

# Public Organizations under Pressure:

## Impact on Dysfunctional Patterns in Service Delivery



Karoline Larsen Kolstad

**Public Organizations under Pressure:  
Impact on Dysfunctional Patterns in Service Delivery**

PhD Dissertation

Politica

©Forlaget Politica and the author 2023

ISBN: 978-87-7335-321-9

Cover: Svend Siune

Print: Fællestrykkeriet, Aarhus University

Layout: Annette Bruun Andersen

Submitted September 15, 2023

The public defense takes place November 17, 2023

Published November 2023

Forlaget Politica

c/o Department of Political Science

Aarhus BSS, Aarhus University

Bartholins Allé 7

DK-8000 Aarhus C

Denmark

# Table of Contents

<b>Acknowledgements</b>	<b>7</b>
<b>Preface</b>	<b>9</b>
<b>1 Introduction</b>	<b>11</b>
1.1 A Foundational Street-level Narrative . . . . .	11
1.2 Aim of the Dissertation . . . . .	12
1.3 Revisiting the Narrative . . . . .	12
1.4 Pointing towards Organizational Solutions . . . . .	13
<b>2 Theoretical Framework</b>	<b>15</b>
2.1 Dysfunctions in Service Delivery . . . . .	15
2.2 Pressures on Public Organizations . . . . .	17
2.3 Core Assumptions . . . . .	18
<b>3 Research Designs and Data</b>	<b>21</b>
3.1 Confounding and Contextual Complexity . . . . .	21
3.2 Quasi-experimental Designs . . . . .	21
3.3 Unemployment Services and Primary Schools . . . . .	23
3.4 Administrative Records and Interviews . . . . .	25
3.5 Overview of the Four Studies . . . . .	26
<b>4 Results</b>	<b>29</b>
4.1 Does High Workload Lead to Discrimination? . . . . .	29
4.2 Does High Workload Lead to Poor Service Quality? . . . . .	31
4.3 Does Performance Pressure Lead to Disciplining? . . . . .	32
4.4 Does Performance Pressure Lead to Gaming and Effort Substitution? . . . . .	34
<b>5 Discussion and Conclusion</b>	<b>37</b>
5.1 Expanding the Street-Level Narrative . . . . .	37
5.2 Practical Implications . . . . .	39
5.3 Future Avenues for Research . . . . .	39
<b>Bibliography</b>	<b>41</b>
<b>English Summary</b>	<b>47</b>
<b>Dansk Resumé</b>	<b>49</b>



## Acknowledgements

Writing a PhD dissertation can be a lonely endeavor, but I have been lucky to be surrounded and supported by many good people on the way. First, I want to thank my two supervisors, Morten Jakobsen and Simon Calmar Andersen. You have patiently read and commented on ideas and drafts over and over again. You have provided me with valuable insights about doing research and communicating my findings. From the start, you have challenged me by questioning my research ideas and by identifying the weakest link in my arguments. Even though this has been tough at times, I truly appreciate your genuine efforts and your push for ambitious research projects. I have learned a lot from your supervision.

I also want to thank the four persons with whom I have shared an office throughout the years. Anne Kirstine Rønn, Amalie Trangbæk, Rikke Overvad Kolding and Anders Leth Nielsen, each of you has been an invaluable support at different stages of my PhD. AK, thank you for helping me take my first insecure steps in the academic world and for making me feel part of the PhD group. Amalie, thank you for always being there. Your friendship and our endless talks on everything in life got me through the pandemic. Rikke, you had to deal with a very pregnant version of me. Thank you for making me laugh and for bearing with my immense appetite and need for naps in the office. Anders, thank you for listening to all my nonsense during the last stretch of my PhD. I have enjoyed our many talks on public administration research, work-life-balance and apple-cider production. In addition, I wish to give a special thanks to Sarah Yde Junge. We may not have shared an office together, but I have learnt so much from you and our talks about natural experiments, registry data and coding strategies. Co-authoring a piece together has been a huge privilege.

Moving on to the department, I want to acknowledge the public administration section for the academic discussions and the sincere interest you take in your colleagues' work. I have learned a lot from listening to your feedback at our weekly meetings throughout the years. Thank you for demonstrating that it is possible to deliver to the point comments while creating a safe and including environment. The PhD group also deserves special thanks for ensuring that there was more than work to PhD life at the department. I am especially grateful to have shared my PhD time with Anna Christine Hjuler Dorph, Aske Halling, Clara Siboni Lund, Emily Rose Tangsgaard, Jakob Holm, Jørgen Skovhus Haunstrup, Kirstine Karmsten, Louise Halberg Nielsen, Mads Pieter van Luttervelt, Marie Fly, Marie Kaldahl Nielsen, Mathias Østergaard-Nielsen, Mathilde Albertsen Winsløw, Mathilde Cechinni, Matilde Jeppesen, Nadine Malmberg Hollmann, Nanna Vestergaard Ahrensberg, Nikolaj K. Andersen, Thomas Artmann Kristensen, Valentin Dauer and William Burchardt Egendal. Thank you very much for all the coffee runs and distractions. They have been invaluable.

I would also like to thank Helene Helboe and Christoffer Green-Pedersen for your sincere interest in the well-being of the PhD students. We can always count on you. In addition, the whole TAP group deserves huge recognition for their tireless help in accounting, language revisions and in keeping the coffee machines running. In specific, I wish to express sincere gratitude

to Annette Bruun Andersen, Ruth Ramm, Olivia Elsebeth Belling-Nami and Njall Beuschel. You have helped me out with all sorts of problems in these past years.

Despite COVID-19, I was fortunate enough to be able to visit Utrecht University School of Governance in the fall of 2021. A special thank you to Noortje de Boer for making it possible and for bringing new perspectives to the discussion of my work. As one of the few young female scholars in political science, you have been a true role model.

In addition, I wish to thank Joe Soss, Sharon Ghilad and Poul Aes Nielsen for being part of the committee and for reading and commenting on my dissertation. Your work has been a big source of inspiration to me, and your genuine engagement with my dissertation means a lot to me. I look forward to discussing it all with you!

Last, but definitely not least, I want to thank my family. Kasper, we started out on this journey together and have shared the many ups and downs of conducting a PhD in parallel. Despite the fact that nano gold particles and street-level bureaucrats do not share many similarities, you have been an invaluable support for me. I appreciate all our talks about our research - be it discussions on probability theory or on how to make a good tweet. Even more so, I appreciate that we haven't been completely caught up in our research but prioritized to make room for much more in life. Inge, thank you for providing me with the strongest incentive to finish work early. The outlook of spending time with you has helped me overcome many challenges in finishing this PhD.

## Preface

My dissertation “Public Organizations under Pressure: Impact on Dysfunctional Patterns in Service Delivery?” consists of this summary report and the four papers listed below. The summary report provides a theoretical framework that cuts across individual papers and discusses important common themes in my work. Details on both methods and theory can be found in the papers. Throughout the summary report, I refer to papers by their letter and short title.

- Paper A (Discrimination Paper)  
Overburdened Bureaucrats: Securing Equal Access to Public Services under COVID-19, *Invited to Revise and Resubmit in Journal of Politics*.
- Paper B (Concept Paper)  
Workload and Street-level Bureaucracy: Clarifying a Key Phenomenon, *Working paper*, co-authored with Sarah Yde Junge.
- Paper C (Disciplining Paper)  
Contextualizing Pressure: Performance Management, Welfare Sanctions and Managerial Buffering of the Street Level, *Under review in Journal of Public Administration Research and Theory*.
- Paper D (Dynamics Paper)  
Changing Performance Management: How Public Service Providers Adapt, *Working paper*.



# Chapter 1

## Introduction

### 1.1 A Foundational Street-level Narrative

Stories about failures in public service delivery frequently appear in the media. Examples include misjudgements in service allocation decisions, poor service quality where citizens are forgotten, and instances of discrimination where some clients are prioritized at the cost of others. These examples of service failures are often explained by the challenging working conditions under which service delivery takes place. Public organizations and public employees are pressured from many sides by factors such as limited resources, high workloads, new regulations, documentation requirements, demanding citizens and politicians with increasing and changing priorities. In this framing, dysfunctional reactions occur not because bureaucrats are mean or incompetent, but because they are pressured and constrained by challenging working conditions.

This story is also present in the public administration literature. In fact, it serves as the foundation of the street-level bureaucracy literature. Lipsky (1980) first described how street-level bureaucrats—public employees in charge of delivering public services to citizens—operate under challenging working conditions. According to Lipsky (1980), street-level bureaucrats work with inadequate resources in circumstances where demand will always increase to meet the supply of services, which means that "...they can never be free from the implications of significant constraints" (81). On this basis, Lipsky (1980) argues that in order to understand street-level bureaucracy, "...one must study the routines and subjective responses street-level bureaucrats develop in order to cope with the difficulties and ambiguities of their job" (82). He thereby points to the important role that conditions of work—including resources, goals and performance measures—play in street-level decision making. In other words, dysfunctional patterns in service delivery are believed to arise because bureaucrats are pressured and constrained by challenging working conditions and not because of bad intent.

The contemporary street-level bureaucracy literature builds on this foundational narrative. Much research documents the existence of macro-level pressures, such as high demand for public services, scarce financial resources and the demands of performance goals (i.e. Lavee 2020; Noordegraaf and Steijn 2013; Brodtkin 2011; Maynard-Moody and Musheno 2020; Zacka 2018; Hupe and Buffat 2014; Watkins-Hayes 2011). Other research identifies patterns in street-level bureaucrats' coping behavior towards citizens (i.e. Tummers et al. 2015; Guul, Pedersen, and Petersen 2020; Jilke and Tummers 2018; Olsen, Kyhse Andersen, and Moynihan 2020; Pedersen and Stritch 2018; Assouline, Gilad, and Ben-Nun Bloom 2021). Far fewer studies combine the two levels of analysis to investigate how pressures at a macro level influence individual encounters between bureaucrat and citizen at a micro level (Rice 2019, 71, Hupe and Buffat 2014, 549, Moynihan 2018). As a result, we have scant knowledge about how pressures on street-level

bureaucracies influence service delivery and whether pressures do in fact cause dysfunctional patterns in the delivery of services. In other words, it is uncertain whether the foundational narrative of the field of street-level bureaucracy holds when put under empirical scrutiny.

## 1.2 Aim of the Dissertation

Street-level bureaucrats deliver services to millions of citizens on a daily basis while navigating pressures from challenging working conditions. Instances of dysfunctional service delivery occur and impose heavy costs on citizens. As a means to alleviate service failures and improve public service delivery, it is important to understand their underlying causes. In this dissertation, I take a step towards doing so by investigating whether the existing “pressure narrative” holds empirically. More specifically, I set out to investigate whether and how different pressures from the political-administrative environment on public service organizations lead to dysfunctional patterns in the delivery of services to citizens.

The dissertation consists of four individual papers that investigate the overall research question from different angles. Two papers look into pressure caused by high workloads and two papers look into pressure that arises from performance management systems. Common to both types of pressure is that they create constraints on street-level organizations by shifting the balance between the goals to be achieved and the resources available to do so. Each of the four papers looks into distinct types of service dysfunctions that undermine either the achievement of efficient service delivery or the achievement of fair service delivery, which constitute two of the core values of public service delivery (Boyne 2003, 14). All four papers rely on quasi-experimental designs that leverage exogenous variation in pressure on public unemployment services and public primary schools. Together with comprehensive micro-level administrative records on service delivery, this approach provides the means to overcome selection while maintaining contextual complexity (Blom-Hansen, Morton, and Serritzlew 2015; Dunning 2012). In this summary report, I focus on drawing conclusions across the four papers instead of going into the specific details of the individual papers.

## 1.3 Revisiting the Narrative

The key claim of the dissertation is that public service organizations and public employees are exposed to substantial pressures stemming from the political-administrative environment, but that they are in fact capable of handling them without turning to dysfunctional behaviors in service delivery. Taken together, the four individual papers demonstrate that pressures from high workloads and performance management systems do not automatically lead to dysfunctional service delivery. The studies provide examples of public employees who work under high workloads and pressure to perform without resorting to discrimination, disciplining, gaming or goal displacement. In fact, the examples demonstrate that public service organizations and public employees are capable of handling substantive pressures, which speaks to the robustness of the public service sector. As a consequence, these examples go against the existing narrative within the literature, where both workload and performance management are believed to be the source of various evils in public service delivery.

The fact that we do not see the expected dysfunctional reactions does not mean that pressure never matters for service delivery. The different papers provide different examples of this. In Paper A (Discrimination Paper), I document that workload pressure does influence service delivery to some extent, but not in the expected dysfunctional form of explicit discrimination. I show this in relation to a substantial increase in workload due to a 20% rise in unemployment caused by the COVID-19 pandemic in Denmark in 2020. I demonstrate that caseworkers did not react to this substantial workload shock by discriminating between ethnic majority and ethnic minority citizens in the delivery of core services. Instead, caseworkers handled the workload by lowering the general service level, which is a functional way of handling an abnormal workload shock. Moreover, I document that ethnic minority citizens who were laid off during the lockdown were stuck in unemployment for a substantially longer period than ethnic Danes. This indicates that unemployment services were not able to overcome existing imbalances between the two groups.

In Paper B (Concept Paper), my coauthor and I also focus on pressure from workload. We document that workload is a multifaceted concept and that dysfunctional reactions only occur when bureaucrats are pressured from a very comprehensive form of high workload. We focus on a schooling context and demonstrate that *caseload*, i.e. number of students in class, does not influence service quality, measured through students' test scores. In contrast, we show that the much more comprehensive concept of *taskload*, which includes the intensity of tasks that each case demands, does have a small negative impact on service quality. This indicates that a very extensive form of high workload, where a high number of clients pose heavy demands, may lead to some dysfunctions.

In Paper C (Disciplining Paper), I move on to investigate how Danish unemployment agencies respond to the pressure associated with a comprehensive performance management system imposed by the Ministry of Employment. I document that the pressure does have a substantial impact on the organization of service delivery, but that the pressure does not lead to increased sanctioning of unemployed clients.

In Paper D (Dynamics Paper), I focus on the same performance management system. I document that local service agencies do not respond to outcome-based accountability by outright gaming, but that the political principal can push service providers towards effort substitution by making it very costly and risky not to do so. Agencies initially pursued "reversed effort substitution," where they prioritized service quality, which is difficult to measure, at the cost of service quantity, which is easy to measure. It is only when the Ministerial principal complicated the system by introducing output-based accountability that service providers were forced to resort to effort substitution, since other strategies became too risky.

Together, these four studies indicate that public service organizations are capable of handling substantive pressures without turning to dysfunctional behaviors.

## 1.4 Pointing towards Organizational Solutions

In general, the dissertation concludes that pressures from high workload and performance management do not lead to the expected dysfunctions. This indicates that decreasing the workload of public employees and revoking the widespread use of performance management systems might not serve as an easy fix in those cases where service failures do occur. In any event, such pres-

ures are often associated with political interest and are consequently difficult to revoke. To some extent, this argument fits well with Lipsky (1980), who argues that public service organizations can never be free from some sorts of pressures and constraints. If this is so, it is not possible to remove the pressures completely, and the solution to dysfunctional service delivery should be found elsewhere. In line with this, the dissertation indicates that we should look for solutions in other parts of the long and complicated chain that exists between the pressures at the macro level and service delivery at the micro level.

Overall, the studies of the dissertation indicate that we should focus more on how pressures play out within an organizational context at the meso level (Hupe and Buffat 2014; Rice 2019) than on the mere existence of the pressure in itself at a macro level. This line of reasoning follows Soss, Fording, and Schram (2011a), who document how pressure tied to a performance management system does lead to dysfunctional service delivery in poorly functioning organizations. In line with this argument, the studies of this dissertation indicate that the opposite scenario also exists. Specifically, the studies indicate that well-functioning organizations with professional bureaucrats, administrative support systems and management that takes its responsibility seriously are equipped to handle the pressures. To some extent, these are characteristics that we can expect from public service organizations. In fact, we construct and design organizations with characteristics such as task division and hierarchical authority so that management is equipped to protect the core task of the organization from external influences (O'Toole and Meier 2010; O'Toole and Meier 2014; O'Toole, Meier, and Nicholson-Crotty 2005; Stazyk and Goerdel 2011; Olsen 2006). The different papers provide examples of this.

First, Paper A (Discrimination Paper) provides an answer to why a substantial workload increase did not act as a catalyst of discrimination in service delivery. Interviews with caseworkers suggest that they felt the pressure from the abnormal workload shock, but that they were equipped to handle it without resorting to discrimination. The Danish unemployment agencies have strong administrative support systems that give caseworkers easy access to comprehensive information about clients. Furthermore, caseworkers have strong professional norms, which means that the objective workload increase did not translate into subjective experiences of greater cognitive load. Together, these factors meant that the workload increase did not force caseworkers to rely on stereotypes about clients and that the mechanisms described by the model of *statistical discrimination* and *unconscious bias* did not play out. This indicates that professional norms and administrative systems that support decision making at the street level are important for how work pressure plays out.

Moving on to the pressure associated with performance management, Paper C (Disciplining Paper) provides some specific answers to why performance pressure did not lead to disciplining as expected. The qualitative part of the study reveals that the pressure from the performance management system never reached the street level, where caseworkers operate and make decisions about welfare sanctions. In fact, the interviews indicate that management intentionally worked to buffer the street level from any such pressure. They do so by translating the general performance pressure into behavioral directives that make sense at the street level. This indicates that management plays a key role in moderating performance pressure, which influences how caseworkers receive and pass along performance pressure toward clients.

Based on these concrete insights from the Danish case, I argue that the solution to dysfunctional service delivery might lie in constructing well-functioning organizational structures that give public service organizations the means to handle the pressures that they will inevitably face.

## Chapter 2

# Theoretical Framework

In this chapter, I present the two main concepts of the dissertation: dysfunctional patterns in service delivery and pressures on public organizations. I then move on to present some of the field's core assumptions about how pressure on public organizations leads to dysfunctional patterns in service delivery. This serves as the theoretical foundation of the dissertation.

### 2.1 Dysfunctions in Service Delivery

Public services are supposed to be delivered in an efficient and fair way (Boyne 2003, 14). Following these two core values of public service delivery, dysfunctions occur if services are delivered in an inefficient or unfair way.

In technical terms, efficiency is the ratio of service inputs to outputs. Inputs are the resources that are put into the organization and outputs are the services actually produced by an organization, such as teaching provided in schools (16). In broader terms, efficiency can be thought of as the extent to which public services are able to produce an intended result with the least possible waste of time, effort and resources. The goal in this sense is to get maximum value for the resources that are invested in the service organizations. Following from this, service delivery is dysfunctional if providers deliver unnecessary services or services that do not help reach the main goal of the organization, such as student learning or reemployment.

Fairness, on the other hand, relates to who gets access to the services delivered. Following this perspective, public services should be delivered in a fair and equitable way. Bureaucrats should grant equal access to all citizens who are entitled to the service (23). It is unfair if individuals or citizen groups are denied access to public services that others with the same entitlement do get access to. The concept of social equity is more extensive and involves the correction of existing imbalances between citizens, i.e., historical, political, social and economic influences that structurally influence different citizens' prospects for opportunity and outcomes (Gooden 2015). Following this logic, it is not enough that bureaucrats deliver the same services to all. Bureaucrats should strive to overcome existing imbalances and prioritize time and resources for those who need them most. Across the four papers, I focus on different forms of dysfunctional patterns that are related to either an efficiency loss or a fairness loss. Table 2.1 includes an overview.

In Paper A (Discrimination), I investigate discrimination in service delivery, which is a dysfunctional pattern associated with a fairness loss. Discrimination refers to allocative exclusion, where bureaucrats allocate public resources differently across groups based on underlying group traits. Several studies show that bureaucrats exclude ethnic minorities from core public services, such as unemployment benefits, medical disability benefits and schooling (Schram et al. 2009; Pedersen and Stritch 2018; Keiser, Mueser, and Choi 2004; Assouline, Gilad, and

Table 2.1: Dysfunctional Patterns in Service Delivery

Inefficiency	Unfairness
Poor service quality	Discrimination in service delivery
Effort substitution	Unequal service outcomes
Gaming	Disciplining
Disciplining	

Ben-Nun Bloom 2021; Olsen, Kyhse Andersen, and Moynihan 2020). Bureaucrats can also discriminate more indirectly by imposing administrative burdens differently across groups. In these cases, all citizens do gain access to services, but some do so under less favorable circumstances. For example, a series of audit studies document that bureaucrats discriminate in the provision of information about government services to citizens (see Olsen, Kyhse Andersen, and Moynihan 2020 for a detailed review). In addition to looking into discrimination in service delivery, I also look into the more encompassing concept of social equity and investigate inequality in service outcomes (Gooden 2015).

In Paper B (Concept Paper), I examine poor service quality. Poor service quality undermines the achievement of the organizational end goal and is thus associated with an efficiency loss. Service quality relates to the standard of service provided (Boyne 2003, 17) and can take many forms. Examples include automatization of services, where bureaucrats rely on rules of thumb instead of treating clients individually (Jewell 2007), and errors, where bureaucrats overlook important aspects of service delivery or fail to guide clients (Van Berkel and Knies 2016).

In Paper C (Disciplining Paper), I look into the disciplining of clients, which is associated with both an efficiency loss and a fairness loss. Public service programs use a range of incentives, monitoring mechanisms, and restrictive regulations and rules aimed at modifying client behaviors. A widely studied tool of discipline is the welfare sanction, which is a financial penalty that reduces or terminates welfare benefits in response to client noncompliance (Fording, Soss, and Schram 2011; Pedersen, Stritch, and Thuesen 2018; Schram et al. 2009; Schram 2008). Disciplining, such as the use of welfare sanctions, can in some instances help bureaucrats obtain the organizational end goal by making clients adhere to program rules. However, research documents that many factors beyond individual noncompliance play into sanctioning decisions and may in some cases be detrimental to reaching the end goal (Soss, Fording, and Schram 2011a). This creates an efficiency loss, because service organizations spend resources on implementing sanctions that do not help attain the end goal. Sanctions are furthermore associated with a fairness loss, because some clients are excluded from benefits to which they are in fact entitled.

In Paper D (Dynamics Paper), I look into gaming and effort substitution, which are widely studied in relation to performance management. Gaming and effort substitution create efficiency losses because both strategies consume resources but produce no or limited value added. Kelman and Friedman (2009, 8) define gaming as “behavior that consumes real resources to manipulate measured performance but produces no genuine value added improvements, even on a measured dimension.” The limiting case of gaming is outright data falsification, but other less clear-cut examples exist: teachers who exclude low-ability students from the official test-

taking pool on the day of the exam (Bohte and Meier 2000; Jacob 2005), or job training providers who manipulate the timing of client exits from training programs (Courty and Marschke 2004).

Effort substitution arises when an organization has more than one dimension to its tasks and when focus is put on the measured dimension at the expense of unmeasured dimensions. In contrast to gaming, effort substitution produces performance improvement along the measured dimension. Effort substitution is often expressed as a focus on quantities of outputs, which are easy to measure, at the cost of quality, which is difficult to measure and thus left out of the measurement system (Holmstrom and Milgrom 1991). Examples include teaching to the test, where teachers focus on test-specific skills at the cost of overall learning (Jacob 2005) and cream-skimming and parking, where employees focus on easy-to-serve clients at the cost of hard-to-serve clients (Koning and Heinrich 2013). The latter example is also associated with a fairness loss, since some clients are prioritized over others.

## 2.2 Pressures on Public Organizations

The term *pressure* implies some sort of demand or constraint on public organizations and public employees' freedom to act (Hupe and Krogt 2013, 60). These constraints may have different sources and take different forms. Consequently, scholars think about and conceptualize pressure differently. For example, Zacka (2018, 123) conceptualizes pressure as the discrepancy between the ideal of public service and the grim realities of everyday work and highlights how the reforms brought about by New Public Management pressure street-level bureaucrats to do even more with less. Hupe and Krogt (2013, 61) distinguish between rule pressure, vocational pressure and societal pressure. Noordegraaf and Steijn (2013) describe how pressures may come from different directions, including above, beneath and sideways. And Newman (2013) distinguishes between existing pressures and new kinds of pressures.

Across different conceptualizations, however, scholars do agree that pressure deals with the relation between a “demand” and a “supply” side, which refers to the difference between what is being asked of and the resources offered to public officials working in contact with citizens. Pressure thus occurs when what is required of street-level bureaucrats exceeds what is provided to them for the fulfilment of their tasks. These demands can take the form of legal requirements, policy goals, organizational objectives and performance targets, but also of norms from peers and expectations from societal actors (Hupe and Buffat 2014; Watkins-Hayes 2011).

In this dissertation, I focus on macro-level pressures that arise from factors in the political-administrative environment. These pressures arise as the result of factors beyond the individual street-level bureaucrat and the individual client. Consequently, I exclude pressures that arise due to individual-level attributes and phenomena such as people's characteristics, beliefs or interactions with others. Instead, I focus on the political-administrative environment that surrounds public service agencies. In particular, I focus on pressure that arises from high workload and pressure that arises from performance management systems.

First, high workload is considered to be inherent to street-level bureaucrats' jobs and is regarded as one of the core pressures on everyday public service delivery (Lipsky 1980; Lavee 2020; McCrea 2022). Workload is often understood as caseload, i.e. the ratio of workers to clients. For example, Lipsky (1980, 29) highlights lower-court judges who are inundated with cases, teachers in overcrowded classrooms, and social services lawyers with immense backlogs.

Workload puts pressure on street-level bureaucrats, since when the number of cases increases, bureaucrats need to process more information and complete more casework. Many factors may influence the workload of street-level bureaucracies, and sometimes these factors are associated with the individual bureaucrat or service organization. Often, however, changes to workload are caused by factors in the political-administrative environment, such as political priorities in the allocation of resources to public service organizations or environmental shocks that influence citizens' demand for services.

Second, performance management systems, which are part of the New Public Management reform movement (Hood 1991; Brodtkin 2011), constitute a more recent source of pressure on street-level bureaucracies. Political principals use performance measurement systems as a means to steer and control public service providers. They do so by defining objectives, setting measurable targets, collecting data and applying incentives (Dooren, Bouckaert, and Halligan 2015; Pollitt 2018). In contrast to workload, performance management is created intentionally by the political principal as a means to gain accountability and efficiency in service delivery. Performance management systems put street-level bureaucracies under pressure, since the systems set performance targets that have to be met and failure to deliver on results is associated with more or less explicit sanctions.

At first sight, high workload and performance management systems appear to deal with two distinct pressures. I argue, however, that both pressures deal with the same underlying phenomenon—that is, the relation between the goals to be achieved and the resources available to do so. Both of the pressures associated with workload and performance management systems distort the balance between the tasks to be done and the resources available to achieve them.

## 2.3 Core Assumptions

It is generally argued that dysfunctional patterns in service delivery arise because street-level bureaucrats are pressured into developing coping mechanisms. The term “coping” captures the prevailing concept for denoting street-level bureaucrats' behaviors and patterns of practice (Hupe and Hill 2019; Tummers and Rocco 2015). The term implies that street-level bureaucrats develop patterns of practice in order to overcome or “cope” with challenging working conditions. As a result, they are also sometimes referred to as survival mechanisms (Lipsky 1980, 187). By characterizing street-level bureaucratic behavior as coping mechanisms, it is presupposed in the conceptualization itself that these patterns of practice must be the result of challenging working conditions. This foundational view is dominant in the literature and has led to several widespread core assumptions about the causes of dysfunctional patterns of practices. In the following four subsections, I will describe each of the assumptions, which together constitute the theoretical basis of the dissertation.

### 2.3.1 Pressured to Discriminate

The leading explanation of discrimination in public service delivery argues that bureaucrats discriminate because they face limited resources and high workloads in public organizations (Keiser and Soss 1998; Schram et al. 2009; Andersen and Guul 2019; Guul, Pedersen, and Petersen 2020; Tummers and Rocco 2015; Jilke, Van Dooren, and Rys 2018; Olson 2016; Assouline, Gilad, and Ben-Nun Bloom 2021; Pfaff et al. 2021). This has been the dominant the-

oretical explanation since Lipsky (1980) first argued that “client differentiation may take place because, confronted with heavy workloads and apparently impossible tasks, bureaucrats [...] attempt to succeed with some clients when they cannot succeed with all” (p. 107) and “differentiating among clients [...] assists bureaucrats in managing their work loads ...” (p. 106).

Models of statistical discrimination and unconscious bias also give us reasons to expect that bureaucrats handle high workloads by discriminating against disadvantaged minority citizens. First, models of statistical discrimination argue that people rely on statistical beliefs because the cost of obtaining more information outweighs the benefit of a more precise evaluation of a client (Phelps 1972, 659). High workloads give bureaucrats less time to gain access to relevant information about each individual and force them to rely on stereotypes about group traits in order to prioritize whom to serve first. Second, models of unconscious bias argue that the implicit reliance on stereotypes depends on the individual’s psychological state and ability to process information. In particular, high cognitive load undermines the capacity for self-regulation and increases reliance on unconscious bias, which may translate into discriminatory behavior (Andersen and Guul 2019; Assouline, Gilad, and Ben-Nun Bloom 2021).

### 2.3.2 Pressured to Deliver Poor Service Quality

High workload also serves as one of the leading explanations of poor service quality. It is generally assumed that high workloads force bureaucrats to focus on quantity at the expense of quality (Van Berkel and Knies 2016; Van Loon and Jakobsen 2018; Bloom, Hill, and Riccio 2003; Hill 2006). More specifically, this can happen through automatization in bureaucrats’ dealings with clients. For example, Jewell (2007) argues that high caseloads threaten to reduce professionals to technicians who rely on informal shortcuts and make them unable to provide responsive services. Similarly, Lipsky (1980, 99) argues that high workloads “... will tend to impose a logic of increasing the quantity of services at the expense of the degree of attention workers give to the individual client.” In addition, Van Berkel and Knies (2016) highlight how high caseloads may increase the probability of workers making errors, such as handing down unjust sanctions, because they spend less time with clients and have less time for decision making.

### 2.3.3 Pressured into Gaming and Effort Substituting

Concerns about dysfunctional reactions to performance management systems have existed as long as the systems themselves. In fact, more attention has been given to dysfunctional reactions than to functional ones—going back to the earliest discussions of the topic in organization theory (Kelman and Friedman 2009). This waste literature generally argues that performance management makes bureaucrats resort to gaming and effort substitution (Courty and Marschke 2004; Bohte and Meier 2000; Jacob 2005; Bevan and Hood 2006; Koning and Heinrich 2013; Wang 2022). It is assumed that gaming arises because service providers wish to maximize benefits from the reward system at the lowest possible cost. Since it takes greater effort to increase value added compared to measured performance, service providers will use their day-to-day familiarity with the technology to increase measured performance (Heinrich and Marschke 2010). For example, it is much easier to exempt some students from the test-taking pool than to improve student learning. In relation to effort substitution, it is argued that the focusing function of performance measures directs the allocation of service providers’ attention toward measured

performance at the cost of service quality. Service providers have no interests in focusing on unmeasured dimensions that are not rewarded (Holmstrom and Milgrom 1991).

### 2.3.4 Pressured to Discipline

Another dysfunctional reaction that has been tied to performance management is the excessive disciplining of clients on welfare benefits. Several scholars theorize and document that the rise of New Public Management has created an organizational environment that is supportive of welfare sanctions (Brodkin 2011; Soss, Fording, and Schram 2011a; Soss, Fording, and Schram 2011b; Caswell and Høybye-Mortensen 2015; Berkel 2020). In particular, performance management and pressure to perform is highlighted as a key driver. In line with the creaming perspective, caseworkers may resort to increased use of welfare sanctions and termination of clients as a means to avoid the risk of negative performance reviews. They use sanctions as a strategy to rid themselves of low-performing clients and thereby restrict their caseloads to “cream” clients who generate positive numbers. This can be characterized as a gaming strategy, where caseworkers wrongfully terminate clients in order to improve their performance metrics.

In contrast to this explanation, Soss, Fording, and Schram (2011a) argue that caseworkers do not sanction clients to advance their narrow-minded pursuit of good performance numbers. Instead, the authors point to a more subtle explanation of how case managers are constrained and disciplined by the environment in which they are embedded. They highlight a conjunction of factors in the Florida Welfare Transition program, namely the distinctive way that this program organizes case management, the specific performance pressures experienced by caseworkers, the limited number of tools available to caseworkers, and caseworkers’ beliefs and frustrations regarding client noncompliance. The increase in use of sanctions is thus the result of a subtle combination of several contextual factors (Soss, Fording, and Schram 2011b, 225).

## **Chapter 3**

# **Research Designs and Data**

In this chapter, I will discuss the challenges associated with answering the research question and how I have chosen to address these in the dissertation. After presenting two specific challenges, I move on to describe the identification strategies, the choice of cases and data sources of the four papers.

### **3.1 Confounding and Contextual Complexity**

The aim of this dissertation is to investigate whether and how different macro-level pressures on public service organizations influence dysfunctional patterns in service delivery at a micro level. Unfortunately, this may be easier said than done. The identification of a real-world causal effect of pressures on service delivery rests on overcoming two key challenges. The first major challenge that needs to be addressed is the problem of confounders. Public service organizations are exposed to different types of pressures and pressure levels for specific reasons. These reasons are very likely linked to the specific way that service organizations go about delivering services to citizens. For example, service providers may experience high levels of workload and high levels of performance pressure because they practice dysfunctional patterns in service delivery and deliver services of poor quality to clients. Background characteristics such as the demographics of the client base and overall managerial strategy may also serve as confounding factors, since they are likely to influence both the level of pressure exposure and the way that services are delivered to citizens.

The second key challenge that needs to be addressed is how to maintain contextual complexity. An important aspect of this dissertation rests on the connection between a macro-level and a micro-level phenomenon, which are often studied separately (Moynihan 2018). In contrast to research questions that focus on the relation between phenomena at the same level of research, this dissertation focuses on macro-level pressures and micro-level service delivery, which are distant concepts that are connected through a long and complicated chain of events. As a result, an important focus of the dissertation rests on identifying whether these macro pressures actually travel all the way through to the micro level. It is therefore important to maintain this complexity in the research design in order to fully understand how macro-level pressures influence micro-level service delivery.

### **3.2 Quasi-experimental Designs**

Both of these challenges need to be addressed in order to credibly answer the dissertation's research question. The confounding challenge calls for randomized controlled trials such as field, survey and laboratory experiments (Gerber and Green 2012). Random assignment of units to

various pressure exposure levels would balance potential confounders and make it possible to identify the causal effect of pressure on service delivery. This approach, however, does not provide a solution to the context challenge. It is difficult to manipulate pressure on public organizations in a credible and authentic way in surveys and laboratory settings, since the pressure has no real consequences in these settings (See Hainmueller and Hangartner 2013; Assouline, Gilad, and Ben-Nun Bloom 2021, Jilke and Van Ryzin 2017, 132). Moreover, manipulation of the pressure component in a lab or survey experimental setting shortens the long and complex chain of events that exists between the macro and micro levels in the real world, with the consequence of losing contextual complexity.

Instead, I choose to rely on quasi-experimental designs (Blom-Hansen, Morton, and Serritzlew 2015; Dunning 2012), in which I exploit exogenous variation in workload and performance pressure that occurs in real-world service contexts. This approach enables me to handle both of the two key challenges, since it provides as-if-random variation in actual pressures in real-world service contexts. The causal inference provided by quasi-experiments is associated with more uncertainty than randomized controlled trials, because the random assignment process is not controlled explicitly by the researcher (Dunning 2012). This uncertainty, however, comes with much greater contextual complexity, since the variation takes place in real-world service encounters, where bureaucrats interact with citizens who are in need of their help. Furthermore, I put a great deal of effort into validating whether the as-if-random assignments hold in each of the individual papers. This strategy provides a way of securing high ecological validity (Kihlstrom 2021), because the experiments take place in real-world public service settings.

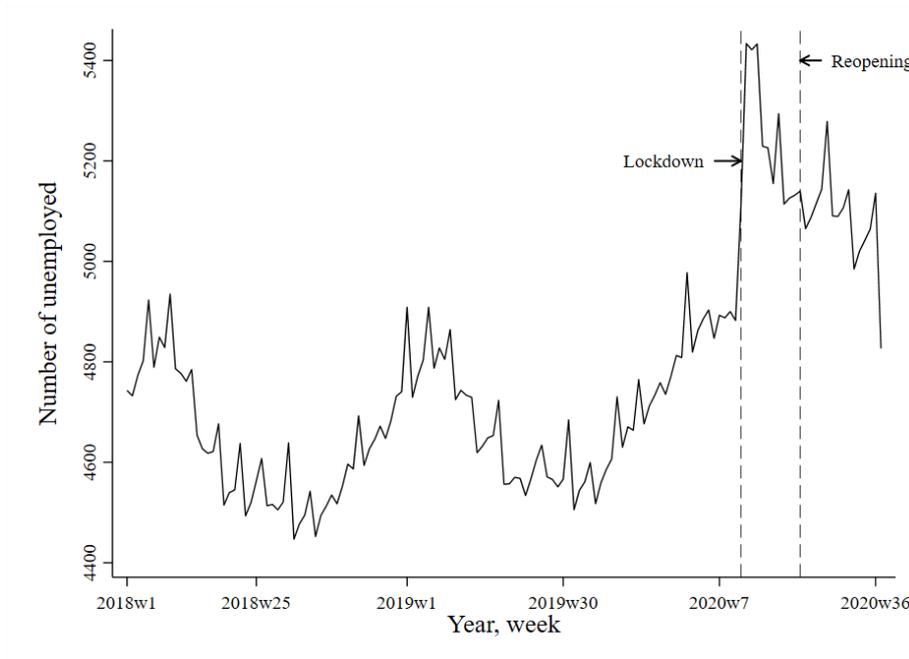
I identify exogenous variation in workload and performance pressure in Danish primary schools and Danish unemployment services. I investigate three sources of exogenous variation in pressures, which are summarized in Table 3.1.

Table 3.1: Sources of Exogenous Variation across Cases

	Unemployment Services	Primary Schools
Workload	Unemployment shock induced by COVID-19 and associated lockdown	Year-to-year variation in cohort size and class size
Performance Management	Danish Ministry of Employment’s benchmark system associated with a discontinuous sanction threat	

First, I look into an unemployment shock induced by COVID-19 and the associated lockdown, which led to an unusual and sudden rise in the caseloads of unemployment agencies. The unemployment shock is illustrated in 3.1.

Second, I look into year-to-year variation in child cohort size, which means that teachers face a varying number of students across classes from year to year (Hoxby 2000). Figure 4.3 illustrates the variation in number of students and number of hard-to-serve students across classes. Hard-to-serve students are students with learning challenges, identified specifically as students with psychiatric diagnoses, immigrant backgrounds, parents with criminal records and parents with no education beyond primary school.



**Figure 3.1: Number of Unemployed**  
 Note: Average Municipal Number of Unemployed, 2018-2020.

Third, I look into one source of as-if-random variation in performance pressure: the Danish Ministry of Employment’s benchmark system. If employment agencies do not satisfy the system’s performance target, agencies are at risk of losing municipal competence over the service area. The system is associated with a discontinuous structure of performance pressure, since agencies performing as expected face no consequences, whereas agencies performing minutely less well than expected face political sanctions. This is illustrated in Figure 3.3.

Overall, these incidences capture variation in pressures with actual and long-term consequences for the functioning of the service organizations. In the individual papers, I argue for why each of these specific incidences serve as relevant interventions for studying the research question. Furthermore, I make great effort to validate whether the assumption of as-if-random assignment holds.

Each of the three sources of variation in pressure entails specific analytical challenges and must be analyzed using different identification strategies. I rely on a diverse set of identification strategies that build on different identifying assumptions. I rely on interrupted time series (ITS), regression-discontinuity design (RDD), difference-in-differences (DiD) and institution fixed effects analysis (Dunning 2012; Angrist and Pischke 2015; Cattaneo, Idrobo, and Titiunik 2019; Shadish et al. 2002; Hoxby 2000).

### 3.3 Unemployment Services and Primary Schools

I chose to conduct the studies within the setting of Danish primary schools and Danish unemployment services. They both provide good empirical cases for investigating the research question. First, both service areas constitute classic examples of street-level bureaucracies, where teachers and caseworkers are responsible for delivering services to citizens under substantial discretion. On top of this, both cases exemplify a substantial citizen-state encounter, where bu-

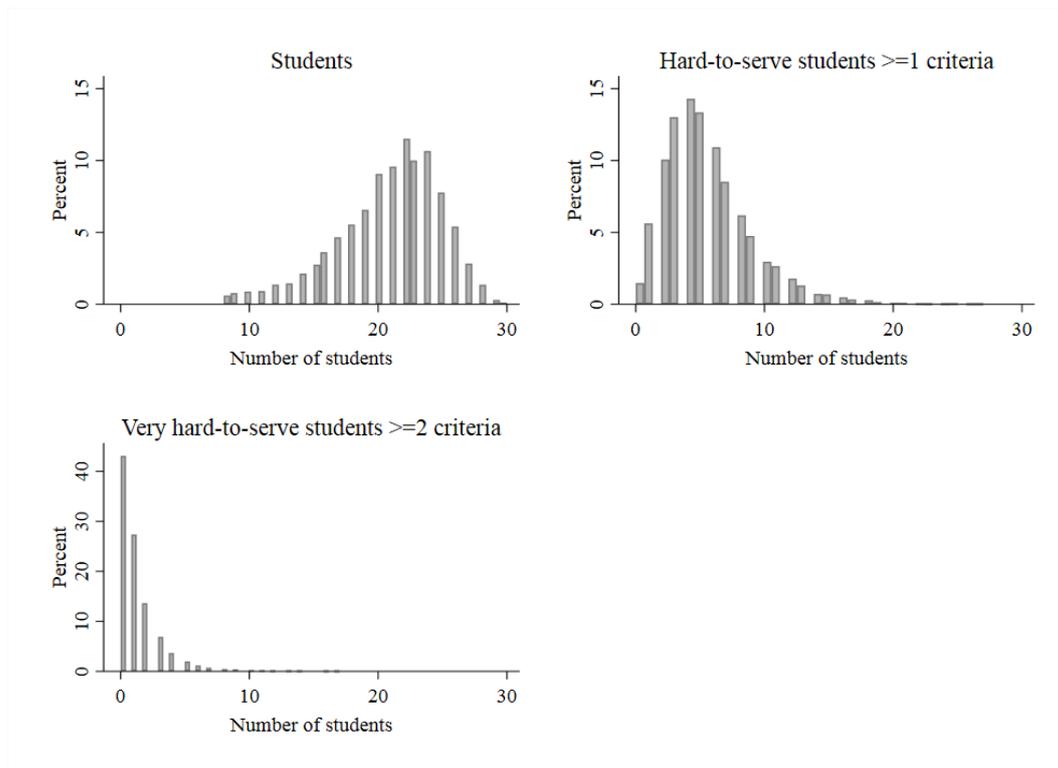


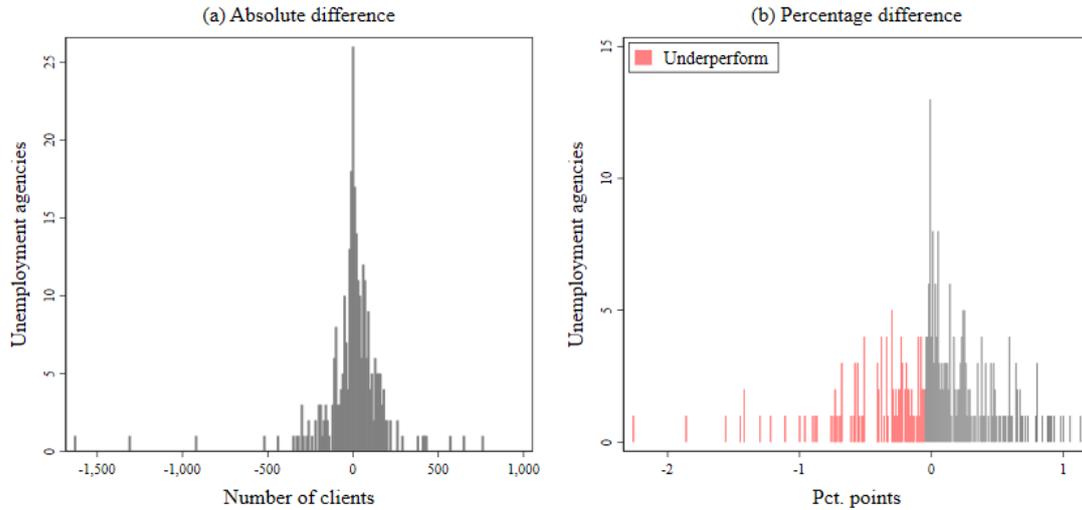
Figure 3.2: Number of Students and Hard-to-serve Students pr. Class

Note: Based on 6,344 classes. Hard-to-serve students fulfill at least one of the four criteria ( $\geq 1$  criterion). Very hard-to-serve students fulfill at least two criteria ( $\geq 2$  criteria).

reacrat and client interact over a long period of time and where the bureaucrat has meaningful influence on the client’s future life prospects—either through support of learning in schools or through influence on chances of getting reemployed in the unemployment service setting. When bureaucrats have a real impact on peoples’ lives, the stakes attached to dysfunctional service delivery increase, which raises the importance of understanding the underlying cause.

Second, both schools and unemployment services are salient cases in terms of investigating reactions to pressure. Both cases are often highlighted as being high-pressure. For example, stories of overcrowded classrooms and long queues of unemployed people waiting to access unemployment benefits appear in the media. Likewise, stories about extensive monitoring and pressure to improve reemployment rates and academic achievement also frequently appear. In addition, earlier studies that investigate how pressure influences service delivery tend to focus on these exact cases. For example, Soss, Fording, and Schram (2011a) investigate how unemployment services react to performance pressure and Andersen and Guul (2019) investigate how workload influences service delivery in a schooling context.

Together, these factors make it relevant to investigate the research question in these two settings. Furthermore, these factors indicate that the two settings constitute most-likely cases (Gerring 2007) in the sense that we should really expect to see that the “pressure narrative” of the street-level bureaucracy literature holds in these cases.



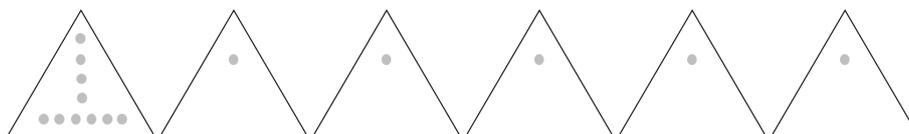
**Figure 3.3: Performance Target Score, Absolute and Percentage Difference**  
 Based on numbers from the Ministry of Employment, pooled across years 2018-2020. Observations in red indicate municipalities that receive a score below zero.

### 3.4 Administrative Records and Interviews

I investigate public service organizations’ reactions to the three sources of pressure through comprehensive individual-level registry data on public service delivery, which taps into different forms of dysfunctional patterns in service delivery. I draw on several different registries from Statistics Denmark (DST), the Danish Agency for Labour Market and Recruitment and the National Agency for IT and Learning. Table 3.2 includes an overview of the different registries that I use in the two cases.

These data sources make it possible to move beyond intentional behavior and investigate actual service delivery. For example, when I study service providers’ use of welfare penalties of unemployed clients, I am not limited to studying caseworkers’ intention to sanction, but can identify patterns in welfare penalties that have been carried out against real-world unemployed people, who are financially dependent on welfare benefits. These data thus represent actual citizen-state encounters that have had a great impact on people’s lives.

In addition to the registries, I conduct 15 interviews with caseworkers and managers from the unemployment service setting. I first undertake an in-depth examination of one local unemployment agency, where I conduct interviews with persons from all levels of the organizational hierarchy. As a second step, I validate the insights from the first interviews by collecting second opinions from agency directors of five other unemployment agencies. The overall structure of the interview collection is shown in Figure 3.4.



**Figure 3.4: Structure of Data Collection**

Table 3.2: Registries

Case	Registry	Content
Unemployment Services	The Danish Ministry of Employment's database: DREAM	Weekly information on unemployment-related public transfers, participation in activation programs and extensive individual background information.
	Records from Danish Agency for Labour Market and Recruitment	Clients' interactions with unemployment services: Welfare sanctions and case-worker meetings.
	Population registry (BEF), Education registry (UDDF)	Individual background information.
Primary Schools	Special education registry (UDSP), Compressed student registry (KOTRE)	Information on schools' enrollment, cohorts and classes.
	Population registry (BEF), Criminal registry (KRIN), Psychological diagnosis registry (PSYKDIAG), Education registry (UDDF)	Individual background information on students and students' parents.
	Records from National Agency for IT and Learning	Students' scores in the national tests.

Substantive and contextual knowledge plays an important role at every stage of natural-experimental research designs (Dunning 2012, 4). I therefore use the interviews to substantiate the claim that assignment to treatment is really as-good-as-random and for interpreting, explaining and contextualizing effects. Interviews are ideal for getting close to the mechanisms behind the general findings, into which the administrative records provide limited insight. In Paper C (Disciplining Paper), I furthermore include the interviews as a data source in their own right as part of a mixed methods study (Hendren et al. 2022; Gilad 2021).

### 3.5 Overview of the Four Studies

Table 3.3 includes an overview of the case, source of exogenous variation, identification strategy and data across the four papers in the dissertation.

Table 3.3 Overview of Case, Source of Exogenous Variation, Identification Strategy and Data

Short Title	Case	Exogenous Variation	Identification Strategy	Data
A. Discrimination Paper	Danish unemployment services	Unemployment shock induced by COVID-19	ITS + DID + Survival analysis	Caseworker meetings
B. Concept Paper	Danish primary schools	Year-to-year variation in child cohort and class size	(Systematic literature review) + Fixed effects estimation	Test scores
C. Disciplining Paper	Danish unemployment services	Benchmark system with a discontinuous sanction threat	RDD	Welfare Sanctions + Qualitative Interviews
D. Dynamics Paper	Danish unemployment services	Benchmark system with a discontinuous sanction threat	Multiple RDDs	Caseworker meetings + Activation measures + Reemployment + Visitation to excluded benefits



## Chapter 4

### Results

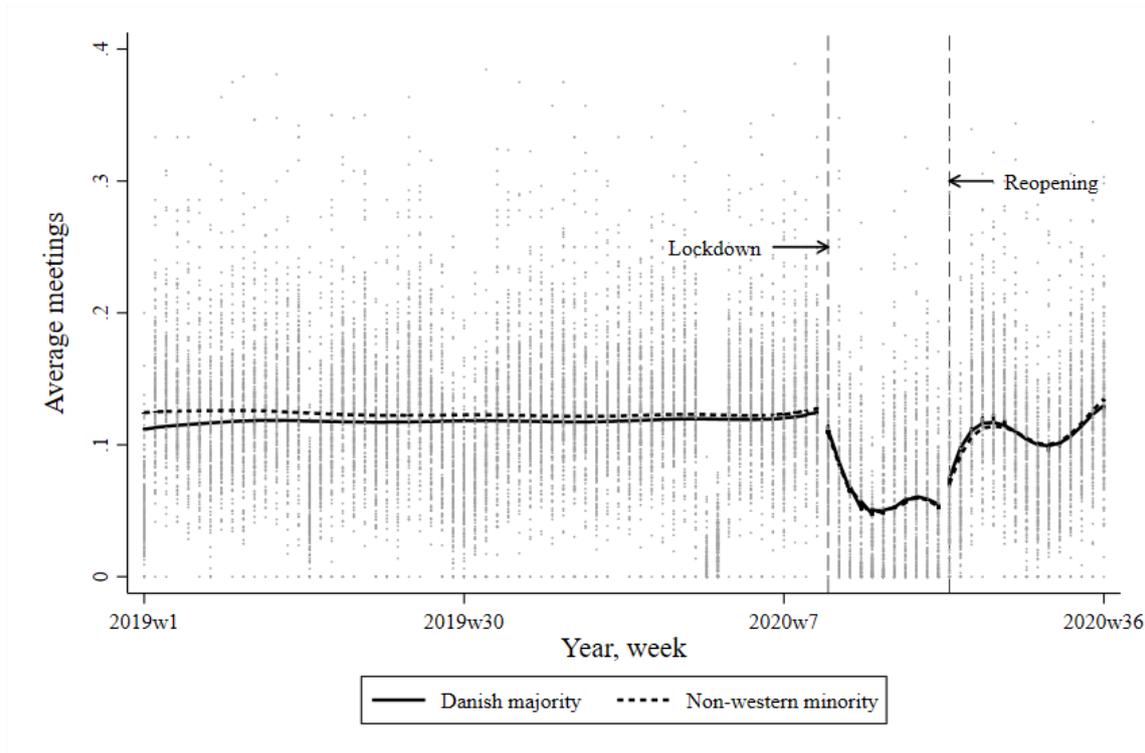
This chapter covers the main conclusions on how macro-level pressure from workload and performance management systems influences service delivery at the micro level. I report the main results from each of the four studies. Overall, the results indicate that public service organizations are capable of handling substantial pressures without turning to dysfunctional behaviors in service delivery. This, however, does not mean that pressure never matters for service delivery. To simplify the presentation, I primarily rely on figures to illustrate the main findings. More detailed regression tables, statistical analysis and robustness tests are available in the individual papers.

#### 4.1 Does High Workload Lead to Discrimination?

In Paper A (Disciplining Paper), I document that the COVID-19 pandemic and lockdown led to a 20% workload increase at Danish unemployment agencies, and that caseworker meetings are a core tool that needed prioritization. In spite of this, Figure 4.1 shows that caseworkers did not respond by discriminating in the delivery of caseworker meetings.

Figure 4.1 shows the average level of caseworker meetings for the ethnic majority and the non-Western ethnic minority aggregated at municipal level. Looking at the solid and the dotted curves, it is clear that the two ethnic groups followed each other closely before the lockdown. The overall trend is clearly disrupted by the COVID-19 pandemic and lockdown. It drops significantly in Week 11 when the lockdown is implemented and increases steeply after Week 22 when the lockdown is partly lifted. In spite of this remarkable disruption, the graph shows no increased difference in meeting trend between the two ethnic groups. The Danish majority (solid line) and non-Western ethnic minority (dotted line) follow each other closely throughout the workload shock. I document that the same pattern exists if we look at municipalities that are exposed to an especially high workload shock, if we focus on minority clients who originate from a MENA country, and if we conduct separate analyses for in-person and digital meetings. This indicates that caseworkers handle the workload by lowering the general service level, but that they do not resort to discrimination.

Moving beyond the allocation of services and taking a look at service outcomes, Figure 4.2 reports the result of a survival analysis. It reveals that ethnic minority citizens who were laid off in the weeks following the lockdown were stuck in unemployment for a substantially longer period than ethnic Danes. Figure 4.2 clearly shows that the two ethnic groups have different survival functions and that the non-Western ethnic minority has a higher probability of staying unemployed. In fact, the non-Western ethnic minority has a 30% lower probability of getting reemployed at any time during the period of analysis compared to the Danish ethnic majority. This reveals that caseworkers were not able to secure equal outcomes for both ethnic groups, and



**Figure 4.1: Weekly Caseworker Meetings, Ethnicity and the Unemployment Shock**

Note: The solid and dotted curves are the predicted level of weekly caseworker meetings per benefit recipient generated by locally weighted (LOWESS) regression with no adjustments for covariates. Caseworker meetings are aggregated as municipal means here to facilitate visualization. For clarity, the y-axis is trimmed and displays the bottom 99 percent of the data.

that ethnic minority citizens are in fact worse off—even though they received the same services, as far as we can see in the observed data.

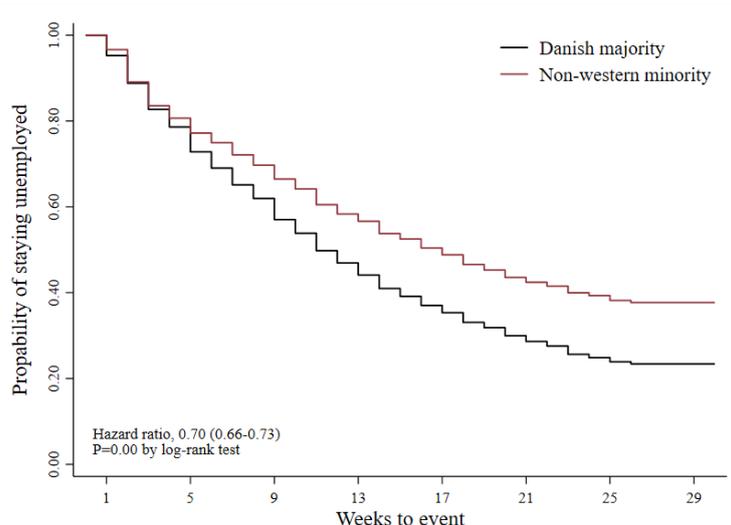


Figure 4.2: Ethnicity and unemployment duration

Note: Kaplan-Meier survival functions for the Danish majority and the non-Western minority. The figure also reports the result of a log-rank test and the hazard ratio, which is calculated based on cox proportional hazard regression model. Based on 37,189 unemployment spells.

## 4.2 Does High Workload Lead to Poor Service Quality?

In Paper B (Concept Paper), my coauthor and I also focus on pressure from workload. We demonstrate that *caseload*, i.e. number of students in class, does not influence service quality, measured through students' test scores. In contrast, we show that the much more comprehensive concept of *taskload* does have a small negative impact on service quality. We take several steps to reach this conclusion.

First, we conceptualize *taskload*. In contrast to the widespread *caseload* concept, taskload includes the intensity of tasks that each case demands. We capture these differences between individual clients by assigning clients with a weight. We define the concept of taskload as the sum of individual clients' weights, which is illustrated in the following equation:

$$Taskload = \sum Weight_i \quad (4.1)$$

Second, we demonstrate how it is possible to measure taskload in an empirical setting. We do so by identifying hard-to-serve clients and by assigning these clients with a weight greater than one. Hard-to-serve students are those with learning challenges, identified specifically as students with psychiatric diagnoses, immigrant backgrounds, parents with criminal records and parents with no education beyond primary school. As a starting point, we assume that hard-to-serve students require double the effort from the teacher compared to the typical student. As a result, we multiply each hard-to-serve student by a weight of 2 in our primary taskload measure. Figure 4.3 demonstrates what different weights mean in practice. The figure illustrates the distribution in the simple caseload measure (i.e. number of students) and in the taskload measure based on different weights for the hard-to-serve ( $\geq 1$  criterion) and the very hard-to-serve ( $\geq 2$  criteria). Focusing on the hard-to-serve ( $\geq 1$  criterion) in Panel A, it is clear that the distribution of the taskload measures is more dispersed than the caseload measure. In fact, what the taskload measure does in this case is to push the distribution to the right. The max-

imum caseload value is 30 since the maximum number of students allowed in class is 30, but the taskload measures exceed this upper limit depending on the choice of weight.

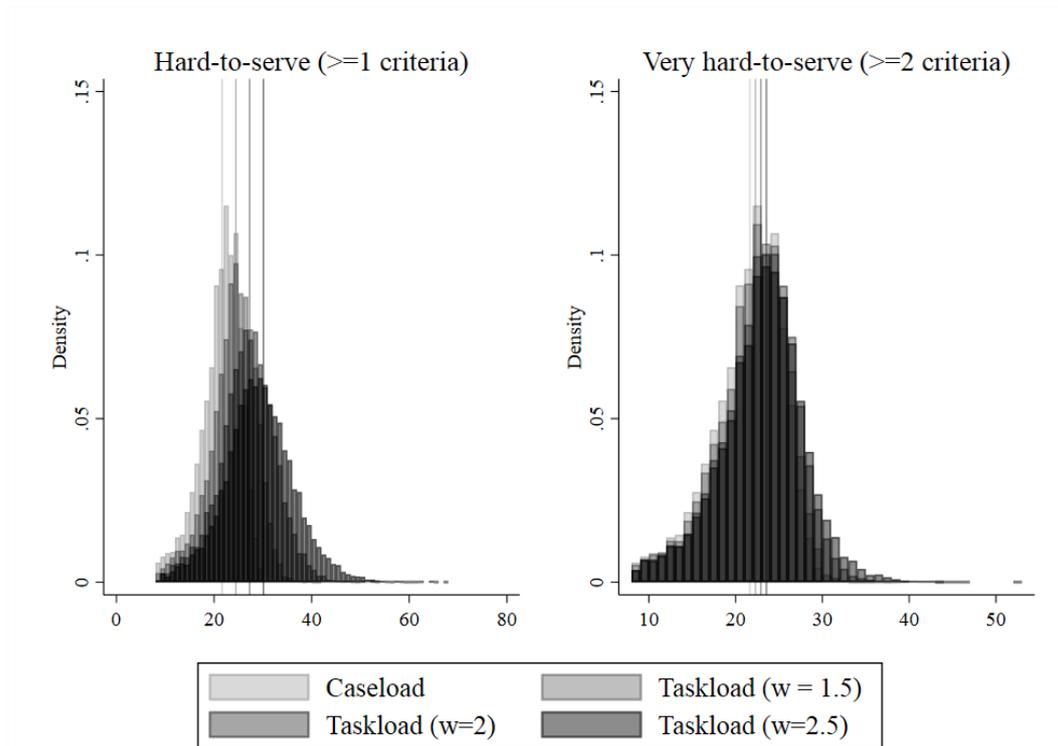


Figure 4.3: Caseload and Taskload with Different Weights  
 Note: Based on 6,344 classes.

Third, we identify the effects of both caseload and taskload on students’ test scores. Table 4.1 reports the results of the analysis. We generally find small effects on service delivery, but we do see that the effect sizes between caseload and taskload vary. Model 1 shows the effect of caseload on students’ academic performance, and it is clear that the effect of an additional student in class is close to zero and statistically insignificant. Moving on to look at Models 2-4 and the effect of taskload, it is clear that the effect here is substantially larger and significant. When accounting for the number of hard-to-serve students in class, one additional student in the teacher’s taskload will cause students to perform 0.3% of a standard deviation less well on their tests.

This implies that while the effect of increased caseload is nonexistent, the effect of taskload is negative and substantial for the average student in Danish schools. As a result, it seems that a very extensive form of workload, where a high number of clients pose heavy demands, may lead to poorer service quality.

### 4.3 Does Performance Pressure Lead to Disciplining?

Paper C (Disciplining Paper) consists of a mixed methods study. The quantitative analysis reveals that Danish unemployment services do not increase the level of welfare sanctions when they are put under increased performance pressure. Figure 4.4 shows the sanction frequency for unemployment agencies with an outcome target score within a bandwidth of  $\pm 0.3$  around the cutoff at zero. Observations to the left of the cutoff, colored with red, represent treated un-

Table 4.1: Caseload, Taskload and Service Quality

	Testscores	Testscores	Testscores	Testscores
Caseload	-0.001 (0.001)			
Taskload, w=1.5		-0.003*** (0.001)		
Taskload, w=2			-0.003*** (0.001)	
Taskload, w=2.5				-0.003*** (0.001)
Testscores, 2. grade	0.672*** (0.002)	0.672*** (0.002)	0.672*** (0.002)	0.672*** (0.002)
Constant	0.044* (0.021)	0.097*** (0.20)	0.119*** (0.019)	0.124*** (0.017)
N	102,825	102,825	102,825	102,825

Note: OLS coefficients. Standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.

employment agencies; that is, agencies at risk of political sanctions. Observations to the right, colored with black, represent agencies that are not at immediate risk. Panel A displays a linear fit, and Panel B displays a quadratic fit between the forcing variable and the sanction frequency. Panel A shows some sign of a drop in sanction frequency at the cutoff, but we should be cautious in drawing conclusions based on the graphical evidence. The drop also seems to disappear when we include a quadratic fit in Panel B. The formal estimates, which are reported in the paper, confirm the null effect. The size and direction of the point estimates are substantially small and insignificant across the different model specifications. The direction of the point estimate also changes across the different models.

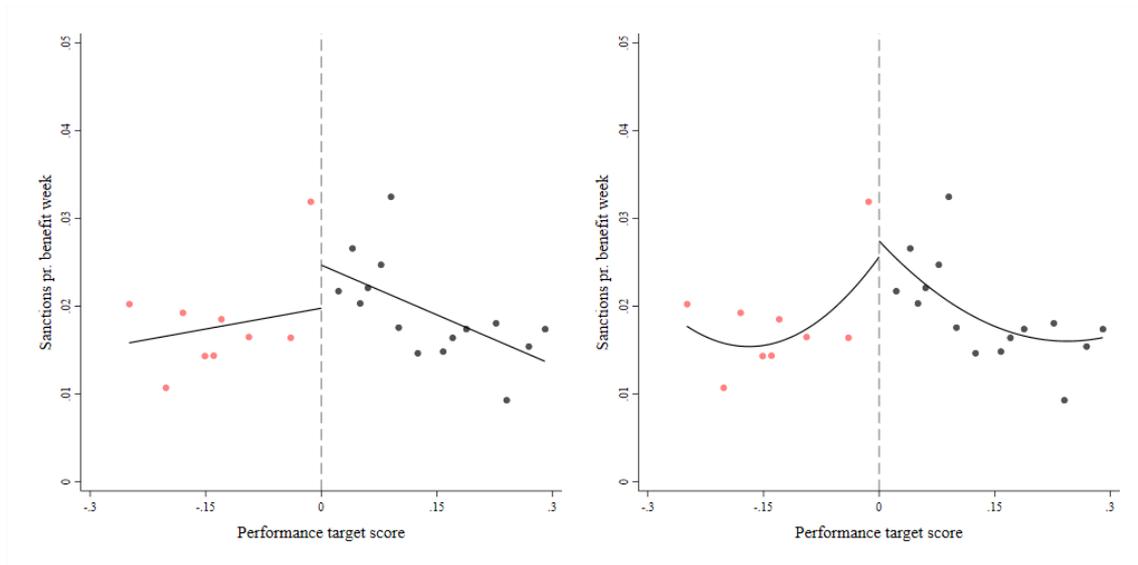


Figure 4.4: Performance Pressure and Sanction Frequency

Note: Binned scatter plot. Each point represents the average value of the dependent variable within evenly spaced bins. Bandwidth =  $\pm 0.3$ . N = 279,384.

This finding stands in contrast to the findings from the United States, and specifically the finding from the register-based analysis conducted by Soss, Fording, and Schram (2011a) in Florida. As a consequence, I take a next step and set out to understand how context matters.

Table 4.2: Overview of differences between the two cases

Theme	Florida, US	DK
Case Work	Reactive and clerical	Proactive and motivational work
Availability of Tools	Limited tools; Sanctioning is the primary tool.	Several supportive tools; Sanctioning is a rarely used last resort
Client noncompliance	Caseworkers get frustrated with clients	Not an area of attention. If so, caseworkers feel sorry for clients
Performance Pressure	Performance pressure is extended to the frontline	Performance pressure is concentrated at top management level and translated downwards

Note: US characteristics are based on the description by Soss, Fording, and Schram (2011a).

This qualitative analysis suggests that Denmark and the state of Florida in the United States differ on important variables (see 4.2). Two things stand out. First, Danish caseworkers hold a fundamentally different role in relation to clients than they do in the United States. The interviews suggest that casework in Denmark is proactive and focused on motivational work. Sanctions are rarely used, and caseworkers have several alternative supportive tools at their disposal that they can use to help people get reemployed.

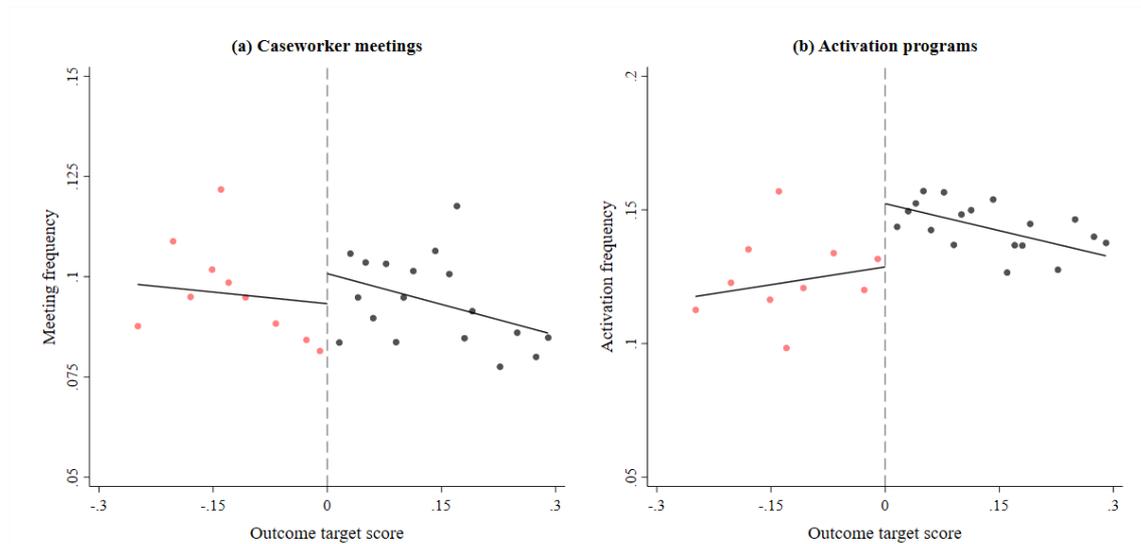
Second, local management at Danish unemployment agencies handles performance pressure in a fundamentally different way than in the United States. The interviews with Danish managers and caseworkers suggest that performance pressure is concentrated at the top management level and that caseworkers do not worry about aggregate performance. In fact, the interviews indicate that management intentionally works to buffer the street level from the pressure. They do so by translating general performance pressure into behavioral directives that make sense at the street level.

## 4.4 Does Performance Pressure Lead to Gaming and Effort Substitution?

In Paper D (Dynamics Paper), I focus on the same performance management system. I document that local service agencies do not respond to outcome-based accountability by outright gaming, but that the political principal can push service providers towards effort substitution by making it very costly and risky not to do so. I exploit the fact that the system changed from being a purely outcome-based system (System 1) into a hybrid system, where service providers are held accountable based on both outputs and outcomes (System 2).

I first document that agencies do not respond to either of the two systems by resorting to strategies of outright gaming, e.g. by exploiting the fact that certain client groups are exempt from the measurement system. I also document that the systems have no influence on agencies'

performance, measured as client exits to reemployment (figures and tables are reported in the paper).



**Figure 4.5: Reaction to Outcome Target, Activation Programs and Caseworker Meetings (System 1)**

Note: Binned scatter plot. Each point represents the average value of the dependent variable within evenly spaced bins. Bandwidth =  $\pm 0.3$ . N = 915,173.

In contrast, I document that they do change the use of core services: caseworker meetings and activation programs. Figure 4.5 shows how agencies react to System 1 by lowering the quantitative level of services. Figure 4.5 shows the activation frequency and meeting frequency across agencies with an outcome target score within a bandwidth of  $\pm 0.3$  around the cutoff. Observations to the left of the cutoff represent treated unemployment agencies; that is, agencies at risk of political sanctions. The graphs show signs of a discontinuity at the cutoff, where treated agencies have a lower meeting and activation frequency. This drop is confirmed in the formal models. Clients who are associated with an unemployment agency that is put under pressure receive 0.02 fewer caseworker meetings per week and participate in 0.03 fewer activation programs per week. At first sight, these effect sizes appear small, but they are in fact substantial. Compared to the activation and meetings frequency in the control group, these drops constitute an approximately 17 percent drop in the activation frequency and an approximately 15 percent drop in meeting frequency.

Survey and interview data indicate that unemployment agencies altered their focus to prioritize quality in services over quantity in services in line with the process-oriented philosophy of case management, which stresses individual assessment and emotional interaction with clients (Dias and Maynard-Moody 2007). This reaction can thus be understood as “reversed effort substitution,” where agencies prioritize service quality, which is difficult to measure, at the cost of service quantity, which is easy to measure.

It is only when the ministerial principal complicates the system by introducing output-based accountability that service providers are forced to resort to effort substitution, since other strategies become too risky. Figure 4.6 graphically displays the result. The graph to the left shows no sign of a discontinuity in meeting frequency at the cutoff. The graph to the right shows signs of a positive jump in activation frequency at the cutoff. This indicates that agencies increase the use

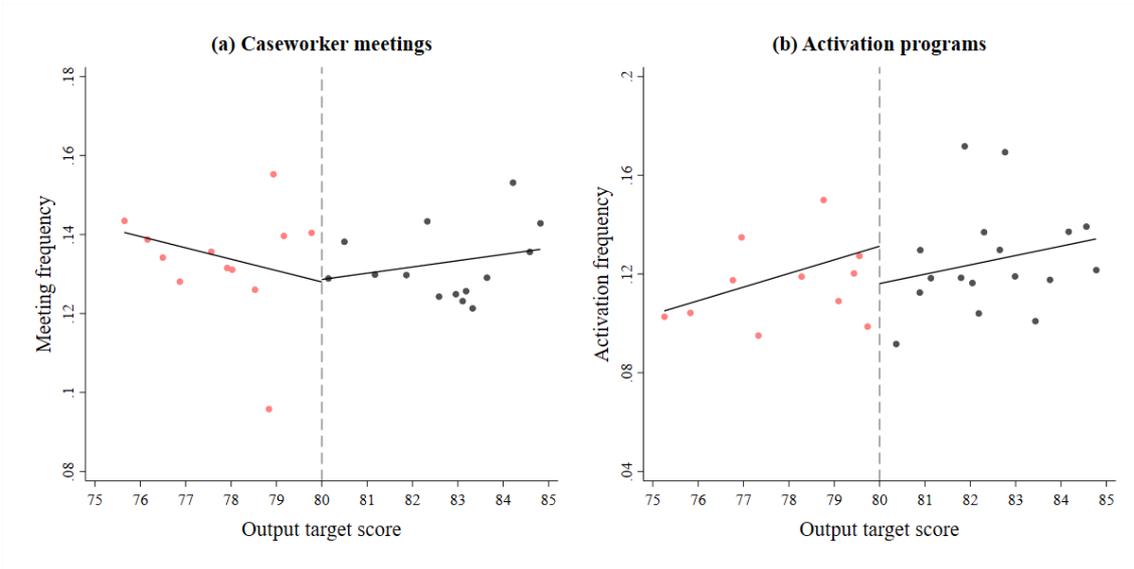


Figure 4.6: Reaction to Output Targets (System 2)

Note: Binned scatter plot. Each point represents the average value of the dependent variable within evenly spaced bins. Bandwidth =  $\pm 5$ .  $N = 156,956$ .

of activation programs when they are performing below the output target that specifically relates to the use of activation programs. This increase is confirmed in the formal models. Clients who are associated with an unemployment agency that performs below the output target receive 0.02 more activation programs per week. Again, this effect size is substantial. Compared to the activation frequency in the control group, this increase constitutes an approximately 13 percent increase in the activation frequency. This strategy can be thought of as effort substitution, where agencies increase the quantified service level at the cost of unmeasured quality.

The analysis thus suggests that performance indicators compete for attention, and that low-risk output targets take precedence when implemented together with high-risk outcome targets. It is much riskier to fulfill abstract outcome targets than concrete output targets, because the distance between decisions and outcomes is much greater than the distance between decisions and outputs. The results thus suggest that the political principal can push service providers towards effort substitution by making it very costly and risky to do otherwise.

## **Chapter 5**

# **Discussion and Conclusion**

In this dissertation, I have set out to answer whether pressures from the political-administrative environment on public service organizations lead to dysfunctional patterns in the delivery of services to citizens. I have done so in four papers that investigate different types of pressures and different types of dysfunctional patterns in service delivery. In this chapter, I will summarize my answer to the research question based on the findings from the four papers and discuss the broader applicability of the findings for the street-level bureaucracy literature. I then go on to outline some of the key implications of the findings. What do they imply for the organization of service delivery, and how can we overcome dysfunctional patterns in public service delivery? Finally, I outline some future avenues for research based on the insights of this dissertation.

## **5.1 Expanding the Street-Level Narrative**

### **5.1.1 Implications and Limitations**

Together, the four studies of the dissertation indicate that public service organizations are capable of handling substantive pressures without turning to dysfunctional behaviors. In particular, the studies provide examples of public employees who work under high workloads and pressure to perform without resorting to discrimination, disciplining, gaming or goal displacement. This does not mean that pressure never matters for service delivery, but that we do not see the specific dysfunctional patterns that we would expect. For example, service providers handle high workloads by decreasing the general service level instead of turning to discrimination.

This insight is based on analyses from primary schools and unemployment agencies. These constitute classic examples of street-level bureaucracy where we would expect to see the “pressure narrative” play out. As a result, these findings are relevant for the broader field of street-level bureaucracy. In fact, the findings challenge the existing narrative within the field, in which both workload and performance management are believed to be the source of various evils in public service delivery. In other words, the findings question some of the core assumptions of the street-level bureaucracy literature, and suggest that we cannot assume that pressures from high workloads and performance management systems automatically lead to dysfunctional patterns in service delivery. This in turn suggests that we should place greater focus on identifying the conditions under which pressures do lead to dysfunctional reactions and on identifying the conditions under which they do not.

The results do however not imply that the “pressure narrative” never holds. Primary schools and unemployment services exemplify service encounters in which bureaucrats mostly rely on deliberative responses. School teachers and caseworkers in the unemployment services generally make informed and long term plans for clients, which they can adjust over time. In other

service contexts, bureaucrats are forced to rely more on quick and intuitive judgments. For example, police officers and emergency physicians have to make quick decisions that are irrevocable. Using the terminology of Kahneman (2011), these bureaucrats are forced to rely more on System 1 thinking than on System 2 thinking. As a result, they may be more sensitive towards pressures from challenging working conditions. Furthermore, bureaucrats navigate many other types of challenging working conditions than high workloads and high-stakes performance targets. For example, Newman (2013) highlights the emergence of new forms of pressures associated with intense media scrutiny and new technologies, which may influence service delivery in a fundamentally different way. This suggests that the “pressure narrative” may hold in other service contexts and in relation to other types of pressures.

### 5.1.2 Well-functioning Organizations

Though not a primary aim of the dissertation, a relevant next step is to consider what we can learn from the dissertation’s positive cases. In doing so, I primarily draw on the qualitative interviews, which complement the administrative records by providing more in-depth knowledge about the underlying mechanisms. These reflections, of course, build on less solid ground and are primarily indicative in nature.

Overall, the qualitative material suggests that the public service organizations under study are equipped to handle certain pressures because of well-functioning organizational structures. For example, Paper A (Discrimination Paper) provides an answer to why a substantial workload increase did not act as a catalyst of discrimination in service delivery. Interviews with caseworkers suggest that professional norms and administrative support systems meant that the workload increase did not force caseworkers to rely on stereotypes about clients and that the mechanisms described by the model of *statistical discrimination* and *unconscious bias* did not play out. Likewise, Paper C (Disciplining Paper) provides some specific answers as to why performance pressure did not lead to disciplining as expected. The qualitative part of the study reveals that management intentionally works to buffer the street level from the pressure by translating general performance pressure into behavioral directives that make sense at the street level.

These examples indicate that well-functioning organizations with professional bureaucrats, administrative support systems and a management that takes its responsibility seriously are equipped to handle the pressures. To some extent, these are characteristics that we can expect from public service organizations. In fact, we construct and design organizations with characteristics such as task division and hierarchical authority so that management is equipped to protect the core task of the organization from external influences (O’Toole and Meier 2010; O’Toole and Meier 2014; O’Toole, Meier, and Nicholson-Crotty 2005; Stazyk and Goerdel 2011; Olsen 2006). At the same time, this also suggests that we cannot assume that these findings will travel to other organizational contexts with different structures. Pressure might lead to dysfunctional reactions in organizations that do not possess these characteristics.

Of course, these propositions need to be tested more thoroughly in future studies. This calls for studies of macro-level pressures in different organizational settings, including both well-functioning and poorly functioning organizations. For example, it would be interesting to test whether high workloads do lead to discrimination in settings where bureaucrats do not have easy access to information about clients and where high workloads may force bureaucrats to rely on stereotypes as a consequence.

## 5.2 Practical Implications

Taking the insights from this dissertation seriously, a cost-efficient politician might wonder whether it is possible to simply increase the pressures on public service organizations without seeing any consequences in terms of poor service quality, discrimination, gaming or effort substitution. For example, is it feasible to put more students in school classes or increase the caseloads given to caseworkers? Can we just demand more from service providers by introducing more outcome-based accountability in the governing of organizations? To some extent, the answer we can provide to such questions based on the findings of the dissertation is yes—public service organizations are capable of handling pressure.

But at the same time, the dissertation also indicates that the full picture is more complicated. First, the studies demonstrate that there are limits to this optimistic conclusion. For example, comprehensive forms of workload pressure do affect service quality, and outcome-based accountability coupled with output-based accountability does lead to effort substitution. This indicates that we cannot increase pressures infinitely. Second, the dissertation suggests that the public service organizations under study are equipped to handle these pressures because of well-functioning organizational structures. This indicates that political principals should be careful not to pressure weak organizations. The study by Soss, Fording, and Schram (2011a) provides a good example of this.

A related question that may arise is how we should react to cases in which dysfunctional service delivery does take place. The dissertation provides some indicative suggestions. Decreasing the workload of public employees and revoking the widespread use of performance management systems might not serve as an easy fix in all cases. Such pressures are often associated with political preferences and are consequently difficult to revoke. Instead, we should examine other parts of the long and complicated chain that exists between the pressures at the macro level and service delivery at the micro level. Taken together, the findings suggest that the solution to dysfunctional service delivery lies in constructing well-functioning organizational structures that give public service organizations the means to handle the pressures that will inevitably arise. To some extent, this argument fits well with Lipsky (1980), who argues that public service organizations can never be free from some forms of pressure and constraints. If this is so, it is not possible to remove pressures completely, and we might be forced to find solutions elsewhere.

## 5.3 Future Avenues for Research

The insights of the dissertation give rise to some general suggestions regarding future avenues for research. Below, I will reflect on some of these.

### 5.3.1 Core Concepts and Generalization across Contexts

The results of the dissertation run counter to a widespread narrative in the public administration literature, which suggests that we should give more attention to some of the core assumptions and concepts in the field. This would imply a break with the tendency to publish research about novel theory, rather than studies that investigate and extend existing theory (Walker, James, and Brewer 2017). In this dissertation, I demonstrate that there is ample reason to do the latter.

For example, my coauthor and I illustrate in Paper B (Concept Paper) that the core concept of workload lacks conceptualization and theorizing. Based on a systematic literature review, we document that very few scholars conceptualize workload, and that those who do disagree on what the concept entails. This suggests that there is a great deal to be gained from revisiting some of the field's classic concepts.

In relation to this, the dissertation indicates that we should be careful when generalizing findings from one national context to another. Many of the influential studies on street-level bureaucracy were conducted in the United States (Lipsky 1980; Maynard-Moody and Musheno 2020; Keiser, Mueser, and Choi 2004; Soss 1999), and it is not certain that these core insights travel to other contexts. Scholars have long pointed to this (O'Toole and Meier 2014; Meier et al. 2015; Pollitt 2013), but few studies actually consider the role of context (Hupe and Buffat 2014). In this dissertation, I demonstrate that we can gain valuable insight from replicating existing studies in new national contexts. Specifically, Paper C (Disciplining Paper) demonstrates how performance pressure interacts with the local organizational setting to produce a different result in Denmark than in the United States (Soss, Fording, and Schram 2011a). This focus on context will only become more important as the field of public administration grows more diverse with the inclusion of perspectives from the Global South.

### 5.3.2 Real-World Service Delivery

The dissertation indicates that it is important to study public service delivery in real-world service contexts. Many of the studies that identify dysfunctional patterns in service delivery rely on survey experimental evidence (i.e. Tummers and Rocco 2015; Schram et al. 2009; Guul, Pedersen, and Petersen 2020; Jilke and Tummers 2018; Pedersen and Stritch 2018; Andersen and Guul 2019; Petersen 2019). These studies have provided valuable insights about patterns in service delivery. However, the heavy reliance on survey experiments within the field of public administration creates an imbalance between survey experimental approaches and studies conducted in more real-world settings. This imbalance may give us a skewed understanding of service delivery, which means that we may misjudge the magnitude of dysfunctional service delivery.

More specifically, we might overlook the role that organizational structures play in supporting decision-making in service delivery. Survey and laboratory experiments tend to isolate bureaucratic decision-making from its organizational environment. Bureaucratic decision-making, however, is not conducted in isolation, but in connection with management, colleagues and administrative systems. The findings from this dissertation indicate that it is important to include these structures when we investigate service delivery. Future studies of service delivery should therefore strive to conduct studies in real-world service settings.

## Bibliography

- Andersen, Simon Calmar, and Thorbjørn Sejr Guul. 2019. "Reducing Minority Discrimination at the Front Line—Combined Survey and Field Experimental Evidence." *Journal of Public Administration Research and Theory*, 429–444.
- Angrist, Joshua David, and Jörn-Steffen Pischke. 2015. *Mastering "Metrics": The Path from Cause to Effect*. 282. Princeton, New Jersey: Princeton University Press. Book. isbn: 9780691152837 9780691152844.
- Assouline, Michaela, Sharon Gilad, and Pazit Ben-Nun Bloom. 2021. "Discrimination of Minority Welfare Claimants in the Real World: The Effect of Implicit Prejudice." *Journal of Public Administration Research and Theory*.
- Berkel, Rik. 2020. "Making welfare conditional: A street-level perspective." *Social Policy and Administration* 54 (2): 191–204.
- Bevan, Gwyn, and Christopher Hood. 2006. "What's Measured is What Matters: Targets and Gaming in The English Public Health Care System." *Public Administration* 84 (3): 517–538.
- Blom-Hansen, Jens, Rebecca Morton, and Søren Serritzlew. 2015. "Experiments in Public Management Research." *International Public Management Journal* 18 (2): 151–170.
- Bloom, Howard S., Carolyn J. Hill, and James A. Riccio. 2003. "Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments." *Journal of Policy Analysis and Management* 22 (4): 551–575.
- Bohte, John, and Kenneth Meier. 2000. "Goal Displacement: Assessing the Motivation for Organizational Cheating." *Public Administration Review* 60:173–82.
- Boyne, George A. 2003. "Sources of Public Service Improvement: A Critical Review and Research Agenda." *Journal of Public Administration Research and Theory* 13 (3): 367–394.
- Brodkin, Evelyn Z. 2011. "Policy Work: Street-Level Organizations Under New Managerialism." *Journal of Public Administration Research and Theory: J-PART* 21 (Supplement 2): i253–i277.
- Caswell, Dorte, and Matilde Høybye-Mortensen. 2015. "Responses from the Frontline: How Organisations and Street-Level Bureaucrats Deal with Economic Sanctions." *European Journal of Social Security* 17 (1): 31–51.
- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge: Cambridge University Press. Book. isbn: 9781108710206.
- Courty, Pascal, and Gerald Marschke. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics* 22 (1): 23–56.
- Dias, Janice Johnson, and Steven Maynard-Moody. 2007. "For-Profit Welfare: Contracts, Conflicts, and the Performance Paradox." *Journal of Public Administration Research and Theory: J-PART* 17 (2): 189–211.

- Dooren, Wouter van, Geert Bouckaert, and John Halligan. 2015. *Performance management in the public sector*. Second edition. Routledge masters in public management. Milton Park, Abingdon, Oxon ;New York, NY: Routledge. Book. isbn: 9780415738095.
- Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Strategies for Social Inquiry. Cambridge: Cambridge University Press. Book. isbn: 1107017661.
- Fording, Richard C., Joe Soss, and Sanford F. Schram. 2011. "Race and the Local Politics of Punishment in the New World of Welfare." *American Journal of Sociology* 116 (5): 1610–57.
- Gerber, Alan S., and Donald P. Green. 2012. *Field experiments : design, analysis, and interpretation*. 1. ed. New York: W. W. Norton. Book. isbn: 9780393979954.
- Gerring, John. 2007. *Case study research : principles and practices*. 265. New York: Cambridge University Press. Book. isbn: 052185928X 0521676568 9780521676564.
- Gilad, Sharon. 2021. "Mixing Qualitative and Quantitative Methods in Pursuit of Richer Answers to Real-World Questions." *Public Performance amp; Management Review* 44 (5): 1075–1099.
- Gooden, Susan T. 2015. "From Equality to Social Equity." In *Public Administration Evolving: From Foundations to the Future*, 1st ed., edited by M.E. Guy and M.M. (Eds.) Rubin, 211–232. Routledge.
- Gul, Thorbjørn Sejr, Mogens Jin Pedersen, and Niels Bjørn Grund Petersen. 2020. "Creaming among Caseworkers: Effects of Client Competence and Client Motivation on Caseworkers' Willingness to Help." *Public Administration Review*.
- Hainmueller, Jens, and Dominik Hangartner. 2013. "Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination." *American Political Science Review* 107 (1): 159–187.
- Heinrich, Carolyn J., and Gerald Marschke. 2010. "Incentives and Their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management* 29 (1): 183–208.
- Hendren, Kathryn, Kathryn Newcomer, Sanjay K. Pandey, Margaret Smith, and Nicole Sumner. 2022. "How qualitative research methods can be leveraged to strengthen mixed methods research in public policy and public administration?" *Public Administration Review*.
- Hill, C. J. 2006. "Casework job design and client outcomes in welfare-to-work offices." *Journal of Public Administration Research and Theory* 16 (2): 263–288.
- Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, Organization* 7 (special): 24–52.
- Hood, Christopher. 1991. "A PUBLIC MANAGEMENT FOR ALL SEASONS?" *Public Administration* 69 (1): 3–19.
- Hoxby, C. M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *The Quarterly Journal of Economics* 115 (4): 1239–1285.
- Hupe, Peter, and Aurélien Buffat. 2014. "A Public Service Gap: Capturing contexts in a comparative approach of street-level bureaucracy." *Public Management Review* 16 (4): 548–569.

- Hupe, Peter, and Michael Hill. 2019. "Positioning street-level bureaucracy research." In *Research Handbook on Street-Level Bureaucracy: The Ground Floor of Government in Context*, edited by P. Hupe, 15–30. Edward Elgar Publishing.
- Hupe, Peter, and Theo van der Krogt. 2013. "4. Professionals dealing with pressures." In *Professionals under Pressure*, edited by Noordegraaf Mirko and Steijn Bram, 55–72. Amsterdam: Amsterdam University Press. isbn: 9789048518302.
- Jacob, Brian A. 2005. "Accountability, Incentives and Behavior : The Impact of High-stakes Testing in the Chicago Public Schools." *Journal of public economics* 89 (5): 761–796.
- Jewell, Christopher J. 2007. *Agents of the welfare state : how caseworkers respond to need in the United States, Germany and Sweden*. New York, NY: Palgrave Macmillan. Book. isbn: 9781403984111.
- Jilke, Sebastian, and Lars Tummars. 2018. "Which Clients are Deserving of Help? A Theoretical Model and Experimental Test." *Journal of Public Administration Research and Theory* 28 (2): 226–238.
- Jilke, Sebastian, Wouter Van Dooren, and Sabine Rys. 2018. "Discrimination and Administrative Burden in Public Service Markets: Does a Public–Private Difference Exist?" *Journal of Public Administration Research and Theory* 28 (3): 423–439.
- Jilke, Sebastian, and Gregg G. Van Ryzin. 2017. "Survey Experiments for Public Management Research." In *Experiments in Public Management Research: Challenges and Contributions*, edited by Gregg G. Van Ryzin, Oliver James, and Sebastian R. Jilke, 117–138. Cambridge: Cambridge University Press. isbn: 9781107162051.
- Kahneman, Daniel. 2011. *THINKING, FAST AND SLOW*. New York: Time Inc. Book.
- Keiser, Lael R., Peter R. Mueser, and Seung-Whan Choi. 2004. "Race, Bureaucratic Discretion, and the Implementation of Welfare Reform." *American Journal of Political Science* 48 (2): 314–327.
- Keiser, Lael R., and Joe Soss. 1998. "With Good Cause: Bureaucratic Discretion and the Politics of Child Support Enforcement." *American Journal of Political Science* 42 (4): 1133–1156.
- Kelman, Steven, and John N. Friedman. 2009. "Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service." *Journal of Public Administration Research and Theory* 19 (4): 917–946.
- Kihlstrom, John F. 2021. "Ecological Validity and "Ecological Validity"." 16 (2): 466–471.
- Koning, Pierre, and Carolyn J. Heinrich. 2013. "Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts." *Journal of Policy Analysis and Management* 32 (3): 461–483.
- Lavee, Einat. 2020. "Who Is in Charge? The Provision of Informal Personal Resources at the Street Level." *Journal of Public Administration Research and Theory*.
- Lipsky, Michael. 1980. *Street-level Bureaucracy: Dilemmas of the Individual in Public Services*. 30. anniversary expanded ed. 275. New York: Russell Sage Foundation. Book. isbn: 9780871545442.
- Maynard-Moody, Steven, and Michael C. Musheno. 2020. *Cops, teachers, counselors : stories from the front lines of public service*. Ann Arbor: University of Michigan Press. Book. isbn: 9780472023875 047202387X.

- McCrea, Austin M. 2022. "Coping with representation: the moderating effect of workload on individual-level representation." *International Public Management Journal* 25 (3): 321–342.
- Meier, Kenneth, Simon Calmar Andersen, Laurence J. O'Toole, Nathan Favero, and Søren C. Winter. 2015. "Taking Managerial Context Seriously: Public Management and Performance in U.S. and Denmark Schools." *International Public Management Journal* 18 (1): 130–150.
- Moynihan, Donald. 2018. "A great schism approaching? Towards a micro and macro public administration." *Journal of Behavioral Public Administration* 1 (1): 1–8.
- Newman, Janet. 2013. "3. Professionals, power and the reform of public services." In *Professionals under Pressure*, edited by Noordegraaf Mirko and Steijn Bram, 41–54. Amsterdam: Amsterdam University Press. isbn: 9789048518302.
- Noordegraaf, Mirko, and Bram Steijn. 2013. "1. Introduction." In *Professionals under Pressure*, edited by Noordegraaf Mirko and Steijn Bram, 11–20. Amsterdam: Amsterdam University Press. isbn: 9789048518302.
- Olsen, Asmus Leth, Jonas Høgh Kyhse Andersen, and Donald Moynihan. 2020. "The Unequal Distribution of Opportunity: A National Audit Study of Bureaucratic Discrimination in Primary School Access." *American Journal of Political Science*.
- Olsen, Johan P. 2006. "Maybe It Is Time to Rediscover Bureaucracy." *Journal of Public Administration Research and Theory* 16 (1): 1–24.
- Olson, Jeremiah C. 2016. "Race and Punishment in American Prisons." *Journal of Public Administration Research and Theory* 26 (4): 758–768.
- O'Toole, Jr Laurence J., Kenneth J. Meier, and Sean Nicholson-Crotty. 2005. "Managing upward, downward and outward." *Public Management Review* 7 (1): 45–68.
- O'Toole, Laurence J., and Kenneth J. Meier. 2010. "In Defense of Bureaucracy: Public managerial capacity, slack and the dampening of environmental shocks." *Public Management Review* 12 (3): 341–361.
- O'Toole, Jr., Laurence J., and Kenneth J. Meier. 2014. "Public Management, Context, and Performance: In Quest of a More General Theory." *Journal of Public Administration Research and Theory* 25 (1): 237–256.
- Pedersen, Mogens Jin, and Justin M. Stritch. 2018. "RNICE Model: Evaluating the Contribution of Replication Studies in Public Administration and Management Research." *Public Administration Review* 78 (4): 606–612.
- Pedersen, Mogens Jin, Justin M. Stritch, and Frederik Thuesen. 2018. "Punishment on the Frontlines of Public Service Delivery: Client Ethnicity and Caseworker Sanctioning Decisions in a Scandinavian Welfare State." *Journal of Public Administration Research and Theory* 28 (3): 339–354.
- Petersen, Niels Bjørn Grund. 2019. "Disciplining the Strong? Discrimination of Service Users and the Moderating Role of PSM and Ability to Cope." *Public Management Review*, 1–21.
- Pfaff, Steven, Charles Crabtree, Holger L. Kern, and John B. Holbein. 2021. "Do Street-Level Bureaucrats Discriminate Based on Religion? A Large-Scale Correspondence Experiment among American Public School Principals." *Public Administration Review* 81 (2): 244–259.

- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62 (4): 659–661.
- Pollitt, Christopher. 2013. *Context in public policy and management : the missing link?* Cheltenham, UK ; Edward Elgar. Book. isbn: 9781781955130.
- . 2018. "Performance management 40 years on: a review. Some key decisions and consequences." *Public Money and Management* 38 (3): 167–174.
- Rice, Deborah. 2019. "Adopting an institutional view in street-level bureaucracy research." In *Research Handbook on Street-Level Bureaucracy: The Ground Floor of Government in Context*, edited by Peter Hupe, 70–85. Edward Elgar Publishing.
- Schram, Sanford F. 2008. *Welfare Discipline Discourse, Governance and Globalization*. Welfare Discipline: Discourse, Governance and Globalization. Philadelphia: Temple University Press. Book. isbn: 1-281-09366-1.
- Schram, Sanford F., Joe Soss, Richard C. Fording, and Linda Houser. 2009. "Deciding to Discipline: Race, Choice, and Punishment at the Frontlines of Welfare Reform." *American Sociological Review* 74 (3): 398–422.
- Shadish, William R., Thomas D. Cook, Donald T. Campbell, and William R. Shadish. 2002. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin. Book. isbn: 9780395615560.
- Soss, Joe. 1999. "Lessons of Welfare: Policy Design, Political Learning, and Political Action." *American Political Science Review* 93:363–380.
- Soss, Joe, Richard Fording, and Sanford F. Schram. 2011a. "The Organization of Discipline: from Performance Management to Perversity and Punishment." *Journal of Public Administration Research and Theory* 21:203–232.
- Soss, Joe, Richard C. Fording, and Sanford Schram. 2011b. *Disciplining the poor : neoliberal paternalism and the persistent power of race*. 368. Chicago studies in American politics. Chicago: The University of Chicago Press. Book. isbn: 9780226768762 0226768767 9780226768779 0226768775.
- Stazyk, E. C., and H. T. Goerdel. 2011. "The Benefits of Bureaucracy: Public Managers' Perceptions of Political Support, Goal Ambiguity, and Organizational Effectiveness." *Journal of Public Administration Research and Theory* 21 (4): 645–672.
- Tummers, L. L. G., V. Bekkers, E. Vink, and M. Musheno. 2015. "Coping During Public Service Delivery: A Conceptualization and Systematic Review of the Literature." *Journal of Public Administration Research and Theory* 25 (4): 1099–1126.
- Tummers, Lars, and Philip Rocco. 2015. "Serving Clients When the Server Crashes: How Frontline Workers Cope with E-Government Challenges." *Public Administration Review* 75 (6): 817–827.
- Van Berkel, Rik, and Eva Knies. 2016. "Performance Management, Caseloads and the Frontline Provision of Social Services." *Social Policy Administration* 50 (1): 59–78.
- Van Loon, Nina Mari, and Mads Leth Jakobsen. 2018. "Connecting governance and the front lines: How work pressure and autonomy matter for coping in different performance regimes." *Public Administration* 96 (3): 435–451.

- Walker, Richard M., Oliver James, and Gene A. Brewer. 2017. "Replication, experiments and knowledge in public management research." *Public Management Review* 19 (9): 1221–1234.
- Wang, Weijie. 2022. "How does Performance Management affect Social Equity? Evidence from New York City Public Schools." *Public Administration Review*.
- Watkins-Hayes, C. 2011. "Race, Respect, and Red Tape: Inside the Black Box of Racially Representative Bureaucracies." *Journal of Public Administration Research and Theory* 21 (Supplement 2): i233–i251.
- Zacka, Bernardo. 2018. *When the State Meets the Street*. 7. edition. Cambridge, MA and London, England: Harvard University Press. Book.

## English Summary

Public service organizations deliver services to millions of citizens on a daily basis. Inevitably, service failures and dysfunctional patterns in service provision sometimes occur. Examples include misjudgements in service allocation decisions, poor service quality and instances of discrimination, where some clients are prioritized at the cost of others. We often explain such service failures by referring to the challenging working conditions of public employees, such as high demands for public services, scarce financial resources and high-stakes performance goals. In this narrative, dysfunctional reactions occur not because bureaucrats are mean or incompetent, but because they are pressured and constrained by challenging working conditions. This is a widespread narrative in the general public and even more so in the academic literature, where it serves as the foundation of the street-level bureaucracy literature. However, we have only limited empirical knowledge about how pressures influence service provision. In other words, it is uncertain whether this foundational narrative holds, and what the underlying causes of service failures are.

This dissertation seeks to remedy this by investigating how pressure from high workloads and performance management systems influences service delivery. Common to both types of pressure is that they create constraints on street-level organizations by shifting the balance between the goals to be achieved and the resources available to do so. The dissertation relies on quasi-experimental designs that leverage exogenous variation in both types of pressures in the context of public unemployment services and public primary schools in Denmark. Together with comprehensive micro-level administrative records on service delivery, this approach gives the means to study the research question in a real-world public service context.

The results demonstrate that public service organizations and public employees are capable of handling substantial pressures without turning to dysfunctional actions in service delivery. In fact, the dissertation provides examples of public employees who work under high workloads and extensive performance targets without resorting to discrimination, disciplining, gaming or goal displacement. As a consequence, these examples run counter to the widespread narrative that says both workload and performance management are believed to be the source of various evils in public service delivery.

The results, however, do not imply that pressure never matters for service delivery, but rather that we do not see the specific dysfunctional patterns that we would expect. For example, service providers handle high workloads by lowering the general service level instead of turning to discrimination. Furthermore, the studies of this dissertation also suggest that there are limits to public organizations' ability to withstand pressure. For example, comprehensive forms of high workload do affect service quality, and outcome-based accountability coupled with output-based accountability does lead to effort substitution. This indicates that we cannot increase pressure infinitely.

Moving one step beyond these findings, the dissertation also indicates that the specific public service organizations under study were equipped to handle the pressures because of well-

functioning organizational structures. Interviews with caseworkers and managers from the Danish unemployment services suggest that professional norms, administrative support systems and managerial buffering of the street level equipped the employees to handle the pressures. This suggests that we should focus more on how pressures play out within an organizational context than on the mere existence of the pressure in itself. Based on this, the dissertation proposes that the solution to dysfunctional service delivery might lie in constructing well-functioning organizational structures that give public service organizations the means to handle the pressures that will inevitably arise.

## Dansk Resumé

Offentlige serviceorganisationer og offentligt ansatte leverer dagligt serviceydelser til millioner af borgere. Nogle gange opstår der fejl og dysfunktionelle mønstre i leveringen af disse ydelser. Eksempler herpå er fejlbedømmelser i beslutninger om ydelsesallokering, ringe servicekvalitet og diskrimination, hvor nogle klienter prioriteres på bekostning af andre. Vi forklarer ofte sådanne svigt med henvisning til de udfordrende arbejdsforhold, som offentligt ansatte arbejder under, såsom knappe finansielle ressourcer, stor efterspørgsel efter service og risikofyldte resultatmål. Ifølge denne fortælling forekommer dysfunktionelle mønstre i serviceleveringen ikke, fordi offentligt ansatte er onde eller inkompetente, men fordi de er under pres og begrænset af udfordrende arbejdsforhold. Det er en udbredt fortælling i den generelle offentlighed og endnu mere i den akademiske litteratur, hvor den tjener som grundlag for den omfattende litteratur om markarbejdere, dvs. offentligt ansatte som har direkte kontakt til borgerne, og som råder over et stort skøn. Vi har dog kun begrænset empirisk viden om, hvordan pres faktisk påvirker leveringen af ydelser. Med andre ord er det usikkert, om denne grundlæggende fortælling holder, og hvad de underliggende årsager til fejl i levering af offentlige ydelser er.

Denne ph.d.-afhandling undersøger, hvordan pres fra en stor arbejdsbelastning og resultatbaserede styringssystemer påvirker leveringen af serviceydelser. Begge typer pres begrænser de offentlige serviceorganisationer ved at ændre balancen mellem de mål, der skal opnås, og de ressourcer, der er til rådighed for at opnå dem. Afhandlingen baserer sig på quasi-eksperimentelle designs, der udnytter eksogen variation i begge typer pres indenfor beskæftigelses- og skoleområdet i Danmark. Sammen med omfattende registerdata på mikroniveau og interviews gør denne tilgang det muligt at undersøge forskningsspørgsmålet i en virkelig servicekontekst.

Resultaterne viser, at offentlige serviceorganisationer og offentligt ansatte er i stand til at håndtere betydeligt pres uden at ty til dysfunktionelle handlinger i leveringen af serviceydelser. Afhandlingen indeholder eksempler på offentligt ansatte medarbejdere, der arbejder under høj arbejdsbelastning og omfattende resultatstyring uden at ty til diskrimination, disciplinering, manipulation eller målforskydning. Disse resultater går imod den udbredte fortælling, der fremhæver både arbejdsbyrde og resultatbaseret styring som årsager til omfattende og mangeartede problemer i offentlig servicelevering. Resultaterne indebærer dog ikke, at pres aldrig betyder noget for levering af offentlige services, men at vi ikke ser de forventede eksplicite dysfunktionelle mønstre. For eksempel håndterer serviceudbyderne høje arbejdsbyrder ved at sænke det generelle serviceniveau i stedet for at ty til diskrimination. Derudover tyder afhandlingens resultater også på, at der er grænser. For eksempel fører omfattende former for arbejdsbyrde til forringelse af servicekvaliteten, og resultatbaseret styring kombineret med aktivitetsbaseret styring fører til målforskydning. Dette indikerer, at vi ikke kan øge presset uendeligt.

Som et næste skridt peger afhandlingen på, at de offentlige serviceorganisationer, som analyserne baseres på, var rustet til at håndtere presset på grund af velfungerende organisatoriske strukturer, så som administrative sagsbehandlingssystemer, stærke faglige normer og en ledelse, der agerer som støddæmper og dermed beskytter frontlinjen, hvor serviceleveringen finder

sted. Det tyder på, at vi bør fokusere mere på, hvordan presset udspiller sig i en organisatorisk kontekst end blot på eksistensen af presset i sig selv. Baseret på dette foreslår afhandlingen, at løsningen på dysfunktionel servicelevering kan findes i opbygningen af velfungerende organisationer, der giver offentlige serviceorganisationer midlerne til at håndtere det pres, som uundgåeligt vil være der.