice (see sections 6.2.4 and 6.2.5). In addition, the animal keepers’ assessments of inspection conversations were assessed through telephone interviews within one week after the inspection (see section 6.2.6). Finally, our intention is to also determine animal keepers’ rate of compliance with the control points selected for inclusion in our study; however, we are still waiting for the necessary data and are therefore not yet able to report any results.

6.2.2. The MI training program

The intervention group consisted of eight women and one man, between 30 and 54 years old (average: 41 years, median: 41 years), of whom one had a high school degree, seven had a university degree in natural sciences and one had a doctoral degree in natural sciences. In the control group there were seven women and two men, between 27 and 48 years old (average: 34 years, median: 29 years), who all had a university degree in natural sciences. The inspectors had worked with inspections and enforcement 5-23 years (average: 12 years, median: 8 years) in the intervention group and 2-9 years (average: 5 years, median: 5 years) in the control group. All had worked with inspections and enforcement during the last twelve months. Among inspectors in the intervention group all had worked mainly with animal welfare inspections. In the control group two inspectors had primarily worked with agricultural inspections, but secondarily worked with animal welfare inspections. In both groups five inspectors had received communicative training, on average 1.5 days in the intervention group and 3 days in the control group. One person in the control group had received MI training for half a day; no one else had any earlier experience of MI.

The inspectors in the intervention group were divided into two groups with four (two from Kronoberg and Västra Götaland, respectively) and six (two from Halland, Södermanland and Västra Götaland, respectively) participants. The purpose of training two smaller groups rather than one large group was to increase the interactivity among inspectors.26 The training program was carried out for six whole days at five occasions spread out over a half year period. The two first training sessions took place on consecutive days. The intervals between the other training days were 1-3 months. The first day was devoted to theoretical orientation and practical applications. The other training days consisted of three hours of theory and practical applications and three hours of feedback on recorded conversations. The training program was implemented in collaboration with the coding laboratory MIC Lab (see section 3.2.2 for more details). At the end of the first day a conversation with a colleague about change was recorded and feedback was given by the trainer on the second day. Before the third to sixth training days each inspector recorded three inspections (either standard controls of other animals

26 Since some exercises are done in pairs, the split into two groups was 6:4 and not 5:5.
or follow-up inspections of cattle), of which two were uploaded to MIC Lab’s homepage to be coded for use as feedback and not to be included in the study. The coders provided written comments explaining the numbers in the protocol in order to facilitate the understanding of feedback. The feedback focused on earlier and current themes. Each inspector made one verbatim transcription of a recorded inspection that was used as a pedagogic tool with the purpose of exemplifying MI concepts and MI skill applications. Table 6.3 below summarizes the themes covered by the MI training program.

<table>
<thead>
<tr>
<th>TRAINING DAY</th>
<th>MI THEME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>To listen and to communicate empathy and collaboration</td>
</tr>
<tr>
<td></td>
<td>– To understand the meaning of MI, four processes</td>
</tr>
<tr>
<td></td>
<td>– To direct towards a change goal</td>
</tr>
<tr>
<td></td>
<td>– To use empathic listening</td>
</tr>
<tr>
<td></td>
<td>– To convey collaboration and equality</td>
</tr>
<tr>
<td></td>
<td>– To emphasize autonomy</td>
</tr>
<tr>
<td></td>
<td>– To avoid MI-non-adherent utterances</td>
</tr>
<tr>
<td>Day 2</td>
<td>To use communication skills</td>
</tr>
<tr>
<td></td>
<td>– To affirm, ask questions, make reflections and summarize</td>
</tr>
<tr>
<td></td>
<td>– To use the coding results as feedback</td>
</tr>
<tr>
<td>Day 3</td>
<td>To strengthen reinforces toward positive behavior</td>
</tr>
<tr>
<td></td>
<td>– To recognize, elicit and strengthen change talk</td>
</tr>
<tr>
<td>Day 4</td>
<td>To make efforts to understand the perspectives of the other</td>
</tr>
<tr>
<td></td>
<td>– To exchange information in a dialogue</td>
</tr>
<tr>
<td></td>
<td>– To understand and implement positive and negative reinforcement</td>
</tr>
<tr>
<td>Day 5</td>
<td>To meet ambivalence and resistance</td>
</tr>
<tr>
<td></td>
<td>– To explore readiness to change and ambivalence</td>
</tr>
<tr>
<td></td>
<td>– To meet and roll with resistance/dissonance</td>
</tr>
<tr>
<td></td>
<td>– To go from eliciting change talk to planning</td>
</tr>
<tr>
<td>Day 6</td>
<td>To use MI in daily inspection practice</td>
</tr>
<tr>
<td></td>
<td>– To summarize the MI training program</td>
</tr>
<tr>
<td></td>
<td>– To form a personal plan for upholding MI-proficiency</td>
</tr>
<tr>
<td></td>
<td>– To use skills in roleplaying</td>
</tr>
</tbody>
</table>

Table 6.3. MI themes for each of the six days of MI training

The choices of MI themes for each training day were made according to common MI practice (Miller & Moyers, 2007) and built on our experiences from the MI training programs of EMT (Forsberg et al., 2014 and 2016) and our recent studies on the application of MI to induce radon gas radiation measurements by property owners (see section 4) and to enhance compliance with waste sorting regulations (see section 5). In addition to the theory and the MI skills described by Miller & Rollnick (2013) the training program also included a brief introduction to fundamental reinforcement principles within psychological learning theory (Sundel & Sundel, 2005). All inspectors participated every training day, except one
inspector who was absent one day due to illness and one inspector who had a higher absence. The latter was not considered in our analysis due to a too low participation rate.

6.2.3. Assessments of inspectors’ MI skills

We employed a pre-post design based on the assessments of all ten inspectors’ MI skills before and after the training program. To that end the intervention group recorded inspections in both stable periods with the informed consent of the farmers. In total there were 493 recorded conversations, 271 before the MI training program started and 222 after it had been completed (see table 6.4). All recordings were made at first inspections during the stable period, i.e. following-up inspections were not recorded. However, one of the inspectors changed to another unit after the first stable period and hence, no longer worked with animal welfare inspections during the second stable period; for this person we used recordings of health safety inspections to evaluate MI skills.

We randomly selected a maximum number of 120 recordings per stable period for coding (see section 3.2.2 for details regarding the coding process). Hence, 12 recordings per inspector were coded before and after the MI training program. However, since two inspectors only recorded 11 inspections during the second stable period, 13 conversations of two other inspectors were coded. Since the animal keepers were identified when recorded inspections were coded, it is possible to link an inspector’s MI skills in an inspection to farmers’ compliance with follow-up measures.

<table>
<thead>
<tr>
<th>Inspector</th>
<th>Recordings stable period 1</th>
<th>Recordings stable period 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Coded</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>40</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>35</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>271</strong></td>
<td><strong>120</strong></td>
</tr>
</tbody>
</table>

Table 6.4. Number of total and coded recordings per stable period

In the present study the first 20 minutes of the inspection was selected for coding or the whole inspection if it was shorter. For every recording the coder needs to know the target behaviors, i.e. the behaviors that the inspector is trying to
influence. Since inspections covered 46 control points in the present study, there were 46 target behaviors that coders had to consider (although a single inspection did not necessarily address all control points, every single control point could potentially be discussed).

6.2.4. Inspectors’ assessments of MI training and inspections

After every training day the inspectors were asked to answer a questionnaire, as described in section 3.2.3. In total 56 questionnaires were submitted during the course of the training program.

Inspectors in both the treatment and the control groups also answered a questionnaire immediately after inspections, for which recording had been permitted and the duration was at least ten minutes (see section 3.2.4 for more details). In total, 657 questionnaires were submitted, 366 before and 291 after the MI training program.

6.2.5. Cattle keepers’ perceptions of inspections

The inspectors in both the intervention and control groups informed all cattle keepers that they would be contacted by phone within one week and asked for permission to answer questions regarding their perceptions of the inspection. These interviews were carried out by one person in stable period 1 and another person in stable period 2. Included in the questionnaire were three questions asking whether it was clear what the CAB expects (rated on a four-grade scale from 1 = “entirely” to 4 = “not at all”), whether measures needed to be undertaken (yes/no) and if so, whether the animal keeper intended to do so (yes/partially/no). The interviewers also asked four questions regarding the perceptions of the inspection, which were rated on a four-graded scale from 1 (“not satisfied”) to 4 (“satisfied”).

The numbers of completed interviews were 159 in the intervention group and 56 in the control group in stable period 1; in stable period 2 the corresponding numbers were 187 and 72 (see table 6.6). A source of uncertainty is that all inspectors have not informed the phone interviewer about all inspections, that the interviewer has not been able to contact all animal keepers and that some animal keepers have declined to participate in the survey.

<table>
<thead>
<tr>
<th></th>
<th>Intervention group (9 inspectors)</th>
<th>Control group (9 inspectors)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stable period 1</td>
<td>Stable period 2</td>
</tr>
<tr>
<td>Reported controls</td>
<td>271</td>
<td>238</td>
</tr>
<tr>
<td>Contacted</td>
<td>220 (81%)</td>
<td>199 (84%)</td>
</tr>
<tr>
<td>Contacted successfully</td>
<td>160 (83%)</td>
<td>191 (96%)</td>
</tr>
<tr>
<td>Accepted interview</td>
<td>159</td>
<td>187</td>
</tr>
</tbody>
</table>

Table 6.6. Summary statistics regarding the animal keepers’ questionnaire
6.2.6. Compliance of cattle keepers

During the inspection compliance with each of the selected 46 control points was assessed. Following previous routines, the inspectors sent an inspection report to the animal keepers within one week. This report summarizes control points for which non-compliance has been established and the measures to deal with these that have been discussed during the inspection. Any demands for measures to attain compliance are followed up, either through an inspection or by requesting the animal keeper to provide documentation regarding the measures that have been undertaken. Alternatively, it suffices to orally agree on the necessary measures at the inspection and having no following-up.

To obtain outcome measures we collected all inspection reports. Information regarding following-up was collected to enable an assessment of the degree to which the animal keepers had undertaken measures dealing with violations in accordance with the inspection reports. In our study we considered measures that could reasonably be dealt with before December 31st in the year when the stable period ends, i.e. 9-14 months after inspection. Hence, measures requiring longer time or not being followed up before December 31st were not taken into account. The following-up that takes place after the second stable period (December 31st, 2016, at the latest) is used to evaluate any difference between animal keepers that were inspected by MI-trained inspectors and animal keepers that were inspected by inspectors in the control group.

For simplicity we only establish whether a control point is complied with or not, i.e. the degree of a violation is not considered. Hence, while we lose some information, our analysis becomes less sensitive to subjective assessments and variations between inspectors that these can give rise to. However, by counting the number of control point violations we obtain an approximate measure for the degree of non-compliance for each animal keeper.

6.2.7. Statistical analysis

The statistical analysis was conducted in the SPSS software environment. All comparisons between means were made by using T-tests for independent groups. Although the inspectors’ MI skills were measured before and after training, we consider the assessments as independent because these were made on different inspections.

6.3. Results

6.3.1. MI skills

Table 6.7 shows the means and standard deviations for the four summary MI skills variables before and after the MI training program. Due to low participation and health problems, one of the inspectors was excluded from the study and hence, we have included measures of nine inspectors in our analysis. At the aggregate level
the inspectors did not improve significantly on any measure; among other measures there was a significant improvement only in the number of MI-adherent utterances from 1.34 to 2.47.

<table>
<thead>
<tr>
<th></th>
<th>Stable period 1</th>
<th>Stable period 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of coded conversations</strong></td>
<td>108</td>
<td>109</td>
</tr>
<tr>
<td><strong>Empathy</strong></td>
<td>2.19 (0.72)</td>
<td>2.12 (0.77)</td>
</tr>
<tr>
<td><strong>Evocation</strong></td>
<td>1.99 (0.65)</td>
<td>2.27 (0.52)</td>
</tr>
<tr>
<td><strong>MI-non-adherent utterances</strong></td>
<td>2.07 (2.59)</td>
<td>2.44 (2.92)</td>
</tr>
<tr>
<td><strong>Ratio reflections/questions</strong></td>
<td>0.95 (0.89)</td>
<td>0.83 (0.67)</td>
</tr>
</tbody>
</table>

Table 6.7. Summary measures of MI skills in stable periods 1 and 2, means (standard deviations) and p-values, for the whole group of nine inspectors

6.3.2. Inspectors’ perceptions of the MI training

Table 6.8 summarizes the responses to the questionnaire that the inspectors were asked to hand in after each training day.

<table>
<thead>
<tr>
<th>MI TRAINING DAY QUESTIONNAIRE</th>
<th>D1</th>
<th>D2</th>
<th>D327</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of training day inspectors</td>
<td>10</td>
<td>10</td>
<td>9/10</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>56</td>
</tr>
<tr>
<td>Number of completed questionnaires</td>
<td>10</td>
<td>10</td>
<td>9/10</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>56</td>
</tr>
<tr>
<td>1. How would you rate the usefulness of feedback on recorded conversations (afternoon)?</td>
<td>5.44</td>
<td>5.00</td>
<td>5.25</td>
<td>5.60</td>
<td>5.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. How would you rate the usefulness of theory reviews and exercises (morning)?</td>
<td>4.80</td>
<td>5.30</td>
<td>5.10</td>
<td>4.89</td>
<td>5.50</td>
<td>5.70</td>
<td>5.22</td>
</tr>
<tr>
<td>3. How would you rate the MI training as a whole?</td>
<td>4.80</td>
<td>5.40</td>
<td>5.20</td>
<td>4.95</td>
<td>5.50</td>
<td>5.60</td>
<td>5.24</td>
</tr>
<tr>
<td>4. Will you be able to apply your new MI-skills in your inspection work?</td>
<td>3.90</td>
<td>5.00</td>
<td>4.90</td>
<td>4.89</td>
<td>5.00</td>
<td>5.40</td>
<td>4.85</td>
</tr>
</tbody>
</table>

Table 6.8. Responses to questions about the MI training days on a six-graded scale, per day and totally, averages for all inspectors28

On all measures there was an increase between days 1 and 2, followed by a decrease until day 4 and a strong increase between days 4 and 6, resulting in higher scores at the end compared to the start of the training program (the usefulness of feedback on recorded conversations was not assessed before day 3). The MI

27 On training day 3 one inspector participated only in the morning.

28 Since training day questionnaires were handed in anonymously, answers from all participating inspectors are included, i.e. also responses from the inspector who was excluded from our evaluations of MI skills due to a too low participation rate.
training as a whole was on average rated at 5.24/6; average ratings were 4.8 after the first day and 5.6 after the final day. The perceived usefulness of the theory reviews and the practical exercises were assessed almost identically every training day and at 5.22/6 on average. The usefulness of feedback on recorded inspections scored 5.32/6 on average and reaching 5.6 after the last training day. As in our previous studies the applicability of MI skills in inspection work was rated at a lower level, 4.85/6 on average, but with a more pronounced increase from 3.9/6 after the first day to 5.4/6 after the final day.

6.3.3. Inspectors’ perceptions of inspections

Immediately after every inspection inspectors in both the treatment and control group answered a questionnaire. In total, 366 inspection questionnaires (259 by the intervention group and 111 by the control group) were submitted before and 291 after the MI training program (209 by the intervention group and 82 by the control group). The results are summarized in table 6.9. Note that questions 4 and 5 were only relevant at some inspections; the number of answers are reported in brackets.

<table>
<thead>
<tr>
<th>INSPECTORS’ PERCEPTIONS OF INSPECTIONS WITH CATTLE KEEPERS</th>
<th>Intervention group (9 inspectors)</th>
<th>Control group (9 inspectors)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SP1</td>
<td>SP2</td>
</tr>
<tr>
<td>Number of submitted questionnaires</td>
<td>259</td>
<td>209</td>
</tr>
<tr>
<td>1. My perception at the start of the inspection was</td>
<td>3.55</td>
<td>3.41</td>
</tr>
<tr>
<td>that the cattle keeper had a positive attitude to it.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. My perception at the end of the inspection was</td>
<td>4.08</td>
<td>4.08</td>
</tr>
<tr>
<td>that the cattle keeper had sufficient knowledge of its</td>
<td></td>
<td></td>
</tr>
<tr>
<td>importance for promoting good animal care.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. My perception is that the cattle keeper clearly showed</td>
<td>4.07</td>
<td>4.01</td>
</tr>
<tr>
<td>that he/she understood the information I wanted to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>convey during the inspection.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. My perception is that the cattle keeper will undertake</td>
<td>4.20</td>
<td>4.16</td>
</tr>
<tr>
<td>measures to achieve compliance. (number of inspectees for</td>
<td>(n=148)</td>
<td>(n=98)</td>
</tr>
<tr>
<td>whom applicable)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. My perception is that the cattle keeper expressed his/</td>
<td>3.42</td>
<td>3.60</td>
</tr>
<tr>
<td>her own reasons and motivation to undertake measures to</td>
<td>(n=153)</td>
<td>(n=98)</td>
</tr>
<tr>
<td>achieve compliance. (number of inspectees for whom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>applicable)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I am satisfied with my own performance during the</td>
<td>3.78</td>
<td>3.77</td>
</tr>
<tr>
<td>inspection.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. My perception is that the cattle keeper had a positive</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>attitude to the inspection at the end of it.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.9. Inspectors’ assessments of inspections of cattle keepers on a five-grade scale in stable periods 1 and 2, averages
There was a significant difference between the intervention and control groups during both stable periods for all seven assessments. Generally inspectors in the intervention group gave lower ratings than the control group. Responses to question 1 displayed a small, but significant deterioration between the first and the second stable period in both groups, i.e. inspectors in both the intervention and the control groups perceived the attitude at the start of the inspection to be more positive in the first stable period.

### 6.3.4. Cattle keepers’ perceptions of inspections

Our analysis of the interviews with animal keepers generally confirm the findings in the preceding section (see table 10). There is no significant difference between the control and the intervention groups, and there is also no significant difference between the two stable periods. However, the intervention group attains slightly higher ratings regarding attitude, competence, efficiency and information both before and after the MI training program.

<table>
<thead>
<tr>
<th>CATTLE KEEPERS’ PERCEPTIONS OF THE INSPECTIONS</th>
<th>Intervention group (nine inspectors)</th>
<th>Control group (nine inspectors)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SP1</td>
<td>SP2</td>
</tr>
<tr>
<td>Number of questionnaires</td>
<td>187</td>
<td>72</td>
</tr>
<tr>
<td>2. Did it during the inspection become clear to you what the CAB expects from you with regard to the control points (1 = “entirely” – 4 = “not at all”)</td>
<td>1.51</td>
<td>1.69</td>
</tr>
<tr>
<td>3.1 Did the inspector during the inspection mention anything that needs to be done? (1 = “yes”, 2 = “no”)</td>
<td>1.48</td>
<td>1.56</td>
</tr>
<tr>
<td>3.2 If yes, do you intend to undertake measures to conform with the issues that the CAB brought to your attention? (1 = “yes”, 2 = “partially”, 3 = “no”)</td>
<td>1.09</td>
<td>1.06</td>
</tr>
<tr>
<td>4. How satisfied are you with the inspector’s attitude? (1 = “not satisfied” – 4 = “satisfied”)</td>
<td>3.89</td>
<td>3.89</td>
</tr>
<tr>
<td>5. How satisfied are you with the inspector’s competence? (1 = “not satisfied” – 4 = “satisfied”)</td>
<td>3.59</td>
<td>3.63</td>
</tr>
<tr>
<td>6. How satisfied are you with the inspector’s efficiency? (1 = “not satisfied” – 4 = “satisfied”)</td>
<td>3.74</td>
<td>3.70</td>
</tr>
<tr>
<td>7. How satisfied are you with the information that you received? (1 = “not satisfied” – 4 = “satisfied”)</td>
<td>3.62</td>
<td>3.55</td>
</tr>
</tbody>
</table>

**Table 6.10.** Cattle keepers’ assessments of the inspections by inspectors in the intervention and the control groups in both stable periods (SP1 and SP2), averages

### 6.4. Concluding remarks

Our preliminary results show that the inspectors who participated in the training program did not improve their MI skills, in sharp contrast to the other studies of the TSFM research program (see sections 4 and 5). Nevertheless, the participating
inspectors’ assessments of the training program were high and increasing over time, although not reaching the very high levels in the two other studies. In particular the perceived usefulness of MI at inspection work increased from 3.9/6 after the first training day to 5.4/6 after the last training day. Meanwhile, cattle keepers rated the inspectors’ attitude very high, at 3.89/4 in the intervention group both before and after the MI training program.

There are several potential reasons for why the present training program did not yield the same positive results as in our other studies. One explanation is that inspectors did not learn the MI method, which might be due to difficulties in adapting MI training to animal welfare inspections.

Cattle inspections differ from the inspection contexts in our two other studies in several respects. In study A1 the MI training program was adapted to inspectors contacting property owners via phone to induce them to carry out radon gas radiation measurements. The MI training program in study A2 targeted inspectors controlling restaurants’ waste sorting system. Both these inspection settings were very specific. In contrast, inspections in the present study covered 46 control points. Hence, it was much more difficult to adapt the MI training skills to all target behaviors.

Due to the high number of control points it was not obvious for inspectors which changes in cattle keepers’ behavior to target at inspections, thus making it more difficult to achieve good results in terms of MI skills. In conversation covering a wide range of issues it is more challenging to be focused, which also has repercussions for the coding results. It is thus conceivable that coders found it hard to identify and classify utterances of inspectees, such that assessments of MI skills through coding became less reliable.

Moreover, these inspections were much longer (sometimes more than two hours) than in our previous studies, sometimes also covering other issues than cattle keeping. As inspections take place at cattle keepers’ homes, conversations often also include non-professional topics (i.e. small talk). Our measurements of MI skills are based on 20 minutes of inspection conversation. Some inspectors indicated which sequence of the conversation should be coded; when such information was missing the last 20 minutes of the recording was coded. In contrast, the first 20 minutes of inspection conversations were coded in our previous studies. Hence, coding in the present studies was less consistent, which might also have affected its reliability negatively. MI is intended for use in focused conversations – at inspections of two hours duration where many issues are discussed it is obviously difficult to identify the sequences that are most relevant for analyzing the inspectors’ MI skills in terms of influencing target behaviors of inspectees.
Generally our results thus indicate that it is harder to apply and assess MI skills in the context of animal welfare inspections. The purpose of the three TSFM MI studies was to explore the adaptability of MI in different inspection settings. Based on the outcomes of these studies we can conclude that the specific inspection conditions are essential for how much inspectors increase their MI skills. If inspection focus on one specific issue, as in our studies on radon gas radiation measurements and waste sorting, MI training leads to improvements in communicative skills that exceed those achieved in a training program targeting different types of inspections, as in the EMT study. If, however, inspections cover a wide range of potential target behaviors (control points) and possibly also other issues as in the present study, it is much harder to pick up MI skills improvements. It is important to emphasize that we do not yet have all results from this study. We do not know whether there are any differences between the two stable periods in terms of the outcome of inspections, i.e. the degree to which measurements have been undertaken to deal with violations in accordance with the inspection reports.

6.5. References


SBA (2014). *Vågledning vid djurskyddskontroller i samband med forskningsprojektet ”Motiverande samtal”* [“Guidance for animal welfare controls in connection with the research project ’Motivational Interviewing’”]. Department for animal welfare and health, Swedish Board of Agriculture (Jordbruksverket).
7. The effectiveness of environmental inspections in oligopolistic markets\textsuperscript{29}

Jonas Häckner and Mathias Herzing\textsuperscript{30}

\textsuperscript{29} This section contains a pre-print version of an article, which was published in \textit{Resource and Energy Economics}, volume 48, pp. 83-97, in May 2017 (https://doi.org/10.1016/j.reseneeco.2017.03.001).

\textsuperscript{30} Department of Economics, Stockholm University, e-mail: Jonas.Hackner@ne.su.se, Mathias.Herzing@ne.su.se
7.1 Introduction

Firms in general have strong incentives to try to reduce marginal cost. Often, this benefits society as a whole and consumers in particular. However, sometimes cost reductions are attained by various illegal activities such as e.g. the use of dirty production technologies. To balance firms’ incentives to pollute environmental inspection agencies monitor firms and punish breaches of legislation. Following Becker (1968), a firm therefore has to weigh the benefits from saving abatement costs against the risk of non-compliance being detected and punished.

If resources available for inspections and enforcement were sufficiently large, fines for violations of legislation were sufficiently severe and inspection technologies were sufficiently efficient, it should be possible to induce compliant behavior in any industry subject to environmental regulation. In reality this is seldom the case. The starting point of our study is that the incentives for violating the law differ across industries and that the inspection agency does not have enough resources to achieve full compliance in all industrial sectors.

We apply the Häckner (2000) model of oligopolistic Cournot competition where the representative consumer’s utility function is quadratic such that demand for differentiated goods is linear. It is assumed that firms have perfect information regarding market conditions and hence, react strategically in case a competitor saves costs by violating the law. Naturally the cost for inducing compliance is related to the gain from unilaterally breaching legislation, which allows us to relate the social benefits of inspections to this cost. More specifically, we account for three potential channels, through which inspection activities affect welfare. First, monitoring of firms is obviously associated with inspection costs. Second, inspections are carried out to achieve compliance with legislation and hence, there are environmental gains. Third, as abatement is likely to increase marginal costs of production, successful inspections will lead to economic losses, because welfare is reduced due to lower market surpluses.\(^{30}\)

Naturally the weight attributed to the economic consequences of inspections may vary across agencies, for example due to the institutional conditions under which they operate. On the one hand, an independent agency may focus solely on the environmental impact of inspections. On the other hand, an agency may be under pressure (e.g. through lobbying) to also consider the market effects of inspections. Given the relative

\(^{30}\)It is conceivable that clean production processes are associated with increases in both fixed and variable costs, but here we focus entirely on the latter effect. Obviously, adding fixed costs would introduce a strong link between the level of these costs and market structure.
weight of surplus losses vis-à-vis environmental gains, the net social benefit in relation to the costs for inducing compliance can be calculated for any market. We thus obtain a measure for the effectiveness of inspections that varies with respect to the following market characteristics:

1) The degree of market concentration;
2) The degree of product differentiation;
3) Shifts in demand;
4) Cost savings from breaching legislation, i.e. abatement costs.

Analyzing the effect of market conditions on the social benefit per dollar spent on inspections enables us to rank sectors with regard to the effectiveness of resources allocated to inducing compliant behavior. Since we assume firms in any given industry to be identical, the agency has to decide between carrying out sufficiently many inspections to deter violations and not monitoring that sector (in which case the environment is harmed, but inspection costs are saved). Hence, our framework can be used as guidance for an agency, which has to determine where to allocate inspection resources that are not sufficient to ensure compliance in all industries subject to environmental legislation.

Our main findings are the following. First, as long as the weight put on economic losses is not too large, inspection agencies should target firms operating on markets with low competitive pressure, i.e. concentrated markets and markets where product differentiation is substantial. Second, inspection agencies should focus on sectors with high abatement costs and weak demand, unless competition is fierce and the relative weight of economic losses is sufficiently small. Third, the higher the weight attributed to surplus losses, the more competitive should the markets targeted by the inspection agency be.

The literature on environmental inspections and enforcement, which has been surveyed by e.g. Cohen (1998) and Heyes (2000), addresses a wide range of issues. Models typically include firms subject to environmental legislation interacting with an agency that optimizes its inspection strategy, e.g. to minimize environmental impact or to maximize social welfare. For example, Malik (1990), Heyes (1994) and Innes (2001) focus on firms’ efforts to avoid detection of non-compliance, while e.g. Nowell and Shogren (1994) model firms’ contesting of fines for violations.

demonstrate that the disclosure of information may lower emissions and enhance social welfare. Langpap (2008) presents a model where self-reporting may enhance private enforcement through citizen suits, such that regulatory and social costs of enforcement are reduced. Other contributions account for the possibility that regulatory outcomes may be influenced by lobbying (see e.g. Cheng and Lai, 2012).

Most inspection models assume that firms are distributed over a continuum regarding one variable (e.g. abatement costs or the detectability of non-compliant behavior). Implicitly it is thus assumed that firms are myopic with respect to market outcomes, i.e. markets are perfectly competitive such that cost savings due to breaches of legislation increase a firm’s profit without influencing market prices and market shares. In contrast, our model accounts for the fact that the degree of competition may vary across markets; hence, non-compliance does not only lower a firm’s costs, but also affects the strategic interaction with its competitors.31

In recent years the efficiency of inspections and the effective use of available resources has received increasing attention by several international organizations. For example, OECD (2014) provides guidelines for regulatory policy practices, emphasizing the importance both of “responsive regulation” of specific businesses (pp.33-35) and of “concentrating resources and efforts where they can deliver the most results” (p.38). These recommendations originate in the insight that resources for monitoring activities are limited and insufficient for achieving full compliance across all industries, which is also the underlying assumption that we make. Moreover, OECD (2014) suggests that “in the absence of comprehensive data and/or fully reliable data, regulatory enforcement agencies should rely on interpreting what data exists (at least to establish which sectors appear to generate the most damage)” (p.28). Our set-up yields some simple rules of thumbs based on a few industry characteristics, which can be applied to allocate inspection resources effectively.

The present model has a focus that is similar to studies that attempt to determine optimal regulatory policy under a budgetary constraint. The impact of different auditing mechanisms on firms’ emissions and self-reporting has recently been analyzed by Oestrich (2015). Cason et al. (2016) and Gilpatric et al. (2011) employ similar models to compare random auditing with tournament audit mechanisms and also present results from experiments that generally confirm their theoretical predictions. Macho-Stadler and Pérez-Castrillo (2006) present a model where firms first choose emissions and then decide the self-

31See also Häckner and Herzing (2012) who analyze how the incentives for unilaterally breaching legislation are influenced by market conditions.
reported emission level, on the basis of which its environmental taxes are determined. Under-reporting emissions saves taxes, but leads to a fine if detected. Firms differ either with respect to the detectability of under-reported emissions or regarding the economic benefit from emissions. For a given budget the inspection agency determines the audit probability for each firm to minimize total emissions. In Bontems and Bourgeon (2005) a regulator designs environmental policy to maximize welfare under asymmetric information regarding firms’ abatement costs. Through costly auditing it is possible to verify whether the self-reported type of a firm is consistent with its emissions, but not to determine whether the firm has reported its type truthfully. It is shown how monitoring alters the incentives of firms to overestimate abatement costs and that, compared to the perfect information case, the optimal policy may be to over- or under-deter firms.

Both Bontems and Bourgeon (2005) and Macho-Stadler and Pérez-Castrillo (2006) focus on the optimal implementation of environmental regulation in the presence of heterogeneous firms, such that each firm can be targeted individually. In contrast, firms in our model are homogeneous, while industries differ with respect to market characteristics. Information regarding firms and markets is perfect, but given that it is not possible to deter non-compliance in all industries due to budgetary constraints an inspection agency will have to decide which type of activities to monitor. Thus, our aim is to draw conclusions regarding the effectiveness of inspections across markets rather than to derive a scheme for the optimal allocation of monitoring resources in a given market.

The paper is organized as follows. In section 2 the basic model is presented and equilibria are characterized. Section 3 presents the costs and benefits of environmental inspections. The impact of market characteristics on the effectiveness of inspections is analyzed in section 4. The paper concludes with some final remarks in section 5. The Appendix includes technical proofs, as well as a list of all parameters and variables that are used in our model (section 7.7.4).

32 See also Macho-Stadler (2008) who focuses on total emissions when different regulatory instruments, i.e., standards, taxes and tradeable permits, are implemented, and Macho-Stadler and Pérez-Castrillo (2010) who analyze how monitoring affects the incentives of firms to adopt cleaner technologies.
7.2 The Model

The basic model is taken from Häckner (2000) where the utility function is of the following type:

\[
U(q, I) = \alpha \sum_{i=1}^{n} q_i - \frac{1}{2} \left( \sum_{i=1}^{n} q_i^2 + 2\gamma \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} q_i q_j \right) + I. \tag{1}
\]

Thus, utility is quadratic in the consumption of \( q \)-goods and linear in the consumption of other goods \( I \). The parameter \( \gamma \in [0, 1] \) measures the substitutability between the products. If \( \gamma = 0 \), each firm has monopolistic market power, while if \( \gamma = 1 \), the products are perfect substitutes.\(^{33}\) Finally, \( \alpha \) is a demand shifting parameter measuring the cutoff price, which for simplicity is assumed to be the same for all differentiated products. Consumers maximize utility subject to the budget constraint \( \sum p_i q_i + I \leq m \), where \( m \) denotes income and the price of the composite good is normalized to one. The first-order condition determining the inverse demand of good \( k \) is

\[
\frac{\partial U}{\partial q_k} = \alpha - q_k - \gamma \sum_{j \neq k} q_j - p_k = 0. \tag{2}
\]

Firms compete in Cournot fashion. For analytical tractability, it is assumed that firms are perfectly symmetric. Under compliance all firms incur the constant marginal cost \( c > 0 \). If a firm violates legislation by using e.g. a low-cost, high-pollution technology, its marginal cost is reduced to \( c_0 < c \). The marginal cost saving \( c - c_0 \) can be thought of as the marginal abatement cost that is saved by a non-compliant firm.

There are basically two ways to analyze the effect of one firm breaching legislation and thereby lowering its constant marginal cost to \( c_0 \). On the one hand, such a unilateral violation may come as a surprise to the competing compliant firms and hence, there will be no strategic response to the violation. On the other hand, considering that real world firms learn over time, it seems reasonable to assume that competing firms are able to infer from observed prices that a violation has taken place and hence, react strategically. For tractability we assume that firms are perfectly informed about market conditions, i.e. there is no uncertainty regarding observed prices, and hence compliant firms will adjust their behavior to a unilateral deviation.\(^{34}\)

\(^{33}\)A negative \( \gamma \) implies that the goods are complementary, a scenario that will not be further analyzed in this study.

\(^{34}\)Naturally firms could respond to a unilateral deviation by also breaching legislation. However, our focus is on the gain from unilaterally violating legislation, which provides a measure for how strong deterrence has to be to induce compliance.
If firm \( k \) unilaterally deviates from compliance, the other firms will adapt by reducing production. There are two possible outcomes, either an interior solution with all firms continuing to produce some quantity, or a corner solution, where it is optimal for the other firms to cease production.

Given that firms \( i \neq k \) do not cease production, the profit maximizing quantity of firm \( k \) equals (see the Appendix, 7.7.1.1, for derivation)

\[
q_k^* = \varphi \left[ (\alpha - c) + (c - c_0)\Psi \right],
\]

where \( \varphi \equiv \varphi(\gamma, n) > 0 \) and \( \Psi \equiv \Psi(\gamma, n) > 0 \). More specifically,

\[
\varphi(\gamma, n) = \frac{1}{2 + (n - 1)\gamma}, \quad \Psi(\gamma, n) = \frac{2 + (n - 2)\gamma}{2 - \gamma}.
\]

It is easily verified that \( \frac{\partial \varphi}{\partial \gamma} < 0 \) for \( n \geq 2 \) (\( \varphi(\gamma, 1) = \frac{1}{2} \)) and \( \frac{\partial \varphi}{\partial n} < 0 \) for \( \gamma > 0 \) (\( \varphi(0, n) = \frac{1}{2} \)). Moreover, \( \frac{\partial \Psi}{\partial \gamma} > 0 \) for \( n \geq 2 \) (\( \Psi(\gamma, 1) = 1 \)) and \( \frac{\partial \Psi}{\partial n} > 0 \). Thus, \( \Psi \) can be interpreted as a measure for the impact of the degree of market competition on the profitability of obtaining a cost advantage. A higher number of firms or a lower degree of product differentiation increase the value of \( \Psi \); a more competitive market increases the scope for taking advantage of unilaterally lowering costs by violating legislation.

The profit maximizing price of firm \( k \) equals (see the Appendix, 7.7.1.1, for derivation)

\[
p_k^* = c_0 + q_k^* = c_0 + \varphi \left[ (\alpha - c) + (c - c_0)\Psi \right],
\]

It immediately follows from (3) and (4) that firm \( k \)'s profit is given by

\[
\pi_k^* = \varphi^2 \left[ (\alpha - c) + (c - c_0)\Psi \right]^2.
\]

An increase in competitive pressure due to a higher number of firms or a lower degree of product differentiation has two opposing effects on profits. On the one hand, the value of \( \varphi \) decreases, which impacts negatively on profits. On the other hand, the incentives to violate legislation are strengthened as \( \Psi \) increases.

The profit maximizing quantity and price of firms \( i \neq k \) equal (see the Appendix, 7.7.1.2, for derivation)

\[
q_i^* = \varphi \left[ (\alpha - c) - (c - c_0)\Theta \right], \quad \Theta(\gamma) = \frac{\gamma}{2 - \gamma} \geq 0.
\]

It is easily verified that \( \frac{\partial \Theta}{\partial \gamma} > 0 \). Lower degrees of product differentiation and market concentration have an unambiguously negative effect on compliant firms' quantities and prices.
Less market power generally lowers profits ($\varphi$ decreases), which in the case of product substitutability is reinforced by the increase in exposure to facing competition from a violating firm ($\Theta$ increases).

From (6) it can be inferred that an interior solution exists whenever $\alpha > \frac{2c - \gamma c_0}{2 - \gamma}$. It is easy to see that this condition is satisfied for all $\gamma$ if $\alpha > 2c$. In what follows it will be assumed that demand is not too weak and hence, the case where a unilateral deviation leads to a monopoly outcome will not be considered.

For notational convenience, we define the variable $t \equiv \frac{\alpha - c}{\alpha - c_0} \in (0, 1]$. Basically, $t$ is an inverse measure of the degree of cost savings obtained by a non-compliant firm; when abatement costs are small, $t$ is close to one, but the lower is $c_0$, the lower is $t$. Moreover, $t$ also increases in demand; the impact of abatement costs is reduced by stronger demand.

Below, we analyze three market outcomes, depending on firm behavior: (a) all firms are compliant; (b) one single firm breaches legislation; and (c) all firms are non-compliant. The second case where one firms violates the law, while all other firms remain compliant provides a measure for how strong the incentives for unilateral deviation are - the larger the gain from violating legislation is, the more intense inspections need to be to maintain compliance in an industry. If the inspection and enforcement regime is too weak to deter violations, all firms will breach legislation, which will lead to higher market surpluses as marginal costs are reduced, but also to a negative impact on environmental quality. Hence, from a policy-making point of view, the cost for inducing compliance has to be weighed against the overall effect on the environment and surpluses. By comparing the outcomes when all firms adhere to the law and all firms are non-compliant, we obtain measures for the impacts of achieving compliance on the environment and on market surpluses.

### 7.2.1 Universal compliance

If all firms comply, the symmetric quantity $q^*_C$, price $p^*_C$ and profit $\pi^*_C$ are obtained by setting $c_0 = c$ in (3), (4) and (5), i.e.

$$q^*_C = \varphi(\alpha - c), \quad p^*_C = c + \varphi(\alpha - c) \quad \text{and} \quad \pi^*_C = \varphi^2(\alpha - c)^2.$$  

The sum of consumers’ and producers’ surplus when all firms are compliant $S_C$ is given by (see the Appendix, 7.7.2, for derivation)

$$S_C = \frac{1}{2} n(\alpha - c)^2 \varphi (1 + \varphi). \quad (8)$$

### 7.2.2 Unilateral non-compliance

If firm $k$ unilaterally breaches legislation, its quantity, price and profit are given by (3), (4) and (5). The gain from unilaterally violating the
law $D$ is thus given by

$$D \equiv \pi^*_k - \pi^*_C = (c - c_0)\varphi^2 \Psi [2(\alpha - c) + (c - c_0)\Psi]. \quad (9)$$

### 7.2.3 Universal non-compliance

If all firms violate legislation, the symmetric quantity $q^*_V$, price $p^*_V$ and profit $\pi^*_V$ are obtained by setting $c = c_0$ in (3), (4) and (5), i.e.

$$q^*_V = \varphi(\alpha - c_0), \quad p^*_V = c + \varphi(\alpha - c_0), \quad \text{and} \quad \pi^*_V = \varphi^2(\alpha - c_0)^2.$$  

The sum of consumers’ and producers’ surplus when all firms are non-compliant $S_V$ is given by (see the Appendix, 7.7.2, for derivation)

$$S_V = \frac{1}{2} n(\alpha - c_0)^2 \varphi (1 + \varphi). \quad (10)$$

### 7.3 The costs and benefits of environmental inspections

In this section we address the issue of which industries should be targeted by environmental inspection and enforcement agencies. More specifically, the goal is to identify the parameter values for cost savings, demand, product differentiation and market concentration that yield the highest gains from conducting inspections.

There are three effects of inspections that need to be taken into consideration. First, costs are incurred from carrying out inspections. Second, there is an environmental gain from achieving compliance. Finally, inducing adherence to legislation leads to higher marginal costs and hence, a loss in market surpluses.

#### 7.3.1 Inspection costs

To assess the inspection costs for achieving compliance, we follow the approach by Becker (1968), according to which the expected penalty has to be larger than the gain from violating the law. Implicitly it is thus assumed that firms have perfect information regarding the inspection agency’s activities. Letting $F > 0$ be the fine for breaching legislation, it is straightforward that compliance can be achieved if and only if $F$ is larger than the gain from unilaterally violating legislation $D$. Given that $F \geq D$, the probability of detecting a violation has to be at least $\frac{D}{F}$ to deter firms from breaching legislation.
It is assumed that the probability of detecting a violation is equal to the inspection frequency, implying that an inspection of a non-compliant firm will lead to its violation being detected with certainty. Thus, if \( F \geq D \), the inspection agency can either choose the minimum inspection frequency \( \frac{D}{F} \) to ensure compliance by all firms, or not to conduct any inspections at all, in which case all firms breach legislation. This implies that there are two equilibrium outcomes to consider, full compliance as described in 2.1 on the one hand, and all firms being non-compliant as described in 2.3 on the other hand.

For simplicity it is assumed that the inspection agency incurs a constant marginal cost of \( \xi > 0 \) per inspection, but no fixed costs. The cost for achieving compliance in an industry is thus given by \( \Delta K \equiv \frac{\eta D}{F} \). Using (9), the cost for inducing compliance can be expressed as

\[
\Delta K = \frac{\xi}{F} (c - c_0) n \varphi^2 \Psi \left[ 2(\alpha - c) + (c - c_0) \Psi \right].
\] (11)

**Lemma 1** The cost for achieving compliance increases linearly in the marginal cost of inspections, decreases in the fine for violations, increases linearly in demand and quadratically in abatement costs, increases in the number of firms and is U-shaped with regard to the degree of product differentiation.

**Proof.** The results for \( \xi, F, \alpha \) and \( c - c_0 \) are straightforward. Since \( n \varphi^2 \Psi \) and \( \Psi \) increase in \( n \), it follows that \( \Delta K \) increases in \( n \). Finally,

\[
\frac{\partial \Delta K}{\partial \gamma} = \frac{2 \varphi}{\varphi_\gamma} + \frac{\partial \Psi}{\varphi_\gamma} + \frac{(c - c_0) \varphi}{2(\alpha - c) + (c - c_0) \Psi}
\]

\[
= -\frac{2(n - 1)}{2 + (n - 1) \gamma} + \frac{2 [(\alpha - c) + (c - c_0) \Psi]}{2(\alpha - c) + (c - c_0) \Psi} \left( \frac{2(n - 1)}{2 + (n - 2) \gamma} > 0 \right)
\]

\[
\Leftrightarrow 2 \frac{2 + (n - 1) \gamma}{2 + (n - 2) \gamma} > \frac{2t + (1 - t) \Psi}{t + (1 - t) \Psi}
\]

The left-hand side increases in \( \gamma \), from 1 to \( \frac{2(n + 1)}{n} \), while the right-hand side decreases in \( \gamma \), from \( 1 + t \) to \( 1 + \frac{t}{1 + (n - 1)(1 - t)} \) (if \( n > 1 \)). Hence, there exists a threshold value \( \hat{\gamma} \in (0, 1) \), such that \( \frac{\partial \Delta K}{\partial \gamma} > 0 \) if and only if \( \gamma > \hat{\gamma} \). \[\blacksquare\]

Thus, sectors with high demand, large abatement costs, fierce competition (high \( \gamma \) or \( n \)) or highly differentiated products (low \( \gamma \)) are associated with high costs for inducing compliance.
7.3.2 Environmental gains from inspections

The ultimate goal of inspections and enforcement is to protect the environment. By inducing firms to be compliant with legislation, environmental quality is enhanced. For analytical tractability the following assumptions are made. First, compliance is associated with no environmental cost. Second, the environmental benefit from avoiding violations increases linearly in abatement costs \( c - c_0 \). Third, the environmental loss under universal non-compliance is proportional to the aggregate output level \( nq^* \).

Given these assumptions, the gain in environmental quality from inducing compliance \( \Delta E \) can be expressed as

\[
\Delta E = \lambda (c - c_0) nq^* = \lambda (c - c_0)(\alpha - c_0)n\varphi,
\]

where \( \lambda \) is the value attached to environmental quality. Note that \( \lambda \) can be seen as both an "objective" parameter value reflecting the environmental effects of an industry, and a "subjective" measure attributed by an agent to improvements in environmental quality. Thus, \( \lambda \) accounts for varying impacts of different types of production and consumption, as well as for variations in how environmental benefits are assessed by different agents. A higher \( \lambda \) may therefore represent a more severe impact of e.g. emissions or a higher valuation of environmental quality by an agent.

Lemma 2 The environmental gain from achieving compliance increases in the value attached to environmental quality, increases linearly in demand and quadratically in abatement costs, increases in the number of firms and decreases as products become more substitutable.

Proof. The results for \( \lambda \), \( \alpha \) and \( c - c_0 \) are straightforward. It is easily verified that \( n\varphi \) increases in \( n \) and hence, \( \Delta E \) increases in \( n \). Since \( \varphi \) decreases in \( \gamma \) (if \( n > 1 \)), \( \Delta E \) decreases in \( \gamma \).

Hence, inducing adherence to legislation benefits the environment most in sectors with high demand, large abatement costs and high aggregate production volumes (i.e. for high degrees of product differentiation or low degrees of market concentration).

7.3.3 Economic losses from inspections

Deterring violations will lead to lower aggregate surpluses as firms’ marginal costs are higher under compliance; this loss is captured by the difference in total surpluses between the equilibrium where all firms violate legislation \( S_V \) and the equilibrium where all firms are compliant \( S_C \).
follows immediately from (8) and (10) that the loss in total surpluses from deterring violations $\Delta S$ is given by

$$\Delta S = S_V - S_C = \frac{1}{2} (c - c_0) [2(\alpha - c) + c - c_0] n \varphi (1 + \varphi). \quad (13)$$

**Lemma 3** The surplus loss from achieving compliance increases linearly in demand and quadratically in abatement costs, increases in the number of firms and decreases as products become more substitutable.

**Proof.** The results for $\alpha$ and $c - c_0$ are straightforward. It is easily verified that $n \varphi (1 + \varphi)$ increases in $n$ and hence, $\Delta S$ increases in $n$. Since $\varphi$ decreases in $\gamma$ (if $n > 1$), $\Delta S$ decreases in $\gamma$.

Thus, the detrimental effect of inspections on surpluses is largest in sectors with high demand, high abatement costs and high aggregate production volumes (i.e. for high degrees of product differentiation or low degrees of market concentration).

### 7.4 Measuring the effectiveness of inspections

In the preceding section it was established how costs, environmental gains and economic losses from inducing compliance with legislation are affected by demand, abatement costs and the degrees of product differentiation and market concentration similarly (with the notable exception of inspection costs increasing in $\gamma$ at sufficiently low degrees of product differentiation). Hence, it is not straightforward how a measure capturing all costs and benefits of inspections responds to changes in any of these variables. Furthermore, there are different ways to construct such a measure.

Our approach is to provide a measure for how effective inspections are, which can be used as a rule of thumb for where limited resources should be allocated. On the one hand, money is spent on inspections with the aim of inducing compliance with legislation. On the other hand, deterring violations of the law is associated with environmental benefits as well as economic losses. The total effect on the environment and surpluses therefore has to be weighed against the costs for carrying out inspections. To account for differences in assessments of the economic impact of inspections, a weight $\mu$ is attributed to surplus losses, such that the total impact of inducing compliance is given by $\Delta E - \mu \Delta S$, which has to be related to the minimal cost for achieving compliance $\Delta K$. The effectiveness of inspections is thus defined as $\Lambda_\mu \equiv \frac{\Delta E - \mu \Delta S}{\Delta K}$, i.e. $\Lambda_\mu$ can
be thought of as the net benefit per dollar spent on inspections in a certain market. Using (11), (12) and (13) it follows that

$$\Lambda_\mu = \frac{F \lambda n(\alpha - c_0)(c - c_0)\varphi - \frac{1}{2} \mu n(c - c_0) [2(\alpha - c) + c - c_0] \varphi (1 + \varphi)}{(c - c_0) n \varphi^2 [2(\alpha - c) + (c - c_0) \Psi]}$$

It is easy to see that the linear first order effects of the number of firms and abatement costs cancel. Further rearrangement of terms yields the following expression:

$$\Lambda_\mu = \frac{F \lambda - \mu \frac{1 + t}{2} (1 + \varphi)}{\xi \varphi \Psi [2t + (1 - t) \Psi]},$$

(14)

It is straightforward that \(\Lambda_\mu\) increases in \(\lambda\) and \(F\) and decreases in \(\mu\) and \(\xi\). Naturally, higher fines and lower marginal inspection costs enhance the effectiveness of inspections. It is also obvious that a higher value attached to environmental quality and a lower weight attributed to surplus losses influence the effectiveness of inspections positively. Thus, the impact of these variables need not be analyzed further.

In what follows we will examine how the effectiveness of inspections is influenced by demand and abatement costs, as well as the degrees of product differentiation and market concentration. First, we evaluate the case where an agent does not take any economic losses of inspections into consideration. In this case the effectiveness of inspections is determined by the relationship between the improvement in environmental quality and the cost for achieving this. We thus obtain a measure for the environmental effectiveness of inspections. Next, we consider the impact on surplus losses vis-à-vis inspection costs, which is referred to as the surplus-loss-to-inspection-cost ratio. Finally, we analyze how the effectiveness of inspections is affected when both environmental benefits and economic costs are accounted for.

### 7.4.1 The environmental effectiveness of inspections

As a benchmark, consider the case where an agent is solely concerned about the environment, i.e. where \(\mu = 0\). This applies to settings where the agent in question represents an independent inspection agency that is free from any pressure to also take the economic consequences of inspections into account. It could be argued that the trade-off between costs and benefits of inducing compliance has already been made when environmental legislation is passed, under the implicit assumption that laws will be enforced by an independent agency. In reality, however, inspection agencies might be subject to some pressure, which we will address in section 4.3.
When no weight is attributed to surplus losses, (14) is reduced to

\[
\Lambda_0 = \frac{F}{\xi \varphi \Psi[2t + (1 - t)\Psi]} \lambda,
\]

which can be seen as a measure of the environmental effectiveness of inspections.

**Proposition 4** The environmental effectiveness of inspections increases in the degrees of market concentration and product differentiation, and increases in demand and decreases in abatement costs if and only if the degrees of market concentration and product differentiation are sufficiently low.

**Proof.** It is easily verified that \(\varphi \Psi\) increases in \(\gamma\) and \(n\). Since \(\Psi\) also increases in \(\gamma\) and \(n\), it follows that \(\Lambda_0\) decreases in \(\gamma\) and \(n\). It is straightforward that \(\Lambda_0\) increases in \(t\) if and only if \(\Psi > 2 (\gamma n > 2)\).

Thus, the environmental effectiveness is smaller in less concentrated markets, although a higher number of firms increases aggregate output and hence, the environmental impact of non-compliance. However, this effect is more than compensated for by the increase in inspection costs as competition becomes fiercer. More specifically, given that the linear first order effects of the number of firms cancel, \(\Lambda_0\) is determined by the ratio of produced quantity per firm under universal non-compliance \(q^*_V\) and the gain from unilaterally breaching legislation \(D\). While \(q^*_V\) unambiguously decreases in \(n\), \(D\) increases in \(n\) if and only if \(t\) is sufficiently large (see Häckner and Herzing, 2012); for small \(t\), \(D\) decreases, but at a lower rate than \(q^*_V\). Hence, \(\Lambda_0\) decreases in the number of firms.

Increasing product substitutability reduces production and hence, the environmental impact of inspections. This effect is reinforced by higher inspection costs as \(\gamma\) increases for sufficiently low degrees of product differentiation; when products are sufficiently differentiated increases in \(\gamma\) also reduce inspection costs, but not sufficiently much to counter the negative effect on environmental gains.

The impacts of demand and abatement costs on the environmental effectiveness of inspections can be inferred from the effect of \(t\). Note that, while both inspection costs \(\Delta K\) and environmental gains \(\Delta E\) increase linearly in \(\alpha\) and quadratically in \(c - c_0\), the impact on the environmental effectiveness depends on \(t\), which increases in \(\alpha\), but decreases in \(c - c_0\), i.e. the effects of demand and abatement costs move in opposite directions, which can intuitively be explained by the fact that an increase in \(\alpha\) reduces cost savings relative to demand.
Basically, the impact of \( t \) on environmental effectiveness is driven by the ratio \( \frac{\alpha - c + (c - c_0)}{2(\alpha - c) + (c - c_0)\Psi} \), which measures how \( q^*_t \) is affected in relation to \( D \) by changes in \( \alpha \) and \( c - c_0 \). It is easy to see that produced quantities when all firms violate legislation (the numerator) are influenced similarly by demand and abatement costs, while the profit from unilateral deviation (the denominator) is more strongly affected by higher demand than by increases in abatement costs if and only if competition is sufficiently weak \( (\Psi < 2) \). This implies that higher demand increases the gain from unilateral deviation and hence, the costs for achieving compliance relatively more than the environmental gain from deterring violations, whereby environmental effectiveness is reduced when competition is weak. In contrast, an increase in abatement costs impacts more strongly on the environmental gain from full compliance than on the gain from unilateral deviation and hence, enhances environmental effectiveness at high degrees of product differentiation and market concentration.

When the competitive pressure is sufficiently strong the reverse is true. Increases in abatement costs influence the profit from unilateral deviation more strongly than higher demand. Thus, higher demand improves environmental effectiveness as the environmental gain from achieving compliance increases relatively more than the gain from unilateral deviation. Conversely, higher abatement costs affect the gain from unilateral deviation more than the environmental gain from inducing compliance, which reduces environmental effectiveness.

To summarize, for an agent targeting environmental effectiveness, inspection resources should be concentrated on sectors where competition is weak (high degrees of market concentration and product differentiation), in which case environmental effectiveness is enhanced by weak demand and large cost savings from non-compliance.

### 7.4.2 The surplus-loss-to-inspection-cost ratio

If some weight is attributed to the impact on surpluses (i.e., \( \mu > 0 \)), we also have to consider the ratio of surplus losses to inspection costs \( \Omega \equiv \frac{\Delta S}{\Delta K} \). More specifically,

\[
\Omega = \frac{F}{\xi} \frac{1 + \frac{1}{2}(1 + \varphi)}{\varphi \Psi [2t + (1 - t)\Psi]}, \tag{16}
\]

**Lemma 5** The surplus-loss-to-inspection-cost ratio increases if the degrees of market concentration and product differentiation increase, if demand increases and if abatement costs decrease.
Proof. By rearranging terms, (16) can be expressed as

\[ \Omega = \frac{F}{\varphi [2 - \frac{2}{1+t} + (\frac{2}{1+t} - 1) \Psi]} . \]

It is easily verified that \( \frac{\varphi \Psi}{1+\varphi} \) increases in both \( \gamma \) and \( n \). Given that \( \frac{2}{1+t} \geq 1 \) and that \( \Psi \) increases in \( \gamma \) and \( n \), it follows that \( \Omega \) decreases in \( \gamma \) and \( n \). Since \( \Psi \geq 1 \), it is straightforward that \( \Omega \) increases in \( t \).

We obtain basically the same results as in proposition 4 (except that \( \Omega \) unambiguously increases in \( t \)). The intuition behind these results is therefore similar. To understand the impact of \( n \) on \( \Omega \), note that the surplus loss from compliance reflects the difference between all firms being non-compliant and all firms being compliant, while the inspection cost is determined by the difference in profits between unilaterally violating legislation and being compliant. Since firm profits increase by more when legislation is breached unilaterally rather than universally, the impact of more firms on inspection costs dominates the effect on surpluses.

Increasing product substitutability reduces production and hence, surpluses when all firms comply and when all firms are non-compliant at the same rate, implying that the surplus loss from compliance decreases. This effect is reinforced by higher inspection costs as \( \gamma \) increases for sufficiently low degrees of product differentiation; when products are sufficiently differentiated increases in \( \gamma \) also reduce inspection costs, but not sufficiently much to counter the negative effect of inspections on surplus losses.

As in the case of environmental effectiveness, demand and abatement costs impact on \( \Omega \) in opposite directions. It suffices to consider the effect on the ratio \( \frac{2(\alpha-c)+(c-c_0)}{2(\alpha-c)+(c-c_0)\Psi} \). Since \( \Psi \geq 1 \), it immediately follows that surplus losses are more strongly affected by higher demand than by increases in abatement costs. Higher demand has a stronger effect on the increases in profits and consumer surplus when all firms are non-compliant than on the profit gain from unilaterally breaching legislation and hence, raises \( \Omega \). In contrast, an increase in abatement costs has a relatively stronger effect on the gain from unilateral deviation than on surplus losses, which reduces \( \Omega \).

To conclude, the surplus-loss-to-inspection-cost ratio is positively affected by stronger demand, lower abatement costs, as well as higher degrees of market concentration and product differentiation.
7.4.3 The effectiveness of inspections in the presence of economic considerations

As demonstrated above, \( \gamma, n \) and \( t \) (when \( \Psi > 2 \)) influence the environmental effectiveness and the surplus-loss-to-inspection-cost ratio similarly. Therefore, it is not straightforward how effective inspections are. The overall measure of effectiveness will crucially depend on the weights attributed to environmental quality and to surplus losses.

7.4.3.1 Market concentration

First, we investigate how the effectiveness of inspections is influenced by the degree of market concentration.

**Proposition 6** If the relative weight attributed to surplus losses is sufficiently small, the effectiveness of inspections increases monotonously in the degree of market concentration. When the relative weight of surplus losses is sufficiently high, the effectiveness of inspections first increases and eventually decreases in the degree of market concentration.

**Proof.** It can be shown that (see the Appendix, A3)

\[
\frac{\partial \Lambda_\mu}{\partial n} > 0 \iff \frac{\mu}{\lambda} > \frac{2}{(1 + t) \left(1 + \varphi + \frac{\psi}{A_n - 1}\right)} \equiv \eta_n,
\]

where \( A_n = \frac{2[t + (1-t)\Psi]}{2(1-t)\Psi} \frac{\varphi}{\psi} \frac{\psi}{\partial \varphi/\partial n} \). The threshold value \( \eta_n \) increases in \( n \) (see the Appendix, 7.7.3), from \( \eta_1 = \frac{2}{1+t} \frac{2+\gamma-(2-\gamma)\frac{t}{t}}{4+\gamma-(2-\gamma)\frac{t}{t}} \) and converging to \( \eta_\infty = \frac{2}{1+t} \). Thus, whenever \( \frac{\mu}{\lambda} \leq \eta_1 \), the above condition is never satisfied and hence, \( \Lambda_\mu \) decreases monotonously in \( n \). If \( \frac{\mu}{\lambda} \in (\eta_1, \frac{2}{1+t}) \), the above condition is satisfied for sufficiently small \( n \), implying that \( \Lambda_\mu \) is inversely U-shaped in \( n \). If \( \frac{\mu}{\lambda} \geq \frac{2}{1+t} \), \( \Lambda_\mu \) increases monotonously in \( n \) - however, in this case \( \Lambda_\mu < 0 \) for all finite \( n \), i.e. the net benefit from inspections is always negative and hence, no inspections should be carried out.\(^{35}\)

When the weight attributed to surplus losses is relatively small, the impact of \( n \) on environmental effectiveness \( \Lambda_0 \) dominates, such that \( \Lambda_\mu \) decreases in \( n \). If the relative weight of surplus losses is sufficiently large, the impact of \( n \) on the surplus-loss-to-inspection-cost ratio \( \Omega \) dominates if and only if \( n \) is sufficiently small; as the competitive pressure from an increase in \( n \) becomes stronger the impact on \( \Lambda_0 \) eventually dominates.

\(^{35}\)More generally, inspections should be carried out if and only if \( \Lambda_\mu > 1 \), i.e. whenever the net benefit is larger than the cost of inspections. However, our main focus is on identifying the most effective way of spending inspection resources.
Note that both $\eta_1$ and $\eta_\infty$ decrease in $t$, and $\eta_1$ increases in $\gamma$, implying that $\Lambda_\mu$ is more likely to increase monotonously in the degree of market concentration when demand is weak, abatement costs are high and the degree of product differentiation is low. The figure below illustrates how the degree of market concentration impacts on the effectiveness of inspections for different $\mu$-values given that $F = 10$, $\xi = 1$, $\lambda = 1$, $t = 0.75$ and $\gamma = \frac{1}{2}$.

![Figure 1. Effectiveness of inspections in relation to firm number](image)

7.4.3.2 Product differentiation

Next, we analyze the impact of the degree of product differentiation on the effectiveness of inspections.

**Proposition 7** If the relative weight attributed to surplus losses is sufficiently small, the effectiveness of inspections increases monotonously in the degree of product differentiation. At intermediate values of the relative weight of surplus losses, the effectiveness of inspections first increases and eventually decreases in the degree of product differentiation. When the relative weight of surplus losses is sufficiently large, the effectiveness of inspections decreases monotonously in the degree of product differentiation.
Proof. It can be shown that (see the Appendix, 7.7.3) 
\[ \frac{\partial \Lambda_\mu}{\partial \gamma} > 0 \iff \frac{\mu}{\lambda} > \frac{2}{(1 + t) \left( 1 + \varphi + \frac{\varphi}{\Lambda_{\gamma - 1}} \right)} \equiv \hat{\gamma}_{\Lambda}, \]
where \( A_\gamma \equiv \frac{2(t + (1 - \psi) \Psi)}{2(t + (1 - \psi) \Psi) \Psi(-\partial \varphi / \partial \gamma)} \). The threshold value \( \hat{\gamma}_{\Lambda} \) increases in \( \gamma \) (see the Appendix, A3), from \( \hat{\gamma}_0 = \frac{2(1-t)}{(1+t)(2-t)} \) to \( \hat{\gamma}_1 = \frac{2}{1 + t} \frac{n(3n+4)-(3n^2+2n-4)t}{n(3n+8)-(3n^2+6n-8)t} \).
Thus, whenever \( \frac{\mu}{\lambda} \leq \hat{\gamma}_0 \), the above condition is never satisfied and hence, \( \Lambda_\mu \) decreases monotonously in \( \gamma \). If \( \frac{\mu}{\lambda} \in (\hat{\gamma}_0, \hat{\gamma}_1) \), the above condition is satisfied for sufficiently small \( \gamma \), implying that \( \Lambda_\mu \) is inversely U-shaped in \( \gamma \). For \( \frac{\mu}{\lambda} \geq \hat{\gamma}_1 \) the above condition is satisfied for all \( \gamma \) and hence, \( \Lambda_\mu \) increases monotonously in \( \gamma \). 

When the weight attributed to the economic consequences of inspections is sufficiently large, the effect of \( \gamma \) on the surplus-loss-to-inspection-cost ratio \( \Omega \) dominates, such that \( \Lambda_\mu \) increases in \( \gamma \). In contrast, the impact of \( \gamma \) on environmental effectiveness \( \Lambda_0 \) dominates and hence, \( \Lambda_\mu \) decreases in \( \gamma \), if the value attached to environmental quality is relatively large. At intermediate levels of the relative weight of surplus losses, the impact of \( \gamma \) on \( \Omega \) dominates if and only if \( \gamma \) is sufficiently small; as the competitive pressure from an increase in \( \gamma \) becomes stronger the impact on \( \Lambda_0 \) eventually dominates.

Note that both \( \hat{\gamma}_0 \) and \( \hat{\gamma}_1 \) decrease in \( t \), and \( \hat{\gamma}_1 \) increases in \( n \), implying that \( \Lambda_\mu \) is more likely to increase in the degree of product differentiation when demand is weak, abatement costs are high and the degree of market concentration is low. The following figure shows how the degree of product differentiation affects the effectiveness of inspections for different \( \mu \)-values given that \( F = 10, \xi = 1, \lambda = 1, t = 0.75 \) and \( n = 3 \).\[36\]

\[36\] Note that when \( \mu \) is sufficiently large, the net effect of inducing compliance is actually negative at high degrees of product differentiation; obviously it would not make sense to enforce compliance in this case.
7.4.3.3 Cost savings and demand

Finally, consider the impact of demand and abatement costs on the effectiveness of inspections.

**Proposition 8** At sufficiently high degrees of market concentration or product differentiation, the effectiveness of inspections increases monotonously in the abatement cost and decreases monotonously in demand for any weights attached to environmental quality $\lambda$ and economic losses $\mu$. When the degrees of market concentration and product differentiation are sufficiently small, the effectiveness of inspections increases monotonously in the abatement cost and decreases monotonously in demand if and only if $\mu$ is sufficiently large in relation to $\lambda$.

**Proof.**

\[ \frac{\partial \Lambda_\mu}{\partial t} = \frac{\partial \Lambda_0}{\partial t} - \mu \frac{\partial \Omega}{\partial t} > 0 \]

\[ \Leftrightarrow \mu < \frac{\frac{\partial \Lambda_0}{\partial \Omega}}{\frac{\partial \Omega}{\partial t}} = \frac{\frac{\kappa \lambda (\Psi - 2)}{\varphi \Psi (2t + (1-t)) \Psi^\varphi}}{\frac{\kappa (1+\varphi)(\Psi - 1)}{\varphi \Psi (2t + (1-t)) \Psi^\varphi}} \Leftrightarrow \mu \frac{\lambda}{t} < \frac{\Psi - 2}{(1+\varphi)(\Psi - 1)} \equiv \tilde{\eta}_t. \]
It is easy to see that $\tilde{\eta}_t \leq 0$ if $\Psi \leq 2$ ($\gamma n \leq 2$). Thus, when the competitive pressure is sufficiently weak, the above condition is never satisfied, implying that $\Lambda_{\mu}$ decreases in $\alpha$ and increases in $c - c_0$ for any values of $\lambda$ and $\mu$. When competition is sufficiently fierce ($\Psi > 2$), $\Lambda_{\mu}$ increases in $\alpha$ and decreases in $c - c_0$ if and only if $\mu$ is sufficiently small in relation to $\lambda$. ■

At high degrees of product differentiation and market concentration the positive effect of an increase in $t$ on the surplus-loss-to-inspection-cost ratio $\Omega$ is reinforced by the negative impact on environmental effectiveness $\Lambda_0$ and hence, $\Lambda_{\mu}$ unambiguously decreases in $t$. If competition is sufficiently strong, the positive effect of $t$ on $\Lambda_0$ dominates the negative impact on $\Omega$, as long as the weight attributed to surplus losses is not too large. This condition becomes less restrictive as the competitive pressure increases, because $\tilde{\eta}_t$ increases in $\gamma$ and $n$.

The figure below illustrates the impact of demand on the effectiveness of inspections for different $\mu$-values given that $F = 10$, $\xi = 1$, $\lambda = 1$, $c = 1$, $c_0 = 0$ and $\gamma = 0.5$, when the competitive pressure is low ($n = 3$) and when it is high ($n = 5$).

![Figure 3. Effectiveness of inspection in relation to demand](image)

The following figure shows how abatement costs affect the effectiveness of inspections for different $\mu$-values under weak ($n = 3$) and fierce
(n = 6) competition given that F = 10, ξ = 1, λ = 1, c = 1 (such that abatement costs c − c_0 take on values in the interval [0, 1]), α = 5 and γ = 0.5.

Figure 4. Effectiveness of inspections in relation to abatement costs

7.5 Concluding remarks

There are some general findings that an inspection agency can use without having access to detailed firm level data. Given that the agency has some information regarding abatement costs, demand and competitive pressure, our results indicate that the following rules of thumb apply.

I. When little or no weight is attributed to surplus losses, inspection agencies should target sectors with low competitive pressure (high degrees of market concentration or product differentiation).

II. Inspection agencies should target sectors with high abatement costs and weak demand, except when competition is fierce and the relative weight of economic losses is sufficiently small.

III. The higher the relative weight of economic losses, the more competitive should the markets targeted by the inspection agency be.
It is important to emphasize that what might seem to be intuitive results on market structure and product differentiation (claim I) are restricted to the case where inspection authorities pay little attention to economic losses. As the weight attributed to surplus losses increases, the focus of inspections should shift to markets with lower degrees of market concentration and product differentiation (claim III).

Note also that the results on abatement costs and demand (claim II) do not depend on the weighting of surplus losses vis-à-vis environmental quality when the degrees of product differentiation and market concentration are high. Intuitively one might expect that firms under pressure from weak demand or high costs for complying with legislation should not be targeted. However, it needs to be emphasized that claim II crucially depends on stronger demand and lower abatement costs enhancing the environmental gains from full compliance relatively more than the cost of deterring non-compliance if and only if the competitive pressure is sufficiently strong. The surplus-loss-to-inspection-cost ratio increases unambiguously when demand rises and abatement costs are lowered, as economic losses from enforcing compliance increase to a higher degree than the gain from unilaterally breaching the law. Hence, when competition is fierce, both the environmental effectiveness and the surplus-loss-to-inspection-cost ratio increase in demand and decrease in abatement cost; as long as the weight attributed to surplus losses is not too large, the impact on environmental effectiveness dominates. Under low competitive pressure, the impacts of changes in demand and abatement costs on environmental effectiveness are reinforced by the effects on the surplus-loss-to-inspection-cost ratio. Thus, if there are two sectors with little competition, the priority should be to inspect the sector with weaker demand (where the effect of less costly deterrence dominates) or the sector with higher abatement costs (where the environmental benefit from inducing compliance carries more weight).

Even though it may be presumed that inspection agencies are instructed not to take market effects into account, it is important to understand how a full-fledged welfare analysis could result in recommendations that differ from the ones advocated by an authority only considering environmental gains. This is of particular interest given that the institutional conditions for environmental inspections and enforcement differ across and within countries, but also with regard to the type of environmental impact that is monitored.

Our analysis is novel as it focuses on differences between sectors and hence, our results should be seen as complementary to the findings of contributions concerning the optimal allocation of limited inspection resources across firms in one sector. Given these differences in method-
ological approach, it is therefore difficult to contrast our conclusions to those generated by such models. Interestingly, however, Macho-Stadler and Pérez-Castrillo (2006) find that, among a continuum of heterogenous firms, those that are easy to deter and those that value pollution less should be targeted. One of our main results (claim II) is that firms with low abatement costs, for which the gain from breaching environmental legislation is small, are more effective to inspect if the number of firms is large enough and the agency is mainly focusing on the environmental consequences of inspections. Hence, when conditions in our framework are comparable to the Macho-Stadler and Pérez-Castrillo model, we obtain a result that is qualitatively similar.

To empirically test our conclusions an option would be to exploit the decentralized structure of environmental inspections and enforcement in Sweden, which has given rise to considerable variation in the organization and implementation of these activities. By comparing the conduct of inspections and enforcement with the prevailing institutional conditions in different municipalities it might be possible to assess whether observed patterns correspond to our findings. For example, the priorities of a municipal administration solely in charge of the environment and public health might differ from those of an integrated administrative unit whose focus is not only on the environment, but also on e.g. building and planning. A recent survey of Swedish municipal environmental managers found that political influence is less common (suggesting a lower weighting of the economic consequences of inspections in the context of our model) among separate environmental administrations (Jacobsson and Källmén, 2016).

Our model could be modified in several ways, e.g. letting firms interact in Bertrand fashion, or having convex production cost functions.\footnote{See also Häckner and Herzing (2016) who consider several extensions of the present model in the context of evaluating the welfare effects of taxation.}

We also assume that the fine for non-compliance is constant, i.e. it does not depend on the severity of a violation of legislation. It is possible to allow for fines to increase in the environmental impact of production, thus reflecting the degree to which abatement costs are saved.

7.6 References


Innes, R., 2000, “Self-Reporting in Optimal Law Enforcement When
Violators Have Heterogeneous Probabilities of Apprehension", *Journal of Legal Studies* 29, 287-300.


7.7 Appendix

7.7.1 Model solutions

7.7.1.1 Firm $k$

Given that all firms $i \neq k$ comply, the reaction functions of firm $k$ and firms $i \neq k$ are given by (see Häckner, 2000)

$$q_k(q_{-k}) = \frac{1}{2} \left[ \alpha - c_0 - \gamma \sum_{j \neq k} q_j \right], \quad (17)$$

$$q_i(q_{-i}) = \frac{1}{2} \left[ \alpha - \gamma \sum_{j \neq k, j \neq i} q_j - \gamma q_k \right]. \quad (18)$$

The sum of all firms’ reaction functions is given by

$$\sum q_j = \frac{(n - 1)(\alpha - c) + \alpha - c_0}{2 + (n - 1)\gamma}. \quad (19)$$

Combining (17) and (19) yields

$$q_k^* = \frac{1}{2 + (n - 1)\gamma} \left[ \alpha - c + \frac{2 + (n - 2)\gamma}{2 - \gamma} (c - c_0) \right].$$

Defining

$$\varphi(\gamma, n) \equiv \frac{1}{2 + (n - 1)\gamma}, \psi(\gamma, n) \equiv \frac{2 + (n - 2)\gamma}{2 - \gamma},$$

firm $k$’s quantity can be expressed as $q_k^* = \varphi[\alpha - c + \psi(c - c_0)]$. Plugging (17) into (2) we obtain

$$p_k^* = \alpha - q_k^* - \gamma \sum_{j \neq k} q_j^* = \alpha - q_k^* + 2q_k^* - (\alpha - c) = c_0 + q_k^*.$$

7.7.1.2 Firms $i \neq k$

By plugging (19) into (18) and rearranging terms we obtain

$$q_i^* = \frac{1}{2 + (n - 1)\gamma} \left[ \alpha - c - \frac{\gamma}{2 - \gamma} (c - c_0) \right].$$

Defining

$$\Theta(\gamma) \equiv \frac{\gamma}{2 - \gamma},$$

firm $i$’s quantity can be expressed as $q_i^* = \varphi[\alpha - c - \Theta(c - c_0)]$. Plugging (18) into (2) yields

$$p_i^* = \alpha - q_i^* - \gamma \sum_{j \neq i} q_j^* = \alpha - q_i^* + 2q_i^* - (\alpha - c) = c + q_i^*.$$
7.7.2 Surpluses

Total surpluses are determined by aggregate profits and consumer surplus. When all firms have the same marginal cost \( c_0 \), each firm produces \( q = (c_0)^{1/2} \) and charges price \( p = c_0 + q \). Hence, total surplus is given by

\[
S = nq^* - \frac{n}{2} [1 + (n - 1)\gamma] q^{*2} - nq^* p^* \\
= nq^* \left\{ q^* + \alpha - \frac{1}{2} [1 + (n - 1)\gamma] q^* - c^* - q^* \right\} \\
= n\alpha - \frac{1}{2} [1 + (n - 1)\gamma] (\alpha - c^*) \phi = \frac{1}{2} n(\alpha - c^*)^2 \phi (1 + \phi).
\]

Thus, \( S_C = \frac{1}{2} n(\alpha - c^*)^2 \phi (1 + \phi) \) and \( S_V = \frac{1}{2} n(\alpha - c_0)^2 \phi (1 + \phi) \).

7.7.3 Proofs of propositions 7 and 8

For \( \omega \in \{\gamma, n\} \),

\[
\frac{\partial \Lambda_\mu}{\partial \omega} = \frac{\partial \Lambda_0}{\partial \omega} - \mu \frac{\partial \Omega}{\partial \omega} > 0 \\
\Leftrightarrow -\kappa \lambda \left[ \frac{2t + (1 - t)\Psi \frac{\partial (\varphi \Psi)}{\partial \omega} + (1 - t)\varphi \Psi \frac{\partial \Psi}{\partial \omega}}{\varphi^2 \Psi^2 [2t + (1 - t)\Psi]^2} \right] > \mu \kappa \left[ \frac{1 + t}{2} \frac{\varphi \Psi [2t + (1 - t)\Psi] \frac{\partial \varphi}{\partial \omega} - (1 + \varphi) \left\{ [2t + (1 - t)\Psi] \frac{\partial (\varphi \Psi)}{\partial \omega} + (1 - t)\varphi \Psi \frac{\partial \Psi}{\partial \omega} \right\}}{\varphi^2 \Psi^2 [2t + (1 - t)\Psi]^2} \right] \\
\Leftrightarrow \frac{\lambda}{\mu} < \frac{1 + t}{2} \left( 1 + \varphi - \frac{\varphi \Psi [2t + (1 - t)\Psi] \frac{\partial \varphi}{\partial \omega}}{[2t + (1 - t)\Psi] \frac{\partial (\varphi \Psi)}{\partial \omega} + (1 - t)\varphi \Psi \frac{\partial \Psi}{\partial \omega}} \right) \\
= \frac{1 + t}{2} \left( 1 + \varphi - \frac{\partial \varphi / \partial \omega}{\varphi / \partial \varphi + \partial \Psi / \partial \omega + (1 - t)\partial \Psi / \partial \omega} \right) = \frac{1 + t}{2} \left( 1 + \varphi + \frac{\varphi}{A_\omega - 1} \right),
\]

where \( A_\omega \equiv \frac{2t + (1 - t)\Psi}{2t + (1 - t)\Psi \Psi (\partial \varphi / \partial \omega)} \).

To prove proposition 7, note that

\[
\tilde{\eta}_n = \frac{1 + t}{2} \left( 1 + \varphi + \frac{\varphi}{A_n - 1} \right) = \frac{1 + t}{2} \left( 1 + \frac{A_n \varphi}{A_n - 1} \right),
\]

where

\[
A_n = \frac{2 \left[ t + (1 - t)\Psi \right] \left[ 2 + (n - 1)\gamma \right]}{2t + (1 - t)\Psi} \left[ 2 + (n - 2)\gamma \right] = \frac{2 \left\{ t(2 - \gamma) + (1 - t) \left[ 2 + (n - 2)\gamma \right] \right\} \left[ 2 + (n - 1)\gamma \right]}{2t(2 - \gamma) + (1 - t) \left[ 2 + (n - 2)\gamma \right] \left[ 2 + (n - 2)\gamma \right]},
\]

and hence,

\[
\frac{A_n \varphi}{A_n - 1} = \frac{t(2 - \gamma) + (1 - t) \left[ 2 + (n - 2)\gamma \right]}{2t\gamma (2 - \gamma) + (1 - t) \left[ 2 + n\gamma \right] \left[ 2 + (n - 2)\gamma \right]}.
\]
Since
\[
\frac{\partial \left( \frac{A_\gamma \varphi}{A_{\gamma-1}} \right)}{\partial n} = 2 \left( \frac{\{2t\gamma(2-\gamma) + (1-t)(2+n\gamma)[2+(n-2)\gamma]\} \gamma(1-t)}{\{2t\gamma(2-\gamma) + (1-t)(2+n\gamma)[2+(n-2)\gamma]\}^2} - \frac{\{t(2-\gamma) + (1-t)[2+(n-2)\gamma]\} [4 + 2(n-1)\gamma] \gamma(1-t)}{\{2t\gamma(2-\gamma) + (1-t)(2+n\gamma)[2+(n-2)\gamma]\}^2} \right)
\]
\[
\quad = -2\gamma [2 + (n-2)\gamma] (1-t) \frac{2(2-\gamma)t + [2 + (n-2)\gamma] (1-t)}{\{2t\gamma(2-\gamma) + (1-t)(2+n\gamma)[2+(n-2)\gamma]\}^2} < 0,
\]
it immediately follows that \(\frac{\partial \eta_\gamma}{\partial \gamma} < 0\).

To prove proposition 8, note that, \(\frac{2[1+(1-t)\Psi]}{2t(1-\Psi)}\) increases in \(\gamma\), because \(\Psi\) increases in \(\gamma\). Moreover,
\[
\frac{\varphi \frac{\partial \Psi}{\partial \gamma} - \frac{\partial \varphi}{\partial \gamma} \Psi}{2 - \gamma} = \frac{2[2 + (n-1)\gamma]}{2 - \gamma} \cdot \frac{2}{2 - 2 + (n-2)\gamma} = \frac{2}{2 - \gamma} \left( 1 + \frac{\gamma}{2 + (n-2)\gamma} \right),
\]
which increases in \(\gamma\) and hence, \(A_\gamma\) increases in \(\gamma\). Since \(\varphi\) decreases in \(\gamma\), it immediately follows that \(\tilde{\eta}_\gamma = \frac{1+t}{2} \left( 1 + \varphi + \frac{\varphi}{A_{\gamma-1}} \right)\) increases in \(\gamma\).

### 7.7.4 List of parameters and variables

- **\(\alpha\)** demand shifting parameter
- **\(\gamma\)** degree of product substitutability
- **\(n\)** number of firms
- **\(c\)** constant marginal cost of compliant firm
- **\(c_0\)** constant marginal cost of violating firm
- **\(t = \frac{\alpha-c}{\alpha-c_0}\)** degree of cost savings from violations
- **\(\varphi = \frac{2^+(n-1)\gamma}{2^+(n-2)\gamma}\)** factor impacting on produced quantities
- **\(\Psi = \frac{2^+(n-2)\gamma}{2^+\gamma}\)** factor impacting on gain from unilateral violation
- **\(\Theta = \frac{\gamma}{2^+\gamma}\)** factor impacting on loss from being exposed to unilateral violation
- **\(\lambda\)** value attached to environmental quality
- **\(\mu\)** weight of surplus losses
- **\(F\)** fine for violations
- **\(\xi\)** constant marginal cost of inspections
- **\(D\)** gain from unilaterally breaching legislation
- **\(\Delta K\)** minimum agency cost of inducing compliance
- **\(\Delta E\)** environmental gain from compliance
- **\(\Delta S\)** surplus loss from compliance
- **\(\Lambda_0 = \frac{\Delta E}{\Delta K}\)** environmental effectiveness of inducing compliance
- **\(\Omega = \frac{\Delta E - \mu \Delta S}{\Delta K}\)** surplus-loss-to-inspection-cost ratio
- **\(\Lambda_\mu = \frac{\Delta E - \mu \Delta S}{\Delta K}\)** effectiveness of inducing compliance
- **\(\eta_\alpha\)** threshold for \(\frac{\mu}{\lambda}\), below which \(\frac{\partial \lambda}{\partial \alpha} > 0\)
- **\(\eta_\gamma\)** threshold for \(\frac{\mu}{\lambda}\), below which \(\frac{\partial \lambda}{\partial \gamma} > 0\)
- **\(\tilde{\eta}_t\)** threshold for \(\frac{\mu}{\lambda}\), below which \(\frac{\partial \lambda}{\partial t} > 0\)
8. An empirical analysis on the effects of inspections on the environmental behavior of dry cleaners in Stockholm

Anders Hed, Mathias Herzing, Linda Hoff Rudhult and Adam Jacobsson

8.1. Introduction

To implement environmental policy efficiently and control emissions it is necessary to regularly inspect facilities that emit harmful substances. The impact of environmental inspections and enforcement (EIE) on operators’ compliance with environmental legislation and regulations has received increasing attention in the literature over the last 20 years. Most empirical studies regarding EIE and its impact on environmental behavior have been carried out in North America, e.g. Hanna and Oliva (2010), Laplante and Rilstone (1996), and Magat and Viscusi (1990).

This study aims to contribute to the literature on the effect of EIE by investigating the relationship between inspections and the environmental behavior of dry cleaners in Stockholm who impact negatively on the environment mainly through the use of the toxic solvent perchlorethylene. Hitherto no empirical study focusing on the impact of inspections of environmentally harmful activities has been carried out in Sweden, where EIE to a large extent is carried out by municipal environmental departments. Moreover, the focus is on small operators in contrast to previous empirical studies that have mostly analyzed industries with large firms.

The data used in this study has been retrieved from annual reports of 50 dry cleaners in Stockholm from 2000 to 2013. These annual reports are submitted to the municipal environmental department and contain data on emissions of perchlorethylene. We also have obtained data on when inspections have taken place. A unique feature of the data is that the timing of inspections is not related to

39 We thank Ulrika Iversen at the Environmental and Public Health Safety Department of Stockholm municipality as well as staff at the municipal archive for help in getting access to and interpretation of data.
40 E-mail: andersohed@gmail.com
41 Department of Economics, Stockholm University, e-mail: Mathias.Herzing@ne.su.se
42 E-mail: linda.rudhult@gmail.com
43 Department of Economics, Stockholm University, e-mail: aja@ne.su.se
the environmental behavior of the dry cleaners. We can thus avoid the endogeneity problems that usually arise when the effects of inspections are analyzed.

To estimate the quantitative relationship between inspections and environmental behavior the focus is on two outcome variables, self-reported emissions of perchlorethylene and submissions of annual reports on time. First, we examine dry cleaners’ behavior on the extensive margin, i.e. whether inspections impact on compliance with environmental legislation regarding emissions of perchlorethylene. To that end we employ a linear probability model that investigates the effect of inspections on the likelihood of emissions exceeding target limits.

Having established a statistically significant positive effect of inspections in the preceding period on compliance, we then focus on how dry cleaners’ behavior is influenced on the intensive margin, i.e. to what extent inspections reduce emissions of perchlorethylene. Our regressions indicate a negative, but statistically insignificant effect of inspections on emissions.

Finally, we also examine the effect of inspections on the probability of dry cleaners handing in annual reports on time. We find no statistically significant effect of inspections on the likelihood of annual reports being submitted on time.

The presentation of this study is organized as follows. Section 8.2 contains a brief review of theories regarding operators’ environmental behavior, as well as results from empirical analyses of the effect of inspections on environmental outcomes. Section 8.3 describes how the environmental department of Stockholm municipality conducts EIE of dry cleaners, presents the data and discusses empirical considerations. In section 8.4 the results from our regressions are presented. Section 8.5 concludes.

8.2. Background

8.2.1. Theory

An operator’s choice to comply with legislation and regulations can be explained by economic, normative or social motives. According to Becker’s (1968) theory on crime and punishment, the decision to follow rules is based on a rational choice aimed at maximizing private economic utility. Hence, if both the threat of being inspected and the cost for breaching legislation are high enough, an operator will be induced to incur the cost of being compliant (Gray and Shimshack, 2011). A distinction can be made between specific and general deterrence, where the former concerns the influence of inspections and sanctions on a single operator’s environmental behavior, while the latter refers to the impact on other operators within the same geographical region or the same industry (Gray and Shimshack, 2011). Inspections can also have an implicit deterrent effect (Gunningham et al,
which is associated with operators developing compliant behavior based on previous experience of the actions of the EIE agency.

An operator’s conduct can also be influenced by normative motives. The likelihood of compliance increases if laws and regulations correspond well to an operator’s ethical and ideological convictions, and if other operators in the same industry or in the same region follow norms regarding environmental behavior (Winter and May, 2001). An EIE agency can supply information and advice, thereby internalizing norms that induce environmentally responsible behavior.

Compliant behavior can also be explained by social motives, which refer to a willingness to gain approval and respect from actors with which the operator interacts, e.g. customers or other firms in the same industry (Parker and Nielsen, 2011; Winter and May, 2001). Alternatively, social motives can be understood in terms of fear of establishing a bad reputation. Furthermore, an operator may aim at being socially accepted by the inspector, such that the incentives to meet the expectations of the inspector and the EIE agency are strengthened if the relationship between the operator and the inspector is good (Winter and May, 2001).

With regard to the abovementioned theories compliance can be seen as a strategic decision to avoid the potential costs of violations being detected or as a strategy to avoid bad publicity and deteriorating relations to other actors and customers. It can therefore be argued that economic as well as normative and social motives influence an operator’s decisions. It is important to emphasize that an operator can breach certain rules even if the intention is to be compliant, e.g. if the operator lacks the financial means or knowledge to follow these rules. According to Stafford (2012), EIE methods that use threats of sanctions to induce compliance may be inefficient, if the operator does not have the means to comply with the rules or does not understand which targets should be reached or which rules should be followed.

8.2.2. Previous empirical studies

Previous empirical studies analyzing the effect of EIE on environmental behavior indicate that inspections as well as the threat of inspections lead to a reduction in environmentally harmful emissions. The majority of these studies uses North American observational data from mostly large-scale industrial sectors and focus on rational and economic motives for environmental behavior. One such study is Magat and Viscusi (1990) who analyze the impact of inspections on self-reported emission levels among firms in the U.S. pulp and paper industry. Based on quarterly data from 77 firms during 1982–1985, they find that inspections lead to 20% lower emissions and a higher degree of self-reporting. However, the study identifies endogeneity problems as poorly performing firms tend to receive more frequent inspection visits, suggesting that the effect of inspections is underestimated.
To deal with endogenous inspections, Laplante and Rilstone (1996) model the perceived threat of inspections. Their study covers 46 firms in the Canadian pulp and paper industry during 1985–1990 and analyzes the effect on self-reported emissions of both inspections that have actually been carried out and the expected threat of being inspected. The results indicate that the degree of self-reporting increases and that inspections lead to firms reducing emissions by 28%, which is higher than the effect observed by Magat and Viscusi (1990).

An impact of inspections on environmental behavior is also found in Hanna and Oliva (2010), who examine the effect of inspections on 17200 American manufacturing firms during 1987–2001. They find that emissions on average decrease by 15 per cent during a five-year period following an inspection. It is also shown that inspections have a stronger impact on firms with low costs of reducing emissions. However, in contrast to Laplante and Rilstone (1996), no effect of an expected threat of an inspection on environmental behavior is found.

As inspections tend to be endogenous reverse causality may arise, making it more difficult to establish a causal effect of inspections (Gray and Shimshack, 2011). To deal with this problem Telle (2013) uses an experimental study to estimate the difference in adherence to the law during 2008–2010 among environmentally harmful operations that had been randomly assigned to treatment and control groups. The treated operations were inspected, while the control group carried out self-monitoring. After two years follow-up inspections were carried out in both groups. Telle finds that the number of detected violations is relatively higher among inspected operators, suggesting that self-monitoring firms tend to under-report emissions. It is also established that inspections decrease the likelihood of violations in the following year by 37 per cent. While our study measures the effect of the timing of routine inspections that occur with three year intervals, the Telle (2013) study investigates the effect of “first ever” inspections. Accordingly, we expect the effects of an inspection on inspectee compliance to be lower in our study.

8.3. EIE of dry cleaners in Stockholm

The negative environmental impact of dry cleaners consists mainly of the use of the harmful and carcinogenic solvent perchlorethylene, which is a volatile organic compound that can cause damages to the nervous system, the kidneys and the liver, and is also dangerous for aquatic organisms.\(^4^4\) Since perchlorethylene easily evaporates, emissions may occur when it dries (Kemikalieinspektionen, 2014). It is therefore subject to legal restrictions in its use to avoid it spreading through aerial emissions and waste (Miljöförvaltningen, 2012). Hence, activities that consume

\(^{44}\) Although there exist more environmentally friendly alternatives, e.g. hydrocarbon cleaning or wet cleaning, almost all dry cleaners in our panel used perchlorethylene as a solvent during the whole period 2000 – 2013.
more than one kilogram of organic solvents per year are obliged to notify the municipal environmental department and are subject to EIE (SFS 1998:899).

Presently the environmental department in Stockholm carries out inspections of 45 dry cleaners that are covered by the notification obligation. These operations are usually small, cyclically sensitive and have a high frequency of change in ownership (Miljöförvaltningen, 2010). EIE of dry cleaners in Stockholm is financed through the “polluter pays principle”, i.e. the current annual fee of SEK 4500 that operators are charged is set to essentially cover the costs of conducting inspections (Miljöförvaltningen, 2012). Having a fixed EIE fee implies that all dry cleaners should be treated in an equivalent way by the environmental department in terms of the frequency of inspections.45

EIE of dry cleaners is intended to lower the risk of emissions of chemicals, to promote better maintenance routines, to improve self-monitoring among operators, and to provide support through information and advice (Miljöförvaltningen, 2010). All dry cleaners are inspected by the same inspector, and inspections are usually carried out without prior notification approximately every three years. Usually it is the time since the last inspection that determines when a facility is inspected. An operator being inspected in year $t$ can thus expect the next inspection to take place in year $t+3$, but cannot be sure about the exact date. The incidence of these inspections is thus not endogenously determined by the environmental behavior of the operator. Besides these inspections that are carried out systematically, facilities are also supposed to be inspected whenever new equipment has been reported, there has been a change in ownership or as a consequence of complaints regarding e.g. noise, smell or other types of disturbances. Such unplanned inspections are not covered by the annual EIE charge; in this case operators have to make additional payments on the basis of an hourly fee (Miljöförvaltningen, 2012).

To verify that dry cleaners self-monitor and manage their facilities correctly, these are required to submit a report by January 31st every year. This annual report must contain data on the used quantity of perchlorethylene, the quantity of washed gods, the quantity of removed hazardous waste, and the calculated aerial emissions of perchlorethylene per kilogram laundry (in per cent). According to the Swedish Environmental Protection Agency (SEPA) aerial emissions of perchlorethylene must not exceed 2% per kilogram dry cleaned laundry (SFS 2013:254). To induce a further reduction the environmental department of Stockholm municipality has since 2010 reduced the target value to 1.5% for dry cleaners in Stockholm (Miljöförvaltningen, 2012).

45 In contrast, applying the EIE strategy “Targeting Policy”, which has been examined in previous studies, would lead to limited resources forcing EIE activities to be concentrated to operations with the worst environmental record (Hanna and Oliva, 2010; Rousseau, 2007).
Reporting emissions exceeding the target level does not lead to further inspections or economic sanctions. Instead non-compliant dry cleaners are urged to find the reason for the high emission level and to instantly carry out measures that reduce emissions below the target level.\footnote{This can be achieved by e.g. through improved maintenance of the machines or through longer drying times and tougher boiling out of distilled sludge (Miljöförvaltningen, 2003).}

The environmental department provides written feedback to all operators that submit an annual report and complementary data may be requested if missing. Operators that have failed to submit the annual report by January 31\textsuperscript{st} receive a reminder, followed by a phone call in case the report is still missing. If, thereafter, the annual report has still not been submitted, the environmental department can impose an injunction with a conditional fine. During the period 2000–2013 no dry cleaner was charged a fine. Nor has any dry cleaner in our study been punished with an environmental sanction charge for violating laws of the Environmental Code during the period 2002–2014.

### 8.4. Data

Data for this study has been collected from the environmental department of the municipality of Stockholm. It comprises all 50 dry cleaners that have been inspected during the period 2000–2013. Our data set is based on all submitted annual reports (including submission dates) and records of where and when inspections have been carried out in this period (a total number of 208). Unfortunately some annual reports are missing. The number of dry cleaners has been relatively constant over time, but not all dry cleaners have submitted a report every year. This is partly due to the fact that some of the dry cleaners were established after 2000, while others terminated their activity before 2013. An annual report may also be missing if there has occurred a change in ownership, since the environmental department does not require new owners to report emissions of the previous owner. For three dry cleaners annual reports during 2000–2005 have disappeared from the archive of the environmental department.

Our study does not take into account observations of dry cleaners that have replaced an old dry cleaning machine with a more environmentally friendly one, e.g. a hydrocarbon cleaning machine, since such machines do not use perchlorethylene as a solvent and hence, emission measures are not comparable. This causes some missing values and could therefore lead to selection problems, if these dry cleaners differ significantly from others in the sample.

For dry cleaners that have reported emission losses for more than one dry cleaning machine we have calculated a weighted average of aerial emissions of perchlorethylene per kilogram laundry. Our data set consists of a total of 50 dry
cleaners and 494 observations. However, the regressions only contain 48 dry cleaners and 413 observations, because we use models with lagged variables (see section 5).

As in previous studies we use self-reported data (Laplante and Rilstone, 1996; Magat and Viscusi, 1990; Lin, 2013). According to the environmental department, the data that dry cleaners report in their annual reports is not verified, which may lead to inaccurate reporting. Systematic under-reporting of emissions could lead to biases in the inspection coefficients. However, considering that emission values above the target value have not led to any sanctions, it seems reasonable to assume that the incentives for under-reporting emissions are weak.

Misrepresentation of values may also be due to careless calculations. According to the environmental department it has occurred that dry cleaners report used quantity of solvent in liters and not in kilograms. Since perchlorethylene has a density of 1.62 g/cm³, reported emissions in liters will be lower than the requested levels in kilograms. The environmental department has corrected such mistakes in their file containing data for 2008–2013. However, since this has not been done for 2000–2007, we have to emphasize that there may be some errors in the self-reported data in that period.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported emissions (%)</td>
<td>494</td>
<td>0</td>
<td>12.8</td>
<td>1.38</td>
<td>0.94</td>
</tr>
<tr>
<td>Share of rep. emissions above permitted value</td>
<td>494</td>
<td>0</td>
<td>1</td>
<td>0.16</td>
<td>0.37</td>
</tr>
<tr>
<td>Share of delayed annual reports</td>
<td>494</td>
<td>0</td>
<td>1</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Number of inspections per dry cleaner and year</td>
<td>494</td>
<td>0</td>
<td>3</td>
<td>0.34</td>
<td>0.55</td>
</tr>
<tr>
<td>Days between inspections</td>
<td>158</td>
<td>11</td>
<td>2748</td>
<td>1048</td>
<td>787</td>
</tr>
<tr>
<td>Total number of inspections</td>
<td>208</td>
<td>2</td>
<td>15</td>
<td>5.15</td>
<td>3.03</td>
</tr>
</tbody>
</table>

Table 8.1. Descriptive statistics of dry cleaners 2000-2013, number of observations (N), min, max and mean values, standard deviations

Table 8.1 provides descriptive statistics of our data. Self-reported emissions of perchlorethylene represent on average 1.38% of dry cleaned laundry. There are two reported zero emissions, while the highest reported value is 12.8%. Since our results could be strongly influenced by these extreme values, we also include regressions where these are excluded. The average share of submitted annual reports with self-reported emissions exceeding the target values is 16 per cent per year. On average more than half (54%) of annual reports are submitted after the January 31st deadline. The average number of inspections per dry cleaner and year is 0.34, which suggests that a dry cleaner is inspected every third year. At 18 occasions between 2000 and 2013 a dry cleaner has been inspected more than once in a year, which indicates that inspections are not entirely exogenous. The interval
between two inspections for a dry cleaner is on average slightly less than three years (1048 days), varying between eleven days and more than 7.5 years, also confirming that the three-year interval for inspections is not entirely predictable. On average a dry cleaner has been inspected five times between 2000 and 2013.

Figure 8.1. Number of submitted annual reports and number of inspections

Figure 8.2. Self-reported emissions over time

Figure 8.1 graphically illustrates the number of inspections and the number of submitted annual reports between 2000 and 2013. During the period 2005–2009 the number of inspections is lower, although the number of active dry cleaners remains the same. The number of submitted annual reports is also lower in this period, indicating that there might be a correlation between the number of inspections and
the number of submitted annual reports. Figure 8.2 shows the dispersion of self-reported emissions between 2000 and 2013. We see that there are four observations that strongly exceed the average.

8.5. Methodology

Regressions may be influenced by non-observable factors that differ among dry cleaners but are assumed to be constant over time, e.g. location and size of the operation. Larger dry cleaners can, for example, be assumed to have larger financial capital, which might increase the likelihood of reaching environmental targets (Winter and May, 2001). To control for such non-observable factors we include fixed effects for each dry cleaner, which has also been done by e.g. Hanna and Oliva (2011) and Laplante and Rilstone (1996). Since not all dry cleaners are inspected every year we also include year effects to control for variable factors that can be assumed to impact on the environmental behavior of all dry cleaners, e.g. business cycles, technological progress or general changes of environmental attitudes in society. By considering year effects it is also possible to isolate the correlation between inspections and environmental behavior from other types of EIE (e.g. information campaigns), assuming that these activities influence all dry cleaners similarly. Year effects should also pick up the impact of switching the inspector, assuming that a change of inspector influences all dry cleaners similarly.

Our data lacks information regarding the cause of individual inspections. Hence, we do not know whether an individual inspection has been carried out routinely or due to e.g. a change of owner or a complaint. Although our data suggest that most inspections have been carried out regularly every third year, we include lagged values for inspections, like e.g. Gray and Shimshack (2011), to deal with reverse causality.  

When using data from registries it is hard to measure how dry cleaners perceive being inspected (Gray and Shimshack, 2011), which is central for understanding behavioral changes of operators. Previous studies have dealt with this by estimating the expected probability of being inspected (e.g. Hanna and Oliva, 2011, and Laplante and Rilstone, 1996). According to Gray and Shimshack (2011) lagged inspections can be used as a substitute for the perceived threat of future inspections, under the assumption that the likelihood of future inspections is based on previous experiences of inspections.

---

47 To deal with reverse causality other empirical strategies have also been employed, e.g. the use of instruments (Laplante and Rilstone, 1996, and Lin, 2013) or an experimental set-up where inspections are carried out randomly (Telle, 2013).

48 To avoid problems with multi-collinearity we have chosen not to include lagged values for inspections for more than one year in the past, as dry cleaners, on average, are inspected every third year.

49 Because of the three-year interval of inspections the risk of being inspected in period $t$ can be assumed to be low if an inspection was carried out in period $t-1$. 
Autocorrelation is common in panel data where the value of a variable in period $t$ depends on the values in period $t-1$, which may lead to underestimated standard errors and hence, erroneous significant results. We therefore use clustered standard errors, which allow observations for each dry cleaner to be correlated over time and to also be robust to heteroscedasticity.

8.6. Empirical models and results

We consider two outcomes of inspections, self-reported emissions on the one hand, and submitting reports on time on the other hand. Regarding emissions, we estimate the effects of inspections both on the extensive and the intensive margin. Moreover, we also attempt to measure the impact of inspections over time.

We present the results from our regression models, with and without extreme values. Since we use lagged variables, data from the year 2000 are excluded. Totally there are 413 observations from 48 dry cleaners in the period 2001–2013.

8.6.1. Model 1: extensive margin effects of inspections on emissions

First, we assess how compliance regarding emissions is affected by inspections, i.e. we examine the extensive margin decisions of dry cleaners. To that end we employ a linear probabilistic model to examine how inspections affect the likelihood of self-reporting emissions above the target value.\textsuperscript{50} Linear probabilistic models generate heteroskedastic error terms, which can be dealt with by using clustered standard errors. The model is specified as follows:

$$AL_{i,t} = \beta_1 I_{i,t} + \beta_2 I_{i,t-1} + a_i + y_t + u_{i,t}$$

(1)

$AL_{i,t}$ (above limit) is a binary variable, which takes on value 1 if perchlorethylene emissions of dry cleaner $i$ are equal to or above the target level in year $t$ and takes on value 0 else. The target (threshold) value is 2% for observations between 2000 and 2009 and 1.5% for observations between 2010 and 2013. As controls we use $I_{i,t}$ (inspection in period $t$, which takes on value 1 if dry cleaner $i$ has been inspected at least once in period $t$) and the lagged value for inspection in the previous year $I_{i,t-1}$, as well as dry cleaner and year fixed effects, $a_i$ and $y_t$, and the error term $u_{i,t}$ that includes varying and non-observable factors that influence the environmental behavior of dry cleaners.

\textsuperscript{50} Since the model assumes a linear relationship, there is a risk that predicted probabilities are below zero or above one, which is not a problem, if the aim is to analyze the relationships between explanatory variables and the dependent variable, as in our case.
Each coefficient on the right-hand side can be interpreted as the average estimated change in the likelihood of a dry cleaner exceeding the threshold value, given a one unit change in the independent variable when keeping the other independent variables constant. The effects of inspections on the likelihood of exceeding the target value are presented in table 8.2.

**Table 8.2. The effect of inspections on emissions above the permitted value**

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Inspection}_{i,t}$</td>
<td>-0.029</td>
<td>-0.009</td>
<td>-0.016</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.035)</td>
<td>(0.038)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>$\text{Inspection}_{i,t-1}$</td>
<td></td>
<td></td>
<td>-0.067*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.037)</td>
<td></td>
</tr>
<tr>
<td>$\text{Intercept}$</td>
<td>0.167***</td>
<td>0.160***</td>
<td>0.137**</td>
<td>0.171**</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.010)</td>
<td>(0.068)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Dry cleaner fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustered standard errors</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-test for the exclusion of variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects = 0</td>
<td>1.84</td>
<td>1.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.137)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inspections = 0</td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>0.039</td>
<td>0.046</td>
</tr>
</tbody>
</table>

*p-values are reported in brackets below the F-values. Within R-square used.

As expected, all coefficients for the inspection variables are negative. The coefficients for inspections in period $t$ are not statistically significant in any of the regressions. In column (4) one-year lagged inspections are included. The coefficient is negative and statistically significant at the 10 per cent level, indicating that an inspection in the previous year is associated with the likelihood of exceeding the target value being 6.7 percentage points lower compared to the case where an inspection has taken place in period $t-2$ (when keeping inspections in period $t$ constant and controlling for fixed effects for units and years). However, the F-test in column (4) shows that the hypothesis of the joint effect of the coefficients of the inspection variables to be equal to zero cannot be rejected. Hence, inspections in year $t$ should perhaps be excluded from the regression specification. We have run alternative regression specifications where we leave out the inspections in period $t$. In these specifications we get even stronger effects from...
inspections in period $t-1$ and a higher statistical significance. However, we chose to show the weaker (and more conservative) results in table 2.

We also ran a regression where we, in addition to the specification in column (4), included the lag of the dependent variable, $AL_{i,t-1}$, as a regressor on the right-hand side of the regression equation. We find significant support for the likelihood of exceeding the target value in period $t$ to be positively correlated with the target value having been exceeded in the previous year, when inspections are held constant. However, including a lagged value for the dependent variable on the right-hand side, can lead to biased and non-consistent estimations which is why we refrain from including the regression results in table 2. Still, the regression indicates that dry cleaners emitting above the target value are more likely to continue doing so compared to dry cleaners emitting below the target value.

8.6.2. Model 2: intensive margin effects of inspections on emissions

In the above regressions we established that inspections in the preceding year have a positive effect on compliance. Hence, there seems to be an effect of inspections on the extensive margin. We now focus on the intensive margin decisions of dry cleaners to explore whether there is an effect of inspections on self-reported emissions of perchlorethylene. The model that we test is specified as follows:

$$E_{i,t} = \beta_1 I_{i,t} + \beta_2 I_{i,t-1} + a_{i,t} + y_t + u_{i,t}$$

$E_{i,t}$ represents self-reported emissions of perchlorethylene in per cent of total wash volume. The coefficient $\beta_1$ measures the change in reported emissions in year $t$ from (at least one) inspection in year $t$. As above we include lagged values for inspections and emissions in the previous year where the coefficient $\beta_2$ measures the change in reported emissions in year $t$ from (at least one) inspection the previous year. Finally, we have dry cleaner and year fixed effects and the error term $u_{i,t}$. The effect of inspections on self-reported emissions of perchlorethylene are presented in table 8.3.

Column (1) presents a model with inspection in period $t$ without fixed effects for unit and year. The inspection coefficient is negative but not significant, which is in accordance with results of previous studies (Laplante and Rilstone, 1996). The inspection coefficient hardly changes when fixed effects and clustered standard errors are included in column (2). Neither does the $R^2$ value change substantially, which indicates that fixed effects for dry cleaners do not pick up much of the variation of the dependent variable. The coefficient is weakened and the $R^2$ value increases when year effects are included in column (3). Thus, the share of explained variation increases if year effects are included; however, this does not give us any information regarding the causal effect of year effects.
In column (4) one-year lagged inspections are included. The coefficient of the lagged variable is negative, but not statistically significant. The results from the F-tests for common significance imply that we cannot reject the zero hypothesis of the coefficients for the inspection variables being equal to zero in regression (4). However, the year effects are jointly significant. To conclude, we do not find any statistically significant evidence for inspections being associated with lower reported emissions.

<table>
<thead>
<tr>
<th>Dependent variable: perchlorethylene emissions in per cent of dry cleaned volume</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspection&lt;sub&gt;it&lt;/sub&gt;</td>
<td>-0.032</td>
<td>-0.018</td>
<td>-0.001</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.073)</td>
<td>(0.073)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Inspection&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.051)</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.323&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.320&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.440&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.459&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.022)</td>
<td>(0.121)</td>
<td>(0.137)</td>
</tr>
<tr>
<td>Dry cleaner fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustered standard errors</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-test for exclusion of variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects = 0</td>
<td>3.75</td>
<td>3.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inspections = 0</td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.693)</td>
</tr>
<tr>
<td>Observations</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
</tr>
<tr>
<td>R²</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>0.066</td>
<td>0.067</td>
</tr>
</tbody>
</table>

Significance levels for coefficients: *p<0.10, ** p<0.05, *** p<0.01. 
*p-values are reported in brackets below the F-values. Within R-squared used.

Table 8.3. The effect of inspections on perchlorethylene emissions

These results show that inspections have no significant impact on emissions of perchlorethylene. However, our regressions on compliance indicate that inspections lead to a reduction in emissions above the permitted value in the following year. Hence, there seems to be an effect of inspections on the extensive but not on the intensive margin, suggesting that inspections only reduce emissions of non-compliant dry cleaners. Since the rate of compliance is high (84%, on average), the total effect on emissions is therefore insignificant.

8.6.3. Model 3: inspection effects on the timeliness of submissions of reports

The second outcome variable that we consider is the timeliness of submitting the annual report on time. We employ a linear probabilistic model to analyze the effect of inspections on the likelihood of dry cleaners not submitting their annual reports on time.
\[ LR_{it} = \beta_1 I_{it} + \beta_2 I_{i,t-1} + \alpha_i + \gamma_t + \epsilon_{it} \] (3)

\( LR_{it} \) (late report) is a binary variable that takes on value 1 if the annual report for dry cleaner \( i \) has not been submitted by January 31st in year \( t \) and takes on value 0 else. As in the previous models we include lagged values for inspections in the previous year, dry cleaner and year fixed effects and the error term \( \epsilon_{it} \).

<table>
<thead>
<tr>
<th>Dependent variable: No annual report submitted by January 31(^{\text{st}} ) = 1</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspection(_{i,t})</td>
<td>0.013</td>
<td>0.009</td>
</tr>
<tr>
<td>( \text{(0.045)} )</td>
<td>( \text{(0.051)} )</td>
<td></td>
</tr>
<tr>
<td>Inspection(_{i,t-1})</td>
<td></td>
<td>-0.057</td>
</tr>
<tr>
<td>( \text{(0.053)} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.680***</td>
<td>0.709***</td>
</tr>
<tr>
<td>( \text{(0.064)} )</td>
<td>( \text{(0.067)} )</td>
<td></td>
</tr>
<tr>
<td>Dry cleaner fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Clustered standard errors</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-test for exclusion of variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year effects = 0</td>
<td>7.83</td>
<td>7.35</td>
</tr>
<tr>
<td>( (&lt;0.001) )</td>
<td>( (&lt;0.001) )</td>
<td></td>
</tr>
<tr>
<td>Inspections = 0</td>
<td></td>
<td>0.66</td>
</tr>
<tr>
<td>( (&lt;0.522) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>410</td>
<td>410</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.191</td>
<td>0.194</td>
</tr>
</tbody>
</table>

*p-values are reported in brackets below the F-values. Within R-square used.*

**Table 8.4. The effect of inspections on the timeliness of submitted annual reports**

The effects of inspections on the probability of submitting the annual report on time are presented in Table 8.4. Compared to the previous models the number of observations is lower, because the dates of submission of the annual reports are missing in three cases. Contrary to what we expected, and in contrast to the results of models 1 and 2, the coefficients for inspections in period \( t \) are positive, albeit not statistically significant. The coefficient for inspection in \( t-1 \) is indeed negative, but statistically insignificant. The F-test shows that we cannot reject the hypothesis of a zero joint inspection effect. Hence, we find no statistically significant effects of the inspections variables on the likelihood of annual reports being submitted on time.

### 8.6.4. Results from model 2 without extreme value

When we include lagged variables in our models some data are excluded and three observations with potential extreme values disappear. There is, however, one observation with a reported emission of 8.9 per cent still remaining in our regressions, which deviates strongly from the average. Extreme values in data sets
with few observations is problematic, as extreme values impact strongly on results. We therefore conduct a sensitivity analysis, where we examine how our results from model 2 change when this observation is excluded (both as present and lagged value). The results are presented in table 8.5.

| Dependent variable: Emissions of perchlorethylene in per cent |
|-------------------|-------------|-------------|-------------|-------------|
| **Independent variable** | (1) | (2) | (3) | (4) |
| Inspection | -0.093* | -0.071** | -0.057 | -0.075 |
| | (0.053) | (0.035) | (0.035) | (0.046) |
| Inspection_{t-1} | -0.047 | | | |
| | (0.044) | | | |
| Intercept | 1.323*** | 1.317*** | 1.469*** | 1.493*** |
| | (0.033) | (0.010) | (0.097) | (0.110) |
| Fixed effects for dry cleaner | No | Yes | Yes | Yes |
| Year effects | No | No | Yes | Yes |
| Clustered standard errors | No | Yes | Yes | Yes |
| F-test for the exclusion of variables | | | | |
| Year effects = 0 | 4.66 | 4.58 | | |
| | (<0.001) | (<0.001) | | |
| Inspections = 0 | | | 1.38 | |
| | | | (0.262) | |
| Observations | 412 | 412 | 412 | 412 |
| R² | 0.006 | 0.006 | 0.07 | 0.072 |

Significance levels for coefficients: *p<0.10, **p<0.05, ***p<0.01. p-values are reported in brackets below the F-values. Within R-square used.

Table 8.5. The effect of inspections on emissions of perchlorethylene without extreme value

The results presented in table 5 differ from the results in table 8.3, thus suggesting that the extreme value impacts on the analysis of model 2. Compared to table 3 the coefficient for inspections in the present period on emissions of perchlorethylene is more negative in all regressions, and, more importantly, significant at the 5 per cent level in regression (2) and significant at the 10 per cent level in regression (1). The coefficient for inspections in the previous year does not exhibit any significant correlation.

We also run a regression, not represented in table 5, where we add the lagged dependent variable on the right-hand side of the regression equation in column (4). The R² value in this additional regression then increases from 0.072 to 0.321, which indicates that the variation of emissions in period t can to a large degree be explained by emissions in the previous year. As expected, the coefficient for the lagged value for emissions is significant at the 1 per cent level. This correlation was also established in Laplante and Rilstone (1996). For the same reasons as before we do not show the regression results from this additional specification due
8.7. Discussion

The aim of this study was to examine the effect of inspections on the environmental behavior of dry cleaners in the municipality of Stockholm during the period 2000 – 2013. To that end we have analysed three regression models where we have investigated the relationship between inspections and three outcome measures: (1) the likelihood of reporting emissions above the permitted value, (2) self-reported emissions of perchlorethylene, and (3) the probability of submitting annual reports on time. Model (1) refers to effects on the extensive margin while model (2) refers to effects on the intensive margin.

Analyzing the extensive margin in model 1 yields statistically significant support for inspections in period \( t-1 \) reducing the risk of dry cleaners exceeding the target value in the following year by, on average, 6.7 per cent points compared to having an inspection in \( t-2 \), when all other independent variables are kept constant. A potential explanation for this result is that those dry cleaners whose emissions are already below the permitted value have small incentives for further reducing emissions. This effect may seem small, but we should remember that this is compared to having an inspection in \( t-2 \). Hence, operators seem to benefit from being reminded and the effects from the inspection on the environmental behavior takes until next year to materialize. Naturally, an ideal empirical identification strategy would have been having a treatment group receiving inspections and a control group not having inspections where both groups previously had not received inspections (as in the Telle, 2013, study). However, these ideal conditions are hard to come by.

Looking at the intensive margin in model 2 we find a negative relationship between inspections and self-reported emissions, which is in line with results from previous studies (Hanna and Oliva, 2010, Laplante and Rilstone, 1996, Magat and Viscusi, 1990). However, our inspection coefficients are not statistically significant when an extreme value is included and hence, the results should be interpreted with caution. When the extreme value is excluded we find that the impact of inspections in year \( t \) on reported emissions in that year is negative and statistically significant.

We also find that 54 per cent of annual reports are submitted to the environmental department late, that is, after January 31st. We do not find any support for inspections affecting the probability of dry cleaners submitting their reports on time.
Previous empirical studies regarding EIE have emphasized that improved environmental behavior of operators is due to inspections having a deterrent effect on breaching legislation and regulations (Hanna and Oliva, 2010, Laplante and Rilstone, 1996). These studies have assumed that the environmental behavior of operators can primarily be explained by economic incentives and that the degree of compliance will be higher if both the threat of being inspected and the cost for violations are larger. However, theories on the deterrent effect of inspections seem to be of little relevance in this study, given the fact that the environmental department of Stockholm municipality has not imposed any sanctions or charges on dry cleaners during the period 2000 – 2013, although dry cleaners frequently do not submit annual reports and, on average, 16 per cent have reported emissions above the target value. Nor have higher reported emissions or failing to hand in the annual report on time led to any extra inspections. Hence, the expected cost of a dry cleaner for violating rules seems to be very low. In this setting it is then remarkable that inspections do have an effect at all.

According to the Environmental Code inspections and enforcement can be carried out with several different methods. To reach the Code’s target of reduced emissions methods can be adjusted with respect to industry-specific conditions. The environmental department in Stockholm systematically conducts inspections of dry cleaners every third year, which aim at supporting preventive measures to reduce the risk of perchlorethylene being emitted to the air. Since the inspections focus on providing information and advice regarding environmental issues to enhance improvements in attitudes and behavior, our results seem to render the informative and educational impact stronger support than the deterrent effect of inspections. Based on theories regarding normative motives for the environmental behavior and compliance of operators, the results from the analysis of the extensive margin (emissions below the target value) could be explained by inspections enhancing the internalization of norms regarding environmental responsibility and also possibly as a reminder of what good environmental behavior is.

In contrast to most previous studies concerning inspections and enforcement we have focused on an industry consisting of small enterprises, which implies that the inspection style has to be adjusted accordingly. Theories on the social motives of environmental behavior suggest that compliance increases if the relationship between inspector/authority and the operator is good. It is plausible that the “soft” inspection style with no sanctions that has been implemented by the environmental department has fostered a good relationship between the parties.

This study is based on a small sample of dry cleaners, which naturally limits the external validity of our results. It is therefore difficult to generalize the results to dry cleaners in other regions or to other environmentally hazardous activities. Further research on other types of activities and other types of inspections and enforcement methods in Sweden and other countries is therefore desirable. Other potential research topics include evaluating cost efficiency of inspections and
enforcement. In this study we have used self-reported data, which may give rise to errors in our analysis (as indicated in, for example, Telle (2013)). To more reliably establish causal effects of inspections on the environmental behavior of operators, more experimental studies like Telle (2013) should be carried out.

8.8. References


9. Measuring the effects of feedback from inspections on cleanliness in Swedish preschools – an experimental approach\textsuperscript{51}

Mathias Herzing and Adam Jacobsson\textsuperscript{52}

9.1. Introduction

The aim of this study is to examine the impact of information following an inspection on future behavior of inspectees. To that end an experimental set-up was employed to evaluate whether the provision of feedback regarding cleanliness in preschools led to improvements. The study was carried out in three Swedish municipalities in cooperation with MSL (Miljösamverkan Stockholms län – “Environmental Collaboration in Stockholm county”).

In the spring of 2016 the level of organic residue on children’s toilet door handles were measured. High levels of organic residue on door handles can spread diseases, thus affecting the health of children and their families, as well as of pre-school staff. These so-called ATP (adenosine triphosphate) measurements took place without the knowledge of the preschools in connection with inspections that were conducted as part of a project focusing on children’s exposure to chemicals. Before a second unannounced measurement was carried out in October 2016 all preschools in one of the municipalities (the treatment group) were sent a letter alerting them to the importance of keeping toilet handles clean and containing information on their individual results as well as their performance in relation to the other preschools in that municipality. The preschools in the other two municipalities (the control group) did not receive any feedback on the first measurement and hence, were not aware that the cleanliness of toilet handles had been checked. This set-up thus enabled us to use a difference-in-difference model to capture the effect of the treatment and test the following hypothesis:

\textsuperscript{51} We would like to thank the participating municipalities of Ekerö, Södertälje and Österåker for their kind assistance on this project. We also wish to thank Mahmood Arai and Jonas Vlachos for valuable comments on the manuscript. All errors are, of course, our own.

\textsuperscript{52} Department of Economics, Stockholm University, e-mail: Mathias.Herzing@ne.su.se, aja@ne.su.se
Hypothesis: Pre-schools in the treatment group will improve their cleanliness more than pre-schools in the control group.

After the second round of measurements had been completed we conducted follow-up telephone interviews with the heads of the pre-schools in the treatment municipality to learn more about their reactions to the letters. This enabled us to also test whether pre-schools claiming to have undertaken actions in response to the letter improved their measurement results more than those who had not.

This study is structured as follows. The following section provides a background to our study and a motivation for our contribution. Section 9.3 presents our research design and the basic data. Results are reported in section 9.4. Section 9.5 contains conclusions and a discussion.

9.2. Background

Environmental and public health safety inspections are essential for promoting compliance with legislation and enhancing a better environment. Effective enforcement of laws and regulations has been shown to be necessary for the proper functioning of the legislative instrument (see, e.g., Gunningham, 2011). Nevertheless there is a lack of knowledge regarding the efficiency of different methods and interventions in practical inspection and enforcement work. This paper will contribute by looking at one specific tool, the provision of written feedback information on measurements that were undertaken at an inspection to assess the cleanliness in pre-schools.

Most previous studies on the effects of different inspection measures on environmental and public health safety outcomes have applied econometric techniques to secondary data and/or data from natural experiments. In contrast, the research strategy of this study is to use a simple but clear randomized quasi experimental design where we can select control and treatment groups as well as the intervention. One advantage of this approach is that it generates primary data for analysis with a direct practical relevance for the operational inspection authorities. This strategy is intended to improve the accuracy in our identification of causal relationships.

A major difficulty in assessing actual inspection methods is the lack of data. We therefore need to generate our own data. For this study we have chosen an experimental strategy with the aim of generating exogenous variation allowing us to go beyond mere correlation analysis towards analyzing causality. Since inspections are costly it was necessary for us to embed our study in the day-to-day work of the municipalities. A major challenge is to find municipalities who are able and willing to participate in experiments on different inspection instruments. In
cooperation with MSL we found three municipalities – Österåker, Ekerö and Södertälje – that wanted to participate in our study.

Previous research has indicated that direct regulation with associated environmental inspections and enforcement (EIE) is the most efficient set of policies to promote good environmental behaviour (Gray, Shimshack, 2011; Delmas, Aigner, 2007). Naturally, one size does not fit all as the efficiency of EIE varies across different types of inspectees (e.g. firms or individuals) depending on factors such as industry, market conditions, production technology, firm size and organizational structure. EIE can be classified into two broad classes – control methods such as physical inspections and enforcement on the one hand, and compliance assistance like subsidies, technical assistance and information campaigns on the other hand. Control measures generate deterrence (see, for example Gunningham et al, 2005; Magat, Viscusi, 1990; Laplante, Rilstone, 1996; Hanna, Oliva, 2010; Telle, 2010) while compliance assistance makes it easier for the inspectee to improve environmental behaviour. Stafford (2012) empirically investigates the relationship between control and compliance assistance and shows that both methods can be effective and that it depends on the type of inspectee whether these two methods are complements or substitutes.

Inspectees may not, however, always breach environmental legislation as a result of rational deliberation, as outlined in the rational polluter model (Becker, 1968), but may instead do so inadvertently due to lacking understanding and knowledge of the environmental requirements (Stafford, 2012). If an inspectee belongs to the latter category, the most efficient method of EIE is likely to be to provide information. Typically information affects inspectees’ environmental behaviour in two ways. First, information may simply provide the knowledge necessary to address the environmental problem at hand. Secondly, the type and formulation of the information may be designed to provide additional pro-environmental drivers such as, for example, individual and social norms.

Matthies’ (2005) heuristic model, as outlined by Fischer (2008), explains individual environmental behaviour occurring in three stages. In the first norm activation phase the individual becomes conscious about some environmental problem (for example as a result from feedback from an inspection agency), the relevance of his/her own behavior and his/her sense of control (over the environmental behavior). During the next phase the individual then weighs and evaluates the situation using different motivations such as personal and social norms or other motives such as economic ones. In the final phase the individual decides what actions to take, if any. The model predicts that feedback will be effective if it captures the individual’s attention, if it links specific actions to specific effects, and finally, if it activates different motives for change such as cost savings, resource conservation, emission reduction and competition (with peers). Naturally, the motives that are supposed to be activated must also be tailored for the specific target group of individuals. We should note that even though the
current paper concerns activities under the purview of the public health safety
legislation (and not the environmental code), the activity of cleaning door handles
for the benefit of a better indoor environment is close enough in character for
Matthies’ model to be relevant for our purposes.

Based on these theoretical findings, Fischer (2008) assesses five review studies and
21 original papers, all in all covering 26 projects from 10 countries, with the
purpose of finding which type of feedback is successful in the specific case of
reducing household electricity consumption. The author finds that the most
successful feedback combines the following factors: it is given frequently and over
a long time, it has an appliance-specific breakdown, it is presented in a clear and
simple way, and it uses computerized and interactive tools.

One specific motivator is peer comparison which activates social norms thus
creating a “norm to conform” (Ayres et al, 2012, page 993). On the one hand, an
individual that receives information explaining that he/she is underperforming in
relation to his/her peers is likely to experience social pressure to improve. On the
other hand, an individual that performs above the average (in terms of positive
environmental behaviour) of his/her peers may instead change his/her behaviour in
an environmentally undesirably way. Fischer’s (2008) review paper finds that peer
comparison feedback information did not have any effect on average electricity
usage. According to Fischer, these results are due to lower-than-average energy
consumers who, in response to the favourable comparison, increase energy usage.
Ayres et al (2012) refer to this phenomenon as a so called “Boomerang effect”.
This highlights the importance of careful framing of information in general and
peer comparison information in particular for obtaining positive environmental
behavioural effects.

Ayres et al, 2012, analyse two large randomized field experiments on a total of 170
000 household customers of two utilities, the Sacramento Municipal Utility District
(SMUD) and Puget Sound Energy (PSE), where a randomized subset of these
households received carefully framed peer comparison information in the shape of
“home energy reports” that compared their energy consumption to that of similar
households. The authors found average reductions in energy consumption in the
households that received reports of 1.2% (PSE) and 2.1% (SMUD) with the effects
lasting 7 and 12 months, respectively. Ayres et al (2012) conclude their paper by
stating that peer-comparison feedback might be an effective tool in other
environmental settings and that this field is in need of further research which we
aim to do in the current paper.

It is clear that the above specific setting of influencing electricity consumers differs
from ours where the receivers of information are heads of pre-schools under
inspection. Still, it is our belief that peer effects are relevant in our setting as well.
9.3. Research design and data

All three municipalities participated in a MSL project called “Chemicals in preschools” where inspectors visit pre-schools in the municipality during spring 2016 to assess the prevalence of dangerous chemicals in e.g. toys. We thus had the opportunity to add an experimental design element at low extra cost to the municipal environmental and public health safety departments. During the inspections of chemicals the inspectors also carried out an ATP (adenosine triphosphate, which is contained in all living cells) measurement to assess the level of organic residue on the childrens’ toilet door handles. This was done by swabbing the handles on the inside of the toilets and then using an ATP meter to obtain a RLU (relative light units, a measure of bioluminescence) number. The direct correlation between total ATP and the number of bacteria is weak, but there is a strong correlation to hygiene, such that a high ATP value indicates a bad hygienic environment.

Österåker was chosen as the treatment municipality because it lies far from Ekerö and Södertälje such that the risk of information regarding the letters spreading to the other municipalities was minimized. This is also why we did not randomly draw individual pre-schools from all three municipalities to the control- and treatment groups but instead picked whole municipalities. The ATP measurements were done in parallel in all three municipalities between February and June 2016. The measurement instruments that were used were two Hygiena SystemSURE Plus (Food Diagnostics) in Södertälje, one Lumitester PD-20 (Kikkoman) in Ekerö, and one Hygiena SystemSURE Plus in Österåker.

On 1 September a letter was posted to all pre-schools in Österåker (the treatment group). The letter (see appendix 1 for a translated version) informed about the ATP measurements that had been carried out during the chemicals inspection, the pre-school’s individual RLU value, the average and median RLU values in Österåker, as well as the percentage of pre-schools in the municipality with a lower RLU value (i.e. a higher degree of cleanliness). Furthermore, it contained information on what RLU values are considered acceptable on freshly cleaned surfaces within the health care sector, and the importance of keeping toilet door handles clean to avoid the spread of infections. Finally, it was emphasized that the study will present the results on an aggregate level so that no individual pre-schools can be identified.

The letter was intended to induce the pre-schools in the treatment group to improve cleaning routines resulting in a lower RLU value in the follow-up ATP measurement in October 2016. There were several mechanisms contained in the letter that we believe would achieve this. Firstly, the RLU value would in most cases be relatively high in comparison to the scale provided from the health care sector, indicating that the handles were quite dirty (even though it was stated that the values apply to freshly cleaned surfaces within the health care sector which is not the case here). This should alert the reader that there is a health safety problem,
thus activating a personal norm to comply with guidelines. Secondly, the pre-schools that are in the higher percentiles (many other pre-schools have lower and thus better results) may feel pressure to improve their status relative to their peers. This constitutes a peer-comparison mechanism and activates social norm pressures to conform. Thirdly, the combined effect of informing about a high RLU value and its potential consequences might have a general “awareness effect” regarding the importance of good hygiene and good cleaning routines. This underscores the relevance of improving the health safety behavior. Finally, the letter could be interpreted as a “signal” from the environmental and public health safety department that cleanliness has received more attention and may therefore be inspected more in the future. This last mechanism possibly creates deterrence on the lines of the literature initialized by Becker (1968).

In order to compare the treatment group (pre-schools in Österåker) to the control group (pre-schools in Ekerö and Södertälje) a second round of ATP measurements was conducted between 10 and 21 October 2016. This time the inspectors made unannounced inspections at all pre-schools. We are thus able to compare the difference in differences between the treatment and control groups’ ATP-measurements. The results are presented in section 9.4.

<table>
<thead>
<tr>
<th></th>
<th>Ekerö</th>
<th>Södertälje</th>
<th>Österåker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pre-schools included in study</td>
<td>28</td>
<td>41</td>
<td>34</td>
</tr>
<tr>
<td>Share of municipal pre-schools (not private)</td>
<td>0.64</td>
<td>0.61</td>
<td>0.53</td>
</tr>
<tr>
<td>Average number of children per pre-school (std. dev.)</td>
<td>24.14 (8.72)</td>
<td>65.34 (35.78)</td>
<td>65.38 (32.42)</td>
</tr>
<tr>
<td>Average number of children per toilet (std. dev.)</td>
<td>10.27 (2.44)</td>
<td>12.41 (4.16)</td>
<td>6.53 (3.47)</td>
</tr>
<tr>
<td>First measurement period</td>
<td>8 Mar – 14 Jun</td>
<td>16 Feb – 30 May</td>
<td>22 Feb – 19 May</td>
</tr>
<tr>
<td>Average date</td>
<td>18 May</td>
<td>9 April</td>
<td>4 April</td>
</tr>
<tr>
<td>Median date</td>
<td>31 May</td>
<td>12 April</td>
<td>7-8 April</td>
</tr>
<tr>
<td>Average date</td>
<td>17 October</td>
<td>18 October</td>
<td>18 October</td>
</tr>
<tr>
<td>Median date</td>
<td>17 October</td>
<td>18 October</td>
<td>18 October</td>
</tr>
</tbody>
</table>

Table 9.1. Summary statistics of the pre-schools included in the study

The second round of measurements took place within a much closer time frame as it was important to avoid information starting to spread. The number of pre-schools included in our study is similar in the three municipalities (see table 1 above). In Ekerö and Österåker all pre-schools were covered, whereas 50% of pre-schools in Södertälje were inspected as part of the chemicals project – inclusion was determined by the time since last inspection (pre-schools that had not received an inspection during the last 3-4 years were included). The timing of inspections as well as the selection of the included pre-schools was random. The share of privately run pre-schools is similar in the three municipalities at around 40 per cent, but somewhat higher in Österåker. Pre-schools were on average much smaller
in Ekerö, with 24 children per pre-school in comparison to 65 children per pre-school in the other two municipalities. The number of children per toilet for which ATP measurements were carried out differs considerably, being 6.5 in Österåker, 10.3 in Ekerö and 12.4 in Södertälje. Round one inspections are spread over a period from 16 February to 14 June while round two inspections are much more concentrated to a period between 10 and 28 October. The table above indicates that Ekerö differs from the other two municipalities in terms of first measurements having been carried out later and also being more concentrated towards the end of the period.

The overlap was not perfect. In round one, inspections in Österåker (the treatment municipality) and Södertälje are more or less evenly and similarly distributed across the period, while Ekerö has a more skewed distribution with more inspections occurring at the very end of the period. In round two Södertälje and Österåker conducted all measurements within one week whereas Ekerö’s measurements took place over a three weeks period. Another factor that might affect the measurements is the time of day when the measurements are done. We have examined the data and could not see any systematic relationship between the time of measurement and the RLU values.

During the follow-up telephone interviews with the treatment group we asked the pre-schools whether they had taken specific actions as a result of the information letter. Out of the 34 participating pre-schools in the treatment municipality we managed to have telephone conversations with 29 of them. In two cases the respondents were not able to answer our questions since the head of the pre-school during the time of the study had quit, thus leaving us with 27 answers. Out of these 24 claimed that they had received the information letter. Two pre-schools claimed that they did not know if the letter had been read since the head of those two pre-schools had been on sick leave and the temporary head during the time of the study had by now retired. Hence, we could in the end verify that 22 pre-schools had both received and read the information letter. Out of these 11 claimed that they had taken actions as a result from reading the information letter. These results are summarized in table 9.2.

| Number of pre-schools in the treatment municipality in the study. | 34 |
| Number of pre-schools who could answer our questions. | 27 (79%) |
| Number of pre-schools that claimed that they had received the information letter. | 24 (71%) |
| Number of pre-schools that claimed that they had received and read the information letter. | 22 (65%) |
| Number of pre-schools that claimed that they had taken actions as a result from reading the information letter. | 11 (32%) |

Table 9.2. Results from the telephone follow-up interviews.
There is thus some uncertainty concerning the observations where we could not verify that the pre-school in question had read the feedback letter. However, those pre-schools may still have received and read the letters. We will at some points in the analysis below exclude these observations to see if and how they might affect the results. Unless otherwise specified, these observations are included in the analysis.

As handles in two toilets in each pre-school were swabbed each inspection we created two variables for the average results per pre-school at the two inspections: \( \text{average\_result\_time1} \) and \( \text{average\_result\_time2} \). We also created the variable \( \text{avdiff} \), measuring the difference between the average results (\( \text{average\_result\_time2} - \text{average\_result\_time1} \)). Finally, in order to better illustrate the relative change, we created the variable \( \text{avdiff\_percent} \). Summary statistics of these variables can be found in table 9.3 below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OBSERVATIONS</th>
<th>MEAN (STD. DEV.)</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Söd</td>
<td>Eke</td>
<td>Öst</td>
<td>Söd</td>
</tr>
<tr>
<td>( \text{av_time1} )</td>
<td>41</td>
<td>28</td>
<td>34</td>
<td>633</td>
</tr>
<tr>
<td>( \text{av_time2} )</td>
<td>41</td>
<td>27</td>
<td>34</td>
<td>370</td>
</tr>
<tr>
<td>( \text{avdiff} )</td>
<td>41</td>
<td>27</td>
<td>34</td>
<td>-263</td>
</tr>
<tr>
<td>( \text{avdiff_percent} )</td>
<td>41</td>
<td>27</td>
<td>34</td>
<td>-0.82</td>
</tr>
</tbody>
</table>

Table 9.3. Summary statistics of RLU measurements, by municipality

We can see that RLU values are generally quite high and vary substantially, between 27 and 18857. We also note that the average difference between rounds 1 and 2 varies between -14154 and 13065. Average RLU values are about 10 times larger in Ekerö than in Södertälje and Österåker, which is probably related to the fact that Ekerö uses a different measurement instrument. We can also not be sure that the relationship between Ekerö’s and the other two municipalities’ measuring equipment is linear. Further, Ekerö deviates substantially concerning the measurement periods as outlined above. For these reasons, we suspect that the measurements are not quite comparable between Ekerö and the other two municipalities and hence, we decided to henceforth drop Ekerö from the analysis.

In order to analyse the effects of the feedback letter and how this treatment interacts with other characteristics of the pre-schools we set up the following regression equation:

\[
\text{avdiff}_t = \beta_0 + \beta_1 \text{treat}_t + \beta_2 \text{middle}_t + \beta_3 \text{top}_t + \beta_4 \text{middle}_t \times \text{treat}_t + \beta_5 \text{top}_t \times \text{treat}_t + \beta_6 X_t + u_t
\]  

(1)
where $\text{avdif}_i$, the dependent variable, is the difference in RLU values for pre-school $i$ between measurements 2 and 1. Among the exogenous variables, $\beta_0$ is a constant while $\text{treat}_i$ is a dummy variable that equals one if pre-school $i$ is in the treatment group (Österåker) and zero otherwise. $\text{middle}_i$ is also a dummy variable that takes the value one if pre-school $i$ has a RLU value-percentile between 34 and 66 and zero otherwise, while the dummy variable $\text{top}_i$ is likewise defined but for the interval of 67-100. The interval 0-33 is left out. To capture the interaction effect of treatment and belonging to a particular RLU value-percentile we construct the terms $\text{middle}_i \times \text{treat}_i$ and $\text{top}_i \times \text{treat}_i$. Finally, $X_i$ is a vector of controls. We expect negative coefficients on the interaction terms as social norm pressures can be expected to be stronger on the pre-schools with high initial RLU values (that is, the dirtiest ones).

9.4. Results

The figure below illustrates average measurement results during the spring and during the autumn in the two municipalities Österåker and Södertälje. We see that the Södertälje pre-schools reduced their average ATP measurements by 42%, while the average ATP measurement remained almost the same in Österåker, in contrast to what we would have expected. However, at the municipal level we obtain results that indicate that notification of measurement results have an impact on future performance.

![Figure 9.1. Municipal average measurements, spring and autumn 2016](image-url)

The following two figures contain scatter plots of measurements in the spring and in the autumn for all pre-schools, with 95% confidence intervals, per municipality. In Södertälje the correlation between measurement results in the spring and in the fall is 0.58 and statistically significant at the 1% level. Hence, our results indicate a high degree of persistence. In Österåker the correlation coefficient is -0.25,
suggesting that pre-schools have been influenced by information regarding the results from the ATP measurements carried out in the spring, but the coefficient is not statistically significant. However, the difference between Österåker’s and Södertälje’s correlation coefficients is significant at the 1% level (see appendix 2 for the statistical test).

**Figure 9.2.** Scatter plot of measurement results in Södertälje

**Figure 9.3.** Scatter plot of measurement results in Österåker
Hence, there seems to be some systematic differences between Södertälje and Österåker, which we will now investigate in more detail in the regression analysis based on equation 1. Our results are presented in table 9.4 below.

<table>
<thead>
<tr>
<th>Dependent variable: $avdiff_i$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$treat_i$</td>
<td>256.3</td>
<td>263.1</td>
<td>688.1***</td>
<td>696.6***</td>
<td>712.8***</td>
<td>712.75***</td>
</tr>
<tr>
<td></td>
<td>(193.6)</td>
<td>(199.2)</td>
<td>(249.8)</td>
<td>(253.0)</td>
<td>(249.7)</td>
<td>(9.38)</td>
</tr>
<tr>
<td>$middle_i$</td>
<td>2.549</td>
<td>-7.746</td>
<td>-9.522</td>
<td>-9.522</td>
<td>(244.6)</td>
<td>(248.5)</td>
</tr>
<tr>
<td></td>
<td>(240.0)</td>
<td>(243.2)</td>
<td>(222.0)</td>
<td>(15.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$top_i$</td>
<td>-776.8***</td>
<td>-785.3***</td>
<td>-786.8***</td>
<td>-786.8***</td>
<td>(240.0)</td>
<td>(243.2)</td>
</tr>
<tr>
<td>$treat_i \times middle_i$</td>
<td>-399.3</td>
<td>-402.9</td>
<td>-254.3</td>
<td>-254.3***</td>
<td>(365.7)</td>
<td>(368.4)</td>
</tr>
<tr>
<td>$treat_i \times top_i$</td>
<td>-865.4**</td>
<td>-859.2**</td>
<td>-992.0**</td>
<td>-992.0***</td>
<td>(353.3)</td>
<td>(356.2)</td>
</tr>
<tr>
<td>$child_{per_loo_i}$</td>
<td>4.311</td>
<td>6.218</td>
<td>7.290</td>
<td>7.290</td>
<td>(26.17)</td>
<td>(20.43)</td>
</tr>
<tr>
<td>$Intercept$</td>
<td>-263.3***</td>
<td>-316.8</td>
<td>1.143</td>
<td>-69.86</td>
<td>-82.10</td>
<td>-82.10</td>
</tr>
<tr>
<td></td>
<td>(130.4)</td>
<td>(350.4)</td>
<td>(169.7)</td>
<td>(289.2)</td>
<td>(267.3)</td>
<td>(131.3)</td>
</tr>
</tbody>
</table>

Table 9.4. Regression results

Looking at models 1 and 2 (columns 1 and 2) we find that the estimated coefficient for the treatment variable ($treat$) to be positive, implying that Österåker pre-schools, on average, become dirtier than Södertälje pre-schools. Still, the coefficient is statistically insignificant. Further, column 1 has a negative and statistically significant intercept, indicating that pre-schools in Södertälje, on average, improve by 263 RLU units (see figure 3). When we control for the number of toilets per child in model 2 we find a positive, but statistically insignificant coefficient for this control variable, and the significance for the intercept disappears.

Model 3 includes all the interaction variables while model 4 also includes the control for the number of toilets per child. These model specifications allow for the feedback letter to have differentiated effects on RLU values depending on the percentile score of individual pre-schools. Looking at model 4 we find that pre-
schools in Österåker with RLU values in the percentile interval 0-33, on average, all else equal, increase their RLU values by 696.6 units more than pre-schools in Södertälje within the same RLU percentile interval. This is an increase of more than 100%, which indicates that pre-schools receiving favourable feedback (that they are among the top performing 33% pre-schools in the municipality) will, on average, “start slipping” considerably concerning their cleaning routines compared to pre-schools in the same percentile range that did not receive any feedback.

The effect of being in the top percentile (that is, 67-100% of the pre-schools in your municipality have a lower (=better) RLU score than you do) is, as expected, quite large in absolute terms: -785. Hence, the dirtiest pre-schools would on average tend to become cleaner. In Österåker we would hypothesize that this is because of the feedback letter but why could this be the case in Södertälje?

Assuming that there is a certain degree of stochasticity in the RLU measurements, receiving a very high measurement the first time would imply that the probability of receiving a lower measurement the next time would be quite high.

The coefficient on the interaction term \( \text{treat} \times \text{top} \) is even larger (-859) in absolute terms than the coefficient for just being in the top percentile. This means that the Österåker top percentile pre-schools will improve much more than pre-schools in Södertälje in the corresponding percentile thus indicating that the feedback letter has a strong and statistically significant effect on badly performing pre-schools.

According to model 4 the expected difference in the change of RLU value between an Österåker and a Södertälje pre-school in the top percentile would, on average, be

\[
E(\text{avdiff}|\text{treat}_i = 1, \text{top}_i = 1) - E(\text{avdiff}|\text{treat}_i = 0, \text{top}_i = 1) = \beta_1 + \beta_5 = 696.6 - 859.2 = -162.6
\]

In model 5 we exclude the Österåker observations where we could not verify that the feedback letter had been read and in model 6 we also cluster the standard errors by municipality. Our results from model 4 are shown to be robust against models 5 and 6.

Given our telephone follow-up of the Österåker pre-schools we can further investigate the mechanisms behind their behaviour. Firstly, can we observe any effects on the RLU levels from taking action? That is, did the pre-schools that took action improve relative to the pre-schools who did not take action? To answer this question, we run a simple regression where the dependent variable is the difference in RLU value between inspections 1 and 2 expressed in percent and the independent variable is a dummy variable for taking action as a result of the information letter. As we are primarily interested in the effects of the information letter we only include those pre-schools who claimed to have read the letter in the regressions. Model (1) in table 9.5 contains the results of this regression.
The coefficient for the dummy variable *Actions* is -35.55 and suggests that a pre-school that takes action will on average reduce the RLU value by 35.55 percentage points more than a pre-school that did not take actions which seems like a reasonable result. Also, the result is statistically significant at the 10% level. Thus, it seems that taking action does pay off.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>(1) Dep. var.: avdiff_percent</th>
<th>(2) Dep. var.: actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actions</td>
<td>-35.55* (20.31)</td>
<td></td>
</tr>
<tr>
<td>Percentile</td>
<td>0.008** (0.003)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>51.32*** (14.36)</td>
<td>0.130 (0.181)</td>
</tr>
<tr>
<td>F-test model</td>
<td>3.06* (0.095)</td>
<td>5.88** (0.025)</td>
</tr>
<tr>
<td>Observations</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Adj R²</td>
<td>0.089</td>
<td>0.189</td>
</tr>
</tbody>
</table>

*Table 9.5. F-values (p-values) of regressions on the effect of taking action and who takes actions. Significance levels for coefficients: *p<0.10, **p<0.05, ***p<0.01.*

Can we say anything about which pre-schools are more likely to undertake actions than others? According to the theory underpinning the so-called “Boomerang effect” pre-schools with a high percentile (a large share of the pre-schools in the municipality has a lower RLU value in inspection 1), that is, a bad result, should on average be more prone to take action due to norm pressures. One way of testing this is to regress the dummy variable for having taken actions on the percentile score of the pre-school from inspection 1. Model 2 in table two suggests that the pre-schools do indeed behave according to our prediction. The coefficient of the *Percentile* variable is 0.008 and is statistically significant at the 5% level. Hence, a one percentage point increase in the percentile score in inspection 1 increases the average probability of a pre-school to take action by 0.8 percentage points. Hence, the worse the pre-school’s RLU result is relative to other pre-schools in the municipality the higher is the probability of taking action. For example, assuming a linear effect, a pre-school which has the worst result (a percentile score of 100%) is 40 percentage points more likely to take action compared to the median pre-school (a percentile score of 50%).

At the end of the telephone interviews the heads of the pre-schools were asked to comment on their reaction to the information letter. A number of heads said that the letter was informative and an “eye-opener”. Pre-schools with lower RLU values, on the other hand, commented that they did not need to take any further action as their cleaning routines seemed to work.
9.5. Discussion

The aim of this study was to examine how feedback information following an inspection affects future behavior of inspectees. Our hypotheses was that the treatment group would be induced to improve their cleanliness more than the control group. As the average result for the treatment group actually deteriorated while it improved for the control group we could find no support for our hypothesis at the municipality level.

However, when analyzing the pre-schools at the individual level we found a positive and statistically significant correlation between RLU levels at the two inspections among pre-schools in Södertälje (that is, persistence), while we found a negative, but not statistically significant, correlation in the treatment municipality of Österåker. We could also show that the correlation coefficients of the control and treatment groups were statistically significantly different.

We then proceeded to investigate how the effects from the feedback letter might be different on pre-schools belonging to different percentiles (top, middle and low) as suggested by the “Boomerang theory”. Our analysis showed us that pre-schools with a very low first RLU value (the “low” group) tended to be less clean in the second inspection. Given a certain level of stochasticity this would be expected, but we could also show that the treatment group increased their RLU values by much more than the control group. This suggests that favorable feedback can, as indicated in Ayres et al (2012), lead to a diminished level of pro-environmental behaviour. At the other end of the distribution, that is within the “top” group, we found the opposite effect. Again, given the stochastic nature of the measurements, this would be expected. However, we could show that the treatment group reduced their RLU values much more than the control group. Hence, our results indicate that feedback does matter and that it is important for EIE agencies to think about what feedback to communicate, and, most importantly, to whom as feedback can have both positive and harmful effects.

A more detailed analysis of the pre-schools in Österåker showed that pre-schools with high percentile scores (relatively dirty pre-schools) were, not so surprisingly, more likely to take action than pre-schools with lower percentile scores. Further, we could find that pre-schools that took action reduced their RLU values more than pre-schools who did not take action.

9.6. References


9.7. Appendix

9.7.1. Appendix 1: Feedback letter to pre-schools in Österåker

Stockholm, 30 August 2016

Hello!
During a previous inspection within the project "Chemicals in pre-schools", the environmental and public health safety department of Österåker municipality conducted a measurement of the amount of organic residue on the toilet door handles, a so-called ATP-measurement. During an ATP measurement the surface in question is swabbed and the ATP meter then indicates how clean the swabbed surface is.

The following value (measured in RLU) was measured at the pre-school "Name of pre-school":

XXXX

As a comparison, the pre-schools in Österåker municipality had the following values:

Average value: 697
Median value: 372

Among the other pre-schools XX% obtained a better result (= lower amounts of organical residue on the toilet door handles).

Within the health care sector, the following scale is used for freshly cleaned surfaces:

0 - 50 = approved,
51 - 100 = approved with a remark,
above 100 = not approved.

A high ATP value can, for example, be due to poor cleaning or to the toilets being used by many children. As freshly cleaned surfaces can quickly become dirty it is important to have regular cleaning routines. Maintaining a high level of hygiene on e.g. toilet door handles is important to prevent spread of infectious diseases.

The ATP-measurements are a part of a wider study conducted by Stockholm University. We would like to stress that the results will only be presented at an aggregate level where individual pre-schools will not be identifiable. In order not to jeopardize the study we cannot provide you with more information at this time.

Sincerely,
Mathias Herzing (Mathias.Herzing@ne.su.se, 08-163043)
Adam Jacobsson (aja@ne.su.se, 08-161756)
Department of Economics
Stockholm University
9.7.2. **Appendix 2: Statistical test of correlation coefficients.**

Testing whether we have statistically significant differences between Södertälje’s and Österåker’s correlation coefficients we calculate the Fisher z-value and associated P-values:

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Z-value</th>
<th>P-value (one tailed)</th>
<th>P-value (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Södertälje vs Österåker</td>
<td>3.78***</td>
<td>0.0001</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

**Table 9.6.** Testing the significance of the differences between correlation coefficients. ***=significance at the 1% level, **=significance at the 5% level.
Inspections and enforcement as instruments for enhancing environmental behavior

Final report

The research program "Inspections and enforcement as instruments for enhancing environmental behavior" has focused on evaluating the effectiveness of inspections and enforcement.

The results presented in this report build on the previous research program "Efficient environmental inspections and enforcement" (Swedish EPA, Report 6558). In total, six studies have been carried out. In three studies, the communicative method "Motivational Interviewing" has been tested in specific inspection contexts with the purpose of studying and improving the interaction between inspector and the inspectee. In addition, a theoretical, an empirical and an experimental study on the effectiveness of inspection and enforcement activities have been conducted.

The aim has been to provide a wide range of analytical approaches and practical applications that can be used for both guidance and development of operational inspections and enforcement methods.