

UNDERSTANDING NEGATIVE PERFORMANCE MANAGEMENT
IN U.S. FEDERAL AGENCIES

by

JUNGYEON PARK

(Under the Direction of George A. Krause)

ABSTRACT

Although negative performance in public organizations is perceived as having relatively greater saliency and attention than positive performance, there has been little formal attention to negative performance management. By focusing on the negative performance of both individual employees and organizations, this study seeks to answer under what conditions government organizations can better manage their performance. Specifically, this dissertation is comprised of three empirical individual chapters that investigate the following questions: first, the underlying causes of differences in supervisors' perception of the effectiveness of the Performance Improvement Period (PIP); second, the role of administrative leaders in the magnitude of payment errors; and finally, the impact of administrative leader types on the direction of payment errors.

Several key findings of this dissertation include the following. In the first empirical chapter, the results demonstrate that the validity and usefulness of performance standards and organizational support have a direct positive impact on the perceived effectiveness of the PIP, as well as an indirect effect through the influence on the confidence in delivering performance feedback. The analysis also supports that supervisors' authority to remove poor performers also positively affects the perceived effectiveness of the PIP. For the second

empirical chapter, the findings suggest that programs administered by career executives or non-acting executives produce lower total payment errors than programs administered by political executives or acting executives. Finally, the third empirical chapter further investigates payment errors by separately analyzing overpayment and underpayment errors. The results indicate that career executives are better suited for managing both overpayment and underpayment errors which are potentially competing performance goals. These three empirical chapters provide helpful information about when supervisors are more or less willing to use the tool for managing poor-performing employees and under which administrative leader types better manage program payment errors.

INDEX WORDS: Public Administration, Public Management, Performance Management, Negative Performance, Government Performance, Performance Improvement Period, Improper Payments, Payment Errors

UNDERSTANDING NEGATIVE PERFORMANCE MANAGEMENT
IN U.S. FEDERAL AGENCIES

by

JUNGYEON PARK

B.A., Yonsei University, South Korea, 2015

M.P.P., Seoul National University, South Korea, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

© 2022

Jungyeon Park

All Rights Reserved

UNDERSTANDING NEGATIVE PERFORMANCE MANAGEMENT
IN U.S. FEDERAL AGENCIES

by

JUNGYEON PARK

Major Professor:	George A. Krause
Committee:	Bradley E. Wright
	Hal G. Rainey
	Donald P. Moynihan

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2022

ACKNOWLEDGEMENTS

This dissertation would not have been possible without continuous support and encouragement from my mentors, family, and friends. First and foremost, I would like to express my deepest appreciation to my mentor George Krause for his thoughtful guidance and unwavering support. From the moment he became my advisor, he showed me his professionalism and tremendous care. Whenever I needed his advice, he was willing to spare his time and answer my questions in his office, at Molly's, and at his home. A special thanks to LeeAnne for graciously welcoming me into your home. George, thank you for believing in me more than I believed in myself. I will remember his passion for research and education.

I also would like to extend my sincere thanks to the remaining members of my dissertation committee. I would like to express my gratitude to Bradley Wright for his thorough and thoughtful comments and advice. Besides this project, he enthusiastically helped me to develop another research project. He taught me how to do research and how to tell stories to the audience. I am deeply grateful to Hal Rainey. He has been a great supporter all these years, and his handwritten markup on my dissertation essays always helped me to think harder and develop my study. I enjoyed having conversations with him in his office, and his sense of humor often brightened my day. I also wish to extend my appreciation to Donald Moynihan. His great suggestions make me think about how my studies relate to the bigger theoretical picture. All committee members provided important contributions to the completion of this dissertation.

Along with my dissertation committee members, I also feel grateful to Dave Lewis for generously sharing his dataset for my dissertation. Without his generosity, this dissertation would have taken more time to collect the data and complete. I also would like to thank him for reading my studies and giving me thoughtful comments. I'm also appreciative of Ed Kellough for serving as my comprehensive exam committee member.

My thanks go to many of my colleagues in the Public Administration department at the University of Georgia. I want to thank my cohorts and especially my Korean colleagues for their support and warm friendship. It has been great to be surrounded by wonderful people during my academic journey.

I also need to thank all my friends in South Korea for their love and support. Though some say dissertation writing can be a lonely pursuit, I was fortunate enough to have incredible friends who have always cheered me up. I especially thank Minji Choi who has been with me every step of the way. Without her encouragement, this dissertation would not have been possible.

Lastly, no words can describe my sincere gratitude to my family. My parents have always believed in me and given me unconditional love and support my entire life. Thanks for always being by my side. I also would like to thank my brother, sister-in-law, and beautiful nephews and niece, Jouho, Dain, and Teho, for their untiring support. Their home in New Jersey has always been my happy place to visit during holidays. Lastly, I thank my grandma, who is now in heaven, for teaching me perseverance and tenacity. Though she did not get to see me complete this journey, I know she is proud of me.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION.....	1
2 THE PERFORMANCE IMPROVEMENT PERIOD (PIP): WHEN DO SUPERVISORS PERCEIVE THIS TOOL AS EFFECTIVE?	7
Introduction	7
Managing Poor-Performing Employees and The PIP.....	10
Hypotheses.....	14
Empirical Design and Strategy	21
Results and Discussion	27
Conclusion	31
3 HOW ADMINISTRATIVE LEADERS AFFECT ADMINISTRATIVE PERFORMANCE: EVIDENCE FROM IMPROPER PAYMENTS IN U.S. FEDERAL PROGRAMS	39
Introduction	39
Improper Payments of U.S. Federal Programs	42

Differences in Administrative Leader Types and Its Implications for Improper Payments of Federal Programs	44
Data, Measures, and Empirical Strategy	48
Statistical Findings	58
Discussion	65
4 DISENTANGLING IMPROPER PAYMENTS IN U.S. FEDERAL PROGRAMS: DISTINGUISHING BETWEEN TYPE I VERSUS TYPE II PAYMENT ERRORS	77
Introduction	77
Disentangling Improper Payments of U.S. Federal Programs: Overpayment Errors and Underpayment Errors	79
Data, Measures, and Empirical Strategy	90
Empirical Findings	103
Discussion	108
5 CONCLUSION	123
REFERENCES	128
APPENDICES	
A Chapter 2 Supplementary Appendix	149
A.1 Construction of Variables	149
A.2 Goodness of Fit Indices for Measurement Model and SEM	152
A.3 Direct Effects, Indirect Effects, and Total Effects of Study Variables ..	153
A.4 Sensitivity Analysis	155

B	Chapter 3 Supplementary Appendix	158
	B.1 Descriptive Statistics and Correlations	158
	B.2 Reduced Model Estimates	160
	B.3 Alternative Baseline Model: IP Rate (%) as Dependent Variable	161
	B.4 Alternative Augmented Model: IP Rate (%) as Dependent Variable	163
	B.5 Reduced Model for Alternative Dependent Variable, IP Rate (%)	165
	B.6 Model Fit Comparisons	166
	B.7 Subsample Analysis of Executive Bureaus (N*T=642)	167
	B.8 Subsample Analysis of Independent Agencies (N*T=259)	169
	B.9 Robustness Check for Analyzing the Moderating Effect of Policy Complexity	171
	B.10 Robustness Check for Analyzing the Moderating Effect of Executive Tenure	172
	B.11 Robustness Check Controlling Different Presidential Administrations	173
C	Chapter 4 Supplementary Appendix	174
	C.1 Descriptive Statistics of Variables in Analysis	174
	C.2 Bivariate Correlations of Variables in Analysis	175
	C.3 Alternative Censored Quantile Regression (CQR) Estimates for U.S. Federal Program Overpayment Errors	177
	C.4 Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors (Estimated with Bootstrapped Standard Errors)	178
	C.5 Alternative Regression Model Estimates for U.S. Federal Program Underpayment Errors (Estimated with Bootstrapped Standard Errors) ...	180

C.6 Average Differential Marginal Effects of Career Executives Versus Political Executives in Underpayment Errors and Overpayment Errors Using Alternative Regression Model with Bootstrapped Standard Errors	182
C.7 Average Differential Marginal Effects of Non-Acting Executives Versus Acting Executives in Underpayment Errors and Overpayment Errors Using Alternative Regression Model with Bootstrapped Standard Errors	183
C.8 Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables)	184
C.9 Alternative Regression Model Estimates for U.S. Federal Program Underpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables)	186

LIST OF TABLES

	Page
Table 2.1: Construction of Study Variables	34
Table 2.2: Descriptive Statistics, Correlations, and Reliabilities	35
Table 3.1: Baseline Model Estimates	70
Table 3.2: Augmented Model Estimates	72
Table 3.3: Average Differential Effects in Improper Payments by Administrative Leaders.	74
Table 4.1: Types of Decision Errors in Improper Payments	115
Table 4.2: Regression Model Estimates for U.S. Federal Program Overpayment Errors ..	116
Table 4.3: Regression Model Estimates for U.S. Federal Program Underpayment Errors	118
Table 4.4: Average Differential Marginal Effects in Overpayment and Underpayment Errors by Administrative Leader Types	120

LIST OF FIGURES

	Page
Figure 2.1: Conceptual Model	36
Figure 2.2: Confirmatory Factor Analysis of Study Variables	37
Figure 2.3: Structural Equation Model (SEM) Results	38
Figure 3.1: The Total Estimated Amounts of Improper Payments in U.S. Federal Government (\$ in Billions, Adjusted to 2015 Constant Dollars)	75
Figure 3.2: The Total Size of Improper Payments in Federal Programs (Per Reporting Year)	76
Figure 4.1: Average Differential Marginal Effects of Career Executives Versus Political Executives in Underpayment Errors and Overpayment Errors	121
Figure 4.2: Average Differential Marginal Effects of Non-Acting Executives Versus Acting Executives in Underpayment Errors and Overpayment Errors	122

CHAPTER 1

INTRODUCTION

There has been a consistent effort in government organizations to improve their performance. The historical origins of the government's effort to improve organizational performance can be dated back to the Progressive Era. The main argument of Progressive Era scholars was a scientific approach to administration in order to avoid waste and improve efficiency. By proposing a politics-administration dichotomy, Goodnow (1900) suggests studying bureaucracy in a scientific manner. Similarly, Taylor (1911)'s scientific management and Gulick (1937)'s POSDCROB (Planning, Organizing, Staffing, Directing, Coordinating, Reporting, and Budgeting) methodologies seek to present the principles or laws of management science that can be universally applied to any setting. Through the 1980s and 1990s, the rise of neoliberalism and neoclassical economics shifted the traditional perspective of public administration into a business-like and market-based perspective. This new perspective causes the movement of "New Public Management (NPM) (Hood 1991; 1995)." With the New Public Management (NPM) doctrine, the idea of 'performance management' has received more attention from practitioners and scholars. The major characteristic of NPM can be summarized as a focus on efficiency and results rather than progress (Goldfinch and Wallis 2010; Hood 1998; Osborne and Gaebler 1992). Along with the 'reinventing government' movement (Osborne and Gaebler 1992), NPM emphasizes business-like competition such as contracting-out and privatization. Since the late 20th century, governments have adopted performance management systems in various

functional areas, including personnel management, budgeting, and procurement and contracting (Abramson, Bruel, and Kamensky 2006). Throughout history, although there has been a consistent debate on how to make the government more efficient and perform well, there seems to be a clear consensus among scholars and practitioners that the government should be efficient and manage performance. In this sense, performance management in various functional areas has received continuous attention.

Until recently, many studies of performance management in the public sector predominantly have analyzed the determinants of positive performance outcomes relating to the attainment of organizational goals or citizen satisfaction (e.g., Brewer and Selden 2000; Lee and Whitford 2013; Meier and O'Toole 2002; Rainey and Steinbauer 1999; Van Ryzin 2006). However, there has been relatively little research and approach focusing on negative performance management. In the real world, government organizations and public officials get added pressure from many stakeholders to respond to negative performance. Specifically, the U.S. Office of Personnel Management (OPM) and the U.S. Merit Systems Protection Board (MSPB) seek to find adequate ways to address poor performers in government organizations and make suggestions to federal agencies, and congressional committees such as the house committee on oversight and reform consistently keep an eye on government operations. Congress also enacted the Inspector General Act of 1978 to ensure efficiency and integrity in government, and thus the office of inspector general within each federal agency serves as an oversight division to detect any mismanagement of agency operations. Still, we can easily find many instances where news media and the public criticize government performance. For example, it has been two years since Covid-19 was first declared a global pandemic, and citizens have been very critical of the government's response to this pandemic. According to the public opinion survey from Pew Research Center conducted in 2022, Americans' criticism of the government's response to

Covid-19 increased significantly (Tyson and Funk 2022). Besides managing unprecedented challenges like Covid-19, people expect the government to be competent in daily operations, but a number of unsuccessful stories criticizing government performance have appeared in the headlines of news articles (e.g., *“1.4 Billion in Stimulus Funds Sent to Dead People, Watchdog Finds”* in The New York Times on July 17, 2020; *“How a Social Security Program Piled Huge Fines on the Poor and Disabled”* in The Washington Post on May 20, 2022; *“Report: U.S. Government Wasted Up to \$1 Billion on Charter Schools and Still Fails to Adequately Monitor Grants”* in The Washington Post on March 25, 2019). From a growing literature on a negativity bias in the use of performance data in the public sector (e.g., George et al. 2020; Hong 2019; Hong and Kim 2019; Marvel 2016; Nielsen and Moynihan 2017; see also Weaver 1986), we also learn that negative performance results draw more undesired attention and criticism from news media, political decision-makers, and the public compared to positive or neutral ones. Despite the possibility that managing poor performance can be the first step to changing the public’s general perception of government and improving government performance, public administration and management scholarship has paid little attention to this aspect.

In this sense, this dissertation presents three empirical essays on negative performance management both at the individual level and the organizational level. Through these essays, this dissertation aims to provide insights for practitioners and policy-makers to better understand personnel management strategies for managing negative performance. Chapter 2, the first empirical essay of this dissertation, explores the Performance Improvement Period (PIP), a tool for dealing with poor-performing employees within government organizations. Since the PIP aims to provide poor-performing employees with an extra opportunity by giving them time and resources to improve their performance before agencies make decisions to demote or remove them, the PIP is considered as a

government managerial tool to have managerial flexibility in the personnel system while providing public employees' job protection at the same time. Although there are many managerial tools that agencies employ to improve employees' performance, such as mentoring, rewarding, empowering, or providing workplace flexibilities, the PIP is designed exclusively for managing poor-performing employees and thus provides a useful opportunity to advance our understanding of negative performance management. This essay raises a question – *When will supervisors perceive the PIP as effective and use this tool?* In this regard, this essay investigates organizational conditions that may affect supervisors' willingness to use this tool.

I expect that this chapter can advance our understanding of individual employees' poor performance management in several ways. First, despite the fact that supervisors are the ones who actually employ tools to deal with poor-performing employees, and their perceptions are more likely to depend on specific structural and operational aspects of the tools (Lin and Kellough 2019), there is little attention to supervisors' perception of performance tools. By focusing on supervisors' perception and evaluation of the PIP, this study can fill the gap in the public administration and public management scholarship. Also, by investigating when supervisors are willing to use performance information to improve individual performance, I expect that this study will contribute to the literature on performance information where most studies investigate the use of performance information at the organization or program levels.

While the first essay focuses on individual employees' negative performance and explores a management tool, the next two essays focus on government organizations' negative performance and investigate under which administrative conditions organizations improve their performance. In the next two empirical essays, I will examine payment errors

of U.S. federal programs, more specifically the magnitude (accuracy) of payment errors and the direction (bias) of payment errors, respectively. Given that improper payments in U.S. federal programs cause severe problems ranging from costing taxpayers significant amounts of money, to reducing public trust in government, and to hampering the fiscal sustainability of government, analyzing the underlying conditions of payment errors can provide some important implications for scholars and practitioners in public management. Exploring payment errors of federal programs also offers a crucial advantage in studying negative performance management for the following reasons. First, by looking at the size of payment errors, we can precisely understand the extent of negative performance. Many performance measures used in government organizations, including the Performance and Accountability Report (PAR) indicator, do not provide much information about how far agencies do not meet their goals or how far they exceed their goals. Another advantage of exploring payment errors is that the task of payment is relatively universal across government agencies, thus making it reasonable to investigate the variations in performance across agencies. In other words, while many performance goals can represent very different things for different agencies, this study can provide implications with more precision by looking at the universal and comparable task, the payment errors.

In this regard, the second empirical essay of the dissertation, chapter 3, examines why some federal programs yield more payment errors than others by focusing on the types of administrative leaders. A growing body of research has explored the relationship between the appointment type and bureaucratic performance (e.g., Gilmour and Lewis 2006; Lewis 2007; Miller 2015), and this study revisits the classic public administration question of whether careerists or political appointees better manage government programs. This study also explores the negative impact of leadership vacancies on government performance by examining how government program performance differs under the

leadership of acting managers. I expect that the findings of this study provide some implications for developing appointment strategies to manage negative performance.

While chapter 3 focuses on the magnitude of total payment errors of government programs, chapter 4, the third empirical essay of the dissertation, disaggregates payment errors into two different types and further explores how the types of administrative leaders affect the direction of payment errors. Improper payments, which represent incorrect amounts of government payments given to recipients, consist of overpayments and underpayments. While the size of total payment errors represents the level of accuracy in program payments, these two distinct types of payment errors imply potential tradeoffs. That is, efforts to reduce overpayment errors can increase underpayment errors either intentionally or unintentionally, or vice versa. As a theoretical discussion, this chapter borrows its insights from Type I and Type II decision-making errors to understand the relations of these two types of errors and investigates which types of administrative leaders are better suited to manage programs in these conditions. By distinguishing between overpayment and underpayment errors and explicitly considering potential tradeoffs between these two errors, I believe this study provides a crucial opportunity to investigate the relationship between the types of administrative leaders and competing measures of performance.

The final chapter summarizes the major findings of the three preceding empirical essays and describes their implications for theory and practice. Also, this chapter addresses some limitations of each essay and identifies the next steps for future research. Finally, this chapter discusses why understanding negative performance management is critical to the broader endeavor of performance management.

CHAPTER 2

THE PERFORMANCE IMPROVEMENT PERIOD (PIP): WHEN DO SUPERVISORS PERCEIVE THIS TOOL AS EFFECTIVE?

Introduction

In recent decades, public management reforms mainly focusing on managerial effectiveness and increasing organizational performance have been sweeping the globe (Brewer and Walker 2013; Kettl 1997; Pollitt and Bouckaert 2004). As a part of the movement, public administration scholars and practitioners have consistently drawn attention to human resource management, specifically how to deal with poor-performing employees in the workplace. Although the ratio of poor-performing employees to total federal employees may be small, even a small number of poor-performing employees can have a disproportionately large impact on the productivity of the workplace (U.S. Merit Systems Protection Board (MSPB) 2003). Through the survey, MSPB (2003) reveals that respondents in a work unit with a poor-performing employee gave twice as lower productivity ratings as employees in a work unit with no poor performers. Moreover, previous studies reveal that the way supervisors deal with poor-performing employees adversely impacts other employees' job satisfaction, motivation, and ultimately organizational performance (e.g., Choi and Whitford 2017; Ko and Hur 2014; Rubin 2009).

Performance Improvement Period (PIP), also known as the Performance Improvement Plan, has been adopted in the federal government as a means of dealing with

poor-performing employees.¹ The PIP refers to a formal opportunity given to poor performers to improve before a removal or demotion action can be taken based on unacceptable performance in a critical element (U.S. Office of Personnel Management (OPM) n.d.(b)). By providing poor-performing employees a certain period of time-set and assistance, the PIP is intended to give poor performers an additional chance to improve their performance to an acceptable level. However, as in the case of many other performance appraisal tools, it is very hard to achieve its intended goal, and thus, it is more likely to be just another personnel constraint considering that the PIP requires intensive resources and time. Or it will just be a tool to reduce the legal liabilities of organizations when conducting adverse action against a poor-performing employee. In the 2016 MPS survey, for the survey questionnaire asking whether they agree that the PIP often results in the employee improving to an acceptable level or better, more than 21% of supervisor respondents disagreed with the statement. Also, for the question asking whether they think that poor performers make a serious effort to use the PIP to improve their performance, more than 33% of supervisory respondents disagreed with the statement. As can be seen in these survey results, there are some variances in supervisors' perceptions of the PIP.

However, little is known regarding how organizational factor shapes the effectiveness of the PIP. To my best knowledge, there are very few articles investigating the PIP, and these articles often treat the PIP as one of many methods to deal with poor performers and neglect when and what conditions supervisors perceive this tool as effective.

¹ The terms Performance Improvement Period and Performance Improvement Plan are often used interchangeably in the government website (Office of Personnel Management and Merit System Protection Board) and the Merit Principles Survey. Hereafter, the acronym PIP will be used.

One case study by Robisch (1996) investigates the reason behind supervisors' reluctance to employ formal procedures for poorly performing employees and suggests the myriad of procedural requirements may impede the implementation of formal procedures, including the PIP. This study suggests that procedural complexity may negatively affect supervisors' use of formal tools to deal with poor performers.

In 2015, MSPB presented several elements of effective PIP, such as stating what performance is expected from the employee in a clear manner, specifying the assistance the agency will provide, designating a person responsible for providing assistance, stating the length of the PIP, and many others. Similarly, Sahoo and Mishra (2012), in their viewpoint article, recommend providing support and resources, having a plan for providing feedback, and stating a clear performance level for a successful PIP. These provide us with some guidelines for the effective implementation of PIP, but these elements have not been empirically tested to analyze whether these actually affect the effectiveness of the PIP.

This study will analyze the determinants for supervisors' perceived effectiveness of the PIP and aim to fill the gap in the extant studies by unraveling the underlying causes of differences in supervisors' perceptions of this issue. By employing the survey data of supervisors' actual use of the PIP, this study expects to provide more reliable evidence of its effectiveness. Also, this study contributes to the broader literature on performance information use. Most of the literature on the use of performance information looks at the program or organizational performance (e.g., Kroll 2015b; Moynihan and Ingraham 2004; see also Kroll 2015a). By studying the PIP, a management tool to improve individual performance, this study seeks to investigate organizational conditions that make supervisors willing to use performance information to improve individual performance. Lastly, this study offers practical lessons for managers interested in adopting the PIP to deal with poor performers. For the effective implementation of the PIP, this study

underscores the importance of providing supervisors with useful performance standards, sufficient organizational assistance, and the authority to exercise power when necessary.

This study proceeds as follows. First, it discusses the government’s consistent attention on managing poor-performing employees and how the PIP works. Next, it proposes hypotheses that encompass the validity and usefulness of performance standards, organizational support, confidence in delivering performance feedback, and authority to remove poor-performing employees. It then presents the empirical design and strategy used to test the hypotheses and demonstrates the empirical findings from the Merit Principles Survey (MPS) 2016 – Path L. This study ends with drawing major conclusions from the findings, including practical implications.

Managing Poor-Performing Employees and the PIP

The issue of managing poor performers in the public service has long been emphasized in civil service reforms and throughout history. Facing the criticism that poor-performing employees in the public sector have little risk of losing their jobs, efforts to strengthen firing and disciplinary procedures have been persistent in the federal government for decades from at least the Nixon administration (the so-called “Malek Manual”).² The Carter administration’s Civil Service Reform Act (CSRA) of 1978, which is considered the most comprehensive reform since the Pendleton Act of 1883, also emphasized the need to fire poor performers. The Carter administration advanced the

² “Malek Manual” is also known as an unofficial guide for Nixon political appointees about sidelining troublesome civil servants and installing political supporters within the bureaucracy (Moynihan 2004).

reform “as a means of motivating federal workers to work harder and of making it easier to fire lazy ones.” (Rainey, Fernandez, and Malatesta 2021: 102). The search for a useful means to address poor-performing employees has continued until recently, as can be seen in Executive Order 13839, “Promoting Accountability and Streamlining Removal Procedures Consistent with Merit System Principle,” which was issued in May 2018.

Despite consistent efforts to manage poor performers in the public service, there are still some concerns over this issue in the federal government. The large federal surveys that the Office of Personnel Management (OPM) and the Merit Systems Protection Board (MSPB) have conducted for years have often found a low percentage level of agreement on the survey item related to managing poor-performing employees within organizations. For example, in the 2020 Federal Employee Viewpoint Survey (FEVS) conducted by the OPM, the survey item that shows the lowest percentage level of agreement was “In my work unit, steps are taken to deal with a poor performer who cannot and will not improve” (42% level of agreement). Similarly, in the 2016 MPS, only 32% of survey respondents agreed with the same above survey item. As can be witnessed in these survey results and the history of civil service reforms, managing poor-performing employees has been a long-standing and serious problem in the Federal workforce. This issue is more complicated in the public sector since there are complex legal requirements for due process for government employees who have the property right in the job. The Fifth and Fourteenth Amendments to the U.S. Constitution ensure that governments cannot deprive any person of “life, liberty, or property without due process of law,” and this requires at least a prior notice and an opportunity for government employees to respond to charges before discharge occurs in termination proceedings (Kellough and Nigro 2006; see also *Cleveland v. Loudermill*, 470 U.S. 532 (1985)).

The Performance Improvement Plan (PIP) was created under the 5 U.S.C. § 4303 in the Civil Service Reform Act (CSRA) of 1978 as a means of dealing with poor-performing employees. This procedure is based on progressive discipline, which gives the employee opportunities to correct his or her behavior while providing the employer with justification for terminating employment if no change occurs in the behavior (Gomez-Mejia, Balkin, and Cardy 2009).

If an employee's job performance is considered unacceptable under a critical job element, supervisors are required to address this issue with the employee. There are two formal procedures supervisors can choose in addressing the unacceptable performance of their employees: Chapter 43 and Chapter 75 of Title 5 of the U.S. Code (U.S. Merit Systems Protection Board (MSPB) 2009).³ While placing the employee on a PIP is required under Chapter 43, it is not required under Chapter 73.⁴ Usually, supervisors work closely with the

³ Chapter 43 and Chapter 75 are the legal procedures in the U.S. Code Title 5, "Government Organization and Employees." Under Chapter 43, the use of PIP offers a lower burden of proof (*substantial evidence*) to supervisors, while performance-based action under Chapter 75 must be supported by *a preponderance of the evidence*. It is shown that federal agencies have different preferences regarding the extent to which Chapter is more often used (MSPB 2009). MSPB (2009) suggests that the nature of the employee's work makes the most difference in which authority is used. To be specific, the use of PIP under Chapter 43 is less appealing for supervisors whose employees perform repetitive daily tasks (mostly blue-collar workers), which require daily-basis supervision when implementing the PIP.

⁴ Except for the requirement for a PIP, most procedures for performance-based actions under Chapter 43 and Chapter 75 are the same or similar (MSPB 2009). More detailed information about the similarities and differences between Chapter 43 and Chapter 75 can be found in MSPB (2009) report.

Human Resources Office Employee Relations staff to determine the appropriate procedure and whether to provide the PIP with poor-performing employees (OPM n.d. (c)).

Once a supervisor decides to provide the PIP, the supervisor should work with the Employee Relations staff to write the draft of the PIP, and the PIP should include the following aspects: “(1) The critical job element(s) being performed unacceptably; (2) Examples of the employee’s unacceptable performance under each critical element (not required but recommended); (3) An explanation of the minimally acceptable level of performance required; (4) An explanation of what the employee must do to demonstrate acceptable performance; (5) Specific tools and support to be provided to help the employee improve during this period as required under 5 CFR 432.104; (6) How long the PIP will remain in effect; and (7) Consequences of failure to improve performance to an acceptable level” (OPM n.d. (c): 5).

During the PIP period, the supervisor has a responsibility to document assignments and assistance provided to the employee and to monitor the employee’s performance to determine whether it reaches an acceptable level. If the employees’ performance does not improve to an acceptable level until the end of the PIP, the supervisor may take the following options: reassigning the employee or proposing the employee’s demotion or removal. Before the final decision of demotion or removal is made, multiple steps should be followed, such as providing the prior written notice of the proposed action, providing an opportunity for the employee to be represented by an attorney, and giving him or her a chance to answer orally and in writing (5 U.S.C. § 4303 (b)). In this section, this study seeks

(<https://www.mspb.gov/mspbsearch/viewdocs.aspx?docnumber=445841&version=446988&application=ACROBAT>)

to delineate the U.S. government's persistent effort to deal with poor-performing employees and the process of the PIP. Next, the main hypotheses of this study will be presented.

Hypotheses

Validity and Usefulness of the Performance Standards

Performance standards, formal procedures that supervisors employ to assess the performance of subordinate employees, are an essential management tool (Lin and Kellough 2019). Supervisors use this tool in almost every decision regarding personnel management issues, including promotion, reassignment, demotion, and removal. In the case of PIP, when determining whether employees are performing poorly in one or more critical elements in their positions, the identification of performance criteria is required. Not just before putting employees into a formal PIP, performance standards are essential to evaluate poor-performing employees' improvement after the completion of the PIP and to determine the final action for poor-performing employees. Supervisors are supposed to let subordinates know the elements and standards that are required to meet when sending the notification document of the PIP (OPM 2017). The fact that the PIP heavily relies on performance standards can make it challenging for supervisors to conduct the PIP (MSPB 2009). Previous studies have suggested that only when procedures or standards are accepted by assessors (supervisors) and assesseees (subordinate employees), performance standards can be considered as a useful tool (Iqbal, Akbar, and Budhwar 2015; Levy and Williams 2004). However, developing performance standards that everyone can agree on is very difficult. Moreover, perfectly objective performance criteria are impossible to identify for most jobs, and for this reason, performance appraisals almost always rest on the subjective decisions of supervisors (Bell, Cantarelli, and Belardinelli 2017; Lin and Kellough 2019). This inherently subjective nature of performance standards makes the

performance appraisal system unreliable and negatively affects employees' perceptions of the appraisal system and organizational management (Iqbal, Akbar, and Budhwar 2015). Just as importantly, supervisors are in a difficult situation where they have no choice but to make various decisions based on their employees' performance, and supervisors are the ones who have the responsibility and the burden to provide evidence of whether their subordinates are performing well or poorly. From the standpoint of supervisors, if they consider the performance standards as invalid and do not think these standards fairly reflect their subordinates' performance, it is hard to expect that supervisors will think the PIP is effective in improving poor-performing employees' performance. Therefore, this study expects that validity and usefulness of the performance appraisal system is a necessary condition for the effective use of the PIP.

In addition to its direct effect on the perceived effectiveness of the PIP, this study expects that supervisors' perceived validity and usefulness of performance standards have indirect effects through their influence on their confidence in delivering performance feedback to subordinates. Performance feedback can be defined as information provided to individuals about past performance-related behaviors with respect to established standards of behavior (Cooper, Heron, and Heward 2007). Given that performance feedback is generated by using the data based on performance standards, performance standards are crucial for supervisors to use. Providing performance feedback can be understood as providing information about individual employees' current status to achieve goals. According to the goal-setting theory of motivation, goal specificity can improve performance since it can serve to affect the direction of attention and effort (Locke and Latham 1990; Steers and Porter 1974). Specific and clear goals for individuals' jobs can reduce search behaviors and help individuals to understand the relationship between performance and subsequent consequences. It is particularly important in the case of the PIP since the

decision on poor-performing employees' removal heavily depends on the performance improvement appraisal. Although having clear and specific standards for individual performance is very difficult, it will ease the supervisors' difficulty in delivering the performance feedback. On the other hand, supervisors who have a lower level of validity in performance standards are more likely to have difficulty delivering objective performance feedback. The following hypotheses test these linkages:

H1a: Supervisors' perceived validity and usefulness of performance standards have a direct positive effect on their perceived effectiveness of the PIP.

H1b: Supervisors' perceived validity and usefulness of performance standards have an indirect positive effect on their perceived effectiveness of the PIP through its influence on their confidence in delivering performance feedback to subordinates.

Organizational Support

Extant studies have suggested that organizational support is an important factor for implementing a program and generating a good outcome (Julien, Somerville and Culp 2011; Ko, Hur and Smith-Walter 2013; Lee et al. 2010). Organizational support for a management practice helps legitimize the practice and provides an impetus for the improvement effort (Chong, White, and Prybutok 2001). Organizational support also can be viewed as employees' resources in the workplace of an organization since organizational support can take various forms, such as providing necessary assistance, training, resources, and informal communication. In the case of PIP, assistance from the human resources department is especially essential for implementing the program since their assistance can provide a big help throughout the process of the PIP. From writing the draft of the performance improvement plan, supervisors work with the Employee Relations staff (OPM n.d. (c)). Since performance-based actions often do not run smoothly, getting technical

assistance or working together with the human resource department may facilitate the process when some issues occur (OPM n.d. (b)). To be specific, the Human Resources Office can help the supervisor with the specific regulatory requirements of the process, collective bargaining agreements, and when considering available options. Timely training and education on the PIP program for supervisors can also be considered as support from the organizational level, and this will help to implement the program in a meaningful way. In addition to the support from the human resources department, training can help supervisors to improve communication skills when discussing performance with poor performers. In addition, training can send the signal to the supervisors that the organizations support providing the PIP to poor-performing employees. Therefore, this study presumes that there is a positive relationship between perceived organizational support and the perceived effectiveness of the PIP.

Besides hypothesizing the direct effect of supervisors' perceived organizational support on their perceived effectiveness of the PIP, this study expects that part of this effect is caused by affecting supervisors' confidence in delivering performance feedback, which can be understood as a partial mediation effect. The PIP can also be seen as the supervisors' opportunity to explicitly express their expectations to their subordinates, as well as the poor-performing employees' opportunity to improve their performance (OPM n.d.(a)). Supervisors who get organizational support on implementing the PIP are more likely to have an opportunity to learn communication skills either through assistance from the human resource team within the organization or through training sessions, which will help to deliver performance feedback. In addition, the organization's support signal in implementing the PIP will help to deliver negative feedback. Providing unfavorable feedback is equally difficult as receiving unfavorable feedback since unfavorable feedback evokes defensiveness or denial in feedback recipients (Steelman and Rutkowski 2004).

Larson (1984) also suggests that supervisors' reluctance to give negative feedback to subordinates is due to the possibility of negative interpersonal repercussions. Therefore, supervisors are less likely to suffer difficulty in delivering performance feedback when they have a high level of organizational support. The following hypotheses test these linkages:

H2a: Supervisors' perceived organizational support has a direct positive effect on their perceived effectiveness of the PIP.

H2b: Supervisors' organizational support has an indirect positive effect on their perceived effectiveness of the PIP through its influence on their confidence in delivering performance feedback to subordinates.

Confidence in Delivering Performance Feedback

In order to improve performance during the PIP and to avoid removal from the workplace, it is crucial for poor-performing employees to understand their shortcomings and how to improve their performance effectively. One of the most effective sources to get this information can be direct supervisors, who are the assessors after the completion of the PIP. Previous studies in the field of organizational management have demonstrated the crucial role of feedback as an instrument to enhance the performance of both individuals and organizations (e.g., Favero, Meier, and O'Toole 2016; Garnett, Marlowe, and Pandey 2008; Ilgen, Fisher, and Taylor 1979; Podsakoff and Farh 1989). Performance feedback can positively influence individual and organizational performance in two ways. First and most importantly, it plays a substantial role in behavior regulation in which feedback "helps employees to keep work-related activities directed toward desired goals" (Larson 1989:408). Related to the goal-setting theory (Locke and Latham 1990), this is more likely when the feedback is clear and specific since it will help individuals to reduce search behaviors on

what they need to do to improve their performance. Another vital role of feedback from supervisors is to build rapport between supervisors and subordinates, increasing subordinates' acceptance and commitment to their performance goals (Campion and Lord 1982; Favero, Meier, and O'Toole 2016). However, unfavorable feedback usually creates an uncomfortable situation between supervisors and subordinates. Studies have found that although negative feedback results in negative attitudes of employees, this relationship can be moderated by contextual characteristics such as feedback quality or feedback delivery (e.g., Steelman and Rutkowski 2004). This implies that the way of delivering feedback and the quality of feedback may affect how employees accept the feedback. As a managerial effort, delivering helpful feedback will signal that poor-performing employees are taken seriously in the organization. Ultimately, this effort can increase the possibility of poor performers' performance improvement. Accordingly, supervisors will be more likely to have a favorable perception of the PIP when they are more confident in communicating with subordinates about their shortcomings and how to deal with them. Therefore, I posit the following hypothesis:

H3: Supervisors' confidence in delivering performance feedback to subordinates has a positive effect on their perceived effectiveness of the PIP.

Authority to remove poor-performing employees

If employees fail to meet the requirement after the PIP, federal agencies are authorized to demote or remove employees for the reason of "unacceptable performance" (5 U.S.C. § 4303 (a)). The PIP guideline also clearly indicates that employees may be demoted or removed if they fail to improve their performance during the PIP. Although supervisors are formally entitled to conduct an adverse action for poor-performing employees, if supervisors have no actual authority to take a specific action to deal with poor-performing

subordinates, such as removing or demoting them from the workplace even after the completion of the PIP, supervisors are more likely to perceive the PIP merely as a useless management tool. Not just formal constraints, public managers are often constrained by informal constraints such as organizational culture (i.e., shared informal rules regarding the proper behaviors for managers) and the lack of management ability (Ban 1995; Kellough 1998). OPM (n.d. (a): 4) also explicitly states that although government-wide regulations allow supervisors to choose options (reassignment, demotion, or removal), each agency may have “some internal rules about considering reassignment before the other choices.”

Granted, when employees substantially improve their performance during this provided period, direct supervisors’ actual ability to remove poor performers may not be a significant issue. However, when poor-performing employees do not improve their performance or even show worse performance during the PIP, failing to remove these employees can deteriorate organizational performance and send a bad signal to other employees in the organization in the long run. Previous studies have also suggested that the difficulty or inability to remove poor-performing employees may harm organizations’ performance since it will not “let managers manage,” create delays in management, and demotivate other employees in the workplace (Brewer and Walker 2013; Kettl 1997). According to the U.S. Government Accountability Office (GAO) (2015), it usually takes 50-110 days to complete the steps associated with the PIP. That is, if supervisors do not have the ability to remove poor performers even after spending 2-4 months of the PIP period, it is a waste of managers’ valuable time and organizational resources.

Expectancy theory (Vroom 1964) also provides the underlying rationale for why supervisors need real authority to conduct an adverse action for poor performers in order to consider the PIP as effective. According to expectancy theory, an individual’s motivation to

behave in a specific way is a function of the expected consequences of that behavior (Vroom 1964). Put simply, people are more likely to behave in a certain way when they believe that behavior will lead to the intended performance and eventually bring the desired results. In the case of the PIP, supervisors will be more likely to consider the PIP as effective if there is a higher probability of effective outcomes, which will be a demotion or removal of an employee showing still-inadequate performance. Therefore, if supervisors are not confident that they have the actual capacity to demote or remove poor-performing employees, it is hard to expect that they will perceive PIP as effective. **Figure 2.1** below provides a conceptual model explaining the relationship between study variables.

H4: Supervisors' perceived authority to remove poor performers has a positive effect on their perceived effectiveness of the PIP.

Empirical Design and Strategy

Data

The data used in this study are taken from a large-N survey, Merit Principles Survey (MPS) – Path L, that the Merit System Protection Board (MSPB) conducted in 2016. The MPS has periodically assessed merit system principles and government-wide civil service issues, such as prohibited personnel practices, employee engagement, and performance management practices, among many others (MSPB 2016). 2016 MPS – Path L provides a unique and valuable opportunity to analyze the use and the effectiveness of the PIP in the eye of supervisors since this survey was only directed to people in a supervisory status in 24 federal agencies. Also, in the survey, supervisors who had actually provided the PIP to their poor-performing employees could only respond to the survey questions related to the PIP. Therefore, it is expected that supervisors' actual experience in PIP may provide more reliable evidence of its effectiveness. However, since the analyses of this study are

solely based on the data of supervisors who actually used the PIP before, there may be a potential concern about selection bias. This study conducts a sensitivity test to see if there are any systematic differences between supervisors who used the PIP and those who did not use the PIP before, and the result suggests that there is no systematic difference in responding to key survey items between these two groups.⁵ The survey was conducted electronically to stratified random samples of permanent, full-time civilian federal employees in 24 agencies and attained a response rate of 40%. By only employing supervisor respondents who had ever used the PIP in the last ten years, the main analysis reported here uses 2,939 observations.

Measures

The perceived effectiveness of the PIP, which is the dependent variable of this study, is defined as the perceived level of effectiveness of the performance improvement plan (PIP). Considering that the ultimate goal of the PIP is to improve poor-performing employees' performance by providing a formal opportunity, the effectiveness of the PIP can be understood as whether a poor-performing employee's performance has improved to an acceptable level or not. The variable is measured using two survey items, asking individual

⁵ Among key covariates, survey items related to the validity and usefulness of performance standards were asked to both groups (supervisors who used the PIP before and supervisors who did not use the PIP before). The differences in means test reveals that supervisors who have no experience in the PIP (Group 2) perceive performance standards as slightly less valid and useful than supervisors who used the PIP before (Group 1) but misses statistical significance (Mean Group 1: 7.93 vs Mean Group 2: 8.21, $p=0.1856$), thus suggesting that there is no meaningful systematic difference between two groups.

supervisors the level of agreement with the following statements: (1) Poor performers make a serious effort to use the PIP to improve their performance; and (2) The PIP often results in the employee improving to an acceptable level or better. The response scale of both survey items is a five-point Likert-style scale.

One of the key independent variables, *validity and usefulness of performance standards*, is defined as the level of each supervisor's perceived validity and usefulness of the performance standards. This study uses the following two survey items to create this latent variable: (1) The performance standards for my subordinates accurately reflect the critical elements necessary for success in their jobs; and (2) The performance standards for my subordinates have measurements that I can use accurately to determine success in the critical elements. Both survey items are measured with a five-point Likert-style scale.

Perceived organizational support is defined as the extent of perceived support getting from the organization in using the PIP. The variable is measured using two survey items (five-point Likert-style scale), asking individual supervisors the level of agreement with the following statements: (1) Human resources staff provide competent assistance regarding PIPs; and (2) Supervisors receive timely training to help them effectively implement a meaningful PIP.

The binary variable, *authority to remove poor-performing employees*, is defined as the level of supervisors' certainty in their ability to remove subordinates who show poor performance even after completing the PIP. It is measured by using a single survey item asking respondents whether they are confident that they would be able to remove that employee if a subordinate employee was deficient in a critical performance element after completion of a PIP. This variable is coded 1 when survey respondents answered 'yes' to the question, and 0 when they answered 'no' or 'not sure.'

The variable, *confidence in delivering performance feedback*, is defined as the level of each supervisor's confidence in communicating with subordinates regarding their performance. It is measured by using a single survey item asking respondents how challenging each respondent finds the following statement: Communicating with employees about how they can improve their performance. Reverse scoring was conducted so that lower values indicate greater difficulty in delivering performance feedback to subordinates. This item is measured with a seven-point Likert-style scale. **Table 2.1** presents the survey items of the study variables constructed for this study.

In order to examine whether the effect of key study variables would hold when taking into other potentially confounding variables, this study includes several control variables.⁶ Although the 2016 MPS-Path L was only directed to each federal agency's supervisors, supervisors have different supervisory statuses (e.g., team leader, supervisor, manager, executive). Previous research has shown that differences in managerial levels affect organizational performance perceptions (e.g., Choi and Rainey 2010; Pitts 2009). Therefore, this study controls for the impacts of supervisory status on the dependent variable. In the survey, the supervisory status variable is coded 1 if a respondent is a team leader, 2 for a supervisor, 3 for a manager, and 4 for an executive. With this survey item, this study creates separate dummies and employs 'executive' as a reference group. Also, this study controls for the ratio of unsatisfactory subordinates whom each supervisor directly supervises. For supervisors, the larger ratio of unsatisfactory subordinates can be a bigger burden of supervisory responsibility in managing individual subordinates'

⁶ This study acknowledges that many important control variables such as demographic variables of supervisors are missing due to data availability. This issue is also discussed in the DISCUSSION section.

performance, leading to less willingness to put their poor performers into the PIP. This control variable is measured by using a survey item asking the approximate percentage of the employees that respondents currently supervise who are less than fully successful without outright failing at any critical elements of the job. It is coded 1 when a respondent answered 0 (no employees that are less than successful without outright failing), 2 for 1-10%, 3 for 11-20%, 4 for 21-30%, and 5 for over 30%. Lastly, this study anticipates that unobserved agency characteristics can account for some variations in the dependent variable. For instance, organizational stance on poor performers can vary by agency. While some agencies may view poor performers as potential assets for development and treat them seriously to improve their performance, other agencies may instead view poor performers as free riders who should be removed or demoted from the workplace immediately. Therefore, agency-fixed effects are included to account as much as possible for underlying heterogeneity arising from agency characteristics, using 23 agency dummies with the Department of Veterans Affairs as the reference category. The **Appendices** section provides detailed information on the survey items and coding scales of all the variables included in this study (see A.1 Construction of Variables).

Table 2.2 below presents descriptive statistics, a correlation matrix, and scale reliabilities for all variables analyzed in the statistical model. Reliability estimates (the Cronbach's alphas) ranged from 0.71 to 0.86, higher than the suggested critical value of 0.70, supporting the use of all three latent variables (Nunnally and Bernstein 1994). Also, the correlation matrix of the variables shows evidence that the measures have discriminant validity. The range of bivariate correlations is from .02 to .30, and the largest bivariate correlation (between perceived effectiveness of PIP and organizational support) indicates that multicollinearity does not seem to be a significant problem in our model.

Several confirmatory factor analyses are conducted to evaluate both the discriminant and convergent validity of the measures. This research first performs a confirmatory factor analysis of a single-factor model to examine whether the single factor can explain the data well. The model fit of the single-factor model is then compared with that of the three-factor hypothesized model.⁷ According to model fit statistics, the results show that the data fit significantly better with the three-factor hypothesized model than the single-factor model. This study investigates the model fit by comparing the values of the standardized root mean square (SRMR) and the coefficient of determination (CD), the only goodness-of-fit parameters available in STATA software when using cluster-robust standard errors (StataCorp 2019).⁸ While the single-factor model does not have acceptable model fit scores (SRMR = 0.175 and CD = 0.865), the three-factor model achieves an acceptable model fit (SRMR = 0.004 and CD = 0.990). The value of SRMR for the three-factor model falls far below the cutoff value of 0.06 (Hu and Benlter 1999), which suggests that the residuals between observed and model-implied correlation are small. Thus, it implies that the measurement model is well-specified. Also, the higher value of CD for the hypothesized model indicates that the hypothesized three-factor model has a larger proportion of explained variation to total variation, suggesting a better model fit. In sum,

⁷ Among five study variables, three variables (Effectiveness of PIP, Performance Standards, Organizational Support) are measured with multiple survey items. By comparing the one-factor model and three-factor model, this study investigates which model fits the data better.

⁸ This study ran the model without cluster-robust standard errors for the sole purpose of checking other model fit statistics, and the results consistently show that the data fit better with the hypothesized model. Detailed information on other model fit statistics is presented in **Appendices A.2 Goodness of Fit Indices for Measurement Model and SEM.**

the model fit parameters indicate that the study items tap multiple dimensions instead of a single dimension, and the hypothesized measurement model provides a good fit to the data.

For convergent validity of the measures, Kline (2011: 116) posits that all indicators to measure latent variables should “have relatively high standardized factor loadings on that factor” and suggests 0.70 as the critical value. **Figure 2.2** presents the result of confirmatory factor analysis of the hypothesized model, showing that high proportions of variance in survey items, between 70.5% and 91.8%, are accounted for by the theoretically hypothesized constructs. Thus, the result provides support for the convergent validity of the measures. Also, as shown in **Figure 2.2**, the range of estimated correlations between latent variables—validity and usefulness of performance standards, organizational support, and the perceived effectiveness of the PIP—is from 0.21 to 0.41, providing evidence that the sets of measures are not highly correlated and discriminated from each other. The average variance extracted (AVE) scores for validity and usefulness of performance standards, organizational support, and the perceived effectiveness of the PIP are 0.549, 0.762, and 0.605, respectively. Since the square root of AVE for each construct was greater than the correlation between the measures, it indicates the measures have discriminant validity (Fornell and Larcker 1981). The composite reliability (CR) of latent variables ranges from 0.708 to 0.865, indicating adequate internal consistency. A series of diagnostic checks demonstrate that the measurement model provides strong evidence for construct validity.

Results and Discussion

After confirming that the hypothesized measurement model fits the data well, this study tests a structural equation model (SEM) to evaluate both the direct and indirect hypotheses advanced in this study. A way to analyze mediation effects is to use SEM since it allows for investigating the direct and indirect effects of independent variables on the

dependent variable in the same model. By jointly estimating both a measurement model and a structural model, SEM also makes it possible to explicitly assess the measurement error and take the measurement error into account when estimating latent variables via observed variables (Byrne 2013; Kline 2011). The overall model fit of the hypothesized structural model reveals that it is a good representation of the underlying theoretical structural model. According to model fit parameters, the results show that the structural model provides a good fit to the data (SRMR = 0.026, CD = 0.973).⁹ It shows that the residuals between observed and model-implied correlations in the SEM model are quite small, and the model explains a considerable amount of variation.

Figure 2.3 above presents the standardized parameter estimates for the structural equation model. The standard errors are clustered at the agency level (24 clusters) to control for intra-agency error correlation, considering that supervisors' responses within each agency may not be independent. The SEM results confirm that the key study variables, validity and usefulness of performance standards, organizational support, and authority to remove poor performers, have direct positive effects on the perceived effectiveness of the PIP (standardized $\beta = .054$, $p < .05$, standardized $\beta = .358$, $p < .001$, standardized $\beta = .112$, $p < .001$, respectively), supporting Hypotheses H1a, H2a, and H4. It is notable that the organizational support for the PIP has much larger leverage on the dependent variable, the perceived effectiveness of the PIP. The results also indicate that

⁹ As discussed earlier, SRMR and CD are the only available model fit indices in the STATA program when using cluster-robust standard errors. This study separately ran the model without using robust standard errors for the sole purpose of checking other model fit statistics, and the result shows a good model fit. Detailed information can be found in the **Appendices** section (A.2 Goodness of Fit Indices for Measurement Model and SEM).

supervisors' confidence in delivering performance feedback positively affects their perceived effectiveness of the PIP (standardized $\beta = .101$, $p < .001$), supporting Hypothesis 3. In addition to the direct effects, this study hypothesizes that trust in performance standards and organizational support have indirect positive effects on the perceived effectiveness of the PIP through their influence on the confidence in delivering performance feedback (Hypotheses 1b and 2b). According to Kline (2011), indirect effects are estimated as the product of the coefficients for the constituent paths. As expected, the results indicate that these factors contribute to improving the PIP's perceived effectiveness in an indirect way as well. The indirect effect of validity and usefulness of performance standards is .015 ($.164 \times .101$), and that of organizational support is .003 ($.047 \times .101$). All direct, indirect, and total effects of the study variables appear in the **Appendices** section (A.3 Direct Effects, Indirect Effects, and Total Effects of Study Variables).

The analysis also includes controls for supervisory status and the ratio of unsatisfactory subordinates. Effects of the control variables show that supervisors who have a smaller percentage of unsatisfactory subordinates are more likely to have a favorable perception of the PIP's effectiveness. Also, the results show that different supervisory status affects the perception of PIP effectiveness. Compared to people in the executive position, supervisors and managers show a favorable perception of the PIP (see **Figure 2.3**).

Even after controlling for likely alternative explanations, the findings of this study confirm the importance of validity and usefulness of performance standards and organizational support in understanding the perceived effectiveness of the PIP. These two variables not only have direct positive effects but also have indirect positive effects on the perceived effectiveness of the PIP through their influence on the confidence in delivering performance feedback. Additionally, the analysis supports that supervisors' certainty in

their authority to remove poor performers and confidence in delivering performance feedback have positive relationships with the perceived effectiveness of the PIP. When taken together, these variables explain almost one-fourth of the variance in the perceived effectiveness of the PIP.

In addition, the results of the main analysis are evaluated for robustness by conducting a sensitivity analysis by excluding respondents who answered 'not sure' to the question regarding the certainty in the capacity of removing poor performers. By excluding these respondents, the sample size reduces from 2,939 to 2,494. This study evaluates whether the reported coefficient estimates are sensitive to the exclusion of a subset of respondents. While this study expects the hypothesized relationships to run in the predicted direction for this additional analysis, the strength of these relationships may be different.

The results of the sensitivity analysis are quite consistent with the main model (see details in **Appendices A.4 Sensitivity Analysis**). The results show that the path analyses for expected relationships are all in the predicted direction, but some changes can be found in the significance level. While the direct effect of organizational support on the confidence in delivering performance feedback shows a change in the significance level from .05 to .001, the direct effect of validity and usefulness of performance standards on the perceived effectiveness of the PIP falls short of statistical significance. In sum, the overall results of sensitivity analysis provide additional support for the theoretical model. The following section will discuss the implications of these results for theory building and management practice.

Conclusion

“Employees should be retained on the basis of the adequacy of their performance, inadequate performance should be corrected, and employees should be separated who cannot or will not improve their performance to meet required standards.”

Sixth Merit System Principles, 5 U.S.C. 2301(b)(6).

In the above phrase, we find the government’s aspiration to keep competent and talented employees in the workplace. However, at the same time, the government has an equally powerful tradition of maintaining regulations that make it difficult to remove or demote poor performers (Bozeman and Rainey 2000; Feeney and Rainey 2010; Rainey, Facer, and Bozeman 1995; Rainey, Pandey, and Bozeman 1995). In this sense, the PIP is a managerial tool that shows the government’s aspiration to have managerial flexibility in the personnel system while protecting public employees’ jobs at the same time. If implemented as intended, the PIP can be a win-win tool for both the agency and the poor-performing employees since it can preserve the investments that both parties have already made (GAO 2015). If not, however, the PIP will turn out to be another red tape to remove poor performers and a waste of government resources. This study seeks to provide an introduction to unravel the underlying causes of different perceptions of its effectiveness.

The statistical evidence compiled in this study corroborates the role of trust in performance standards and organizational support in influencing the perceived effectiveness of the PIP. The results show that supervisors who perceive a higher level of validity and usefulness of performance standards or a higher level of organizational support are more likely to have a favorable perception of the PIP’s effectiveness. The results also indicate that both have indirect effects on the perceived effectiveness of the PIP through its influence on the confidence in delivering performance feedback. Lastly, it is shown that the

supervisors' authority to remove poor performers has a positive relationship with their perceived effectiveness of the PIP. The findings demonstrate that organizational support has the strongest statistically significant positive association with the perceived effectiveness of the PIP, followed by the authority to remove poor performers.

Some important limitations to this study should be addressed. First of all, this study was unable to include potentially important control variables such as the job tenure or demographic information of supervisors or the job types of poor-performing employees. The reliance on single or two-item measures can be problematic, leaving the possibility of missing important aspects of constructs. A common method bias problem may exist since data on study variables were collected through the same method. If separate data on the actual outcomes, such as the percentage of employees who stayed or were removed after the PIP, were possible, this study would have more reliable results. However, by employing structural equation modeling (SEM), this study directly models measurement error and allows for confirmation of the construct validity of measures before these measures are incorporated into the variance-based model. To be specific, this study conducted confirmatory factor analyses (CFA) to examine whether all used items can be explained by one single factor, and the result showed a very poor fit indicating that common method bias was not a severe problem. While SEM has been recommended as a possible approach to deal with the common method bias problem (Favero and Bullock 2015; Williams, Hartman, and Cavazotte 2010), the results of this study still should be viewed with caution. Focusing on supervisors' perceptions rather than objective measures of outcomes can be another limitation of this study. However, since the objective of this study is to learn under what conditions supervisors perceive the PIP as an effective tool to improve individual performance, focusing on supervisors' perceptions provides some valuable evidence, such as the importance of organizational support and the authority to remove poor performers.

Lastly, analyzing other related actors' perceptions, such as perceptions of subordinates who received the PIP before or Employee Relations staffs who helped implement the PIP within agencies, remain beyond the scope of this study due to data availability. Compared to the analysis of supervisors' perceptions, analysis of subordinates' perceptions may reveal a stronger positive relationship between the validity of performance standards and the effectiveness of the PIP since subordinates are the ones who receive the PIP due to their poor performance and thus, the extent to which their performance is adequately measured may be crucial in evaluating the effectiveness of the PIP. Meanwhile, Employee Relations staffs may perceive the PIP as more effective than supervisors when there is strong organizational support from the human resources office for the PIP because they may consider their role more essential in implementing the PIP by providing training and assistance for supervisors. If the data are available, one avenue for future research is to examine how the results might differ across various actors' perceptions.

Despite the above limitations and scope conditions, this study contributes to the performance information literature by investigating organizational conditions that may affect managers' perceptions of the PIP and ultimately their willingness to use this information. The findings of this study should not be considered definitive but take one step forward in understanding what conditions supervisors perceive this tool as effective. By focusing only on the supervisors who have ever used the PIP, this study seeks to provide some reliable evidence. The findings of this study provide several practical implications. Efforts should be undertaken to establish valid and useful performance standards for evaluating individual employees so that supervisors can use them when delivering performance feedback. Also, for the effective implementation of the PIP, this study underscores the importance of providing supervisors with sufficient assistance and the authority to exercise power when necessary.

Table 2.1 Construction of Study Variables

Study Variables	Survey Items
Perceived Effectiveness of the PIP	Two survey items (5-Likert scale) <ul style="list-style-type: none"> – Poor performers make a serious effort to use the PIP to improve their performance. – The PIP often results in the employee improving to an acceptable level or better.
Validity and Usefulness of Performance Standards	Two survey items (5-Likert scale) <ul style="list-style-type: none"> – The performance standards for my subordinates accurately reflect the critical elements necessary for success in their jobs. – The performance standards for my subordinates have measurements that I can use accurately to determine success in the critical elements.
Perceived Organizational Support	Two survey items (5-Likert scale) <ul style="list-style-type: none"> – Supervisors receive timely training to help them effectively implement a meaningful PIP. – Human resources staff provide competent assistance regarding PIPs.
Confidence in Delivering Performance Feedback	Single survey item (7-Likert scale) <ul style="list-style-type: none"> – Communicating with employees about how they can improve their performance.
Authority to Remove Poor-performing Employees	Single survey item (binary variable) <ul style="list-style-type: none"> – If a subordinate employee was deficient in a critical performance element after completion of a PIP, are you confident that you would be able to remove that employee?

Table 2.2 Descriptive Statistics, Correlations, and Reliabilities

(N= 2,939)	M	SD	Scale Range	Correlations and Reliabilities									
				1	2	3	4	5	6	7	8	9	10
Study Variables													
1. Effectiveness of PIP	6.22	1.81	2–10	(.71)									
2. Validity and Usefulness of Performance Standards	7.93	2.15	2–10	.16*	(.86)								
3. Organizational Support	5.87	2.42	2–10	.30*	.29*	(.75)							
4. Authority to Remove Poor Performers	.28	.45	0–1	.16*	.14*	.23*	–						
5. Confidence in Delivering Performance Feedback	5.34	1.47	1–7	.14*	.17*	.09*	.05*	–					
Control Variables													
6. Team Leader	.01	.08	0–1	–.03	–.02	–.03	.01	.01	–				
7. Supervisor	.49	.50	0–1	.07*	–.06*	.00	–.03	.08*	–.08*	–			
8. Manager	.41	.49	0–1	–.04*	0.01	–.02	–.01	–.04*	–.07*	–.82*	–		
9. Executive	.09	.29	0–1	–.04*	0.09*	.04*	.07*	–.08*	–.03	–.31*	–.27*	–	
10. Ratio of Unsatisfactory Subordinates	1.53	.88	1–5	–.14*	–.19*	–.12*	–.09*	–.05*	.04*	.08*	–.05*	–.07*	–

Note: All latent variables (Effectiveness of PIP, Performance Standards, Organizational Support) are treated as the multi-item summed index in this table. Cronbach's alpha is in parentheses. * p < .05

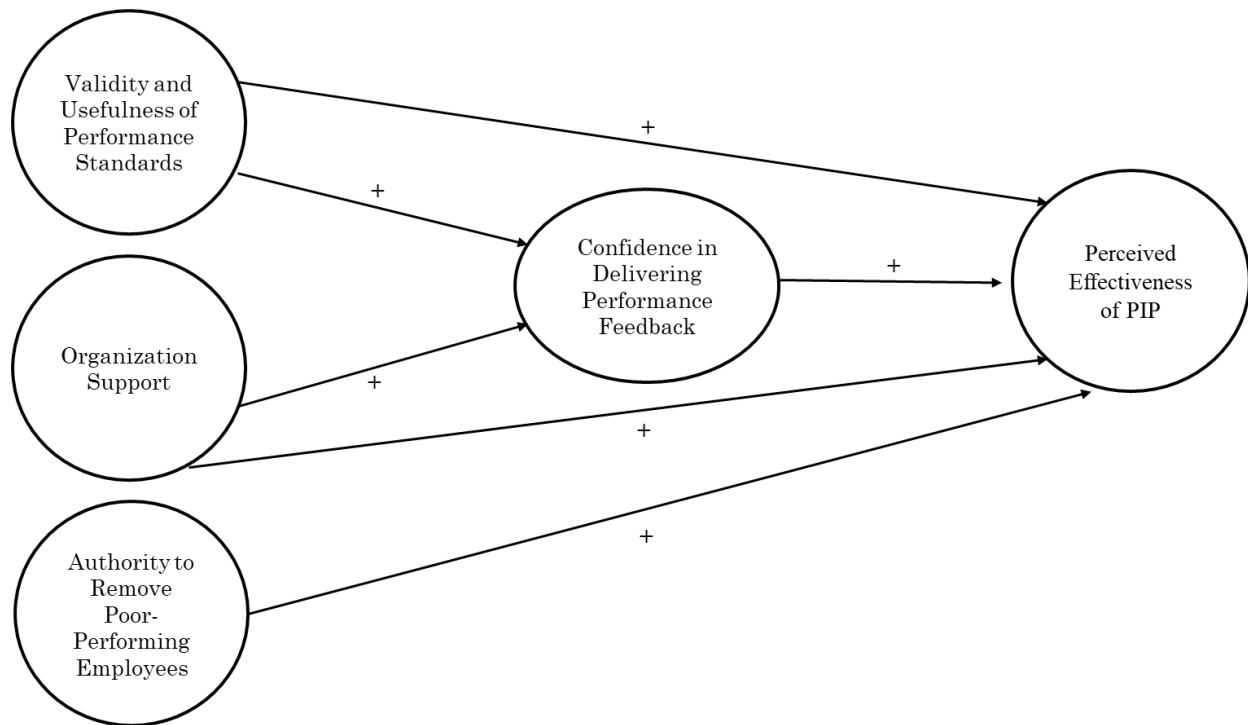


Figure 2.1 Conceptual Model

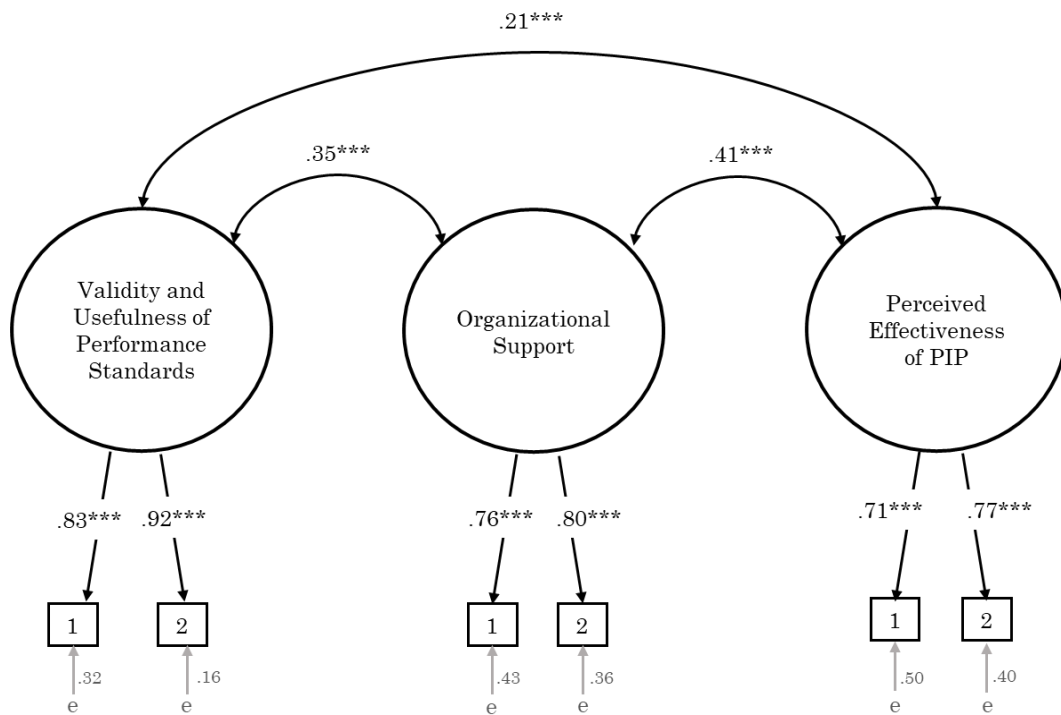


Figure 2.2 Confirmatory Factor Analysis of Study Variables

Note: standardized parameter estimates. * $p < .05$ ** $p < .01$ *** $p < .001$

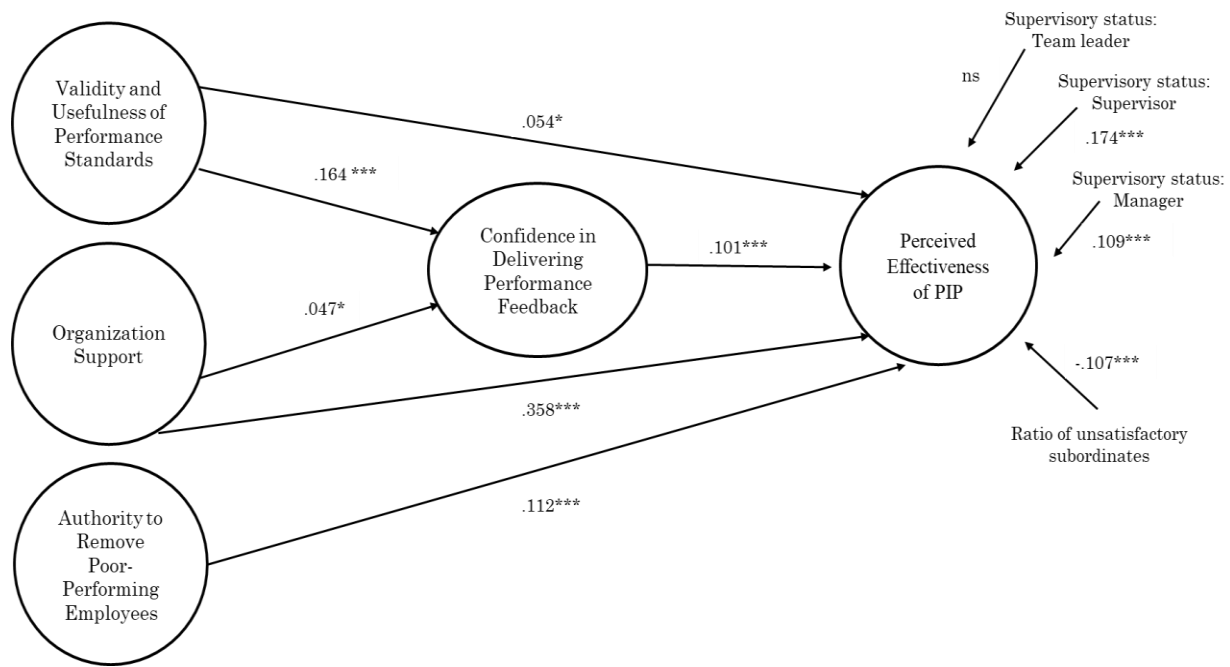


Figure 2.3 Structural Equation Model (SEM) Results

Note: Standardized parameter estimates. * $p < .05$ ** $p < .01$ *** $p < .001$.

CHAPTER 3
HOW ADMINISTRATIVE LEADERS AFFECT
ADMINISTRATIVE PERFORMANCE:
EVIDENCE FROM IMPROPER PAYMENTS IN U.S. FEDERAL PROGRAMS

Introduction

Improper payments in U.S. federal programs represent a longstanding and widespread problem, costing taxpayers significant amounts of money and hampering the fiscal sustainability of the government. This issue also harms the integrity and reputation of the federal government and compromises public trust (U.S. Government Accountability Office (GAO) 2011). Improper payments refer to payments in an incorrect amount (either overpayments or underpayments), payments that should not have been made at all, or payments made to an ineligible recipient or for an ineligible purpose (Congressional Research Service 2016).

Fewer payment errors contribute to better program performance by improving efficiency and reflecting greater organizational responsibility. It is because fewer improper payments represent a program's ability to accomplish its task with the least amount of wasted money and suggest careful stewardship of taxpayer funds. The numbers show a wide variation in improper payment rates and amounts across federal agencies. For instance, in FY2019, the Flood Risk Map and Risk Analysis program reported improper payment estimates of only 0.01% and \$10,000 (2015 constant-dollar), while the Medicare

Fee-for-Service program reported payment errors of 7.25%, totaling \$27.5 billion (2015 constant-dollar).

While reducing improper payments of federal programs has been one of the top priorities in U.S. government, it is also true that paying right amounts to eligible recipients is relatively more routinized and simpler task with less policy discretion than other government performance goals. This study seeks to investigate whether the difference in leadership type matters even with somewhat routinized tasks. Very few studies have analyzed improper payments in federal programs (e.g., Bullock 2014; Greer and Bullock 2018; Lee 2021); of these, most highlight one specific program. For instance, Greer and Bullock (2018) focus on the Unemployment Insurance program in the Department of Labor at the state level, arguing that the implementation of information technology tools and increased communication among stakeholders can lower improper payments. Similarly, Lee (2021) analyzes the impact of simplified reporting on payment errors related to the Supplementary Nutrition Assistance Program (SNAP) in the Department of Agriculture at the state level. While both studies help understand how innovative and well-designed strategies can reduce payment errors, relatively little is known about the variances between federal programs and how the type of leadership shapes the size of improper payments in federal programs. Bullock (2014) seeks to investigate the variations across federal programs in eighteen agencies by focusing on the impact of program designs (e.g., previous performance and size of programs) and various employees' perceptions (e.g., job satisfaction, pay satisfaction) of improper payments, but the role of leaders in managing improper payments, which can be a key of the problem, has not yet been explored.

This study seeks to analyze whether leader type differences can make a difference in performance management. Specifically, this study hypothesizes that programs managed

by career executives or non-acting executives will be more likely to make smaller improper payments compared to those managed by political executives and acting executives.

These propositions were empirically evaluated using data on 901 U.S. federal programs, covering 56 agencies/subagencies in 2002 – 2019 time period.¹⁰ Statistical evidence reveals that programs managed by career executives and non-acting executives are more likely to show better performance via smaller payment errors. These findings contribute to existing research comparing the management effectiveness of career executives and political appointees-managed programs. More importantly, this study investigates how program performance differs under the leadership of an acting manager, a mostly overlooked phenomenon (see Kinane 2021). The findings provide important implications for both researchers and practitioners, suggesting that acting officials cannot fully fill the “leadership vacuum” created by leadership vacancies, resulting in poor organizational performance.

The paper proceeds as follows. The first section discusses the nature of improper payments by federal programs and identifies trends. The next proposes two testable hypotheses to understand the impact of the leader types within the organizational unit on the payment errors in federal programs. The data, variables, and methodology are then presented, followed by the empirical findings. This study concludes by discussing its theoretical and practical implications and contributions to research.

¹⁰ Although improper payments of U.S. federal programs are required to report since FY2004, some programs’ actual activity periods start in 2002. More information will be discussed in Data, Measures, and empirical strategy section.

Improper Payments of U.S Federal Programs

*“Effective stewardship of taxpayer funds is a critical responsibility
of the Federal Government.”*

— U.S. Government Official Website, *PaymentAccuracy.gov*

Addressing improper payments in federal programs has been a major priority of the U.S. government since the 1990s, beginning with concerns from the Government Accountability Office (GAO) (Dodaro 2013). Overpayments and payments made to ineligible recipients represent the loss of taxpayer money otherwise available for critical, under-resourced programs such as education or child welfare policy. Underpayments matter as well: these payments degrade the integrity of federal programs and undermine the public’s trust in government (Malenich 2019). For these reasons, Presidents and the legislative branch have put efforts to reduce federal programs’ improper payments to save taxpayers money by reducing fraud, waste, and abuse in government programs.

The George W. Bush administration continued these efforts to improve financial performance by reducing improper payments through executive orders and laws, including the Improper Payments Information Act of 2002 (IPIA). The IPIA, the cornerstone legislation for improper payments, required executive branch agencies to identify programs and activities susceptible to significant improper payments, estimate annual amounts improperly paid, and report these estimates and actions taken to reduce them (Congressional Research Service 2010). With the motto of “waste not, want not,” the Obama administration also emphasized responsible stewardship by working to reduce improper payments better spent on other under-resourced policies such as education or student loans (Marr 2010). Accordingly, the Improper Payments Elimination and Recovery Act of 2010

(IPERA) was enacted, amending IPIA and expanding requirements for recovering overpayments across a broad range of federal programs. IPERA retained the core provisions of the IPIA while requiring improvements in agency improper payment estimation methodologies and reduction plans. It also significantly expanded the scope and reporting requirements of recovery audit programs. Finally, Congress passed the Improper Payments Elimination and Recovery Improvement Act (IPERIA) of 2012, intensifying its efforts to identify, prevent, and recover improper payments made from federal program funds. Put simply, with IPIA as a cornerstone law, the government has sought to address improper payment problems by improving requirements and methodologies in new legislations.

Fifteen years after IPIA, reducing government-wide improper payments continues to be a top priority. The Trump administration introduced an initiative called “Getting Payments Right,” naming the improper payment problem a key element of the Cross-Agency Priority (CAP) goal of the President’s management agenda.

Despite consistent efforts to reduce improper payments, these problems persist. Although seemingly a small percentage of a program’s total outlays, the amounts of improper payments of U.S. federal programs are nontrivial. **Figure 3.1** displays the sum of the total constant dollar value of improper payments in the U.S. federal government between FY2004 and FY2019. The largest amounts of improper payments are reported in FY2019 (\$163 billion), up from their lowest levels in FY2005 (\$46 billion).¹¹ Improper payments of

¹¹ The largest amounts of improper payments in FY2019 may be due to greater outlays of programs. Government-wide total program outlays for each year cannot be obtained due to data availability. In the statistical analysis, however, this study controls program outlays to account for scale effects in predicting the volume of improper payments.

U.S. federal programs have been estimated to total almost \$1.7 trillion government-wide from FY2003 through FY2019 (GAO 2020).

Differences in Administrative Leader Types and Its Implications for Improper Payments of Federal Programs

Improper payments of the U.S. federal programs are a serious problem that both congress and Presidents have paid attention to, as can be witnessed in a series of congressional legislations, executive orders, and the President’s management agenda. Top administrative officials are the ones who are accountable to the President and Congress for their performance in managing improper payments. Not only the accountability that leaders hold themselves for program performance, but they also have fiduciary responsibilities for program management. According to the Federal Managers’ Financial Integrity Act of 1982, it is clearly stated that each agency head must establish controls to reasonably ensure the effectiveness and efficiency of financial operations. With regard to improper payment problems, making payment decisions may be a relatively routinized task done by employees within the agencies, but the source and the remedy of the problem are rooted in the leaders. To address improper payment problems, GAO (2001b) emphasizes that the role of a leader is to create a control environment overarching the following four aspects: monitoring, risk assessment, control activities, and information and communication. In general, top-level officials can “instill a culture of accountability by adopting a positive and supportive attitude toward improvement and the achievement of established program outcomes” by “setting and maintaining the ethical tone, delegating roles and responsibilities, and implementing human capital initiatives clearly communicated the need for change” (GAO 2001b: 12). Relatedly, the Partnership for Public

Service (2018: 3) also emphasizes the role of leaders in addressing improper payment problems by suggesting that “agency leaders must engage relevant stakeholders, look holistically at the process, and use timely and accurate data to understand what is driving improper payments.” This study seeks to explore whether the types of administrative leaders, career executives versus political executives and non-acting executives versus acting executives, make a difference in program performance.

First, presidents use an appointment strategy to control bureaucracy and policy implementation (Moe 1982; Moe 1985). By filling key positions with people who will be responsive to presidential direction, presidents seek to influence bureaucracy (e.g., Bertelli and Feldmann 2007; Hollibaugh, Horton, and Lewis 2014; Lewis 2008; Parsneau 2013). For instance, Lewis (2008) shows that presidents seek to improve the odds of achieving policy goals by creating more slots for loyal political appointees or by shifting key management positions from careerists to appointed positions. Scholars have consistently examined how this appointment strategy affects organizational management, by comparing the performance of political executive- and career executive-administered programs, and suggesting an appropriate mix of the two (e.g., Gallo and Lewis 2012; Gilmour and Lewis 2006; Lewis 2007; also see Krause, Lewis, and Douglas 2006). In terms of managing performance and management effectiveness, Lewis (2008) finds that career executives have more direct agency experience and longer tenures, and both of these factors will improve careerists’ knowledge and expertise of their agencies. In other words, greater reliance on political executives at the expense of career executives may indicate the loss of managerial knowledge and expertise, as they lack “the subject area background or public management qualifications of their careerist counterparts” (Lewis 2012a: 589). Put simply, career executives exhibit more substantive expertise in the program area (Cohen 1998; Gilmour

and Lewis 2006; Lewis 2007; 2012a), leaving them better able to oversee programs and identify and solve problems. For leaders to create the control environment to manage improper payment problems, program knowledge and expertise become even more crucial (GAO 2012). Beyond the greater program expertise and experiences of career executives when compared to political executives, their differing motives may also affect their behaviors in managing improper payments. Specifically, political executives, regarded as “short-timers” or “in-and-outers” due to their relatively limited tenures, bring a short-term perspective and endure political pressure (Ban and Ingraham 1990; Hecló 1977; Pfiffner 1987). Although both political executives and career executives will have the same goal to reduce errors, political executives may not have the same impetus to improve their agency performance as career executives since they have different time horizons and often set their goals to further advancement, either to higher-level political positions or to the private sector (Brewer and Maranto 2000). In addition, this different motive of political executives may affect other employees within the organization. Richardson (2019) shows that civil servants who perceive greater politicization will reduce their investment in expertise, suggesting politicization may reduce agency policy expertise.

Conversely, career executives, who often outlast political appointees, have a long-term perspective unaffected by short-term political pressures (Pfiffner 1987; Lewis 2012a). Instead of short-term goals to benefit stakeholders for partisan and electoral purposes, career executives, insulated from politics and evaluated systematically via a merit system, prioritize competent, long-term program management. Since they spend most of their careers in a single organization, career executives are more likely to “be dedicated to their agency’s health, mission, and clientele” (Brewer and Maranto 2000: 72; see also Hecló 1977). Career executives’ dedication to their organizations will lead them to work towards a

reduction in payment errors, as managing improper payments is one of the top priorities of federal agencies. According to this mechanism, I posit the following hypothesis:

H1: *Federal programs managed by career executives will yield **lower** improper payments compared to programs managed by political executives.*

The absence of a permanent manager may also adversely affect program performance. Acting managers are those who temporarily assume leadership responsibility of the administrative unit. Leadership vacancy or filling in with acting officials are often consequences of turnover of existing leadership or presidential nomination or confirmation delays (Chang, Lewis, and McCarty 2001; O’Connell 2009; O’Connell 2020; Ostrander 2016). Previous studies have documented that leadership vacancies can lead to poor performance (e.g., Mendelson 2015; O’Connell 2009; 2015); employee uncertainty of the administration’s policy direction can create confusion within the organization (O’Connell 2009). Also, leadership vacancies can foster agency inaction since employees expect that priorities will eventually change with new permanent leadership (Mendelson 2015; O’Connell 2009). Furthermore, acting officials are temporary substitutes without the full authority of the office that accompanies Senate confirmation (Dull and Roberts 2009; O’Connell 2009). Besides the lack of formal authority, the lack of perceived authority creates more obstacles for acting leaders to manage their organizations. Acting managers question their own right to make long-term decisions, and they are not perceived by subordinates as having the full authority of the manager (Partnership for Public Service 2020). Robert Bonner, a former head of the U.S. Drug Enforcement Administration (DEA) and U.S. Customs and Border Protection Agency, explained:

[I]t is enormously important that the people that work for you in that agency view you, not as an ‘acting,’ but as somebody who is going to be around for a while....

You’re not going to be able to command as much respect and attention from your own people and from other agencies whose cooperation is important. (Tau 2019: 3).

With regard to designing and implementing actions to reduce improper payments, McCoy Williams, GAO managing director of financial management and assurance, emphasizes “communication, coordination, and cooperation among federal agencies and the states” (Jochum 2008: Paragraph 7). Without sufficient formal and perceived authority, an acting official, even one with substantial agency experience, may not only struggle to direct subordinates effectively and solve long-term problems within their agency (O’Connell 2009), but may also face challenges in communicating and coordinating with other agencies or the states. Therefore, it is reasonable to surmise that federal programs administered by acting officials would yield more payment errors than those administered by permanent leaders. The second hypothesis associated with the consequences of acting officials applied to improper payments of federal programs is as follows:

H2: *Federal programs managed by non-acting executives will yield lower improper payments compared to programs managed by acting executives.*

Next, the data, measures, and empirical strategy for evaluating these hypotheses are discussed.

Data, Measures, and Empirical Strategy

These hypotheses are evaluated using data on the size of improper payments in 166 federal programs administered by 56 U.S. federal agencies/subagencies during the 2002 –

2019 period.¹² The unit of analysis is program-level, with 901 total observations. Various data sources are leveraged in this study, including each agency’s annual Performance and Accountability Report (PAR) or Agency Financial Report (AFR) for improper payment data (outcome variable), OPM personnel data on employment-based measures used in this study, and *Federal Yellow Books* and *Leadership Connect* search engines for appointment type and background information on top officials administering programs.

Dependent Variable

To assess improper payments of U.S. federal programs, this study focuses on program-level improper payment rates and dollar amounts. Each agency is required to report improper payment estimates for programs identified as susceptible to “significant” improper payments, based on the guidance provided by the Office of Management and Budget (OMB).¹³ Improper payment amounts are the sum of the absolute value of

¹² Although federal agencies have reported their programs’ improper payment data annually since FY2004, fiscal calendar cycles for each program vary. For instance, while the Senior Companion Program administered by the Corporation for National and Community Service (CNCS) has a payment period from April 2016 to March 2017, payment period of the Risk Management Agency (RMA)’s Federal Crop Insurance Corporation (FCIC) program is from July 2014 to June 2015 (both reported in FY2017 Agency Financial Report). Working with these archival data requires great care, especially when handling many programs in different agencies across time.

¹³ Definitions of programs susceptible to “significant” improper payments have changed over time. During FY2004–2010, the Act defined “significant” as improper payments that exceed both \$10 million and 2.5% of program outlays. From FY2011–FY2012, IPERA added one more criterion and defined “significant” as either (1) improper payments that exceed both \$10 million and 2.5% of program or activity outlays; or (2) improper payments in excess of \$100 million (Congress Research

overpayments and underpayments. These data are reported by agencies through *Paymentaccuracy.gov* and can also be obtained from each agency's annual Performance and Accountability Report (PAR) or Agency Financial Report (AFR).¹⁴ This study takes the natural logarithm to adjust for a skewed and leptokurtic distribution of the dependent variables.¹⁵ Studies in various disciplines suggest that using a ratio-transformed variable can be problematic, especially when the denominator varies widely since it can produce inaccurate parameter estimates and thus provide evidence for a spurious relationship (e.g., Aldama and Bisbee 2018; Certo et al. 2020; Kronmal 1993; Liermann et al. 2004). Therefore, in addition to analyzing improper payment rates, this study employs improper payment dollar amounts as a dependent variable and adds the total outlay of federal programs (the denominator of the ratio-transformed variable, improper payment rate) as a control. Improper payment amounts are measured as the 2015 constant-dollar value. Although this study reports statistical results of both dependent variables (improper payment rates and dollar amounts), the interpretation and discussion in the Empirical

Service (CRS) 2016). Since FY2013, the 2.5% threshold dropped to 1.5%. This study accounts for this change by including a control variable (please see the discussion of *control covariates*).

¹⁴ PAR or/and AFR are obtained from each agency's website. The full list of agency website data sources can be obtained from the author.

¹⁵ This study compares the model fit between models using natural logged dependent variable and models using unlogged dependent variable, and the results reveal the superiority of models using natural logged dependent variable. Full results of model fit comparisons appear in **Appendices B.6**.

Findings section will focus on the statistical estimates of the improper payment dollar amounts.¹⁶

Data covering this study's sample period in **Figure 3.2** display the sum of all improper payments from the sample of data reported between FY2004–FY2019. This figure offers a conservative portrait of payment errors since some programs administered by multiple agencies or outside of the one-year cycle were excluded from the sample for the sake of comparability issues. In the sample, the largest total size of reported improper payments is \$103.03 billion (2015 constant-dollar) in FY2014, while the smallest total size is \$29.78 billion (2015 constant-dollar) in FY2007. **Figure 3.2** also shows a significant increase in the total size of improper payments from FY2009 and FY2010.¹⁷

Key Covariates

The first primary covariate, *Career Executive*, is defined as the appointment type of an administrative manager managing a given program in a year, where 0 indicates political executives while 1 indicates career executives. Here, an administrative manager is defined as a top official in a given administrative unit (agencies/bureaus) managing a given federal program. This study creates a weighted measure by employing more granular data in cases of leadership turnover, in which two or more individuals manage the program in a given

¹⁶ Statistical findings employing improper payment rate as a dependent variable can be found in the **Appendices** section (Full Model in **B.3 and B.4**, and Reduced Model in **B.5**). A reduced model is presented to ensure that the findings are not the result of overfitting the data.

¹⁷ One possible explanation for this increase may be the improvement of improper payment estimation methodology. As Congress enacted IPERA in 2010, a more sophisticated methodology was introduced to estimate improper payments in federal programs. This study seeks to account for the effect that the new legislation may produce by adding control variables in the statistical model.

year. For instance, if an executive A (political appointee) served for three months (25%) while an executive B (career executive) served for nine months (75%) for the same program, this measure is calculated as 0.75 ($0.25 \times 0 + 0.75 \times 1 = 0.75$). According to **H1**, this covariate should be negatively associated with the size of improper payments of federal programs.

The second primary covariate, *Non-Acting Executive*, is defined as whether an administrative manager assumes the responsibility of a higher-level position on a permanent basis within a given program in a year, where 0 indicates acting appointment (acting political executives and acting career executives) while 1 indicates regular appointment (permanent political executives and career executives).¹⁸ As with the *Career Executive* covariate, in cases with two or more executives in a given year, a weighted measure is employed. **H2** predicts that this covariate is also negatively related to improper payments.

Control Covariates

In addition to key covariates of interest, other executive-specific or program/agency-specific considerations may affect the (in)accuracies of federal program payments. These controls account for systematic variations in improper payments that are not attributable to the main covariates and thus are important to consider as alternative explanations. First, several covariates are considered to account for an executive's managerial expertise derived from prior experience. This study controls whether the executive had *Previous Department Experience*. If an individual had work experience in the department before

¹⁸ The correlation between *Career Executive* variable and *Non-Acting Executive* variable was -.34, indicating a weak negative correlation.

becoming a leader, this variable is measured as 1. Otherwise, it is measured as 0. This indicator is created as a weighted measure in cases where more than one executive administered a federal program in a given year, with the minimum value of 0 for no prior experience in the current department and the maximum value of 1 with prior experience in the current department. Similar to prior department-specific experience, *Executive Tenure*, a count of the number of months the executive has been in charge of the bureau leadership, is also included in the statistical models. Each variable is posited to be negatively related to improper payments since executives possessing more managerial agency/program-specific experience are more likely to make proper payments by utilizing their managerial experiences and knowledge. *Executive Turnover* in a given year as a dummy variable is also included in the statistical model. Since leadership turnover can introduce numerous detrimental effects to the organization, ranging from leadership vacuums to a lack of continuity and to working relationship disruption among related inter- and intra-agency groups (Gilmour and Lewis 2006), a positive coefficient sign is hypothesized.

Additional covariates are considered to account for program-specific reporting requirements. First, a binary indicator (*Disaster Fund*) indicating whether the program received supplemental disaster relief administrative funds is included. For FY2014 and beyond, the Disaster Relief Appropriations Act 2013 (Pub. L. No. 113-2) required all programs receiving funds to be deemed susceptible to significant improper payments and report improper payments. This measure should be inversely correlated with improper payments since the affected programs had to report regardless of meeting high-risk criteria. Also, a *Section 57* binary indicator variable is included, measured as 1 for programs under the former Section 57 of OMB Circular No. A-11, and 0 otherwise. Since these programs were originally designated as high-risk programs by OMB in FY2001, and OMB required

agencies to report these programs' improper payments, these programs are expected to yield higher improper payments.¹⁹

In addition, this study controls whether the program *stopped* reporting improper payments in the next fiscal year by showing evidence on low improper payments or *reentered* to report improper payments in the current fiscal year after its relief from annual reporting requirements and being indicated that the program is again susceptible to improper payments, via the separate binary indicators, *Exit* and *Reentry*. The decisions to both stop and reenter reporting improper payments are made by the Office of Management and Budget (OMB) (OMB Memorandum M-15-02, Appendix C to Circular No. A-123 2014). Each variable is posited to be negatively related to improper payment size since the *Exit* variable directly indicates that these programs produced sufficiently lower improper payments in the current period (and previous years) to get relief from reporting. Also, even though the *Reentry* variable indicates these programs are again susceptible to significant improper payments, these programs will be more likely to show lower improper payments, considering their historical evidence on low improper payments in the past when compared to programs that had never been granted relief from reporting. An ordinal variable indicating the number of programs reporting improper payments within a given administrative unit (agencies/subagencies) is also included (*Number of Programs* variable), with the expectation that administrative managers with more program responsibilities to

¹⁹ Programs listed in the former Section 57 of OMB Circular A-11 for FY 2001 are not permanently subject to reporting improper payments (OMB 2006). If a program has documented two consecutive years of improper payments under \$10 million annually, the agency may request relief from annual reporting requirements for this program. The *Section 57* variable represents whether the program was originally under the former Section 57.

report improper payments will face more challenges in managing multiple programs, and thus lead to larger improper payments.

In addition, since most federal programs calculated their improper payments through samples instead of tallying all the payments made in a given year, this study seeks to account for possible systematic variations in improper payment size driven by sampling methods by using separate binary indicators. By employing the ‘actual’ group (100% of actual payments) as a reference group, four separate binary indicators are included in the statistical model, ranging from the least robust sampling method to the most robust sampling method.²⁰ These covariates are predicted to have an ambiguous relationship with improper payments, since the least robust sampling method may serve to exploit the data and report fewer errors (i.e., a negative relationship with improper payments), or draw biased samples resulting in large errors unintentionally (i.e., a positive relationship with improper payments).

Changes in reporting requirements for improper payments under different Acts during the sample period are also considered. The three major pieces of legislation on improper payments between FY2004 and FY2019, IPIA 2002, IPERA 2010, and IPERIA

²⁰ Four separate binary indicators are as follows: 1) sampling1: +/- 3% or better margin of error at the 95% confidence interval (+/- 2.5% or better margin of error at the 90 % confidence interval); 2) sampling2: a wider than +/- 3% or margin of error at the 95% confidence interval (a wider than +/- 2.5% margin of error at the 90% confidence interval); 3) sampling3: statistical method, but not enough information; 4) sampling4: no information about sampling method. The number of programs that fit in each “sampling” category is as follows: 29 programs (100% payments as a reference group), 545 programs (sampling1), 83 programs (sampling2), 239 programs (sampling3), and 5 programs (sampling4).

2012, changed the definitions of key terms (such as ‘significant improper payment’) and reporting requirements.²¹ By creating separate binary indicators, this study seeks to account for the effect of the new legislation. By employing the *IPERIA* years (programs reporting improper payments in FY2013-FY2019) as a reference group, an *IPIA* binary variable indicating programs reporting improper payments in FY2004-FY2010, and an *IPERA* binary variable indicating programs in FY2011-2012 were included in the statistical model. Lastly, when employing the natural log of improper payment dollar amount as a dependent variable, the natural log of total outlays ($\ln[outlay]$) is included as a control and expected to produce a positive relationship: programs with larger outlays are likely to make relatively more errors in payments.²² The total outlay is also measured as the annual 2015 constant-dollar total outlay of the program. In the **Appendices** Section, **B.1** reports descriptive statistics and correlations for all variables.

Empirical Strategy

This study employs OLS regression analysis to evaluate the change in the conditional mean of the dependent variable. For modeling varying sizes of improper payments, the hypotheses are further evaluated utilizing the quantile regression estimation strategy originally developed by Koenker and Bassett (1978) and extended to account for heteroskedasticity and intra-cluster correlation by Parente and Silva (2016). Quantile regression is an alternative to linear regression approaches which produce estimates of the conditional mean value of a dependent variable at a given value of predictor variables.

²¹ Please check Footnote 13.

²² When using the natural log of improper payment rate as a dependent variable, the natural log of total outlays is excluded in the statistical model to avoid the double counting problem (Please see **Appendices B.3, B.4, and B.5**).

While the conventional regression approach presents “a grand summary for the averages of the distributions corresponding to the set of x 's” (Mosteller and Tukey 1977:266), it poorly estimates the conditional mean for the extreme cases in the sample (Koenker and Hallock 2001). Quantile regression allows the researcher to consider the entire distribution of an outcome variable and to analyze any quantiles in the full conditional. Quantile regression makes it possible to assess the effects of key explanatory variables across different sizes of improper payments without truncating the sample by segmenting the outcome variable as in the subgroup analysis. By employing quantile regression methods, this study seeks to answer a more nuanced question: are the effects of determinants the same for federal programs with small improper payments (belonging to the lower quantiles) as for those with large improper payments (belonging to the higher quantiles)? The hypotheses that this study is testing are not conditional hypotheses where the effects are expected to change with varying sizes of improper payments, but this study seeks to explore whether the effects of administrative leader types persist across different sizes of improper payments since there is a wide variance in improper payment amounts. For instance, the improper payment amount of \$1 million within a program ($\tau = 0.21$) is substantively different from the improper payment amount of \$300 million within a program ($\tau = 0.90$). Considering that approximately 69 percent of improper payment amounts were concentrated in three program areas (Medicaid, Medicare, and Earned Income Tax Credit) in the fiscal year 2019 (GAO 2020), investigating local estimates at specific quantiles of the outcome variable is beneficial to this study after looking at the average effects across the entire distribution. Also, since there is a high possibility that observations within each program may not be independent, and this could lead to residuals that are not independent, robust standard errors are cluster-adjusted by programs in each model to account for arbitrariness within-

program. Parente and Silva (2016) propose a specification test for detecting the presence of intra-cluster correlation, and the test results reject the null hypothesis of independently distributed observations within programs and suggest that intra-cluster correlation should be considered at all estimated quantiles in this study.

Statistical Findings

The full set of the baseline OLS regression result and quantile regression results appears in **Table 3.1**.²³ In the baseline model, executive tenure and previous department experience variables are excluded in the model since these variables may partial out some of the variations of the key covariates of interest. In addition to the OLS estimates, the quantile regression models are estimated for the following quantile of the distribution of improper payment dollar amounts corresponding to the sample (and year listed in parentheses): $\tau = 0.0921$: \$0 (\$ in Millions) — Border Security Fencing program administered by U.S. Customs and Border Protection (Reporting Year: FY2016); $\tau = 0.2109$: \$1 (\$ in Millions) — Milk Income Loss Contract program administered by Farm Service Agency (Reporting Year: FY2014); $\tau = 0.3818$: \$10 (\$ in Millions) — Child and Adult Care Food program administered by U.S. Food and Nutrition Service (Reporting Year: FY2010); $\tau = 0.5483$: \$40 (\$ in Millions) — State Home Per Diem Grants programs administered by Veterans Health Administration (Reporting Year: FY2012); $\tau = 0.6759$: \$100 (\$ in Millions) — Head Start program administered by Administration for Children and Families

²³ The statistical findings of the reduced model (omitting control variables) can be also found in the **Appendices (B.2)**. Reduced model ensures that the findings are not the result of overfitting the data.

(Reporting Year: FY2007); $\tau = 0.80$: \$ 300 (\$ in Millions) — Military Retirement Benefits program administered by Defense Finance Accounting Service (Reporting Year: FY2018); $\tau = 0.96$: \$11,805 (\$ in Millions) — Medicare Fee-for-Service administered by Centers for Medicare and Medicaid Services (Reporting Year: FY2008). These quantiles are delineated based on the criteria for significant improper payments under the law. First, \$10 million of improper payments in a program ($\tau = 0.3907$ in the sample) is a threshold for significance if the improper payment *rate* of that program also exceeds 1.5% [Criterion 1].²⁴ In addition, if the improper payment amount of a program exceeds \$100 million ($\tau = 0.6870$ in the sample), that program is deemed high-risk regardless of the improper payment rate size [Criterion 2]. By using these criteria, this study seeks to assess the effects of key explanatory variables across different sizes (small, moderate, large) of improper payment programs.

The findings related to some of the control variables are worth noting. Program outlays are positively associated with improper payment amounts not just on the conditional mean but also in every quantile, which accounts for scale effects. Executive turnover does not show a significant association with improper payments in most quantiles but shows a strong positive relationship with very large improper payments ($\tau = 0.96$, IP amounts = \$11,805 Million). This finding suggests that leadership turnover within an organization may hurt performance by yielding more payment errors, especially for programs with very large improper payments. As expected, both programs receiving

²⁴ The threshold of improper payment rate has been changed from 2.5% to 1.5% since FY2013. However, the threshold of improper payment amounts employed with improper payment rate [Criterion 1], 10 million dollars, has not been changed.

disaster funds and those granted reliefs from reporting in the next fiscal year have a negative relationship with improper payments on the conditional mean and in most quantiles. Finally, the results show that programs, where their agencies have more programs required to report improper payments, are more likely to produce larger improper payment amounts for moderate IP sizes ($0.3818 \leq \tau \leq 0.6759$, \$10 Million \leq IP amounts \leq \$100 Million) and exceptionally large ones ($\tau = 0.96$, IP amounts = \$11,805 Million).

Next comes an analysis of the estimates of the main hypotheses explaining how the leadership type differences can affect federal programs' payment errors. OLS estimation results focusing on average marginal effects across the entire sampling distribution of improper payment amounts reveal that programs managed by career executives (chief manager) yield lower improper payments compared to programs managed by political executives (**H1**) and are significant at 90 percent confidence level. The results of the quantile regression models also suggest similar findings at moderate IP amounts ($0.3818 \leq \tau \leq 0.5483$, \$10 Million \leq IP amounts \leq \$40 Million). To understand the substantive magnitude of the effects, I calculate and evaluate the exponentiated regression coefficients since the dependent variable is log-transformed due to its skewed and leptokurtotic nature. **Table 3.3** provides average differential effects in improper payments by administrative leaders. When comparing programs managed by political executives and programs managed by career executives, at the mean, we would expect to see lower improper payments (by about 79.5%) in career executive-managed programs. In addition, the result shows that the numerical effects of this covariate decrease when comparing programs at moderate IP size ($\tau = 0.3818$, IP amounts = \$10 Million) and programs at larger IP size ($\tau = 0.5483$, IP amounts = \$40 Million). Specifically, we would expect to see lower improper payments (by about 82.6%) in career executive-managed programs when compared to

political executive-managed programs at moderate IP size, and about 76.9% lower improper payments in career executive-run programs at larger IP size.

Next, the statistical results also show strong support for **H2**, suggesting that programs managed by non-acting executives yield lower improper payments at the conditional mean (OLS result) and at small to fairly large quantiles ($0.2109 \leq \tau \leq 0.80$, $\$1 \text{ Million} \leq \text{IP amounts} \leq \300 Million). The numerical estimates reveal that the largest effect (about 87.8% lower) appears at the conditional mean IP size ($\exp(\beta) = 0.122$). Among quantile regression results, the effect size is the largest at the moderate IP size ($\tau = 0.3818$, IP amounts = \$10 Million: about 84.4% lower) and the smallest at the large IP size ($\tau = 0.80$, IP amounts = \$300 Million: about 63.3% lower).

After conducting the statistical analysis for the baseline model, this study further analyzes the model controlling for executive tenure and previous department experience. By controlling those two variables in the statistical analysis, this study seeks to net out learning effects obtained from the management experiences, from the effects of executive appointment type (political executives vs career executives) on the size of improper payments. The full set of statistical analysis results appears in **Table 3.2**. For control variables that are also included in the baseline model, the results show substantively similar findings in terms of point estimates and statistical inferences. The OLS result reveals no statistical relationship with both previous department experience of executives and longer tenures in leadership positions and improper payment amounts. However, these two covariates have a strong negative association with improper payment amounts for very large IP sizes ($\tau = 0.96$, IP amounts = \$10,400 Million). Given that larger sizes of improper payments often indicate program complexity, the result implies that the tenures and

previous department experiences of executives become more important and beneficial to reduce improper payments when managing more complex programs.

At the conditional mean, *Career Executive* variable misses statistical significance, but even after controlling for executive tenure and previous department experience, the results suggest that programs managed by career executives (chief manager) yield much smaller improper payments at moderate IP amounts ($0.3818 \leq \tau \leq 0.5483$, \$10 Million \leq IP amounts \leq \$40 Million), compared to programs managed by political executives (**H1**). As discussed in the hypotheses section, not just the learning effects obtained from previous work experiences and long tenure, different motivations between career executives and political executives may affect program performance in terms of payment errors. With regard to numerical effects, when comparing programs managed by career executives and political executives, we would expect to see lower improper payments (by about 84.5% and 78%) in political executive-managed programs at moderate IP sizes ($\tau = 0.3818$, IP amounts = \$10 Million; $\tau = 0.5483$, IP amounts = \$40 Million, respectively).

As in the baseline model, the results also reveal that programs managed by non-acting executives yield smaller improper payments not just at the conditional mean amounts, but also at both small and large IP amounts ($0.2109 \leq \tau \leq 0.80$, \$1 Million \leq IP amounts \leq \$300 Million) than programs managed by acting executives (**H2**). The numerical effect is largest (about 88.8% smaller) at the conditional mean, and is smaller at the larger IP amounts ($\tau = 0.80$, IP amounts = \$300 Million, about 50.9% smaller) (Please see **Table 3.3** for numerical effects at every estimated quantile).

To ensure that the findings are not the result of overfitting the data, these data are further analyzed by omitting all control variables, except the natural log of the total outlays to account for scale effects in predicting the volume of improper payments. This reduced

model reveals statistical patterns that are substantially similar to those presented in the manuscript with few exceptions. For **H1** (comparing career executive- and political executive-managed programs), omission of control variables produces substantively similar findings in terms of point estimates and statistical inferences, with more precision for programs with fairly small IP amount ($\tau = 0.2109$, IP amounts = \$1 Million) and very large IP amount ($\tau = 0.96$, IP amounts = \$11,805 Million). For **H2** (comparing non-acting and acting executive-managed programs), the reduced model reveals similar findings with less precision for programs with moderate IP size ($\tau = 0.3907$, IP amounts = \$10 Million), but more precision for programs with very large IP size ($\tau = 0.96$, IP amounts = \$10,400 Million).²⁵

This study also conducts subsample analyses comparing programs in bureaus in executive departments (N=642) and programs in independent agencies (N=259) since there is a difference in how proximate the leader is in organization between these two subsamples.²⁶ For the programs in executive bureaus (N=642), the results are consistent with the results of main analyses, with more precision. However, for the programs in independent agencies (N=259), the results are estimated with much less precision. One possible reason of failing to attain statistical significance would be the physical distance between leaders and subordinates caused by organizational structure, or the small sample size.

Furthermore, in order to evaluate whether the impact of leadership type differences on improper payments exacerbates when there is a higher level of policy program

²⁵ For additional details, see **Appendices B.2. Reduced Model Estimates**.

²⁶ The full set of regression results appears in **Appendices B.7 & B.8**.

complexity, this study conducts an additional statistical analysis by incorporating two interaction terms, one combining career executive and policy complexity and the other combining non-acting executive and policy complexity, and evaluates marginal effects based on the interdecile change in policy complexity. In order to measure policy program complexity, this study employs a professional staff ratio within a bureau/agency in a given year as a proxy measure (e.g., Chun and Rainey 2005; Lee, Rainey, and Chun 2010). This measure is obtained by calculating the percentage of “professional” employees in a work unit in a given year from the Central Personnel Data File (CPDF) of the U.S. Office of Personnel Management (OPM).²⁷ The result of this robustness check reveals that there is no consistent evidence that policy complexity measured by professional staff ratio has a moderating effect in the relationship between the leadership type difference and improper payment amounts.²⁸

This study also investigates whether the executive tenure attenuates differences between career executive-managed programs and political executive-managed programs, and between non-acting executive-managed programs and acting executive-managed programs as executive tenure increases. In the appendices section, two interaction terms (one combining career executive and executive tenure and the other combining non-acting executive and executive tenure) are included, and the marginal effects based on the interdecile change in executive tenure are evaluated. The result presents null findings,

²⁷ The CPDF (Central Personnel Data File) categorizes occupational categories as professional, administrative, technical, clerical, other white collars, blue-collar, and unspecified (OPM 2006).

²⁸ The full set of analyses appears in **Appendices B.9 Robustness Check for Analyzing the Moderating Effect of Policy Complexity**.

with few exceptions, which implies that leadership differential effects are not varying with the executive tenure.²⁹

Lastly, this study also estimates the model controlling for different presidential administrations (Republican presidents and Democratic presidents) since there may be some differences in the degree of emphasis on reducing improper payments despite bipartisan support. The result reveals substantively similar findings in terms of point estimates and statistical inferences even after controlling different presidential administrations.³⁰

Discussion

After 15 years of federal oversight to identify and reduce improper payments from federal programs, the issue still remains one of the most critical challenges for the U.S. government. This study has sought to understand why some federal programs make more payment errors than others and explore the central role of the leadership types. As GAO (2002: 8) emphasizes, “the basic or root causes of improper payments can typically be traced to a lack of or breakdown in internal control.” Among various strategies that government should adopt to reduce improper payments, GAO (2011) emphasizes the role of top-level officials in establishing a culture of accountability. From this perspective, this study finds

²⁹ The full set of analyses appears in **Appendices B.10 Robustness Check for Analyzing the Moderating Effect of Executive Tenure.**

³⁰ The full set of analyses appears in **Appendices B.11 Robustness Check Controlling Different Presidential Administrations.**

that the differences in administrative leader types can affect the levels of improper payments.

Although the current study examines improper payments of federal programs, its focus on administrative leaders has broader implications for the study of public management. There has been an argument about whether the program performance is better at the agencies led by political appointees or careerists. While one argues that political appointment strategies employed by presidents to control bureaucracy may harm organizational performance, the other argues that political appointees perform better since they are directly responsive to the President (See Lewis 2012a). At the end of President Trump's term, there was an attempt to hire a new class of political appointees, Schedule F, and convert existing upper-level career officials into political positions by issuing an executive order 13957 (Silver 2021). Although the justification of this executive order that the direct accountability of Presidents will improve government performance is not that strong, there is a chance that a future president waits for the next opportunity and adopts a similar executive order on this basis (Moynihan 2021 a). Therefore, it becomes even more crucial to have credible evidence on the argument of career executives' and political executives' performance. By showing that career executive-administered programs make systematically lower improper payments than political executive-administered programs, this study implies that any strengths of political appointees may not be effective in reducing improper payments. Rather, career executives' long-term perspectives, dedication to their organization, and agency-specific expertise make it possible to better manage organizational performance. This study also emphasizes how the leadership vacuum created by the lack of permanent top officials can affect program performance. Even if acting officials fill in the leadership position, their lack of formal and perceived authority

and temporary position leave them unable to effectively manage performance. By focusing on the performance outcomes of acting positions, this study is expected to contribute to extant performance studies that tend to be restricted to the effects of different types of permanent appointees (e.g., Gallo and Lewis 2012; see also Kinane 2021).

Although the results reveal some evidence that career executive-administered programs make lower payment errors than political executive-administered programs, this study does not suggest that career executives would outperform political executives in every aspect. However, while political executives tend to have better political skills in negotiating with other actors and dealing with the media (Gallo and Lewis 2012), this study implies that these advantages may not improve program management, such as reducing improper payments. As previous research has suggested, agencies can perform better with a mix of these groups (e.g., Krause, Lewis, and Douglas 2006; Suleiman 2003); each appointment type has different strengths that can add value to the organizations.

More importantly, considering the effect of non-acting executives on performance, it is worth asking how to deal with leadership vacuums within organizations. Given the nature (e.g., lack of authority, temporary position) and inevitability of the acting position, it is crucial to designate qualified officials with sufficient agency experience and knowledge as acting executives to mitigate inherent problems.

The contribution of this study should be considered in light of its limitations. First, although this study seeks to account for possible systematic variations in improper payment size driven by sampling methods, this study cannot address the possibility that agencies may not report accurate improper payment estimates or reduce eligible recipients' access to improve performance on improper payments. Additionally, this study cannot rule out the possibility of reverse causality between executive turnover and the size of improper

payments. Also, although this study examines whether the impact of leadership type differences on improper payments exacerbates when there is a higher level of policy program complexity, other program-related characteristics such as the complexity of eligibility criteria are also likely to be related to improper payment amounts. With the data in hand, however, this study could not address these limitations. Lastly, one possible extension of this current study would be looking at the combined effect of different types of administrative leaders: 1) non-acting career executives; 2) acting career executives; 3) non-acting political executives; 4) acting political executives. Based on the mechanism that I employed in this study, programs managed by non-acting career executives will be more likely to yield lower improper payments because non-acting career executives not just have more expertise and knowledge but also have the authority to use their wherewithal to manage programs. Conversely, programs managed by acting political executives will be more likely to yield higher improper payments compared to the other three types of administrative leaders since acting political executives have less expertise and knowledge about the program and lack formal and informal authority because of their acting positions. Therefore, this study suggests the need for an additional study to evaluate under which combinations of administrative leader appointment types better manage programs.

Although additional research is needed to verify these findings, this study has several potentially important implications for the body of work on management and performance. First, and most importantly, this study investigates the impact of different types of leadership on reducing improper payments. With more wherewithal and greater incentives, career executives and non-acting executives are found to be more effective in addressing the problems of improper payments. Second, even for this relatively routinized task, this study finds that the type of leadership is crucial since top administrative officials

have performance accountability consequences. Given that this study analyzes agencies' task of reducing improper payments which is an administrative and routinized task, the differences in the impact of leadership type may be understated. In other words, this study suggests that the impact of leadership type differences may be even bigger for program accountability in areas where agencies are performing policy-specific tasks.

**Table 3.1 Baseline Model Estimates
(Omission of Executive Tenure and Previous Department Experience Controls)**

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$	Quantile $\tau = 0.2109$	Quantile $\tau = 0.3818$	Quantile $\tau = 0.5483$	Quantile $\tau = 0.6759$	Quantile $\tau = 0.80$	Quantile $\tau = 0.96$
IP (\$) in Millions			IP: \$0	IP: \$1	IP: \$10	IP: \$40	IP: \$100	IP: \$300	IP: \$11,805
Career Executive	-	-1.584* (0.956)	-1.349 (2.012)	-0.772 (0.853)	-1.707** (0.792)	-1.466*** (0.487)	-1.286 (0.872)	-0.850 (0.612)	-0.291 (0.303)
Non-Acting Executive	-	-2.107*** (0.683)	-0.977 (1.091)	-1.308* (0.773)	-1.860*** (0.684)	-1.779*** (0.421)	-1.663*** (0.481)	-1.003** (0.413)	-0.224 (0.262)
Executive Turnover	+	-0.149 (0.303)	-0.298 (0.420)	-0.273 (0.333)	-0.234 (0.309)	-0.157 (0.220)	-0.089 (0.322)	-0.035 (0.158)	0.712*** (0.266)
Disaster Fund	-	-3.255*** (1.006)	-7.863** (3.225)	-7.594*** (1.553)	-3.939** (1.693)	-3.552*** (0.921)	-2.163*** (0.793)	-1.423** (0.585)	-0.707 (0.438)
Section 57	+	0.421 (0.766)	1.251 (1.375)	0.225 (0.737)	-0.435 (0.576)	-0.057 (0.393)	0.186 (0.407)	0.052 (0.437)	-0.021 (0.428)
Exit	-	-1.444*** (0.503)	-2.453 (1.584)	-2.057*** (0.486)	-1.955*** (0.490)	-1.544*** (0.439)	-1.359*** (0.407)	-1.372*** (0.391)	-0.301 (0.309)
Reentry	-	-1.427 (1.007)	-3.949 (5.903)	-2.154** (0.952)	-1.658 (1.058)	-1.303 (0.987)	-0.767 (1.211)	-0.340 (0.665)	-0.939*** (0.223)
Number of Programs	+	0.096 (0.062)	0.057 (0.181)	0.085 (0.064)	0.101*** (0.038)	0.088*** (0.031)	0.073** (0.033)	0.042 (0.040)	0.055** (0.026)
Sampling1	+/-	-0.644 (1.328)	-1.527 (2.344)	-0.878 (1.799)	0.625 (3.058)	-0.887 (1.527)	0.140 (1.337)	-0.304 (0.771)	0.212 (0.500)
Sampling2	+/-	1.811 (1.411)	2.154 (2.717)	0.895 (2.004)	2.126 (3.092)	0.563 (1.529)	1.502 (1.344)	0.766 (0.771)	0.738 (0.565)
Sampling3	+/-	-0.364 (1.343)	-1.172 (2.320)	-0.927 (1.916)	0.973 (2.943)	-0.611 (1.502)	0.394 (1.333)	-0.273 (0.766)	0.223 (0.475)

Sampling4	+/-	-0.096 (1.555)	-0.091 (2.455)	-0.566 (2.044)	-0.226 (2.965)	-2.201 (1.537)	-1.272 (1.518)	-2.026 (1.510)	-0.578 (0.634)
IPIA	+/-	-0.720 (0.463)	-1.100 (0.988)	-0.610 (0.463)	-0.405 (0.353)	-0.513* (0.271)	-0.563** (0.258)	-0.202 (0.266)	-0.101 (0.197)
IPERA	+/-	-1.142* (0.589)	-10.620*** (2.994)	-0.434 (0.542)	-0.344 (0.315)	-0.346 (0.287)	-0.563* (0.327)	-0.340 (0.371)	-0.016 (0.204)
Ln (Outlay)	+	1.059*** (0.139)	0.955** (0.414)	1.039*** (0.175)	1.083*** (0.150)	0.927*** (0.128)	0.938*** (0.126)	1.010*** (0.084)	1.048*** (0.073)
R^2		0.4829	0.3548	0.4602	0.4731	0.4776	0.4784	0.4684	0.4495
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T = 11.34 P > T = 0.000	T=12.88 P > T = 0.000	T=18.386 P > T = 0.000	T=19.732 P > T = 0.000	T=23.204 P > T = 0.000	T=20.522 P > T = 0.000	T=18.744 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 3.2 Augmented Model Estimates
(Inclusion of Executive Tenure and Previous Department Experience Controls)**

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$	Quantile $\tau = 0.2109$	Quantile $\tau = 0.3818$	Quantile $\tau = 0.5483$	Quantile $\tau = 0.6759$	Quantile $\tau = 0.80$	Quantile $\tau = 0.96$
IP (\$) in Millions			IP: \$0	IP: \$1	IP: \$10	IP: \$40	IP: \$100	IP: \$300	IP: \$11,805
Career Executive	-	-1.725 (1.053)	-1.311 (2.489)	-1.203 (0.996)	-1.865** (0.849)	-1.512*** (0.532)	-1.102 (0.780)	-0.657 (0.580)	-0.304 (0.273)
Non-Acting Executive	-	-2.188*** (0.773)	-1.200 (1.283)	-1.482* (0.814)	-1.910*** (0.684)	-1.721*** (0.468)	-1.374*** (0.517)	-0.712* (0.401)	-0.313 (0.369)
Executive Tenure	-	0.005 (0.010)	0.015 (0.013)	0.010 (0.007)	0.003 (0.007)	-0.003 (0.005)	-0.007 (0.006)	-0.009 (0.006)	-0.007** (0.003)
Previous Dep. Experience	-	0.196 (0.486)	0.362 (1.075)	0.050 (0.528)	0.198 (0.385)	0.060 (0.371)	0.077 (0.342)	-0.030 (0.400)	-0.521** (0.237)
Executive Turnover	+	-0.124 (0.302)	-0.250 (0.458)	-0.154 (0.345)	-0.236 (0.282)	-0.142 (0.228)	-0.137 (0.302)	0.047 (0.178)	0.414** (0.203)
Disaster Fund	-	-3.327*** (1.023)	-7.755** (3.570)	-7.990*** (1.583)	-3.934** (1.706)	-3.477*** (0.952)	-2.167*** (0.759)	-1.373** (0.661)	-0.404 (0.448)
Section 57	+	0.466 (0.732)	1.185 (1.585)	0.327 (0.728)	-0.387 (0.606)	-0.021 (0.429)	0.227 (0.392)	0.250 (0.464)	-0.286 (0.455)
Exit	-	-1.444*** (0.502)	-2.339 (1.746)	-2.269*** (0.473)	-1.985*** (0.623)	-1.513*** (0.478)	-1.338*** (0.373)	-1.518*** (0.327)	-0.294 (0.306)
Reentry	-	-1.481 (0.994)	-4.258 (6.124)	-2.284** (0.963)	-1.910* (1.061)	-1.356 (1.025)	-0.845 (1.220)	0.220 (0.731)	-0.203 (0.421)
Number of Programs	+	0.099 (0.060)	0.078 (0.177)	0.090 (0.056)	0.112*** (0.041)	0.085** (0.033)	0.062* (0.032)	0.042 (0.044)	0.066** (0.030)
Sampling1	+/-	-0.615 (1.332)	-1.908 (2.683)	-0.975 (1.857)	0.779 (3.783)	-0.901 (1.606)	0.485 (1.175)	-0.173 (0.927)	-0.280 (0.897)

Sampling2	+/-	1.830 (1.421)	1.713 (3.275)	0.888 (2.063)	2.321 (3.792)	0.643 (1.608)	1.874 (1.217)	0.822 (0.902)	0.274 (0.920)
Sampling3	+/-	-0.352 (1.343)	-1.360 (2.707)	-1.121 (2.017)	1.058 (3.616)	-0.607 (1.567)	0.636 (1.167)	-0.145 (0.911)	-0.428 (0.870)
Sampling4	+/-	-0.037 (1.550)	0.004 (2.550)	-0.937 (2.094)	-0.128 (3.649)	-2.164 (1.622)	-1.122 (1.325)	-1.917 (1.208)	-1.857 (1.147)
IPIA	+/-	-0.707 (0.468)	-0.982 (1.108)	-0.429 (0.414)	-0.285 (0.349)	-0.578* (0.301)	-0.581** (0.266)	-0.183 (0.346)	-0.076 (0.181)
IPERA	+/-	-1.083* (0.576)	-10.481*** (3.083)	-0.461 (0.723)	-0.162 (0.302)	-0.406 (0.309)	-0.579** (0.248)	-0.431 (0.417)	-0.368 (0.224)
Ln (Outlay)	+	1.048*** (0.138)	0.960** (0.463)	0.984*** (0.173)	1.064*** (0.166)	0.939*** (0.137)	0.921*** (0.137)	0.981*** (0.111)	1.119*** (0.087)
R^2		0.4833	0.3515	0.4592	0.4724	0.4781	0.4762	0.4655	0.4483
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T=11.003 P > T = 0.000	T=13.23 P > T = 0.000	T=17.238 P > T = 0.000	T=20.062 P > T = 0.000	T=22.132 P > T = 0.000	T=19.301 P > T = = 0.000	T=16.512 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3 Average Differential Effects in Improper Payments by Administrative Leaders

IP (\$) in Millions		OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805
Baseline Model Estimates									
Career Executive	Exp(β) (Factor Change)	0.205* [0.03, 1.35]	0.259 [0.00, 13.46]	0.462 [0.09, 2.47]	0.181** [0.04, 0.86]	0.231*** [0.09, 0.60]	0.276 [0.05, 1.53]	0.427 [0.13, 1.42]	0.748 [0.41, 1.35]
	% Factor Change	-79.5%	-74.1%	-53.8%	-81.9%	-76.9%	-72.4%	-57.3%	-25.2%
Non- Acting Executive	Exp(β) (Factor Change)	0.122*** [0.03, 0.47]	0.377 [0.04, 3.21]	0.270* [0.06, 1.23]	0.156*** [0.04, 0.60]	0.169*** [0.07, 0.39]	0.190*** [0.07, 0.49]	0.367** [0.16, 0.82]	0.799 [0.48, 1.34]
	% Factor Change	-87.8%	-62.3%	-73%	-84.4%	-83.1%	-81%	-63.3%	-20.1%
Augmented Model Estimates (Inclusion of Executive Tenure and Previous Department Experience Controls)									
Career Executive	Exp(β) (Factor Change)	0.178 [0.02, 1.42]	0.270 [0.00, 35.71]	0.300 [0.04, 2.12]	0.155** [0.03, 0.82]	0.220*** [0.08, 0.63]	0.332 [0.07, 1.54]	0.518 [0.17, 1.62]	0.738 [0.43, 1.26]
	% Factor Change	-82.2%	-73%	-70%	-84.5%	-78%	-66.8%	-48.2%	-26.2%
Non- Acting Executive	Exp(β) (Factor Change)	0.112*** [0.02, 0.52]	0.301 [0.02, 3.74]	0.227* [0.05, 1.12]	0.148*** [0.04, 0.57]	0.179*** [0.07, 0.45]	0.253*** [0.09, 0.70]	0.491* [0.22, 1.08]	0.731 [0.35, 1.51]
	% Factor Change	-88.8%	-69.9%	-77.3%	-85.2%	-82.1%	-74.7%	-50.9%	-26.9%

Note: 95% confidence intervals appear inside brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

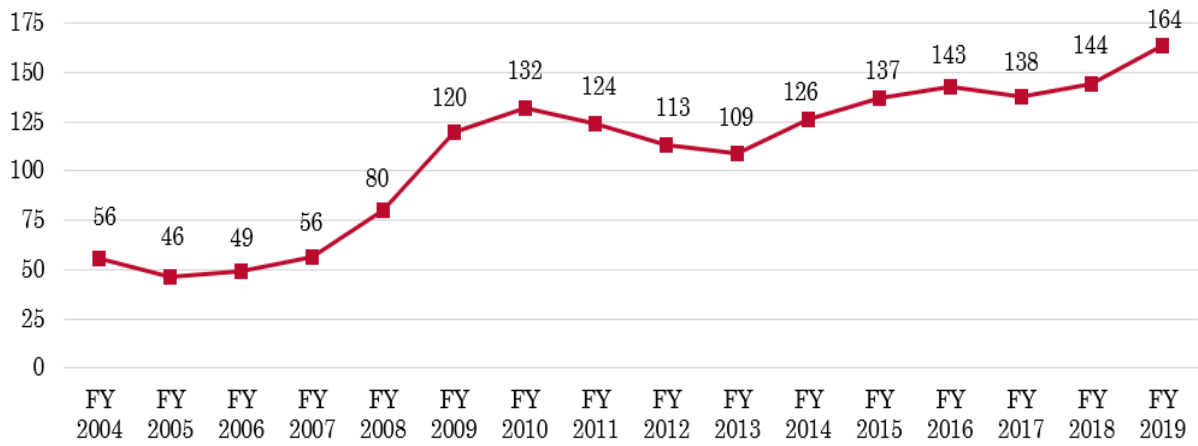
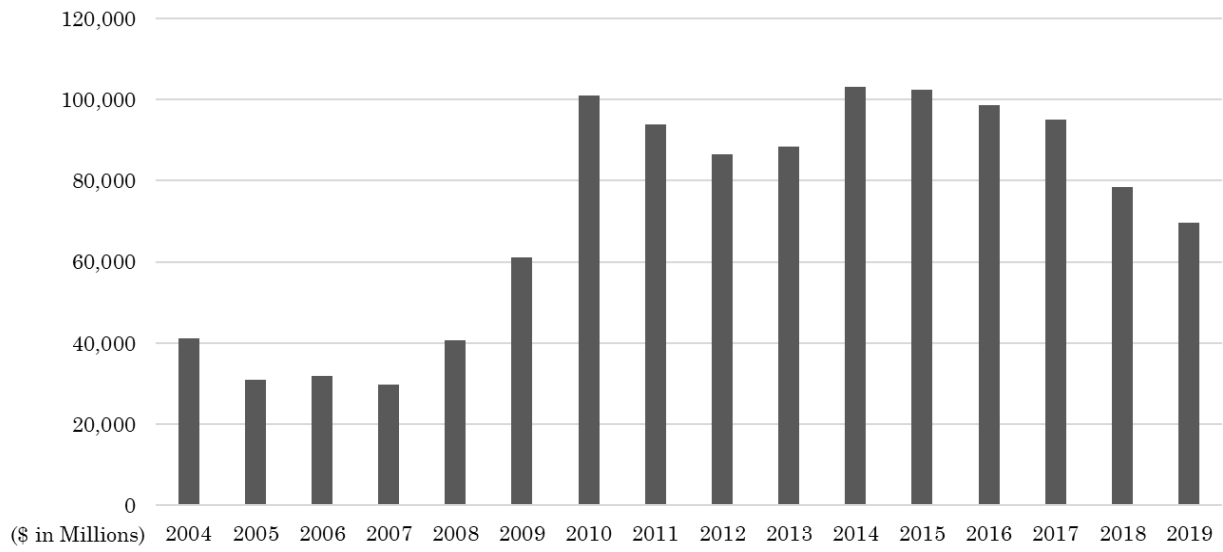


Figure 3.1 The Total Estimated Amounts of Improper Payments in U.S. Federal Government (\$ in Billions, Adjusted to 2015 Constant Dollars)



**Figure 3.2 The Total Size of Improper Payments in Federal Programs
(Per Reporting Year)**

CHAPTER 4

DISENTANGLING IMPROPER PAYMENTS IN U.S. FEDERAL PROGRAMS: DISTINGUISHING BETWEEN TYPE I VERSUS TYPE II PAYMENT ERRORS

Introduction

Improper payments made by federal agencies cost taxpayers significant amounts of money, while also contributing to the fiscal unsustainability of the government (The U.S. Government Accountability Office (GAO) 2012). Since the Bush administration, the efforts to improve financial performance by reducing improper payments have continued. Although a number of executive orders and laws including the Improper Payments Information Act of 2002 (IPIA), the Improper Payments Elimination and Recovery Act of 2010 (IPERA), the Improper Payments Elimination and Recovery Improvement Act of 2012 (IPERIA) have been passed to address the improper payment problem, reducing government-wide improper payments is still one of the key challenges in the U.S. federal government.

While higher payment errors represent poor program performance reflecting the lack of accountability and accuracy, different types of payment errors represent different problems. While overpayment errors represent the monetary loss of taxpayers that could have been used for other critical policy areas, underpayment errors degrade the integrity of federal programs and undermine the public's trust in government (Malenich 2019). Moreover, underpayment errors often make programs less successful since most programs aim to provide adequate financial aid to eligible individuals. One example of an underpayment could be social security recipients who do not get the full amount of aid they

should have received. As these two types of errors create different problems, government efforts to reduce overpayment and underpayment errors imply potentially competing performance goals. While an effort of reducing overpayment errors represents a performance goal of reducing waste and spending, an effort of reducing underpayment errors represents a performance goal of increasing the reach of federal programs for eligible recipients.

This study seeks to evaluate whether the types of administrative leadership make a difference in achieving somewhat conflicting performance goals, which is reducing overpayment errors and underpayment errors even after controlling for the political influence. With performance accountability as well as fiduciary responsibility for program management, top administrative leaders will seek to address improper payment problems by effectively using their capacities. Specifically, programs managed by career executives or non-acting executives are hypothesized as being more likely to yield lower over and underpayment errors compared to those managed by political executives and acting executives. As a logical extension, by applying the logic of Type I and Type II decision errors, this study further evaluates whether these differential effects between career executives and political executives produce the same magnitude for overpayment errors and underpayment errors or asymmetric for the types of errors.

These hypotheses are empirically evaluated using data on 807 total U.S. federal programs covering 56 agencies/subagencies in 2002-2019 time period. By analyzing the different types of payment errors separately, this study seeks to evaluate how administrative leader types affect the direction of payment errors. The statistical evidence reveals that programs managed by career executives are more likely to show better performance via smaller payment errors, both in overpayment errors and underpayment

errors. Also, these findings reveal that programs managed by non-acting (permanent) executives are more likely to yield lower overpayment errors compared to programs managed by acting executives. By separately analyzing overpayment and underpayment errors, this study seeks to consider the tradeoffs between these errors and show which actors are best placed to manage in these conditions.

This essay proceeds as follows. First, this study considers how different types of administrative leaders might affect the overpayment and underpayment errors, and proposes testable hypotheses. Second, the data, variables, and empirical strategies will be discussed. Finally, this study presents and discusses the statistical results and offers concluding remarks regarding the role of administrative leadership types in managing payment decision errors, overpayment and underpayment errors.

Disentangling Improper Payments of U.S. Federal Programs: Overpayment Errors and Underpayment Errors

Improper payments in U.S. federal programs refer to payments made in an incorrect amount, payments that should not have been made at all, or payments made to an ineligible recipient or for an ineligible purpose (Congressional Research Service 2016). Total improper payment amounts refer to the total amounts of overpayment errors and underpayment errors (OMB Circular A-123's Appendix C 2018). Overpayment errors include payments made in incorrect higher amounts and payments that should not have been made at all to ineligible recipients. Medicare, the largest federal health care coverage program provided by the Centers for Medicare and Medicaid Services within the Department of Health and Human Services (HHS), is one clear example of overpayment errors. The report from the HHS inspector general revealed that ineligible recipients

including illegal immigrants and prison inmates received more than \$120 million in Medicare services from 2009 to 2011 (Cheney 2013). Conversely, underpayment errors represent payments made in incorrect lower amounts and payments that should have been made to eligible recipients. For example, in 2005, Congress members expressed concern over the issue of underpayment errors in federal retirement annuities provided by the Office of Personnel Management since retiring federal employees have received annuities much less than what they are owed (Rutzick 2005). Both types of improper payments represent payment decision errors in U.S. federal programs.

Although most payment decisions may be determined by lower-level employees, administrative leaders have a fiduciary responsibility for financial management within their organizations. Stemming from the Accounting and Auditing Act of 1950, the Federal Managers' Financial Integrity Act (FMFIA) of 1982 (Pub. L. No. 97-255, 96 Stat. 814) requires the head of each agency to establish an internal control system to ensure efficient financial operations and "to report annually to the President and the Congress on the condition of agency internal control and accounting systems" (GAO 1989:2). This Act holds administrative leaders accountable to the President and Congress for financial performance. More directly, to manage improper payments in U.S. federal programs, the Improper Payments Elimination and Recovery Act of 2010 (IPERA) clearly specifies the responsibilities of the agency heads in reviewing programs, carrying out the financial management improvement program, and submitting a compliance plan to Congress. In addition, if agencies determine that they do not have sufficient resources to establish and maintain effective internal controls to reduce payment errors, the heads of the agencies can request the resources in their budget submissions (Public Law 111-204 Sec.2. Improper Payments Elimination and Recovery). GAO (2001b: 47) also suggests that the involvement

and support of top-level officials are “especially critical given that an investment of time and money is often needed in these types of efforts.” In other words, administrative leaders have both responsibilities and tools to address improper payment problems. The discretion that administrative leaders have in addressing improper payments in terms of providing training, adopting innovative technologies, and overseeing the overall process is crucial to setting the stage for change. One example in the real world is the case of the Department of Labor’s Office of Unemployment Insurance. To curtail the abuse of unemployment insurance, the administrator decided to create the Unemployment Insurance Integrity Center of Excellence to provide training (e.g., fraud investigation, data analytics, tax auditing) and offer grants for innovative solutions at the state level (The Partnership for Public Service 2018). Gay Gilbert, the administrator at the Department of Labor’s Office of Unemployment Insurance, said in an interview in 2014 that “the federal unemployment insurance program has reduced two of their largest improper payment root cause areas: inaccurate data about why someone is out of work and continued benefits payments after an individual returns to work” since the inception of the center (The Partnership for Public Service 2018: 6). Similarly, the Centers for Medicare and Medicaid Services (CMS) established the Center for Program Integrity (CPI) and the Medicaid Integrity Institute (MII) to provide training and technical assistance for state Medicaid agencies.

Reports from the government and non-profit organizations also emphasize the role of administrative leaders as strategies to manage improper payments. Specifically, in its report “The Buck Stops Here: How Federal Leaders Can Reduce Improper Payments,” The Partnership for Public Service (2018) suggests that agency leaders can bring about change in managing improper payments by using timely and accurate data, engaging relevant stakeholders, and taking a holistic approach. Also, considering a series of GAO reports have

emphasized the role of leaders in creating a control environment to manage improper payments (e.g., GAO 2001a; 2001b; 2020), the key to solving improper payment problems may rest on administrative leaders.

Essay 2 (Chapter 3: “*How Administrative Leaders Affect Administrative Performance: Evidence from Improper Payments in U.S. Federal Programs*”) explores whether different types of administrative leaders, career executives versus political executives and non-acting executives versus acting executives, make a difference in reducing payment errors. Following on from Essay 2, this essay seeks to separately analyze the different types of payment errors and evaluates whether administrative leadership types affect both types of errors, overpayment errors and underpayment errors, in the same way. Since overpayment errors represent the monetary loss of the government and underpayment errors can undermine the public’s trust in the government by degrading the integrity of federal programs, reducing both types of errors is critical for the government. This study posits that programs managed by career executives will yield lower overpayment and underpayment errors than programs managed by political executives since greater program expertise and experience, and different motives that career executives possess will make them better manage the programs. When dealing with improper payments, program knowledge and expertise become even more significant for leaders to create an adequate control environment (GAO 2012). The appointment of political executives in the top position can lead to weakness in the control system within the agency since these “outsiders” lack substantive knowledge about the organization and federal government structures and policies (Moynihan and Roberts 2010). Gailmard and Patty’s model (2007; 2012) of bureaucratic investments in expertise implies that career executives can accumulate knowledge over time, and this knowledge makes them experts in

their management area and thus more effective at achieving goals, assuming they have administrative discretion. Besides the investment in expertise, career executives' job stability also makes them able to consider the long-term effectiveness of the program and statutory objectives and to be responsive to stakeholders not currently in power. Similarly, Lewis (2008) finds that career executives have more direct agency experience and longer tenures, and both factors will improve careerists' knowledge and expertise about their agencies. Since political executives lack "the subject area background or public management qualifications of their careerist counterparts" (Lewis 2012a: 589), greater reliance on political executives at the expense of career executives may indicate the loss of managerial knowledge and expertise. Considering improper payments are payment decision errors, competence embodied in expert knowledge is critical for making key management decisions based on "objective standards" (Bertelli and Lynn 2003; Moynihan and Roberts 2010; see also Kaufman 1956). Career executives, therefore, exhibit more substantive expertise in their agency's assigned programs (Cohen 1998; Gilmour and Lewis 2006; Lewis 2007; 2012a), thus enabling them to be more effective at managing programs, as well as both identifying and resolving problems.

Political executives have different motives than career executives when it comes to managing overpayment errors and underpayment errors. Because political executives are regarded as "short-timers" or "in-and-outers" due to their relatively short tenures, they often bring a short-term perspective and endure political pressure (Ban and Ingraham 1990; Hecl 1977; Pfiffner 1987). Although both political executives and career executives share the same goal to reduce errors, political executives may not have the same impetus to improve their agency's performance as career executives do since they have different time horizons and often focus on further job advancement, either to higher-level political

positions or to the private sector (Brewer and Maranto 2000). These different motives of political executives not only affect their own behaviors but also affect other employees' behaviors within the organization. For instance, Richardson (2019) finds that civil servants who perceive greater politicization will reduce their investment in expertise, suggesting politicization may reduce agency policy expertise. Conversely, career executives have a long-term perspective when dealing with organizational problems and are not affected by short-term political pressures (Pfiffner 1987; Lewis 2012a). Instead of short-term goals intended to benefit stakeholders for partisan and electoral purposes, career executives are insulated from politics and prioritize competent and long-term program management. Since they spend most of their careers in a single organization, career executives are more likely to "be dedicated to their agency's health, mission, and clientele" (Brewer and Maranto 2000: 72; see also Heelo 1977). Career executives' dedication to their organizations will lead them to work towards a reduction in both overpayment and underpayment errors. According to this mechanism, I posit the following pair of corresponding hypotheses:

H1a: *Federal programs managed by career executives will yield lower overpayment errors compared to programs managed by political executives.*

H1b: *Federal programs managed by career executives will yield lower underpayment errors compared to programs managed by political executives.*

The absence of a permanent administrative leader may also adversely affect program performance due to the lack of formal and perceived authority. Acting managers refer to those who temporarily assume the leadership responsibility of the administrative unit. As a consequence of turnover of existing leadership or presidential nomination or confirmation delays, there are cases of leadership vacancies or filling in with acting officials (Chang, Lewis, and McCarty 2001; O'Connell 2008; O'Connell 2020; Ostrander 2016).

Recently, delays in confirming nominations have made acting positions more likely over time. The average Senate confirmation process for presidential appointments takes 112 days under the Obama administration and 117 days under the Trump administration (O'Connell 2021). Also, having temporary office holders appears to be a deliberate strategy for the President as a means to bypass Senate oversight. For instance, President Trump removed many key officials and was open about his preference for temporary (acting) appointees but acting officials not only cannot establish medium-term goals but also cannot offer a vision for the direction of their agency (Moynihan 2021b). Previous studies also have documented that leadership vacancies can lead to poor performance not only because employees' uncertainty of the agency's policy direction can create confusion within the organization, but also because agency inaction can occur with the expectation that agency priorities and directions will eventually change with the permanent leadership (Mendelson 2015; O'Connell 2008; 2015). Their lack of accountability adds more challenge to program management as acting officials temporarily fill the positions and not have been confirmed to the jobs. Furthermore, acting officials are temporary substitutes without the full authority of the office that accompanies Senate confirmation (Dull and Roberts 2009; O'Connell 2008). Besides the lack of formal authority, the lack of perceived authority creates more obstacles for acting leaders to manage their organizations. Acting managers question their own right to make long-term decisions, and they are not perceived by subordinates as having the full authority of the manager (Partnership for Public Service 2020). Without sufficient formal and perceived authority, acting officials may not only struggle to direct subordinates effectively and solve long-term problems within their agency (O'Connell 2008) but may also face challenges in communicating and coordinating with other agencies or other levels of government. Therefore, it is reasonable to surmise that

federal programs administered by acting officials would yield more overpayment and underpayment errors than those administered by permanent leaders. The second set of hypotheses pertain to the consequences of permanent, non-acting versus temporary, acting administrative leaders with respect to program payment errors:

H2a: *Federal programs managed by non-acting executives will yield lower overpayment errors compared to programs managed by acting executives.*

H2b: *Federal programs managed by non-acting executives will yield lower underpayment errors compared to programs managed by acting executives.*

This pair of hypotheses predicts how the degree of stability associated with administrative leadership affects the direction of payment errors by unpacking the aggregated improper payments. In terms of accuracy, overpayment errors and underpayment errors work the same way where both errors undermine the accuracy of government payments. These two types of program payment errors reflect the type of decision errors resulting from public sector organizations (e.g., see Bendor 1985; Landau 1979). **Table 4.1** provides a classification of these types of decision errors applied to improper payments. Overpayment errors in federal programs can be seen as a Type I error since programs assert some recipients as eligible when they are, in fact, ineligible, committing the ‘false positive’ decision errors. Conversely, underpayment errors can be understood as a Type II error since programs assert some recipients as ineligible when they are actually eligible, committing ‘false negative’ decision errors. An extreme strategy of reducing overpayment errors would be paying no one. In that case, there will be no overpayment while the underpayment error amounts will be skyrocketing. An opposite extreme approach would be paying everyone, yielding extremely high overpayment errors and no underpayment. Although these two examples are extreme cases, they imply that

efforts to benefit as many (few) as possible may unintentionally increase the risk of overpayment errors (underpayment errors) and decrease the risk of underpayment errors (overpayment errors). Making eligibility criteria more complex may reduce overpayment errors but may increase underpayment errors while making programs more accessible may increase overpayment errors but decrease underpayment errors. In other words, efforts to reduce these two different types of payment errors imply potentially competing performance goals since there are tradeoffs between these two errors.

As logical extensions of the above hypotheses, **H1a** and **H1b**, this study seeks to further analyze whether there are symmetric or asymmetric average differential effects between career executives and political executives for overpayment and underpayment errors. As a baseline, the *Symmetry Hypothesis* predicts that there will be no difference in the magnitude of the average differential marginal effects for overpayment and underpayment errors. In other words, without any incentives that will increase the risk of committing one type of payment error while decreasing the risk of committing the other type of payment error, the magnitude of average differential effects between careerists and political appointees will be about the same for overpayment errors and underpayment errors because of relatively higher capacity and competence in program management that career executives have.

However, there can be competing explanations that predict the asymmetric differential effects for overpayment and underpayment errors between career executives and political executives since they have different incentive mechanisms. In the centralized government website publishing agencies' reports on improper payments, *PaymentAccuracy.gov*, the government clearly places a priority on dealing with overpayment errors by indicating that "although working to reduce all improper payments

is important, prevention of improper payments resulting in a monetary loss is the federal government's highest priority" (<https://www.paymentaccuracy.gov/about-payment-accuracy/>). In general, political pressures have focused more on reducing overpayment errors which are often framed as "waste, fraud, and abuse" (Herd and Moynihan 2018). For instance, Congress has put pressure on government agencies to be more aggressive in reducing overpayment errors for welfare programs such as social security programs or the Earned Income Tax Credit (EITC) (Rein 2022). News articles reveal that the media also keep an eye on overpayment problems more rigorously. On July 17th, 2020, the New York Times harshly criticized the government for delivering more than a million stimulus payments (worth about \$1.4 billion) to dead people (Rappeport 2020). Considering the heavier burdens that top administrative leaders have for overpayment errors, career executives will have a lower incentive to distribute more funds than the correct amount, which leads to increasing the risk of committing Type II payment errors while decreasing the risk of committing Type I payment errors. In other words, to avoid the blame and being called for congressional investigations for mismanagement, career executives will seek to avoid overpayment errors as much as possible, which may unintentionally increase the risk of underpayment errors.

Although political executives also share the same goal of reducing payment errors, they also have another motive to benefit their constituent groups. Previous studies on distributive policymaking behavior within U.S. federal agencies have shown 'particularistic' and targeted financial resources allocation behaviors for political purposes (e.g., Berry and Gersen 2017; Krause and Zarit 2022; Kriner and Reeves 2015). A political appointment system is thought to be the way to secure a responsive competence since the best way for presidents to ensure achieving their policy goals through bureaucracy is to appoint people

who are loyal to presidents. In other words, presidents can better serve their core constituent groups through the practice of placing political appointees into the top posts in bureaucratic positions (Lewis 2012b; Peters and Pierre 2004).

Therefore, with their motives to appease core constituent groups, political executives will have a greater incentive to distribute more funds, which leads to increasing the risk of Type I payment errors while decreasing the risk of Type II payment errors. That is, their efforts to distribute more money to their constituent groups through federal programs may increase the risk of overpayment errors while decreasing the risk of underpayment errors. This study does not contend that political executives will intentionally and directly instruct employees to make overpayments. However, considering most decisions involve uncertainty, we suggest that decision-makers often have to choose between increasing the risk of overpayment errors and increasing the risk of underpayment errors, and political executives are more likely to increase the risk of Type I payment errors. Because of these different motives that administrative leaders have for underpayment problems (i.e., a greater incentive for career executives to distribute fewer funds versus a lower incentive for political executives to distribute fewer funds), the magnitude of the average differential effects between career executives and political executives will be smaller for underpayment errors. Conversely, since there is a lower incentive for career executives to distribute more funds vis-à-vis a greater incentive for political executives to distribute more funds, the magnitude of the average differential effects between types of administrative leaders will be greater for overpayment errors. This study seeks to evaluate the following competing perspectives:

H3a (Symmetry Hypothesis): *There will be no difference in the magnitude of the average differential effects between administrative leaders for overpayment and underpayment errors.*

H3b (Asymmetry hypothesis): *The magnitude of the average differential effects between administrative leaders for overpayment errors will be greater compared to underpayment errors.*

Data, Measures, and Empirical Strategy

The data used to evaluate the propositions are the size of overpayment errors and underpayment errors in 162 federal programs administered by 56 U.S. federal agencies/subagencies over a 2002–2019 period. The unit of analysis is program-level, with 807 total program-year observations. The sample of programs in this study solely includes programs administered by a single agency and programs reporting improper payment estimates (both overpayment errors and underpayment errors) in a one-year cycle. Nonetheless, these data availability constraints offer the advantage for evaluating the effects of administrative leadership types since the sample of data does not suffer from confounding factors attributable to the programs managed by multiple agencies or/and programs managed by multiple administrative leaders.

Dependent Variables

While Essay 2 (Chapter 3) focuses on the aggregate amounts of improper payments (the sum of the absolute value of overpayment error amounts and underpayment error amounts) to assess the magnitude of total payment errors, this essay focuses on the amounts of overpayment errors and underpayment errors separately to assess the direction

of the payment errors. Not only are required to report the total amounts of improper payments, but federal agencies are also required to report the estimated amounts of overpayment errors and underpayment errors for the programs that are identified as susceptible to “significant” improper payments.³¹ Each agency reports these data either in its annual *Performance and Accountability Report (PAR)* or *Agency Financial Report (AFR)* and is also required to report improper payments information to the Office of Management and Budget (OMB) for publication on the centralized government website, *PaymentAccuracy.gov*. One of the biggest and best-known federal programs covered in this study is Supplemental Security Income (SSI) program administered by Social Security Administration (SSA). In 2019, for example, SSA reported 4.4 billion dollars for overpayment errors and 787 million dollars for underpayment errors in SSI program. There are many smaller federal programs required to report payment errors, such as Milk Income Loss Contract Program administered by Farm Service Agency, Assistance to Firefighters Grant Program administered by Federal Emergency Management Agency, and Telecommunications Relay Service administered by Federal Communications Commission. The dependent variables of interest are measured as the natural logarithm of over and underpayment error dollar amounts (adjusted for 2015 constant dollars) to adjust for the highly skewed and leptokurtic distribution of variables.³² Dependent variables subject to a

³¹ There is no separate reporting requirement (criterion) for overpayments and underpayment errors. Agencies are required to report both overpayment and underpayment errors for programs that are identified as susceptible to “significant” overall improper payments. For the definition of “significant” improper payments, please see Footnote 13.

³² The skewness and kurtosis of overpayment and underpayment errors are respectively as follows: 1) the skewness and kurtosis of overpayment errors: 6.672 and 54.186; 2) the skewness and kurtosis

natural log-transformation also allow for interpreting coefficients as percentage changes associated with the dependent variable based on a unit change in a given covariate (Wooldridge 2010).

Key Covariates

The first primary covariate employed to evaluate the hypotheses centered on administrative leadership is *Career Executive* defined as the appointment type of an administrative top official managing a given program in a given year. To address the cases where two or more top officials manage the program in one year due to leadership turnover, this study creates a weighted measure bounded between 0 and 1. For instance, when a top official A serves 4 months and the other top official serves 8 months during a one-year period, the appointment type (i.e., career executive and political executive) is calculated based on the months that each official serves. While 0 indicates all the administrative managers in a given program in a given year are political executives, 1 indicates the administrative manager(s) as career executives. **H1a** and **H1b** predict that this covariate is hypothesized as being negatively associated with overpayment and underpayment errors, respectively.

The other primary covariate for evaluating these administrative leadership hypotheses, *Non-Acting Executive*, is defined as whether an administrative manager assumes the responsibility of the job position on a permanent basis for a given program in a given year. Along with the *Career Executive* variable, this measure is also a weighted

of underpayment errors: 7.573 and 72.282. After transforming the dependent variables into natural logged forms, the skewness and kurtosis of overpayment errors decrease to -1.362 and 4.048, respectively. Also, the skewness and kurtosis of underpayment errors decrease to -0.129 and 1.374, respectively.

measure bounded between 0 and 1, where 0 indicates all the administrative managers in a given program in a given year are acting executives, while 1 indicates all the administrative managers are non-acting executives. Both **H2a** and **H2b** predict that this covariate should be negatively associated with both overpayment and underpayment errors, respectively.

Control Variables

In addition to the primary covariates of interest, the statistical models control for the external political influence that can account for systematic variations in over and underpayment errors. Since this essay focuses on the direction of the payment errors by separately analyzing overpayment and underpayment errors, political motives can play a crucial role in making different types of errors and are thus important to consider as alternative explanations.

In pursuit of the president's policy goals, presidents often exert control over federal agencies. While some agencies' missions and activities share presidents' priorities and preferences, other agencies do not. When dealing with administrative agencies, presidents have two different motives depending on whether agencies' activities are ideologically aligned with the president or opposed to the president (Hudak 2014; Krause and O'Connell 2019; Krause and Zarit 2022; Lewis 2008; 2012b). With ideologically aligned agencies, presidents seek to strengthen the work of agencies that share similar policy preferences so that presidents can benefit their core partisan constituent groups. For instance, presidents often use executive branch discretion over making the targeted allocation of financial resources, such as federal funds and grants for their purpose of electoral interests (Berry and Gersen 2017; Hudak 2014). In other words, presidents are 'particularistic' in that they favor certain policies that can disproportionately benefit their constituencies (Kriner and

Reeves 2015). This motive is also called a ‘distributive motive’ in that presidents seek to ensure delivering the agencies’ policy benefits to the core constituent groups (Krause and Zarit 2022). In other words, presidents are more likely to use the tools and resources that they have, to exert their agenda control. Conversely, with ideologically opposed agencies, presidents are more likely to offer relatively less support for the organizations since their policy activities are relatively less likely to represent the president’s core constituent groups. Since presidents prioritize the needs of their partisan constituencies over others, a significant difference in presidents’ support is found between agencies (Kriner and Reeves 2015). In other words, presidents have a ‘substitution motive’ in which they seek to shift agencies’ policy directions to their preferred directions (Krause and Zarit 2022). Also, presidents often act as veto players and seek to limit the resources and support in order to control ideologically opposed agencies.

Hudak (2014) also suggests the ideological congruence between an agency and the president as an important condition for bureaucratic responsiveness since ideological alignment eases the process and the production of policy that is consistent with presidential preferences. With presidents’ support on ideologically aligned agencies’ agendas, agencies will aggressively take initiative and pursue its goal. In the case of ideologically opposed agencies, on the other hand, these agencies will be less responsive since agencies do not share the same interests with the president. Without presidents’ support, ideologically opposed agencies will be more likely to conserve their resources by not taking action actively. Based on presidential studies of bureaucratic control, Rezaee (2016) indicates that since the president wants to veto the agency’s initiative if the agency and the president are opposed, agencies will rather conserve their resources. When applying this logic to improper payment cases, programs that are housed in presidential aligned agencies will be more

likely to be responsive to the president's interests and make an effort to provide payments to eligible recipients as many as possible with presidential support. This motivation inevitably bears the risk of overpayment errors and avoids the risk of underpayment errors. On the other hand, it is hard to expect that programs that are housed in presidential opposed agencies will get presidential support, so these programs will be rather less responsive, meaning exerting small efforts to provide payments to eligible recipients. Thus, there is a possibility that this movement inevitably bears the risk of underpayment and avoids the risk of overpayment. For this reason, *President-Agency Alignment* variables are included to ensure that these political characteristics are not confounding the effects of key covariates. To measure whether an agency's activities are compatible with the president in ideological terms (*Aligned Agency*), neutral (*Neutral Agency*), or at odds with the president (*Opposed Agency*), 95% Bayesian credibility estimates for the agency ideology scores developed by Clinton and Lewis (2008) are used.³³ For example, *Aligned Agency* variable equals 1 if the president is a Republican and the agency mission is conservative or if the president is a Democrat and the agency mission is liberal, and otherwise 0. By employing *Neutral Agency* group as a reference group, two separate binary variables, *Aligned Agency* variable and *Opposed Agency* variable, are included in the statistical model.

As another means of accounting for the political influence over the bureaucracy, this essay employs the first dimension of Selin's (2015) latent measure of agency insulation from

³³ Clinton and Lewis (2008) used expert surveys and measurement models to estimate agency ideology scores and to categorize agencies by perceived ideology (Liberal, Moderate, and Conservative). Based on Clinton and Lewis (2008) measure, the sample of data is divided as follows: Aligned Agencies (N=185, 22.92%), Neutral Agencies (N=361, 44.73%), and Opposed Agencies (N=261, 32.34%).

politics. This *Agency Insulation* variable refers to the independence level of key decision-makers within agencies determined by organizational structures such as the terms of appointment, agencies' location in the bureaucracy, or constraints on the ability of political principals to remove the top officials (Selin 2015). Higher values indicate greater independence of key decision-makers from political influence. Increases in this covariate should be associated with fewer overpayment errors and underpayment errors since agencies where organizational structures limit the political control are less likely to make improper payments derived from political motives.

In addition to accounting for external political influence, the statistical models control for other program/agency-specific characteristics that may affect both the payment errors in the same direction as employed in Essay 2. First, *Executive Turnover* for a given program in a given year as a dummy variable is included in the statistical model. Considering various pernicious effects of executive turnover on the organization such as leadership vacuums, working relationship disruption (Bertelli and Lewis 2012), or lack of continuity (Gilmour and Lewis 2006), a positive coefficient sign is hypothesized both for overpayment and underpayment errors. Additional covariates are considered to account for program-specific reporting requirements. Programs either receiving disaster relief funds or being originally designated as high-risk programs under the law are required to report the improper payments regardless of meeting "significant" high-risk criteria. To be specific, for FY2014 and beyond, the Disaster Relief Appropriations Act 2013 (Pub. L. No. 113-2) required all programs receiving disaster funds report improper payments. *Disaster Fund*, a binary indicator indicating whether the program received supplemental disaster relief administrative funds is included, and it should be negatively correlated with both overpayment errors and underpayment errors since the affected programs are required to

report regardless of meeting high-risk criteria. *Section 57* binary variable indicating whether programs were originally designated as high-risk programs under the former Section 57 of OMB Circular No. A-11 is included. Since the fact that programs are listed under the former Section 57 represents the high possibility of erroneous payments, these programs are expected to yield both higher overpayment errors and underpayment errors.

Programs' previous improper payment history should affect the current level of improper payments. *Exit* and *Reentry* of the program for a given year are incorporated as separate binary indicators to capture such effects. Based on the previous history, the Office of Management and Budget (OMB) determines whether to give programs a relief from reporting payment errors for the next several years (*Exit* variable) or an obligation to report them again (*Reentry* variable) (OMB Memorandum M-15-02, Appendix C to Circular No. A-123 2014). Both variables are posited to be negatively related to both overpayment errors and underpayment errors since the *Exit* variable directly indicates that these programs have produced sufficiently low payment errors in the current period (and previous years) to get relief from reporting. Also, even though the *Reentry* variable indicates these programs are again considered to produce "significant" improper payments, reentered programs will be more likely to show lower improper payments in both overpayment errors and underpayment errors, considering their historical evidence of low improper payments compared to programs that had never been granted relief from reporting.

In addition, by using separate binary indicators, the statistical model accounts for possible systematic variations in improper payment size (both overpayment errors and underpayment errors) affected by sampling methods since most federal programs calculate their improper payments through samples instead of tallying all the payments made in a given year. Since the accuracy of payment error estimates may vary depending upon the

level of robustness of the sampling methods, controlling for different sampling methods is crucial for obtaining accurate, unbiased estimates of primary covariates. By identifying the ‘actual’ group (100% of actual payments) as a reference group, four separate binary indicators are included in the statistical model ranging from the least robust sampling method to the most robust sampling method.³⁴ The nature of this relationship is ambiguous. Agencies may employ the least robust sampling method to exploit the data and report fewer errors (i.e., negative relationships with both overpayment errors and underpayment errors). Conversely, this sampling method may unintentionally draw biased samples resulting in large errors (i.e., positive relationships with both overpayment errors and underpayment errors).

The statistical models also account for the possibility that the overpayment and underpayment error amounts are influenced by the changes in reporting requirements for improper payments under different Acts during the sample period. The three major pieces of legislation on improper payments between FY2004 and FY2019 are IPIA 2002, IPERA 2010, and IPERIA 2012, which changed the definitions of key terms (such as ‘significant

³⁴ As employed in Essay 2, four separate binary indicators are as follows: 1) sampling1: +/- 3% or better margin of error at the 95% confidence interval (+/- 2.5% or better margin of error at the 90 % confidence interval); 2) sampling2: a wider than +/- 3% or margin of error at the 95% confidence interval (a wider than +/- 2.5% margin of error at the 90% confidence interval); 3) sampling3: statistical method, but not enough information; 4) sampling4: no information about sampling method. The number of programs that fit in each “sampling” category is as follows: 29 programs (100% payments as a reference group), 497 programs (sampling1), 71 programs (sampling2), 207 programs (sampling3), and 3 programs (sampling4).

improper payment³⁵) and reporting requirements.³⁵ By taking the *IPERIA* years (programs reporting improper payments in FY2013-FY2019) as a reference group, an *IPIA* binary variable indicating programs reporting improper payments in FY2004-FY2010, and an *IPERA* binary variable indicating programs in FY2011-2012 are included in the statistical model.³⁶ Considering that agencies with more program management responsibilities will have more difficulties in managing multiple programs, the number of programs reporting improper payments within a given administrative unit (agencies/subagencies) for a given year is also included (*Number of Programs* variable) in the models. This ordinal variable should be positively associated with both overpayment errors and underpayment errors. Lastly, the natural log of total outlays (*ln[outlay]*, adjusted for 2015 constant dollars) is included as a control covariate to account for the scale effects and is expected to produce a positive relationship with both overpayment errors and underpayment errors. In the **Appendix, Table A-1** and **A-2** report descriptive statistics and correlations for all variables.

Empirical Strategy

This essay first follows the same empirical strategies that are employed in Essay 2. First, ordinary least squares (OLS) regression analysis is employed to evaluate the change in the conditional mean of the underpayment and overpayment error dependent variables. Then, the hypotheses are further analyzed based on the varying magnitudes of the

³⁵ For the detailed information about changes in reporting requirements of each legislation, please check Footnote 13.

³⁶ The number of programs that fit in each category is as follows: 524 programs for *IPERIA* (as a reference group), 183 programs for *IPIA*, and 100 programs for *IPERA*.

overpayment and underpayment errors employing a quantile regression modeling approach that accounts for intra-unit correlation among program payment errors via a cluster-based correction of the standard errors (see Parente and Silva 2016). Quantile regression offers a complementary approach to conditional mean regression models by modeling the entire empirical distribution function of the outcome variables. This is a critical issue since OLS models not only offer a biased estimate of the conditional mean in the presence of widely dispersed data such as those from U.S. federal program payment errors but also fail to capture the changing nature of statistical relationships as the dependent variable changes in magnitude (Koenker and Hallock 2001). Therefore, the quantile regression analyses allow for potential heterogeneous administrative leadership effects that substantially vary across highly variable amounts of overpayment and underpayment errors that cannot be gleaned from average values of these outcome measures.³⁷

A series of alternative quantile regression models are employed to evaluate the sensitivity of the primary estimates of interest in this study. Because many program-year observations report neither overpayment error (11.77% of sample observations) nor underpayment error (38.29% of sample observations), these models are also estimated using a censored quantile regression (CQR) estimation technique developed by Chernozhukov and Hong (2002) to account for the substantial fraction of zeroes in the data

³⁷ The conditional means only characterize the 34th quantile of overpayment error amounts and 42nd quantile of underpayment error amounts in the sample of data. Also, the distributions of overpayment errors and underpayment errors in the sample of data widely vary. For instance, while zero overpayment error accounts for 11.77%, zero underpayment error accounts for 38.29%. Furthermore, the 97th percentile of overpayment error represents \$10,265 Million, but the 97th percentile of underpayment error represents \$1,083 Million.

for the dependent variables that may yield biased and inconsistent estimates (Greene 1997). This particular method handles censoring semi-parametrically through the conditional quantile function. The CQR model reveals statistical patterns that are substantially similar to those presented in the manuscript with few exceptions.³⁸ For **H1a** (comparing career executive-run and political executive-run programs for overpayment errors), the CQR model produces substantively similar findings in terms of point estimates and statistical inferences, with less precision (changed from 99% significance level to 95% significance level) for programs with moderate overpayment error amounts ($\tau = 0.55$, overpayment error amounts = \$31 million). For **H2a** (comparing non-acting executive-run and acting executive-run programs for overpayment errors), the CQR model reveals similar findings with less precision for programs with small overpayment error size ($\tau = 0.25$, overpayment error amounts = \$0.88 million), but more precision for programs with very small overpayment error size ($\tau = 0.10$, overpayment error amounts = \$0).

In addition, sensitivity analyses are performed by employing the OLS and quantile regression models with bootstrapped standard errors consisting of 500 resamples of the residuals with replacement. The results, reported in the supplementary appendix, are similar but estimated with greater precision.³⁹ Specifically, for **H1a** (career executive-

³⁸ This robustness check appears in *Appendix Table A-3. Alternative Censored Quantile Regression (CQR) Estimates for U.S. Federal Program Overpayment Errors*. Due to model convergence problems, only the overpayment model obtained model convergence.

³⁹ Please see *Appendix Table A-4. Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors (Estimated with Bootstrapped Standard Errors)*, *Table A-5. Alternative Regression Model Estimates for U.S. Federal Program Underpayment Errors (Estimated with*

managed programs versus political executive-managed programs for overpayment errors), the model produces the same results for point estimates and statistical inferences with more precision at the conditional mean ($p = 0.035$ versus $p = 0.226$) and for programs with small overpayment error amounts ($\tau = 0.25$, overpayment error amounts = \$0.88 million: $p = 0.096$ versus $p = 0.368$). Similarly, for **H1b** (non-acting executive-managed programs versus acting executive-managed programs for overpayment errors), the precision of the estimates is higher for programs with small overpayment error amounts ($\tau = 0.25$, overpayment error amounts = \$0.88 million: 95% significance level versus 90% significance level) and large overpayment error amounts ($\tau = 0.85$, overpayment error amounts = \$512 million: $p = 0.081$ versus $p = 0.153$). Lastly, for the model evaluating the difference between career executives and political executives in reducing underpayment errors (**H2a**), the model with bootstrapped standard errors is estimated with more precision for moderate underpayment error amounts ($\tau = 0.40$, underpayment error amounts = \$3,000: $p = 0.096$ versus $p = 0.244$) and fairly large underpayment error amounts ($\tau = 0.70$, underpayment error amounts = \$4.6 million: $p = 0.006$ versus $p = 0.042$).

Lastly, to evaluate whether different presidential administrations affect the types of payment errors in a different way, alternative analyses are performed by including dummy indicator for *Republican President* (Democratic President: 0 and Republican President: 1) instead of *President-Alignment* variables in the statistical model.⁴⁰ There is a possibility

Bootstrapped Standard Errors), and *Figures A-1 & 2. Alternative Quantile Regression Model Estimates Using Bootstrapped Standard Errors*.

⁴⁰ Please see Appendix *Table A-5. Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables)* and *A-6. Alternative Regression Model Estimates for U.S. Federal*

that overpayment errors may decrease while underpayment errors may increase under republican Presidents because of their particular concern about waste and spending. The results indicate that different presidential administrations do not show a significant association with both overpayment errors and underpayment errors in most quantiles, but show a strong negative relationship with very large overpayment errors ($\tau = 0.97$, OP= \$10.265 billion). This finding suggests that overpayment errors may decrease under republican President for programs with exceptionally large overpayment errors. This alternative model reveals substantially similar statistical patterns to those presented in the manuscript for the key covariates.

Empirical Findings

The regression model estimates based on 807 program-year observations appear in **Table 4.2** for overpayment errors and **Table 4.3** for underpayment errors. Besides OLS estimates, the quantile regression models are estimated for the following quantiles of the distribution of respective overpayment and underpayment error dollar amounts corresponding to the sample: $\tau = 0.10$: *overpayment errors* — \$0, *underpayment errors* — NA⁴¹; $\tau = 0.25$: *overpayment errors* — \$884,550, *underpayment errors* — \$0; $\tau = 0.40$: *overpayment errors* — \$8 million, *underpayment errors* — \$3,000; $\tau = 0.55$: *overpayment errors* — \$31 million, *underpayment errors* — \$550,065; $\tau = 0.70$: *overpayment errors* —

Program Underpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables).

⁴¹ Due to the significant fraction of zeroes in the underpayment errors of the sample (about 38%), this essay does not estimate the effect on the 10th quantile to avoid redundancy.

\$112 million, underpayment errors — \$4.6 million; $\tau = 0.85$: overpayment errors — \$512 million, underpayment errors — \$77 million; $\tau = 0.97$: overpayment errors — \$10.265 billion, underpayment errors — \$1.083 billion. The quantiles are delineated in the same way for both overpayment errors and underpayment errors to assess the effects of primary covariates across the different sizes (small, moderate, large) of payment error types.

Unsurprisingly, the model estimates reveal that program outlays are positively associated with both overpayment errors and underpayment errors not just on the conditional mean but also in almost every quantile, accounting for scale effects. Besides the program outlays, both programs receiving disaster relief funds and programs being granted relief from reporting for the next fiscal year are negatively associated with both overpayment errors and underpayment errors. Specifically, those programs are associated with a significant decrease in overpayment errors in every quantile. For underpayment errors, while programs receiving disaster relief funds are more likely to produce fewer errors for moderate underpayment error sizes ($0.55 \leq \tau \leq 0.70$, $\$550,065 \leq$ underpayment error amounts \leq \$4.6 million), programs being granted relief are negatively associated with moderate underpayment error sizes ($\tau = 0.40$, underpayment error amounts = \$3,000; $\tau = 0.70$, underpayment error amounts = \$4.6 million) and exceptionally large size ($\tau = 0.97$, underpayment error amounts = \$1.083 billion). Further, the statistical results reveal that programs within the agency that has more programs to report improper payments and thus has more management responsibilities yield higher overpayment errors at small to fairly large quantiles ($0.25 \leq \tau \leq 0.70$, $\$884,550 \leq$ overpayment error amounts \leq \$112 million). Conversely, this covariate does not show a significant association with underpayment errors in most quantiles but shows a positive relationship with extremely large underpayment errors ($\tau = 0.97$, underpayment error amounts = \$1.083 billion). As shown in

Essay 2 evaluating overall improper payments, executive turnover also shows a strong positive relationship with very large overpayment errors ($\tau = 0.97$, overpayment error amounts = \$10.265 billion), indicating that leadership turnover within an organization may hurt performance by yielding more overpayment errors, especially for programs with extremely large overpayment errors.

Since this essay separately analyzes overpayment errors and underpayment errors rather than evaluating overall improper payments as in Essay 2, this essay accounts for the external political influence that may affect over and underpayment errors in a different direction. For programs that are housed in presidential aligned agencies, the results show no statistical relationship with overpayment error amounts but as expected, the results reveal a strong negative association with underpayment errors not just at the conditional mean but also at moderate ($\tau = 0.40$, underpayment error amounts = \$3,000) to large underpayment error amounts ($0.70 \leq \tau \leq 0.97$, \$4.6 million \leq underpayment error amounts \leq \$1.083 billion). These results provide partial evidence that programs that are housed in presidential aligned agencies may be more likely to make an effort to provide payments to eligible recipients as many as possible with presidential support, which inevitably avoids the risk of underpayment. However, for programs that are housed in presidential opposed agencies, the results show a positive relationship for large overpayment error amounts ($0.55 \leq \tau \leq 0.70$, \$31 million \leq overpayment error amounts \leq \$112 million), meaning programs that are opposed to presidential ideology yield more overpayment errors. One possible explanation for this finding is that these agencies may have a stronger motive to serve their eligible groups as much as possible through their programs, regardless of ideology alignment with the president. This issue needs to be more explored in the future. Finally, although programs that are housed in the politically insulated agencies (measured

by Selin's 1st dimension) do not show a strong statistical relationship with both overpayment errors and underpayment errors, these programs show an expected negative relationship with small overpayment error amounts ($\tau = 0.10$, overpayment error amounts = \$0) and moderate underpayment error amounts ($\tau = 0.40$, underpayment error amounts = \$3,000).

Next comes the statistical analyses of the main hypotheses explaining how leadership type differences can affect overpayment and underpayment errors in U.S. federal programs. **H1a** and **H1b** predict that programs managed by career executives will yield lower overpayment errors and underpayment errors than programs managed by political executives. Although *Career Executive* covariate misses statistical significance at the conditional mean of both overpayment errors and underpayment errors, the results suggest that programs managed by career executives yield much lower overpayment errors at moderate to large overpayment error amounts ($0.40 \leq \tau \leq 0.70$, \$8 million \leq overpayment error amounts \leq \$112 million) (**H1a**) and also yield lower underpayment errors at fairly large underpayment error amounts ($0.70 \leq \tau \leq 0.85$, \$4.6 million \leq underpayment error amounts \leq \$77 million) (**H1b**), compared to programs managed by political executives.

In order to understand the substantive magnitude of the effects, **Table 4.4** and **Figures 4.1 & 4.2** provide the average differential effects in overpayment errors and underpayment errors by administrative leaders. When comparing programs managed by political executives and programs managed by career executives, the results show that the numerical effects of this covariate decrease when comparing programs at moderate overpayment error size ($\tau = 0.40$, overpayment error amounts = \$8 million) and programs at larger overpayment error size ($\tau = 0.70$, overpayment error amounts = \$112 million). Specifically, we would expect to see lower overpayment errors (by about 83.1%) in career

executive-managed programs when compared to political executive-managed programs at moderate overpayment error size, and about 77% lower overpayment errors in career executive-run programs at larger overpayment error size. Conversely, for underpayment errors, the numerical effect of this covariate is larger at a large underpayment error size. While career executive-run programs yield about 77.2% lower underpayment errors compared to political executive-managed programs at fairly large underpayment error size ($\tau = 0.85$, underpayment error amounts = \$77 million), about 70.5% lower underpayment errors are more likely to occur in career executive-managed programs at 70th quantile of underpayment size (underpayment error amounts = \$4.6 Million).

Next, the statistical results do not provide support for **H2b**, but show strong support for **H2a** indicating that programs managed by non-acting executives yield lower overpayment errors at the conditional mean (OLS result) and at small to fairly large quantiles ($0.25 \leq \tau \leq 0.70$, $\$0.88 \text{ million} \leq \text{overpayment error amounts} \leq \112 million). The numerical estimates reveal that the largest effect (about 85.6% lower) appears at the conditional mean of overpayment size. Among quantile regression results, the effect size is the largest at the moderate IP size ($\tau = 0.55$, overpayment error amounts = \$31 million: about 79.2% lower) and the smallest at the small overpayment error size ($\tau = 0.25$, overpayment error amounts = \$0.88 million: about 66.2% lower).⁴²

Although **Table 4.4** and **Figure 4.1** show the average marginal differential effects in overpayment errors and underpayment errors by administrative leaders, a statistical test is needed to evaluate whether career executives relative to political executives are

⁴² Please see **Figures 4.1 and 4.2** for statistically significant numerical effects at the corresponding estimated quantile.

equally indifferent to overpayment and underpayment errors of similar magnitudes (**H3a** and **H3b**). To test whether the average marginal differential effects for corresponding overpayment error quantiles are the same as the average differential effects for corresponding underpayment error quantiles, the between-model covariances of parameter estimates should be obtained, but the quantile regression model with clustering does not provide these covariance matrices. As an alternative, this study instead uses the quantile regression model with bootstrapped standard errors that are employed for the sensitivity analysis. Wald test result suggests that the average marginal differential effects for corresponding overpayment error quantiles are not different from those for corresponding underpayment error quantiles, which supports **H3a** (Symmetry Hypothesis).⁴³ In other words, the magnitude of the average differential effects between administrative leaders for overpayment errors is not significantly greater compared to underpayment errors. This issue is worth to be explored in future research to see if the results remain the same when using the quantile regression model with clustering.

Discussion

The U.S. federal government has made continuous efforts to identify, report, and reduce improper payments in the government's programs and activities. The Biden administration is also committed to greater accountability and efficiency by improving federal payments in government programs. The recent brief of the U.S. Office of

⁴³ 25th quantile: $\chi^2 \sim (1) = 0.00$, $p = 0.969$; 40th quantile: $\chi^2 \sim (1) = 0.12$, $p = 0.730$; 55th quantile: $\chi^2 \sim (1) = 0.16$, $p = 0.690$; 70th quantile: $\chi^2 \sim (1) = 0.19$, $p = 0.662$; 85th quantile: $\chi^2 \sim (1) = 1.72$, $p = 0.190$; 97th quantile: $\chi^2 \sim (1) = 0.00$, $p = 0.990$

Management and Budget (OMB) in the Biden administration shows the administration's willingness and commitment to address this issue:

“Here at OMB, we’ve also been working closely with the oversight community—including IGs and the Government Accountability Office —alongside agencies across Government to make payment integrity a top priority, particularly when it comes to payment errors that result in monetary loss. (...) President Biden has been very clear that this Administration will take decisive action to ensure our Government delivers results for the American people in the most effective, equitable, and accountable way possible. That means committing to facing the real challenges of combatting payment errors and making Government work better. Sharing this data transparently and being direct about the challenges we face are part of that commitment.” (OMB 2021: Briefing Room: Updated Data on Improper Payments, Paragraphs 7-8).

Despite the persistent efforts, tackling improper payments of U.S. federal programs is still a challenging issue. In Essay 2, the study has sought to understand why some federal programs make more overall payment errors than other programs and explore the central role of the administrative leadership types. The study finds that career executive-run programs and non-acting executive-run programs are more likely to yield lower overall improper payments than political executive and acting executive-run programs. Still, little is known from Essay 2 regarding how the administrative leadership types affect the direction of payment errors. Since there is a potential tradeoff between overpayment errors and underpayment errors (i.e., efforts to reduce overpayment errors can unintentionally increase underpayment errors and vice versa), disaggregating total payment errors into overpayment and underpayment errors can further advance our understanding of improper payments. Therefore, following on from Essay 2 and its logical implications, this essay

separately analyzes overpayment and underpayment errors and seeks to evaluate whether the types of administrative leadership still matter for reducing both types of payment errors even after controlling for external political influence. Since external political influence can affect overpayment and underpayment errors in a different direction, accounting for this influence is crucial for this study. Furthermore, this study seeks to evaluate whether there is a symmetric or asymmetric effect in the magnitude of the average differential effects between career executives and political executives for overpayment and underpayment errors.

Even after controlling for external political influence, the evidence offered in this study provides compelling support that programs managed by career executives yield lower overpayment errors and underpayment errors compared to programs managed by political executives. This finding adds one more crucial piece of evidence to the argument of career executives and political executives' performance management. By showing that programs managed by career executives yield lower overpayment errors and underpayment errors instead of showing partial effects or opposite directions, this study suggests that career executives' capacity in terms of expertise, knowledge, and long-term perspectives are found to be more effective in managing improper payments.

With regard to the administrative leadership type differences between non-acting executives and acting executives, this study also finds partial support in that programs managed by non-acting executives are more likely to yield lower overpayment errors compared to programs managed by acting executives. Given that the overpayment problem is a more prioritized problem in the government because of the nature of the problem (e.g., monetary loss of taxpayers), this finding raises an important question on how we can bridge the performance gap between non-acting executives and acting executives. Since there are

always acting executives at some point due to leadership turnover or vacancies, it is crucial to appoint qualified acting executives with enough experience and knowledge to mitigate the performance gap.

Unfortunately, this study has some limitations. Although the statistical models of this study account for the possible systematic variations driven by sampling methods, there is still a possibility that agencies may not report the accurate amounts of payment errors (both overpayment and underpayment error amounts). One possible way to deal with this problem would be by looking into each agency's annual audit reports, but this strategy cannot fully rule out the possibility of inaccurate information reported by agencies. Additionally, although this study shows that programs managed by career executives and non-acting executives have lower payment errors, this study cannot ascertain the precise manner by which administrative leaders influence program payments. However, from the Federal Managers' Financial Integrity Act (FMFIA) of 1982 and the major pieces of legislation on improper payments (i.e., IPIA, IPERA, and IPERIA), we can infer that agency heads not only have the responsibility of managing improper payments but also have some discretion to use their resources (e.g., hiring the right people, technology adoption, oversight, monitoring) to reduce the errors. Also, several anecdotes have revealed that the type of administrative leaders makes a difference in managing improper payments. For instance, the head of the Office of Cybersecurity, Energy Security and Emergency Response (CESER) within the Energy Department, who was a political executive during 2018-2019, was criticized for her mismanagement of improper payments, and the inspector general pointed out a "lack of established internal controls" as a critical reason (Rubin 2021). The results presented here offer a useful starting point by suggesting that the administrative leadership type does matter for managing improper payments, but it would be desirable to

dig deeper into how the administrators actually influence the payments in future research. One way to explore this question is by looking at agencies' self-reported sources of program errors in their report. The sources of program errors include program design or structure, insufficient documentation, administrative or process errors, or failure to verify the data. If the study can disaggregate payment errors by the above-mentioned sources and evaluate the role of leadership types for each, it can provide richer implications in addressing improper payment problems.⁴⁴ As discussed in Essay 2, another extension of this study would be investigating the combined effect of administrative leader types on the direction of payment errors. Based on the appointment types, there can be four different combined administrative leader types: 1) non-acting career executives; 2) acting career executives; 3) non-acting political executives; 4) acting political executives. It is expected that programs managed by non-acting career executives will yield lower overpayment and underpayment errors while programs managed by acting political executives will yield higher overpayment and underpayment errors than programs managed by other combined administrative leader types. It would be interesting to see if programs managed by acting career executives and programs managed by non-acting political executives have asymmetric effects on overpayment and underpayment errors because of their different motivations that I discussed in this study.

Even with these aforementioned caveats, the empirical evidence contributes to understanding performance management in U.S. public agencies. While existing literature

⁴⁴ This essay could not pursue this strategy since information about each agency's self-reported sources of errors is only available from the FY2015 report, which will reduce the sample size significantly.

on payment errors has mostly focused on the effect of specific technological tools in reducing payment errors for a specific program in one agency (e.g., Greer and Bullock 2018; Lee 2021), this study suggests that administrative leaders may be the remedy of the problem and evaluates how the type of administrative leaders systematically predicts payment errors with the data covering multiple federal programs.

On a broader level, this study reveals that the types of administrative leadership can be crucial for understanding how agencies can effectively pursue their organizational goals. Government programs are often characterized by political salience, goal ambiguity, and competing missions. This essay asks — *which actors in leadership positions are best placed to manage in these conditions?* Exploring this question is not an easy task since it is difficult to find potentially competing performance measures in the same task and tie them with the role of leadership types. Using the example of payment errors, a salient performance indicator that can be either overpayment or underpayment errors, this study seeks to examine under what administrative leadership types these potentially competing performance goals can be achieved. By showing that career executives, relative to political appointees, are better able to reduce both types of payment errors, this study suggests that the expertise and incentives of career executives make them better able to satisfy both the goals that leaders must be immediately responsive to and the statutory goals that may be of minor concern of those leaders. Also, given that paying the right amount is a relatively routinized administrative task where the least discretion is needed, establishing an internal management system should be relatively easier than the other tasks. Conversely, for relatively non-routinized and policy-related tasks where agencies need more expertise and technical knowledge, having internal control should be more difficult and the role of administrative leadership should matter even more. Confirming the findings of Essay 2, by

showing that administrative leadership types matter even in a relatively routinized task of payments, this study emphasizes the importance of leadership in managing government performance.

Table 4.1 Types of Decision Errors in Improper Payments

		The Proper Course of Action	
		Payment	No Payment
Agency Decision	Payment	Correct Decision	Overpayment (Type I Error)
	No Payment	Underpayment (Type II Error)	Correct Decision

Table 4.2 Regression Model Estimates for U.S. Federal Program Overpayment Errors

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.10$ OP=\$0	Quantile $\tau = 0.25$ OP=\$0.88 million	Quantile $\tau = 0.40$ OP=\$8 million	Quantile $\tau = 0.55$ OP=\$31 million	Quantile $\tau = 0.70$ OP=\$112 million	Quantile $\tau = 0.85$ OP=\$512 million	Quantile $\tau = 0.97$ OP=\$10.27 billion
Career Executive	-	-1.263 (1.040)	-0.596 (2.486)	-1.063 (1.180)	-1.777** (0.719)	-1.717*** (0.605)	-1.470** (0.634)	-0.610 (0.648)	-0.242 (0.336)
Non-Acting Executive	-	-1.937*** (0.715)	-2.142 (1.721)	-1.085* (0.583)	-1.304** (0.526)	-1.568*** (0.407)	-1.294*** (0.410)	-0.661 (0.462)	-0.096 (0.183)
Aligned Agency	+	-0.285 (0.682)	-0.258 (1.330)	0.005 (0.631)	0.041 (0.412)	0.265 (0.459)	0.153 (0.445)	0.103 (0.499)	0.426 (0.289)
Opposed Agency	-	0.419 (0.623)	0.671 (1.370)	0.586 (0.634)	0.750 (0.530)	0.808* (0.468)	1.028** (0.469)	0.446 (0.458)	0.312 (0.317)
Agency Insulation	-	-0.719 (0.600)	-3.564*** (1.286)	-1.341 (0.867)	-0.525 (0.539)	-0.335 (0.413)	-0.260 (0.350)	0.020 (0.308)	-0.110 (0.114)
Executive Turnover	+	-0.045 (0.359)	0.212 (0.735)	-0.021 (0.445)	-0.037 (0.265)	0.107 (0.309)	0.213 (0.304)	0.075 (0.198)	0.876*** (0.144)
Disaster Fund	-	-3.147*** (1.037)	-5.464* (3.046)	-7.505*** (1.720)	-3.755*** (1.370)	-3.459*** (1.153)	-2.151*** (0.691)	-1.419** (0.646)	-1.376*** (0.378)
Section 57	+	0.986 (0.956)	3.472 (3.186)	0.522 (0.778)	-0.049 (0.615)	-0.078 (0.595)	0.078 (0.509)	0.409 (0.580)	-0.691 (0.546)
Exit	-	-2.132*** (0.550)	-5.265** (2.383)	-2.435*** (0.497)	-2.643*** (0.496)	-1.866** (0.840)	-1.443*** (0.326)	-1.379*** (0.387)	-0.561** (0.245)
Reentry	-	-1.306 (1.152)	-2.321 (1.865)	-2.419** (1.046)	-0.987 (1.844)	-1.692 (1.082)	-0.800 (1.125)	-1.196 (1.695)	-0.746 (0.535)
Number of Programs	+	0.126 (0.078)	0.186 (0.146)	0.113* (0.066)	0.123*** (0.041)	0.109** (0.045)	0.086** (0.044)	0.034 (0.046)	0.056 (0.035)
Sampling1	+/-	-0.763 (1.362)	-3.795* (2.284)	-1.976 (2.494)	0.389 (2.298)	-0.638 (3.753)	0.228 (0.987)	0.698 (1.289)	-1.379** (0.701)
Sampling2	+/-	2.251 (1.427)	1.418 (4.403)	-0.720 (3.051)	1.926 (2.304)	0.996 (3.611)	1.484 (1.021)	1.703 (1.362)	-1.108 (0.913)
Sampling3	+/-	0.262 (1.340)	0.140 (3.180)	-1.135 (2.596)	1.132 (2.123)	0.005 (3.670)	0.927 (0.964)	0.811 (1.219)	-1.118* (0.609)

Sampling4	+/-	0.796 (1.500)	2.049 (5.446)	-1.764 (3.422)	1.353 (2.437)	-1.054 (3.361)	0.563 (1.489)	0.898 (2.243)	-2.635** (1.235)
IPIA	+/-	-1.192* (0.693)	-2.105** (1.059)	-0.817 (0.638)	-0.417 (0.348)	-0.487* (0.276)	-0.426 (0.376)	-0.193 (0.359)	-0.219 (0.222)
IPERA	+/-	-1.154* (0.664)	-7.193** (3.623)	-0.557 (0.842)	-0.270 (0.426)	-0.321 (0.277)	-0.139 (0.320)	-0.055 (0.274)	-0.137 (0.168)
Ln (Outlay)	+	1.003*** (0.154)	0.906** (0.376)	1.006*** (0.229)	1.079*** (0.155)	0.976*** (0.154)	0.914*** (0.111)	0.905*** (0.130)	1.041*** (0.096)
R^2		0.4721	0.3760	0.4453	0.4561	0.4578	0.4551	0.4421	0.4148
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T = 10.58 P > T = 0.000	T=13.13 P > T = 0.000	T=13.77 P > T = = 0.000	T=18.64 P > T = 0.000	T=20.39 P > T = 0.000	T=17.08 P > T = = 0.000	T=20.06 P > T = 0.000
N		807	807	807	807	807	807	807	807

Note: The dependent variable is the natural log of overpayment error amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions.

OP indicates overpayment error amounts (\$).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.3 Regression Model Estimates for U.S. Federal Program Underpayment Errors

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.25$ UP=\$0	Quantile $\tau = 0.40$ UP=\$3,000	Quantile $\tau = 0.55$ UP=0.55 million	Quantile $\tau = 0.70$ UP=4.6 million	Quantile $\tau = 0.85$ UP=\$77 million	Quantile $\tau = 0.97$ UP=1.083 billion
Career Executive	-	-1.319 (1.361)	-1.122 (1.646)	-2.275 (1.953)	-1.341 (1.291)	-1.221** (0.600)	-1.477*** (0.555)	-0.229 (0.459)
Non-Acting Executive	-	-1.145 (1.235)	-1.930 (2.361)	-1.141 (1.387)	-0.378 (0.798)	-0.588 (0.589)	-0.575 (0.363)	-0.186 (0.254)
Aligned Agency	-	-3.384*** (1.110)	-2.172 (2.252)	-5.276*** (1.958)	-2.319 (1.629)	-1.200* (0.685)	-1.184** (0.504)	-0.844*** (0.245)
Opposed Agency	+	-1.683 (1.181)	-1.629 (2.182)	-1.508 (1.639)	-0.861 (1.110)	-0.437 (0.609)	-0.556 (0.482)	-0.387* (0.204)
Agency Insulation	-	-0.634 (0.806)	0.171 (0.757)	-2.603*** (0.929)	-2.071 (2.047)	-0.691 (0.540)	-0.177 (0.409)	-0.104 (0.130)
Executive Turnover	+	-0.112 (0.571)	-0.731 (1.005)	-0.278 (0.750)	0.255 (0.493)	0.302 (0.323)	0.283 (0.259)	-0.135 (0.166)
Disaster Fund	-	-1.405 (1.566)	0.213 (1.480)	-2.089 (2.217)	-5.196** (2.115)	-6.487*** (1.535)	-2.588 (2.314)	-0.588 (0.662)
Section 57	+	2.377 (1.491)	8.679 (5.877)	2.593 (2.360)	2.006 (1.224)	1.476* (0.813)	1.508** (0.597)	0.784*** (0.246)
Exit	-	-1.265 (0.769)	-0.275 (0.803)	-2.772*** (1.060)	-1.102 (0.908)	-1.127** (0.558)	-0.641 (0.623)	-1.100*** (0.359)
Reentry	-	-1.304 (1.848)	-0.482 (1.648)	-1.668 (2.567)	-0.508 (3.590)	-0.567 (1.286)	-0.029 (0.839)	-0.678*** (0.161)
Number of Programs	+	-0.000 (0.137)	0.002 (0.157)	-0.064 (0.213)	0.014 (0.108)	0.025 (0.059)	0.050 (0.046)	0.034* (0.020)
Sampling1	+/-	-2.123 (1.303)	-1.178 (1.255)	-3.674** (1.849)	-2.935** (1.440)	-2.565** (1.066)	1.342 (2.930)	-0.043 (0.871)
Sampling2	+/-	-1.861 (1.705)	-1.459 (1.806)	-3.513 (2.157)	-2.748 (2.290)	-1.893* (1.148)	1.596 (2.918)	-0.603 (0.860)
Sampling3	+/-	-3.318** (1.420)	-2.598 (2.652)	-4.288** (2.048)	-3.531* (1.916)	-2.747** (1.104)	0.812 (2.972)	-0.966 (0.880)

Sampling4	+/-	-3.056 (3.104)	-3.002 (3.709)	0.810 (2.938)	-2.160 (3.343)	-2.619** (1.316)	-0.288 (3.101)	-2.615*** (1.005)
IPIA	+/-	0.170 (0.917)	-0.565 (1.752)	0.611 (1.124)	0.051 (0.782)	0.308 (0.432)	0.280 (0.393)	0.125 (0.192)
IPERA	+/-	-1.636** (0.747)	-1.317 (1.016)	-3.563** (1.542)	-0.864 (1.131)	-0.456 (0.385)	-0.409 (0.332)	0.241 (0.317)
Ln (Outlay)	+	1.070*** (0.254)	0.392 (0.354)	1.289*** (0.370)	1.149*** (0.238)	0.987*** (0.156)	0.994*** (0.116)	0.992*** (0.062)
R^2		0.3538	0.2487	0.3323	0.3318	0.3113	0.3171	0.3137
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T=20.16 P > T = 0.000	T=18.29 P > T = 0.000	T=18.21 P > T = 0.000	T=16.79 P > T = 0.000	T=18.62 P > T = 0.000	T=11.03 P > T = 0.000
N		807	807	807	807	807	807	807

Note: The dependent variable is the natural log of underpayment error amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Since programs with an absence of underpayment errors account for 38.29% of the sample of data, the 10th quantile was not calculated to avoid redundancy. Key covariates of interest appear in shaded regions. UP indicates underpayment error amounts (\$).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 4.4 Average Differential Marginal Effects in Overpayment and Underpayment Errors
by Administrative Leader Types**

		OLS (Mean)	Quantile $\tau = 0.10$	Quantile $\tau = 0.25$	Quantile $\tau = 0.40$	Quantile $\tau = 0.55$	Quantile $\tau = 0.70$	Quantile $\tau = 0.85$	Quantile $\tau = 0.97$
Overpayment Errors									
Career Executive	Exp(β) (Factor Change)	0.283 [0.04, 2.20]	0.551 [0.00, 72.57]	0.345 [0.03, 3.50]	0.169** [0.04, 0.69]	0.180*** [0.05, 0.59]	0.230** [0.07, 0.80]	0.543 [0.15, 1.94]	0.785 [0.41, 1.52]
	% Factor Change	-71.71%	-44.90%	-65.5%	-83.08%	-82.03%	-77.02%	-45.67%	-21.46%
Non-Acting Executive	Exp(β) (Factor Change)	0.144*** [0.04, 0.59]	0.117 [0.00, 3.44]	0.338* [0.11, 1.06]	0.271** [0.10, 0.76]	0.209*** [0.09, 0.46]	0.274*** [0.12, 0.61]	0.517 [0.21, 1.28]	0.908 [0.63, 1.30]
	% Factor Change	-85.59%	-88.26%	-66.20%	-72.86%	-79.15%	-72.58%	-48.35%	-9.19%
Underpayment Errors									
Career Executive	Exp(β) (Factor Change)	0.267 [0.02, 3.93]	N/A	0.326 [0.01, 8.24]	0.103 [0.00, 4.75]	0.262 [0.02, 3.30]	0.295** [0.09, 0.96]	0.228*** [0.08, 0.68]	0.795 [0.32, 1.96]
	% Factor Change	-73.26%		-67.45%	-89.72%	-73.83%	-70.52%	-77.17%	-20.49%
Non-Acting Executive	Exp(β) (Factor Change)	0.318 [0.03, 3.65]	N/A	0.145 [0, 14.95]	0.320 [0.02, 4.86]	0.685 [0.14, 3.28]	0.555 [0.17, 1.76]	0.563 [0.28, 1.15]	0.830 [0.50, 1.37]
	% Factor Change	-68.18%		-85.48%	-68.05%	-31.48%	-44.48%	-43.71%	-16.95%

Note: 95% confidence intervals appear inside brackets. To avoid redundancy, the 10th quantile was not calculated for underpayment errors since programs with an absence of underpayment errors account for 38.29% of the sample.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

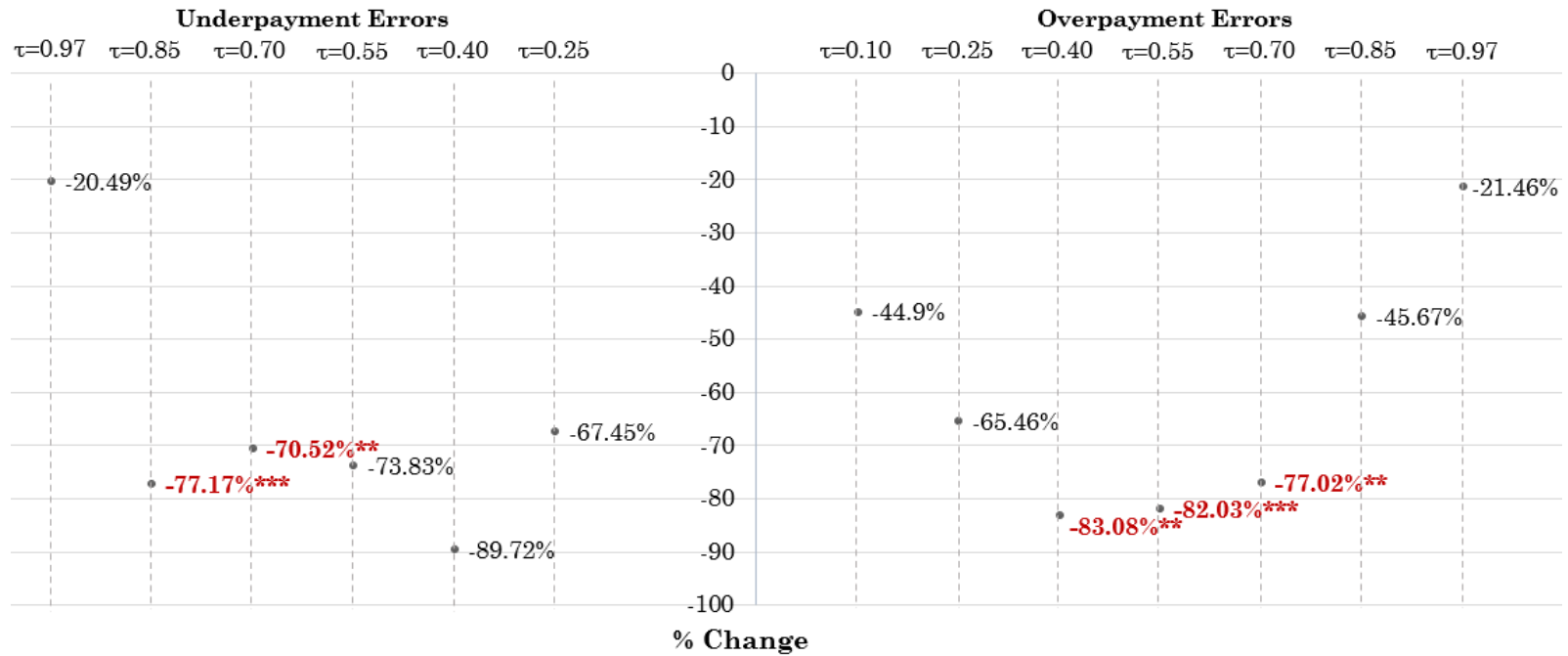


Figure 4.1 Average Differential Marginal Effects of Career Executives Versus Political Executives in Underpayment Errors and Overpayment Errors

Note: Point estimates represent the average differential marginal effects (% factor change) between career executive-managed programs versus political executive-managed programs with respect to overpayment and underpayment errors at each quantile. Since programs with no underpayment errors account for 38.29% of the sample, the 10th quantile was not calculated to avoid redundancy.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

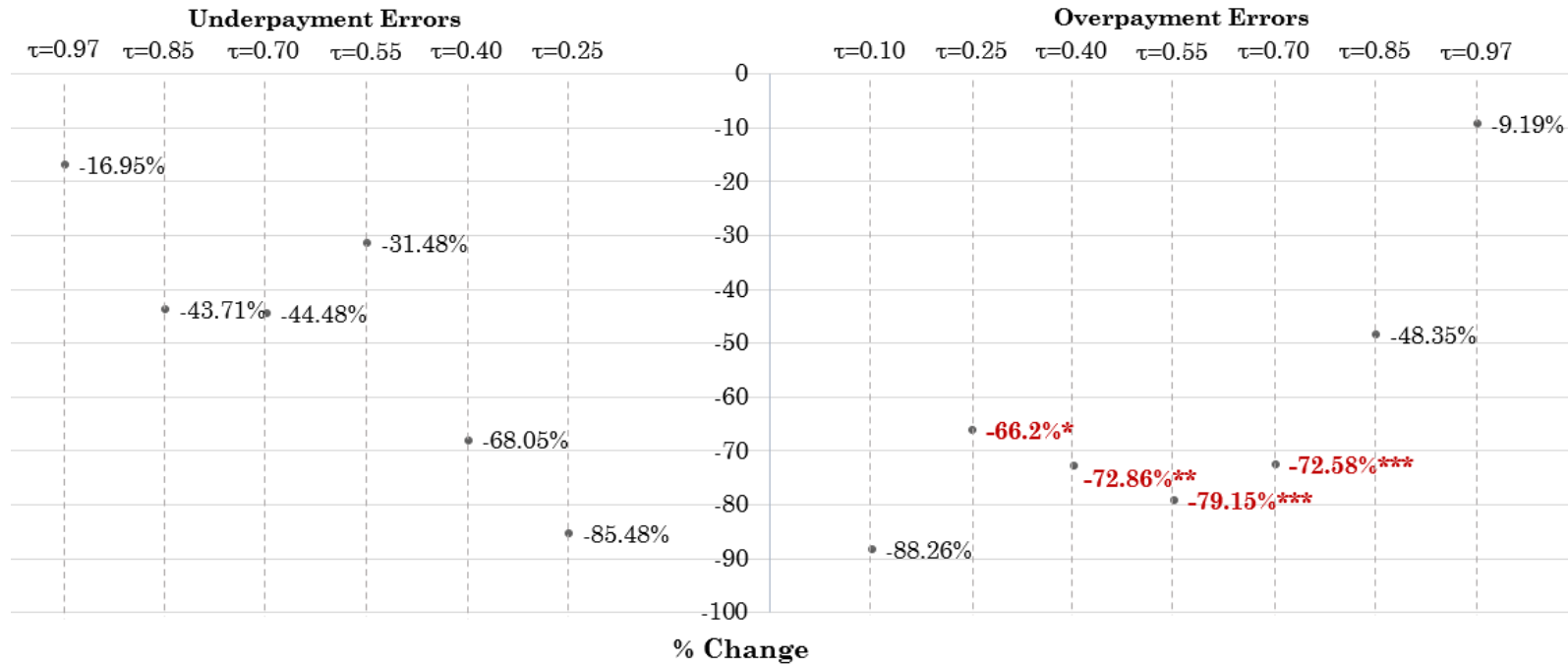


Figure 4.2 Average Differential Marginal Effects of Non-Acting Executives Versus Acting Executives in Underpayment Errors and Overpayment Errors

Note: Point estimates represent the average differential marginal effects (% factor change) between acting executive-managed programs versus non-acting executive-managed programs with respect to overpayment and underpayment errors at each quantile. Since programs with an absence of underpayment errors account for 38.29% of the sample, the 10th quantile was not calculated to avoid redundancy.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

CHAPTER 5

CONCLUSION

Findings and Contributions

This dissertation makes several theoretical and practical contributions. Above all, the topic of performance management has gotten consistent attention throughout history, and as the reach and impact of government programs on people's lives increase, managing government performance has become more crucial. This dissertation urges public administration scholars and practitioners to give attention to managing poor performance. At the individual level, effective management of poor-performing employees is expected to have a bigger positive impact within organizations than rewarding good performers since mismanagement of poor-performing employees can demotivate other employees in the workplace and thus affect organizational climate (Brewer and Walker 2013; Kettl 1997). At the organizational level, this dissertation introduces payment errors in federal programs as a good exemplar for studying negative performance because of its characteristics of universality and quantifiability.

Chapter 2, the first empirical essay of the dissertation, investigates when supervisors perceive the performance improvement period (PIP) as an effective strategy for dealing with poor-performing employees. Even though the percentage of poor performers in public organizations may be small, they can bring a disproportionately large negative impact on organizations (Fox 2011). For this reason, it is crucial to understand how government organizations can effectively manage poor-performing employees. Using the

2016 Merit Principles Survey – Path L, this study highlights the importance of organizational conditions that make supervisors more or less willing to use performance data and this tool. The findings of this chapter have several practical implications. It is recommended that public officials in managerial positions should consider providing supervisors with enough authority and organizational assistance when supervisors use the PIP. Also, although it is extremely difficult, efforts to develop valid performance measures should be made since it will help supervisors to deliver clear performance feedback and ultimately help to implement the PIP.

In chapters 3 and 4, the magnitude and the direction of payment errors are explored, respectively. Chapter 3 suggests that programs managed by career executives or non-acting (permanent) executives make lower payment errors than programs managed by political executives or acting executives. In chapter 3, we have learned the relationship between the types of administrative leaders and overall payment errors, but this chapter cannot provide the answers to the question – *which directions do administrative leaders reduce the errors? Reducing overpayment errors? Underpayment errors? Or both?* Chapter 4, the third empirical essay of this dissertation, further explores these payment errors by separately analyzing overpayment and underpayment errors. Findings of chapter 4 confirm that career executives are better suited to manage program performance in terms of reducing both overpayment and underpayment errors. These consistent findings of chapters 3 and 4 provide important implications for the heated debate over increasing the number of political appointees at the expense of career executives. Advocates of increasing the number of political appointees within government organizations argue that it will improve government responsiveness. Opponents argue that it will decrease agencies' capacity and expertise that are necessary for managing government programs. The findings add

significant evidence that career executives are better at managing program performance compared to political executives, at least for managing payment errors. Also, these studies provide a practical implication that filling a leadership vacuum with qualified officials is critical considering the inherent lack of authority in acting positions.

Finally, chapter 4 has a unique contribution to the study of performance management that is distinct from chapter 3. In this chapter, two different types of payment errors provide a unique opportunity to examine in what ways administrative leaders manage potentially competing performance goals. Even if there is a possibility of increasing underpayment errors, some people may want to reduce overpayment errors, while other people may want to reduce underpayment errors at the risk of increasing overpayment errors. In these tradeoff conditions, the findings imply that career executives have more capacity to balance the attention to both types of errors and satisfy both goals.

Limitations and Next Steps for Future Research

Despite the contributions of this dissertation, there are several limitations. However, these limitations can provide some guidance for future research. Chapter 2, the first essay of the dissertation, can be benefited from the data on non-supervisory personnel's perception of the PIP or their actual behavior. The purpose of this chapter was originally to investigate supervisors' perceptions of the PIP so that this study can advance our understanding of when supervisors will use this performance information. However, exploring subordinates' perceptions of this tool will help to get a balanced assessment and thus better understand the effectiveness of the PIP. It is possible that subordinates have different views on the PIP since they are the ones receiving performance evaluations. For instance, subordinates may view the PIP as a pressure to voluntarily leave the work unit

rather than a genuine managerial effort to help poor-performing employees. In addition, information on poor-performing employees' outcome behaviors of whether their performance actually improves after the PIP will make the study stronger and better.

One major drawback of the next two empirical essays on payment errors is that these studies cannot provide a precise direct mechanism on how administrative leaders actually influence program payments. Borrowing from the previous literature, these studies build the case based on the logical foundations where career professionals have long-term perspectives of their agencies, knowledge, and expertise in specific programs compared to political appointees (e.g., Gilmour and Lewis 2006; Lewis 2008; 2012). Similarly, non-acting (permanent) executives have a vision for agencies' long-term mission and possess an informal authority as well as formal authority compared to acting executives. By using their wherewithal, this study suggests that career executives and non-acting executives are more likely to establish robust internal control systems (e.g., monitoring, overseeing, and hiring highly qualified employees) to better manage program payments. Even though statistical analyses also show support for this relationship, these studies can make a stronger argument if the evidence of what specific tools administrative leaders use and how these tools actually influence payment errors is added. More investigations in this aspect are necessary as the next step for future research.

Although the specific research questions of the three empirical essays in this dissertation are distinct, the ultimate aim of this dissertation is to develop a theoretical and practical understanding of personnel managerial strategies to manage negative performance. These three empirical essays provide collective insights on why some government organizations face more persistent problems in performance and how to manage negative performance. Understanding negative performance management is

critical to the broader endeavor of performance management since managing negative performance is more likely to be the first step and a more pressing task for government organizations to improve their performance. Also, examining negative performance can provide richer information, such as how far organizations are from achieving goals or how effective a strategy is to improve poor performance, which can offer valuable insights into the study of broader performance management.

REFERENCES

- Abramson, Mark A., Jonathan D. Breul, and John M. Kamensky. 2006. *Six Trends Transforming Government*. Washington, D.C.: IBM Center for the Business of Government.
- Aldama, Abraham, and James Bisbee. 2018. "When Political Scientists Should Avoid Ratio Regressions." (Working Paper) Retrieved July 10, 2021.
https://s18798.pcdn.co/abrahamaldama/wp-content/uploads/sites/6464/2018/09/ratio_manuscript.pdf.
- Andrews, Rhys, and George A. Boyne. 2010. "Capacity, Leadership, and Organizational Performance: Testing the Black Box Model of Public Management." *Public Administration Review* 70(3): 443-454.
- Andrews, Rhys, and Gene A. Brewer. 2013. "Social Capital, Management Capacity and Public Service Performance: Evidence from the U.S. States." *Public Management Review* 15(1): 19-42.
- Avellaneda, Claudia N. 2016. "Government Performance and Chief Executives' Intangible Assets: Motives, Networking, and/or Capacity?" *Public Management Review* 18(6): 918-947.
- Ban, Carolyn, and Patricia W. Ingraham. 1990. "Short-Timers: Political Appointee Mobility and Its Impact on Political-Career Relations in the Reagan Administration." *Administration & Society* 22(1): 106-124.

- Ban, Carolyn. 1995. *How do Public Managers Manage? Bureaucratic Constraints, Organizational Culture, and the Potential for Reform*. San Francisco, CA: Jossey-Bass.
- Bellé, Nicola, Paola Cantarelli, and Paolo Belardinelli. 2017. "Cognitive Biases in Performance Appraisal: Experimental Evidence on Anchoring and Halo Effects with Public Sector Managers and Employees." *Review of Public Personnel Administration* 37(3): 275-294.
- Bendor, Jonathan B. 1985. *Parallel Systems: Redundancy in Government*. Berkeley, CA: University of California Press.
- Berry, Christopher R., and Jacob E. Gersen. 2017. "Agency Design and Political Control." *Yale Law Journal* 126 (4): 1002-1049.
- Bertelli, Anthony M., and Laurence E. Lynn, Jr. 2003. "Managerial Responsibility." *Public Administration Review* 63(3): 259-268.
- Bertelli, Anthony M., , and Sven E. Feldmann. 2007. "Strategic Appointments." *Journal of Public Administration Research and Theory* 17(1): 19-38.
- Bertelli, Anthony M., and David E. Lewis. 2012. "Policy Influence, Agency-Specific Expertise, and Exit in the Federal Service." *Journal of Public Administration Research and Theory* 23(2): 223-245.
- Boylan, Richard T. 2004. "Salaries, Turnover, and Performance in the Federal Criminal Justice System." *Journal of Law and Economics* 47(1): 75-92.
- Brewer, Gene A., and Robert A. Maranto. 2000. "Comparing the Roles of Political Appointees and Career Executives in the U.S. Federal Executive Branch." *The American Review of Public Administration* 30(1): 69-86.

- Brewer, Gene A., and Sally Coleman Selden. 2000. "Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies." *Journal of Public Administration Research and Theory* 10(4): 685-712.
- Brewer, Gene A., and Richard M. Walker. 2013. "Personnel Constraints in Public Organizations: The Impact of Reward and Punishment on Organizational Performance." *Public Administration Review*. 73(1): 121-131.
- Bullock, Justin B. 2014. A Theory and Empirical Examination of Bureaucratic Errors. (Doctoral Dissertation. University of Georgia).
- Byrne, Barbara M. 2013. *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. 2nd Edition. New York, NY: Routledge.
- Campion, Michael A., and Robert G. Lord. 1982. "A Control Systems Conceptualization of the Goal-Setting and Changing Process." *Organizational Behavior and Human Performance*. 30(2): 265-287.
- Certo, S. Trevis, John R. Busenbark, Matias Kalm, and Jeffery A. LePine. 2020. "Divided We Fall: How Ratios Undermine Research in Strategic Management." *Organizational Research Methods* 23(2): 211-237.
- Chang, Kelly, David E. Lewis, and Nolan McCarty. 2001. "The Tenure of Political Appointees." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.
- Cheney, Kyle. 2013. "Medicare improperly paid \$120M." Politico, January 24. Retrieved May 9, 2022. <https://www.politico.com/story/2013/01/reports-medicare-paid-120m-in-illegal-care-086676>.

- Chernozhukov, Victor, and Han Hong. 2002. "Three-Step Censored Quantile Regression and Extramarital Affairs." *Journal of the American Statistical Association* 97(459): 872-882.
- Choi, Sungjoo, and Hal G. Rainey. 2010. "Managing Diversity in U.S. Federal Agencies: Effects of Diversity and Diversity Management on Employee Perceptions of Organizational Performance." *Public Administration Review*. 70(1): 109-121.
- Choi, Sungjoo, and Andrew B. Whitford. 2017. "Employee Satisfaction in Agencies with Merit-Based Pay: Differential Effects for Three Measures." *International Public Management Journal* 20(3): 442-466.
- Chong, Hyonsong, Richard E. White, and Victor Prybutok. 2001. "Relationship among Organizational Support, JIT Implementation, and Performance." *Industrial Management & Data Systems* 101(6): 273-280.
- Chun, Young Han, and Hal G. Rainey. 2005. "Goal Ambiguity and Organizational Performance in U.S. Federal Agencies." *Journal of Public Administration Research and Theory* 15(4): 529-557.
- Clinton, Joshua D., and David E. Lewis. 2008. "Expert Opinion, Agency Characteristics, and Agency Preferences." *Political Analysis* 16(1): 3-20.
- Cohen, David M. 1998. "Amateur Government." *Journal of Public Administration Research and Theory* 8(4): 450-497.
- Congressional Research Service. 2016. *Improper Payments Legislation: Key Provisions, Implementation, and Selected Proposals in the 114th Congress*. CRS Report.
- Cooper, John O., Timothy E. Heron, and William L. Heward. 2007. *Applied Behavior Analysis*. 2nd ed. Upper Saddle River, NJ: Merrill / Prentice Hall.

- Cyert, Richard. M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ, 2(4): 169–187.
- Dodaro, Gene L. 2013. “Addressing the Long-Term Fiscal Challenge.” *Public Administration Review* 73(6): 782-783.
- Dull, Matthew, and Patrick S. Roberts. 2009. “Continuity, Competence, and the Succession of Senate-Confirmed Agency Appointees, 1989-2009.” *Presidential Studies Quarterly* 39(3): 432-453.
- Dunn, Delmer D. 1997. *Politics and Administration at the Top: Lessons from Down Under*. Pittsburgh, PA: University of Pittsburgh Press.
- Favero, Nathan, and Justin B. Bullock. 2015. “How (Not) to Solve the Problem: An Evaluation of Scholarly Responses to Common Source Bias.” *Journal of Public Administration Research and Theory* 25(1): 285-308.
- Favero, Nathan, Kenneth J. Meier, and Laurence J. O’Toole Jr. 2016. “Goals, Trust, Participation, and Feedback: Linking Internal Management with Performance Outcomes.” *Journal of Public Administration Research and Theory* 26(2): 327-343.
- Feeney, Mary K., and Hal G. Rainey. 2010. “Personnel Flexibility and Red Tape in Public and Nonprofit Organizations: Distinctions Due to Institutional and Political Accountability.” *Journal of Public Administration Research and Theory* 20(4): 801–26.
- Fornell, Claes, and David F. Larcker. 1981. “Evaluating Structural Equation Models with Unobservable Variables and Measurement Error.” *Journal of Marketing Research* 18(1): 39-50.
- Fox, Tom. 2011. “Poor Performers – Dealing with Performance Issues in Federal Government.” October 17. Retrieved June 10, 2022.
<https://www.washingtonpost.com/blogs/ask-the-fedcoach/post/poor-performers-->

dealing-with-performance-issues-in-federal-government/2011/03/04/gIQAPD0NrL_blog.html.

Gailmard, Sean F., and John W. Patty. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science* 51(4): 873-89.

Gailmard, Sean, and John Patty. 2012. *Learning While Governing: Information, Accountability, and Executive Branch Institutions*. Chicago: University of Chicago Press.

Gallo, Nick, and David E. Lewis. 2012. "The Consequences of Presidential Patronage for Federal Agency Performance." *Journal of Public Administration Research and Theory* 22(2): 219-243.

Garnett, James L., Justin Marlowe, and Sanjay K. Pandey. 2008. "Penetrating the Performance Predicament: Communication as a Mediator or Moderator of Organizational Culture's Impact on Public Organizational Performance." *Public Administration Review* 68(2): 266-281.

George, Bert, Martin Baekgaard, Adeliën Decramer, Mieke Audenaert, and Stijn Goeminne. 2020. "Institutional Isomorphism, Negativity Bias and Performance Information Use by Politicians: A Survey Experiment." *Public Administration* 98(1): 14-28.

Gilmour, John B., and David E. Lewis. 2006. "Political Appointees and the Competence of Federal Program Management." *American Politics Research* 34(1): 22-50.

Goldfinch, Shaun, and Joe Wallis. 2010. "Two Myths of Convergence in Public Management Reform." *Public Administration* 88(4): 1099-1115.

Gomez-Mejia, Luis R., David B. Balkin, Robert L. Cardy, and Kenneth P. Carson. 2009. *Managing Human Resources*. Upper Saddle River, NJ: Pearson Prentice Hall.

- Goodnow, Frank. 1900. *Politics and Administration: A Study in Government*. New York: Macmillan.
- Greene, William H. 2002. *Econometric Analysis*. New York, NY: Prentice Hall.
- Greer, Robert A., and Justin B. Bullock. 2018. "Decreasing Improper Payments in a Complex Federal Program." *Public Administration Review* 78(1): 14-23.
- Gulick, Luther. 1937. "Notes on the Theory of Organization." In L. Gulick & L. Urwick (Eds.), *Papers on the Science of Administration*. (pp. 1-46). New York: Institute of Public Administration.
- Heclo, Hugh. 1977. *A Government of Strangers: Executive Politics in Washington*. Washington, DC: Brookings Institution.
- Herd, Pamela, and Donald P. Moynihan. 2018. *Administrative Burden: Policymaking by Other Means*. New York, NY: Russell Sage Foundation.
- Hollibaugh Jr, Gary E., Gabriel Horton, and David E. Lewis. 2014. "Presidents and patronage." *American Journal of Political Science* 58(4): 1024-1042.
- Hong, Sounman. 2019. "A Behavioral Model of Public Organizations: Bounded Rationality, Performance Feedback, and Negativity Bias." *Journal of Public Administration Research and Theory* 29(1): 1-17.
- Hong, Sounman, and Youngrok Kim. 2019. "Loyalty or Competence: Political Use of Performance Information and Negativity Bias." *Public Administration Review* 79(6): 829-840.
- Hood, Christopher. 1991. "A public Management for All Seasons?" *Public Administration* 69(1): 3-19.
- Hood, Christopher. 1995. "The "New Public Management" in the 1980s: Variations on a Theme." *Accounting Organisations and Society* 20(2): 93-109.

- Hood, Christopher. 1998. *The Art of the State: Culture, Rhetoric, and Public Management*. The Clarendon Press.
- Hu, Li-tze, and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6(1): 1-55.
- Hudak, John. 2014. *Presidential Pork: White House Influence Over the Distribution of Federal Grants*. Washington, DC: Brookings Institution Press.
- Ilgen, Daniel R., Cynthia D. Fisher, and M. Susan Taylor. 1979. "Consequences of Individual Feedback on Behavior in Organizations." *Journal of Applied Psychology* 64(4): 349-371.
- Ingraham, Patricia W., and Amy Kneedler Donahue. 2000. "Dissecting the Black Box Revisited: Characterizing Government Management Capacity." In *Governance and Performance: New Perspectives*. Chapter 10: 292-318.
- Iqbal, Muhammad Zahid, Saeed Akbar, and Pawan Budhwar. 2015. "Effectiveness of Performance Appraisal: An Integrated Framework." *International Journal of Management Reviews* 17(4): 510-533.
- Jochum, Elizabeth N. 2008. "Agencies Report More Improper Payments, Take Steps to Address Them." *Government Executive*, January 24. Retrieved November 2, 2021. <https://www.govexec.com/federal-news/2008/01/agencies-report-more-improper-payments-take-steps-to-address-them/26158/>.
- Julien, Mark, Karen Somerville, and Neil Culp. 2011. "Going Beyond the Work Arrangement: The Crucial Role of Supervisor Support." *Public Administration Quarterly* 35(2): 167-204.

- Kaufman, Herbert. 1956. "Emerging Conflicts in the Doctrines of Public Administration." *American Political Science Review* 50(4): 1057-1073.
- Kellough, J. Edward. 1998. "The Reinventing Government Movement: A Review and Critique." *Public Administration Quarterly* 22(1): 6-20.
- Kellough, J. Edward, and Lloyd G. Nigro. 2006. "Dramatic Reform in the Public Service: At-Will Employment and the Creation of a New Public Workforce." *Journal of Public Administration Research and Theory* 16(3):447-466.
- Kettl, Donald F. 1997. "The Global Revolution in Public Management: Driving Themes, Missing Links." *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management* 16(3): 446-462.
- Kinane, Christina M. 2021. "Control without Confirmation: The Politics of Vacancies in Presidential Appointments." *American Political Science Review* 115(2): 599-614.
- Kline, Rex B. 2011. *Principles and Practice of Structural Equation Modeling*. 3rd Edition. New York, NY: The Guilford Press.
- Ko, Jaekwon, SeungUk Hur, and Aaron Smith-Walter. 2013. "Family-Friendly Work Practices and Job Satisfaction and Organizational Performance: Moderating Effects of Managerial Support and Performance-Oriented Management." *Public Personnel Management* 42(4): 545-565.
- Ko, Jaekwon, and SeungUk Hur. 2014. "The Impacts of Employee Benefits, Procedural Justice, and Managerial Trustworthiness on Work Attitudes: Integrated Understanding Based on Social Exchange Theory." *Public Administration Review* 74(2): 176-187.
- Koenker, Roger, and Gilbert Bassett Jr. 1978. "Regression Quantiles." *Econometrica: Journal of the Econometric Society* 46(1): 33-50.

- Koenker, Roger, and Kevin F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15(4): 143-156.
- Krause, George A., David E. Lewis, and James W. Douglas. 2006. "Political Appointments, Civil Service Systems, and Bureaucratic Competence: Organizational Balancing and Executive Branch Revenue Forecasts in the American States." *American Journal of Political Science* 50(3): 770-787.
- Krause, George A., and Anne Joseph O'Connell. 2019. "Loyalty–Competence Trade-offs for Top U.S. Federal Bureaucratic Leaders in the Administrative Presidency Era." *Presidential Studies Quarterly* 49(3): 527-550.
- Krause, George A., and Matthew Zarit. 2022. "Selling Out? Contingent Politicization and Contracting Risk in US Federal Procurements, 2001–2016." *Journal of Political Institutions and Political Economy* 2(4): 509-535.
- Kriner, Douglas L., and Andrew Reeves. 2015. *The Particularistic President: Executive Branch Politics and Political Inequality*. New York, NY: Cambridge University Press.
- Kroll, Alexander. 2015a. "Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research." *Public Performance & Management Review* 38(3): 459-486.
- Kroll, Alexander. 2015b. "Exploring the Link Between Performance Information Use and Organizational Performance: A Contingency Approach." *Public Performance & Management Review* 39(1): 7-32.
- Kronmal, Richard A. 1993. "Spurious Correlation and the Fallacy of the Ratio Standard Revisited." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 156(3): 379-392.

- Landau, Martin, and Russell Stout Jr. 1979. "To Manage is Not to Control: Or the Folly of Type II Errors." *Public Administration Review* 39(2): 148-156.
- Larson Jr, James R. 1984. "The Performance Feedback Processes: A Preliminary Model" *Organizational Behavior and Human Performance* 33(1): 42-76.
- Larson Jr, James R. 1989. "The Dynamic Interplay Between Employees' Feedback-Seeking Strategies and Supervisors' Delivery of performance feedback." *Academy of Management Review* 14(3): 408-422.
- Lee, DonHee, Sang M. Lee, David L. Olson, and Soong Hwan Chung. 2010. "The Effect of Organizational Support on ERP Implementation." *Industrial Management & Data Systems* 110(2): 269-283.
- Lee, Jung Wook, Hal G. Rainey, and Young Han Chun. 2010. "Goal Ambiguity, Work Complexity, and Work Routineness in Federal Agencies." *The American Review of Public Administration* 40(3): 284-308.
- Lee, Sabinne. 2021. "Impact of Administrative Burden on Inappropriate Payment Error: a Blessing or a Curse?" *International Review of Public Administration* 26(1): 18-40.
- Lee, Soo-Young, and Andrew B. Whitford. 2013. "Assessing the Effects of Organizational Resources on Public Agency Performance: Evidence from the U.S. Federal Government." *Journal of Public Administration Research and Theory* 23(3): 687-712.
- Levy, Paul E., and Jane R. Williams. 2004. "The Social Context of Performance Appraisal: A Review and Framework for the Future." *Journal of Management* 30(6): 881-905.
- Lewis, David E. 2007. "Testing Pendleton's Premise: Do Political Appointees Make Worse Bureaucrats?" *The Journal of Politics* 69(4): 1073-1088.
- Lewis, David E. 2008. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*. Princeton, NJ: Princeton University Press.

- Lewis, David E. 2012a. "The Contemporary Presidency: The Personnel Process in the Modern Presidency." *Presidential Studies Quarterly* 42(3): 577-596.
- Lewis, David. 2012b. "Presidential Politicization of the Executive Branch in the United States." In M. Lodge & K. Wegrich (Eds.), *Executive Politics in Times of Crisis* (pp. 41-62). London: Palgrave Macmillan UK.
- Liermann, Martin, Ashley Steel, Michael Rosing, and Peter Guttorp. 2004. "Random Denominators and the Analysis of Ratio Data." *Environmental and Ecological Statistics* 11(1): 55-71.
- Lin, Yu-Chun, and J. Edward Kellough. 2019. "Performance Appraisal Problems in the Public Sector: Examining Supervisors' Perceptions." *Public Personnel Management* 48(2): 179-202.
- Locke, Edwin A., and Gary P. Latham. 1990. *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall.
- Malenich, Larry. 2019. "Stop the Check: Steps to Reduce Uncle Sam's Improper Payments." Government Executive, March 29. Retrieved July 15, 2021.
<https://www.govexec.com/management/2019/03/stop-check-steps-reduce-uncle-sams-improper-payments/155872/>.
- Marr, Kendra. 2010. "Obama Signs Waste Crackdown." *Politico*, July 22. Retrieved July 15, 2021. <https://www.politico.com/story/2010/07/obama-signs-waste-crackdown-040111>.
- Marvel John D. 2016. "Unconscious Bias in Citizens' Evaluations of Public Sector Performance." *Journal of Public Administration Research and Theory* 26(1):143-158.
- Meier, Kenneth J., and Laurence J. O'Toole Jr. 2002. "Public Management and Organizational Performance: The Effect of Managerial Quality." *Journal of Policy*

- Analysis and Management: The Journal of the Association for Public Policy Analysis and Management* 21(4): 629-643.
- Mendelson, Nina A. 2014. "The Uncertain Effects of Senate Confirmation Delays in the Agencies." *Duke Law Journal* 64(8): 1571-1606.
- Miller, Susan M. 2015. "The Relationship Between Short-Term Political Appointments and Bureaucratic Performance: The Case of Recess Appointments in the United States." *Journal of Public Administration Research and Theory* 25(3): 777-796.
- Moe, Terry M. 1982. "Regulatory Performance and Presidential Administration." *American Journal of Political Science* 26(1): 197-224.
- Moe, Terry M. 1985. "The Politicized Presidency." In J. E. Chubb & P. E. Peterson (Eds.), *The New Direction in American Politics* (pp. 235-271). Washington, DC: Brookings Institution.
- Mosteller, Frederick and John Tukey. 1977. *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA: Addison-Wesley.
- Moynihan, Donald P., and Patricia Wallace Ingraham. 2004. "Integrative Leadership in the Public Sector: A Model of Performance-Information Use." *Administration & Society* 36(4): 427-453.
- Moynihan, Donald P., and Alasdair S. Roberts. 2010. "The Triumph of Loyalty over Competence: The Bush Administration and the Exhaustion of the Politicized Presidency." *Public Administration Review* 70(4): 572-581.
- Moynihan, Donald P. 2015. "Uncovering the Circumstances of Performance Information Use Findings from an Experiment." *Public Performance & Management Review* 39(1): 33-57.

- Moynihan, Donald P. 2021a. "Public Management for Populists: Trump's Schedule F Executive Order and the Future of the Civil Service." *Public Administration Review* 82(1): 174-178.
- Moynihan, Donald P. 2021b. "Populism and the Deep State: The Attack on Public Service Under Trump." In Michael Bauer, B. Guy Peters, Jon Pierre, Kutsal Yesilkagit, and Stefan Becker. *Democratic Backsliding and Public Administration: How Populists in Government Transform State Bureaucracies*. Cambridge: Cambridge University Press.
- Nielsen, Poul A., and Donald P. Moynihan. 2017. "How Do Politicians Attribute Bureaucratic Responsibility for Performance? Negativity Bias and Interest Group Advocacy." *Journal of Public Administration Research and Theory* 27(2): 269-283.
- Nunnally, Jum C., and Ira H. Bernstein. 1994. *Psychological Theory*. New York, NY: McGraw-Hill.
- O'Connell, Anne Joseph. 2008. "Vacant Offices: Delays in Staffing Top Agency Positions." *Southern California Law Review* 82: 913-1001.
- O'Connell, Anne Joseph. 2020. "Actings." *Columbia Law Review* 120(3): 613-728.
- O'Connell, Anne Joseph. 2021. "The Biden Administration Still Has a Lot of Vacant Positions. It's a Growing Problem." The Washington Post. October 28. Retrieved June 10, 2022. <https://www.washingtonpost.com/politics/2021/10/28/biden-administration-still-has-lot-vacant-positions-its-growing-problem/>.
- Osborne, David, and Ted Gaebler. 1992. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. Plume.
- Ostrander, Ian. 2016. "The Logic of Collective Inaction: Senatorial Delay in Executive Nominations." *American Journal of Political Science* 60(4): 1063-1076.

- Parente, Paulo M.D.C., and João M.C. Santos Silva. 2016. "Quantile Regression with Clustered Data." *Journal of Econometric Methods* 5(1): 1-15.
- Parsneau, Kevin. 2013. "Politicizing Priority Departments: Presidential Priorities and Subcabinet Experience and Loyalty." *American Politics Research* 41(3): 443-470.
- Partnership for Public Service. 2020. *The Replacements: Why and How "Acting" Officials Are Making Senate Confirmation Obsolete*. Retrieved on June 20, 2021.
<https://ourpublicservice.org/wp-content/uploads/2020/09/The-Replacements-1.pdf>.
- Peters, B. Guy, and Jon Pierre. 2004. "Politicization of the Civil Service: Concepts, Causes, Consequences", in B. Guy Peters and Jon Pierre (eds), *Politicization of the Civil Service in Comparative Perspective*. London: Routledge, 1–13.
- Pfiffner, James P. 1987. "Political Appointees and Career Executives: The Democracy Bureaucracy Nexus in the Third Century." *Public Administration Review* 47(1): 57-65.
- Pitts, David. 2009. "Diversity Management, Job Satisfaction, and Performance: Evidence from U.S. Federal Agencies." *Public Administration Review* 69(2): 328-338.
- Podsakoff, Philip M., and Jiing-Lih Farh. 1989. "Effects of Feedback Sign and Credibility on Goal Setting and Task Performance." *Organizational Behavior and Human Decision Processes* 44(1): 45-67.
- Pollitt, Christopher, and Geert Bouckaert. 2004. *Public Management Reform: A Comparative Analysis*. New York, NY: Oxford University Press.
- Rainey, Hal G., Rex Facer, and Barry Bozeman. 1995. "Repeated Findings of Sharp Differences between Public and Private Managers' Perceptions of Personnel Rules." Paper presented at the Annual Meeting of the American Political Science Association, Chicago, August 31–September 3.

- Rainey, Hal G., Sanjay Pandey, and Barry Bozeman. 1995. "Research Note: Public and Private Managers' Perceptions of Red Tape." *Public Administration Review* 55(6): 567–74.
- Rainey, Hal G., and Paula Steinbauer. 1999. "Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations." *Journal of Public Administration Research and Theory* 9(1): 1-32.
- Rainey, Hal G., and Barry Bozeman. 2000. "Comparing Public and Private Organizations: Empirical Research and the Power of A Priori." *Journal of Public Administration Research and Theory* 10(2): 447–69.
- Rainey, Hal G., Sergio Fernandez, and Deanna Malatesta. 2021. *Understanding and Managing Public Organizations*. (6th Edition). San Francisco, CA: Jossey-Bass Inc.
- Rappeport, Alan. 2020. "\$1.4 Billion in Stimulus Funds Sent to Dead People, Watchdog Finds." The New York Times. July 17. Retrieved May 8, 2022.
<https://www.nytimes.com/2020/06/25/us/politics/coronavirus-stimulus-dead-people.html>.
- Rein, Lisa. 2022. "How a Social Security Program Piled Huge Fines on the Poor and Disabled." The Washington Post. May 20. Retrieved June 10, 2022.
<https://www.washingtonpost.com/politics/2022/05/20/social-security-fraud-penalties/>.
- Rezaee, Janna Marie King. 2016. *Essays on Presidential and Congressional Oversight of the Executive Branch*. Ph.D. Dissertation. UC Berkeley.
- Richardson, Mark D. "Politicization and Expertise: Exit, Effort, and Investment." *The Journal of Politics* 81(3) (2019): 878-891.
- Rosen, Bernard. 1983. "Effective Continuity of U.S. Government Operations in Jeopardy." *Public Administration Review* 43(5): 383-392.

- Rubin, Ellen V. 2009. "The Role of Procedural Justice in Public Personnel Management: Empirical Results from the Department of Defense." *Journal of Public Administration Research and Theory* 19(1): 125-143.
- Rubin, Jennifer. 2021. "Opinion: Yet Another Reason We Need Fewer Political Appointees." *The Washington Post*, August 1. Retrieved August 2, 2021.
<https://www.washingtonpost.com/opinions/2021/08/01/heres-why-we-need-fewer-political-appointees/>.
- Rutzick, Karen. 2005. "Lawmakers Rap OPM on Underpaid Retirement Annuities." *Government Executive*. August 8. Retrieved May 9, 2022.
<https://www.govexec.com/pay-benefits/2005/08/lawmakers-rap-opm-on-underpaid-retirement-annuities/19843/>.
- Sahoo, C.K. and Sukanta Mishra. 2012. "Performance Management Benefits Organizations and Their Employees." *Human Resource Management International Digest* 20(6):3-5.
- Selin, Jennifer L. 2015. "What Makes an Agency Independent?" *American Journal of Political Science* 59(4): 971-987.
- Silver, Hilary. 2021. "Trump's Assault on the Federal Government isn't Over." *The Hill*, January 19. Retrieved October 10, 2021.
<https://thehill.com/opinion/white-house/534795-trumps-assault-on-the-federal-government-isnt-over>.
- Simon, Herbert A. 1947. *Administrative Behavior*. New York: Free Press of Glencoe.
- StataCorp. 2019. *Stata: Structural Equation Modeling Reference Manual. Release 16*. Retrieved June 30, 2020. <https://www.stata.com/manuals/sem.pdf>.
- Steelman, Lisa A., and Kelly A. Rutkowski. 2004. "Moderators of Employee Reactions to Negative Feedback." *Journal of Managerial Psychology* 19(1): 6-18.

- Steers, Richard M., and Lyman W. Porter. 1974. "The Role of Task-Goal Attributes in Employee Performance." *Psychological Bulletin* 81(7): 434–52.
- Suleiman, Ezra N. 2003. *Dismantling Democratic States*. Princeton: Princeton University Press.
- Tau, Byron. 2019. "Half of 10 Biggest Federal Law Agencies Lack Permanent Chiefs." *The Wall Street Journal*, May 16. Retrieved June 30, 2021.
<https://www.wsj.com/articles/half-of-10-biggest-federal-law-agencies-lack-permanent-chiefs-11558006200>.
- Taylor, Frederick W. 1911. *The Principles of Scientific Management*. New York: Harper.
- Tyson, Alec, and Cary Funk. 2022. "Increasing Public Criticism, Confusion over Covid-19 Response in U.S." February 9. Retrieved June 10, 2022.
<https://www.pewresearch.org/science/2022/02/09/increasing-public-criticism-confusion-over-covid-19-response-in-u-s/>.
- U.S. General Accounting Office (GAO). 2001a. *Financial Management: Poor Internal Controls Expose Department of Education to Improper Payments* (GAO-01-1151). Retrieved May 1, 2022. <https://www.gao.gov/assets/gao-01-1151.pdf>.
- U.S. General Accounting Office (GAO). 2001b. *Strategies to Manage Improper Payments: Learning from Public and Private Sector Organizations* (GAO-02-69G). Retrieved May 1, 2022. <https://www.gao.gov/assets/gao-02-69g.pdf>.
- U.S. Government Accountability Office (GAO). 2011. *Improper Payments: Recent Efforts to Address Improper Payments and Remaining Challenges* (GAO-11-575T). Retrieved May 30, 2020. <https://www.gao.gov/assets/gao-11-575t.pdf>.

- U.S. Government Accountability Office. 2012. *Strategies to Manage Improper Payments: Learning from Public and Private Sector Organizations* (GAO-02-69G). Retrieved May 1, 2021. <https://www.gao.gov/assets/gao-02-69g.pdf>.
- U.S. Government Accountability Office (GAO). 2015. *Federal Workforce: Improved Supervision and Better Use of Probationary Periods Are Needed to Address Substandard Employee Performance*. Retrieved April 8, 2020. <https://www.gao.gov/assets/gao-15-191.pdf>.
- U.S. General Accounting Office. 2020. *Payment Integrity: Federal Agencies' Estimates of FY2019 Improper Payments* (GAO-20-344). Retrieved May 21, 2021. <https://www.gao.gov/assets/gao-20-344.pdf>.
- U.S. Merit Systems Protection Board (MSPB). 2003. *The Federal Workforce for the 21st Century: Results of the Merit Principles Survey 2000*. Retrieved April 8, 2020. <https://www.mspb.gov/MSPBSEARCH/viewdocs.aspx?docnumber=253631&version=253918&applic>.
- U.S. Merit Systems Protection Board (MSPB). 2009. *Addressing Poor Performers and the Law*. Retrieved April 8, 2020. <https://www.mspb.gov/mspbsearch/viewdocs.aspx?docnumber=445841&version=446988&application=ACROBAT>.
- U.S. Merit Systems Protection Board (MSPB). 2015. *The How and Why of an Effective Performance Improvement Plan*. Retrieved November 11, 2021.
- U.S. Merit Systems Protection Board (MSPB). 2016. *Merit Principles Survey 2016 Data and Methodology*. Retrieved September 29, 2020. [https://www.mspb.gov/foia/Data/MSPB MPS2016 MethodologyMaterials.pdf](https://www.mspb.gov/foia/Data/MSPB_MPS2016_MethodologyMaterials.pdf).

- U.S. Office of Management and Budget. 2006. *Memorandum for Heads of Executive Departments and Agencies* (M-06-23). Retrieved June 05, 2021. <https://georgewbush-whitehouse.archives.gov/omb/memoranda/fy2006/m06-23.pdf>.
- U.S. Office of Management and Budget (OMB). 2018. *OMB Circular A-123's Appendix C*. Washington, DC: Author. <https://www.whitehouse.gov/wp-content/uploads/2018/06/M-18-20.pdf>.
- U.S. Office of Management and Budget (OMB). 2021. *Updated Data on Improper Payments*. December 30. Retrieved May 8, 2022. <https://www.whitehouse.gov/omb/briefing-room/2021/12/30/updated-data-on-improper-payments/>.
- U.S. Office of Personnel Management (OPM). n.d.(a). *Providing an Opportunity to Improve*. Retrieved April 8, 2020. <https://www.opm.gov/policy-data-oversight/employee-relations/reference-materials/the-performance-improvement-period.pdf>.
- U.S. Office of Personnel Management (OPM). n.d.(b). *Employee Relations Practitioners Performance Issues Facts*. Retrieved April 8, 2020. <https://www.opm.gov/policy-data-oversight/employee-relations/employee-rights-appeals/performance-based-actions/facts.pdf>.
- U.S. Office of Personnel Management (OPM). n.d.(c). *Managing Federal Employees' Performance Issues or Misconduct*. Retrieved January 3, 2022. <https://www.opm.gov/policy-data-oversight/employee-relations/reference-materials/managing-federal-employees-performance-issues-or-misconduct.pdf>.
- U.S. Office of Personnel Management. 2014. *The Guide to Data Standards. Part A: Human Resources*. Retrieved January 31, 2021. <https://www.opm.gov/policy-data-oversight/data-analysis-documentation/data-policy-guidance/reporting-guidance/part-a-human-resources.pdf>.

- U.S. Office of Personnel Management (OPM). 2018. *FedScope: Federal Human Resources Data*. Washington, DC: OPM. Retrieved January 7th, 2022.
https://www.fedscope.opm.gov/datadefn/aeMRI_sdm.asp.
- Van Ryzin, Gregg G. 2006. "Testing the Expectancy Disconfirmation Model of Citizen Satisfaction with Local Government." *Journal of Public Administration Research and Theory* 16(4): 599-611.
- Vroom, Victor Harold. 1964. *Work and Motivation*. New York, NY: John Wiley.
- Weaver, R. Kent. 1986. The Politics of Blame Avoidance. *Journal of Public Policy* 6(4): 371-398.
- Williams, Larry J., Nathan Hartman, and Flavia Cavazotte. 2010. "Method Variance and Marker Variables: A Review and Comprehensive CFA Marker Technique." *Organizational Research Methods* 13(3): 477-514.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

APPENDICES

A Chapter 2 Supplementary Appendix

A.1 Construction of Variables

Dependent Variable

– Perceived effectiveness of the PIP (Cronbach's alpha = .71)

Two survey items asking the effectiveness of implementing the PIP to poor performers (5-Likert scale)

- Poor performers make a serious effort to use the PIP to improve their performance.
- The PIP often results in the employee improving to an acceptable level or better.

Key Independent Variables

– Validity and usefulness of performance standards (Cronbach's alpha = .86)

Two survey items asking the level of validity and usefulness of performance standards (5-Likert scale)

- The performance standards for my subordinates accurately reflect the critical elements necessary for success in their jobs.
- The performance standards for my subordinates have measurements that I can use accurately to determine success in the critical elements.

– Organizational support (Cronbach's alpha = .75)

Two survey items asking (5-Likert scale)

- Supervisors receive timely training to help them effectively implement a meaningful PIP.
- Human resources staff provide competent assistance regarding PIPs.

– Confidence in delivering performance feedback

Single survey item asking how challenging respondents find communicating with employees about how they can improve their performance (7-Likert scale: Reverse coding)

- Communicating with employees about how they can improve their performance

– Authority to remove poor-performing employees

Single survey item asking certainty to remove subordinates who show poor performance even after the completion of a PIP.

- If a subordinate employee was deficient in a critical performance element after completion of a PIP, are you confident that you would be able to remove that employee?

0 = No / Not sure

1 = Yes

Control variables

– Supervisory status dummies (3 dummies with the executive position as a reference category)

- Team leader
- Supervisor
- Manager

- Executive

– Ratio of unsatisfactory subordinates

Single survey item asking the approximate percentage of the employees that respondents currently supervise directly (with no intermediate supervisors) who are less than fully successful without outright failing at any critical elements of the job. A critical element is a task for which management has told the employee that success in the task is necessary for success in the position.

1 = zero (no employees that are less than successful without outright failing)

2 = 1–10%

3 = 11–20%

4 = 21–30%

5 = over 30%

– Agency dummies (23 agency dummies with the Department of Veterans Affairs as the reference category)

The list of federal agencies: Department of Air Force, Department of Agriculture, Department of Army, Department of Commerce, Department of Defense, Department of Justice, Department of Labor, Department of Energy, Department of Education, Environmental Protection Agency, Federal Deposit Insurance Corporation, General Services Administration, Department of Homeland Security, Department of Housing and Urban Development, Department of Interior, The National Aeronautics and Space Administration, Department of Navy, Office of Personnel Management, U.S. Securities and Exchange Commission, Department of State, The U.S. Social Security Administration,

A.2 Goodness of Fit Indices for Measurement Model and SEM

Model	χ^2	CFI	CFI (Standard)	RMSEA	RMSEA (Standard)	SRMR	SRMR (Standard)
1-factor CFA	$\chi^2(9) = 2321.40$.575	>.90	.296	<.08	0.175	<.06
3-factor CFA	$\chi^2(6) = 5.43$	1.00	>.90	.000	<.08	0.004	<.06
SEM	$\chi^2(177)=824.47$.897	>.90	.035	<.08	0.026	<.06

Note: RMSEA = root mean square error of approximation; CFI: comparative fit index. The above goodness of fit indices are derived from the model without using cluster-robust standard errors. Standards are based on Hu and Bentler (1999).

When comparing the measurement models, the hypothesized model provides better model fit statistics than the one-factor model. The hypothesized model has a higher value of the comparative fit index (CFI), which is greater than the 0.90 value, the conventional value used to suggest a good model fit. Also, the value of the root mean square error of approximation (RMSEA) is lower than the conventional value of 0.08, indicating a good fit for the data. The goodness of fit indices of the SEM analysis show similar results to those of the measurement model.

A.3 Direct Effects, Indirect Effects, and Total Effects of Study Variables

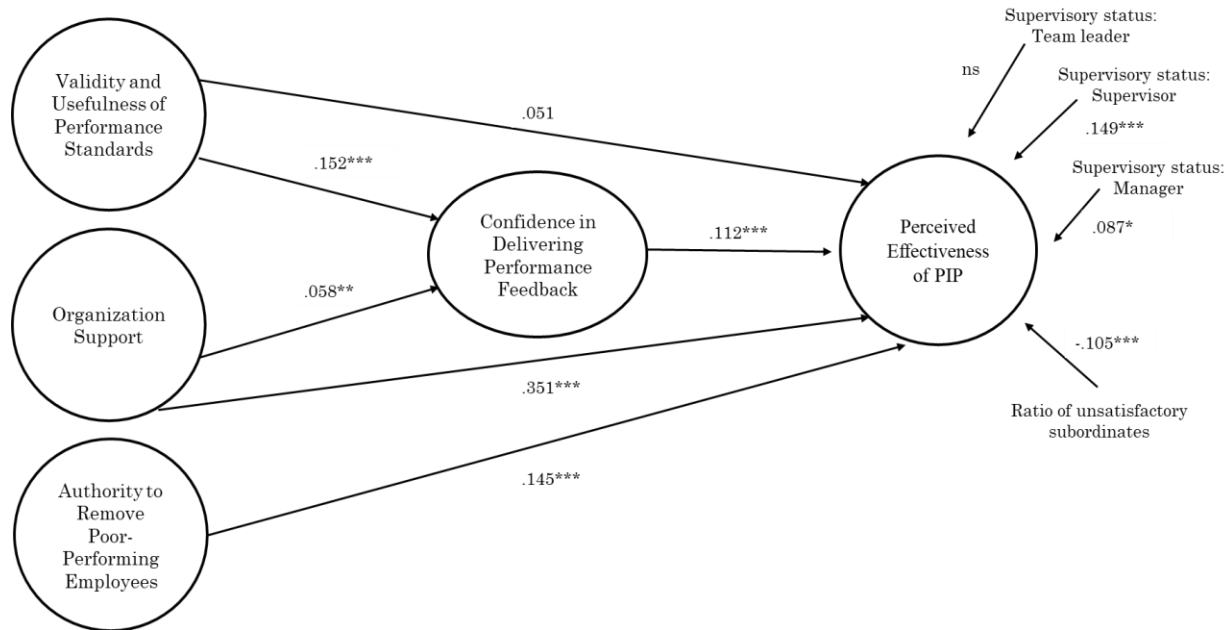
(N = 2,939)	Unstandardized β	Standardized β
Direct effects		
Validity and Usefulness of Performance Standards → Confidence in Delivering Performance Feedback	.272*** (0.041)	.164*** (0.024)
Organizational Support → Confidence in Delivering Performance Feedback	.065* (0.028)	.047* (0.020)
Validity and Usefulness of Performance Standards → Perceived Effectiveness of PIP	.048+ (0.025)	.054* (0.027)
Organizational Support → Perceived Effectiveness of PIP	.265*** (0.024)	.358*** (0.029)
Confidence in Delivering Performance Feedback → Perceived Effectiveness of PIP	.054*** (0.014)	.101*** (0.028)
Authority to Remove Poor-Performing Employees → Perceived Effectiveness of PIP	.195*** (0.044)	.112*** (0.025)
Indirect effects		
Validity and Usefulness of Performance Standards → Perceived Effectiveness of PIP	.015*** (0.004)	.017
Organizational Support → Perceived Effectiveness of PIP	.003** (0.001)	.005
Total effects (Direct + Indirect effects)		
Validity and Usefulness of Performance Standards → Confidence in Delivering Performance Feedback	.272*** (0.041)	.164

Organizational Support		
→ Confidence in Delivering Performance Feedback	.065* (0.028)	.047
Validity and Usefulness of Performance Standards		
→ Perceived Effectiveness of PIP	.062* (0.024)	.071
Organizational Support		
→ Perceived Effectiveness of PIP	.268*** (0.024)	.362
Confidence in Delivering Performance Feedback		
→ Perceived Effectiveness of PIP	.054*** (0.014)	.101
Authority to Remove Poor-Performing Employees		
→ Perceived Effectiveness of PIP	.195*** (0.044)	.112
Agency Fixed Effects		INCLUDED

Note: Robust standard errors are reported in parentheses. Standard errors are adjusted for 24 agency clusters. Standard errors of the standardized effects for indirect effects and total effects are not reported (StataCorp 2019). Due to limited space, this table does not display agency-fixed effects. + p < .10 * p < .05 ** p < .01 *** p < .001

A.4 Sensitivity Analysis

Empirical analysis excluding respondents who answered ‘not sure’ to the question asking the confidence in the capacity of removing poor performers (N=2,494)



Note: Agency fixed effects are included. Standard errors adjusted for 24 agency clusters. Standardized parameter estimates. * $p < .05$ ** $p < .01$ *** $p < .001$

Direct Effects, Indirect Effects, and Total Effects of Study Variables

(N = 2,494)	Unstandardized β	Standardized β
Direct effects		
Validity and Usefulness of Performance Standards → Confidence in Delivering Performance Feedback	.251*** (0.045)	.152*** (0.026)
Organizational Support → Confidence in Delivering Performance Feedback	.079** (0.025)	.058** (0.019)
Validity and Usefulness of Performance Standards → Perceived Effectiveness of PIP	.045 (0.030)	.051 (0.032)

Organizational Support → Perceived Effectiveness of PIP	.254*** (0.030)	.351*** (0.036)
Confidence in Delivering Performance Feedback → Perceived Effectiveness of PIP	.060*** (0.015)	.112*** (0.030)
Authority to Remove Poor- Performing Employees → Perceived Effectiveness of PIP	.245*** (0.052)	.145*** (0.030)
Indirect effects		
Validity and Usefulness of Performance Standards → Perceived Effectiveness of PIP	.015*** (0.003)	.017
Organizational Support → Perceived Effectiveness of PIP	.005** (0.002)	.007
Total effects (Direct + Indirect effects)		
Validity and Usefulness of Performance Standards → Confidence in Delivering Performance Feedback	.251*** (0.045)	.152
Organizational Support → Confidence in Delivering Performance Feedback	.079** (0.025)	.058
Validity and Usefulness of Performance Standards → Perceived Effectiveness of PIP	.060* (0.030)	.068
Organizational Support → Perceived Effectiveness of PIP	.259*** (0.030)	.357
Confidence in Delivering Performance Feedback → Perceived Effectiveness of PIP	.060*** (0.015)	.112
Authority to Remove Poor-	.245***	.145

Performing Employees (0.052)
→ Perceived Effectiveness of
PIP

Agency Fixed Effects	INCLUDED
----------------------	----------

Note: Standard errors are adjusted for 24 agency clusters. Standard errors of the standardized effects for indirect effects and total effects are not reported (StataCorp 2019). Due to limited space, this table does not display agency-fixed effects. + p < .10 * p < .05 ** p < .01 *** p < .001

B Chapter 3 Supplementary Appendix

B.1 Descriptive Statistics and Correlations

(N= 901)					Correlations								
	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9
Study Variables													
1. Ln (IP Amount)	1.89	5.94	-13.82	10.76	1								
2. Career Executive	0.15	0.32	0	1	-0.05*	1							
3. Non-Acting Executive	0.82	0.32	0	1	-0.11*	-0.34*	1						
Control Variables													
4. Executive Tenure	20.97	21.85	0	135	-0.02	0.08*	0.29*	1					
5. Previous Dep. Experience	0.55	0.47	0	1	-0.01	0.33*	-0.25*	-0.09*	1				
6. Executive Turnover	0.32	0.47	0	1	0.03	0.25*	-0.53*	-0.28*	0.21*	1			
7. Disaster Fund	0.10	0.30	0	1	-0.52*	0.01	-0.08*	-0.04	0.08*	0.03	1		
8. Section 57	0.44	0.50	0	1	0.39*	-0.08*	-0.08*	-0.00	-0.16	-0.02	-0.30	1	
9. Exit	0.11	0.31	0	1	-0.30*	-0.02*	0.03*	0.01	0.01	-0.03	0.27*	-0.21*	1
10. Reentry	0.02	0.12	0	1	-0.05*	0.00	0.01	-0.02	0.06	0.01	-0.04*	-0.06	0.07*
11. Number of Programs	6.49	4.74	1	14	-0.03*	-0.18*	0.11*	-0.10*	-0.14*	-0.00	-0.07*	-0.19*	0.04
12. Sampling 1	0.60	0.49	0	1	0.02	0.21*	-0.11*	-0.01	0.09*	0.01	-0.09*	-0.13*	-0.01
13. Sampling 2	0.09	0.29	0	1	0.09*	0.06	-0.03	-0.09*	0.06	0.08*	-0.08*	-0.20*	-0.04
14. Sampling 3	0.27	0.44	0	1	0.06*	-0.21*	0.09*	0.05	-0.15*	-0.06	-0.04	0.33*	-0.05
15. Sampling 4	0.01	0.07	0	1	-0.00	-0.04*	0.04*	-0.04	0.01	-0.02	-0.03	-0.01	0.02
16. IPIA	0.30	0.46	0	1	0.13*	-0.12*	0.14*	-0.03	-0.04	-0.02	-0.22*	0.23*	-0.10*
17. IPERA	0.12	0.32	0	1	0.01	0.01	0.03	-0.12*	-0.04	-0.03	-0.12*	0.08*	-0.03
18. Ln (Outlay Amount)	7.60	2.99	-3.22	13.76	0.65*	0.05*	-0.10*	0.02	0.00	-0.01	-0.65*	0.60*	-0.33*

B.1 Descriptive Statistics and Correlations (cont.)

(N= 901)	Correlations									
	10	11	12	13	14	15	16	17	18	
Control Variables										
10. Reentry	1									
11. Number of Programs	0.06	1								
12. Sampling 1	-0.08*	-0.09*	1							
13. Sampling 2	0.08*	0.19*	-0.39	1						
14. Sampling 3	0.01	-0.03	-0.74*	-0.19*	1					
15. Sampling 4	0.11*	0.04	-0.09*	-0.02	-0.04	1				
16. IPIA	0.02	0.05	-0.16*	0.17*	0.10*	0.05	1			
17. IPERA	0.07*	-0.01	-0.02	-0.07*	0.10*	-0.03	-0.23*	1		
18. Ln (Outlay Amount)	-0.06	-0.24*	0.14*	-0.09*	0.09*	-0.01	0.18*	0.09*	1	

B.2 Reduced Model Estimates

Omitting all control covariates, except the natural log of the total outlays in a given program-year to account for ‘scale’ effects in predicting the volume of improper payments.

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$	Quantile $\tau = 0.2109$	Quantile $\tau = 0.3818$	Quantile $\tau = 0.5483$	Quantile $\tau = 0.6759$	Quantile $\tau = 0.80$	Quantile $\tau = 0.96$
IP (\$) in Millions			IP: \$0	IP: \$1	IP: \$10	IP: \$40	IP: \$100	IP: \$300	IP: \$11,805
Career Executive	–	–1.971** (0.978)	–1.774 (1.688)	–1.810** (0.747)	–1.891** (0.748)	–2.174*** (0.559)	–1.184 (1.034)	–0.533 (0.627)	–0.921** (0.380)
Non-Acting Executive	–	–1.695** (0.714)	–0.800 (1.081)	–0.635 (0.838)	–0.786 (0.581)	–1.488*** (0.345)	–0.994* (0.532)	–0.619* (0.343)	–1.123*** (0.407)
Ln (Outlay)	+	1.273*** (0.077)	1.763*** (0.188)	1.456*** (0.107)	1.185*** (0.166)	1.110*** (0.106)	1.065*** (0.087)	1.070*** (0.061)	1.052*** (0.054)
R^2		0.4315	0.4281	0.4282	0.4293	0.4307	0.4304	0.4262	0.4296
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T=16.767 P > T = 0.000	T=18.685 P > T = 0.000	T=23.958 P > T = 0.000	T=25.724 P > T = 0.000	T=25.948 P > T = 0.000	T=25.799 P > T = 0.000	T=21.475 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by programs appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Alternative Baseline Model: IP Rate (%) as Dependent Variable

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$ IP: 0%	Quantile $\tau = 0.2919$ IP: 0.25%	Quantile $\tau = 0.5671$ IP: 1.53%	Quantile $\tau = 0.6459$ IP: 2.50%	Quantile $\tau = 0.7580$ IP: 5.01%	Quantile $\tau = 0.8801$ IP: 10%	Quantile $\tau = 0.9612$ IP: 20.26%
Career Executive	-	-1.922** (0.851)	-0.961 (1.423)	-1.073** (0.441)	-1.669*** (0.391)	-1.549** (0.633)	-0.846 (0.702)	-0.275 (0.352)	-0.522** (0.235)
Non-Acting Executive	-	-2.106*** (0.661)	-0.868 (0.702)	-1.309** (0.571)	-1.872*** (0.408)	-1.456*** (0.394)	-1.231** (0.514)	-0.428 (0.305)	-0.402 (0.275)
Executive Turnover	+	-0.071 (0.286)	0.362 (0.369)	-0.376 (0.265)	-0.120 (0.274)	0.021 (0.286)	-0.012 (0.244)	0.179 (0.175)	0.674*** (0.238)
Disaster Fund	-	-4.950*** (1.001)	-9.947*** (1.287)	-11.427*** (1.135)	-2.692** (1.274)	-2.251*** (0.735)	-1.616** (0.712)	-1.234** (0.476)	-1.053*** (0.383)
Section 57	+	0.467 (0.572)	1.302 (1.366)	0.071 (0.444)	-0.121 (0.401)	-0.080 (0.416)	0.064 (0.471)	0.083 (0.336)	0.077 (0.284)
Exit	-	-1.414*** (0.521)	-1.258* (0.761)	-1.855*** (0.528)	-1.467*** (0.418)	-1.320*** (0.475)	-1.408*** (0.367)	-0.549 (0.578)	-0.367 (0.267)
Reentry	-	-1.731* (0.917)	-7.941*** (1.422)	-1.391* (0.836)	-1.280* (0.718)	-0.796 (2.463)	-0.663 (1.047)	-0.811* (0.487)	-1.206*** (0.212)
Number of Programs	+	0.095* (0.053)	-0.002 (0.110)	0.088** (0.037)	0.091*** (0.033)	0.076** (0.036)	0.059 (0.038)	0.019 (0.031)	0.038 (0.024)
Sampling1	+/-	0.827 (1.757)	-0.074 (0.873)	0.265 (1.183)	-0.694 (1.695)	0.096 (1.049)	-0.247 (0.764)	0.059 (0.510)	0.448 (0.431)
Sampling2	+/-	2.927 (1.817)	3.883*** (1.444)	2.055 (1.352)	0.740 (1.678)	1.552 (1.107)	0.935 (0.807)	0.808 (0.565)	0.919* (0.480)
Sampling3	+/-	1.024 (1.763)	0.010 (1.143)	0.437 (1.319)	-0.336 (1.725)	0.365 (1.103)	-0.058 (0.730)	-0.010 (0.552)	0.399 (0.441)
Sampling4	+/-	1.092 (2.008)	0.917 (1.641)	-0.048 (1.447)	-2.085 (1.711)	-1.281 (1.118)	-1.771 (1.339)	-0.278 (0.830)	-0.101 (0.531)

IPIA	+/-	-0.632 (0.392)	-0.675 (0.754)	-0.342 (0.332)	-0.463 (0.287)	-0.628** (0.263)	-0.416* (0.252)	-0.138 (0.196)	-0.090 (0.176)
IPERA	+/-	-0.942* (0.492)	-8.905*** (1.938)	-0.115 (0.334)	-0.270 (0.312)	-0.586* (0.314)	-0.280 (0.324)	0.010 (0.279)	-0.020 (0.177)
<i>R</i> ²		0.2263	0.1355	0.1903	0.1990	0.2003	0.1917	0.2020	0.1641
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T=9.627 P > T = 0.000	T=15.856 P > T = 0.000	T=20.457 P > T = 0.000	T=23.455 P > T = 0.000	T=20.436 P > T = 0.000	T=17.49 P > T = 0.000	T=17.13 P > T = 0.000
<i>N</i>		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment rates (%). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. * p < 0.10, ** p < 0.05, *** p < 0.01.

B.4 Alternative Augmented Model: IP Rate (%) as Dependent Variable

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$ IP: 0%	Quantile $\tau = 0.2919$ IP: 0.25%	Quantile $\tau = 0.5671$ IP: 1.53%	Quantile $\tau = 0.6459$ IP: 2.50%	Quantile $\tau = 0.7580$ IP: 5.01%	Quantile $\tau = 0.8801$ IP: 10%	Quantile $\tau = 0.9612$ IP: 20.26%
Career Executive	-	-2.022** (0.935)	-1.220 (1.327)	-1.416** (0.653)	-1.660*** (0.522)	-1.316** (0.701)	-0.788 (0.647)	-0.529 (0.475)	-0.153 (0.455)
Non-Acting Executive	-	-2.218*** (0.724)	-1.129 (0.936)	-1.512** (0.687)	-1.749*** (0.527)	-1.313*** (0.444)	-0.904** (0.414)	-0.548 (0.371)	-0.289 (0.378)
Executive Tenure	-	0.006 (0.009)	0.014 (0.011)	0.006 (0.006)	-0.003 (0.005)	-0.007 (0.006)	-0.010 (0.008)	-0.007* (0.004)	-0.008** (0.003)
Previous Dep. Experience	-	0.045 (0.436)	0.110 (0.730)	0.224 (0.420)	0.095 (0.353)	-0.046 (0.329)	-0.080 (0.334)	-0.201 (0.274)	-0.603** (0.281)
Executive Turnover	+	-0.028 (0.286)	-0.262 (0.379)	-0.289 (0.251)	-0.215 (0.237)	-0.031 (0.246)	-0.034 (0.200)	0.108 (0.153)	0.372** (0.164)
Disaster Fund	-	-4.965*** (1.005)	-10.257*** (1.103)	-11.184*** (0.904)	-2.716 (1.677)	-2.144*** (0.713)	-1.644** (0.641)	-1.173* (0.683)	-0.909** (0.382)
Section 57	+	0.465 (0.558)	1.010 (1.127)	0.153 (0.451)	-0.135 (0.387)	-0.032 (0.433)	0.058 (0.434)	0.063 (0.385)	-0.133 (0.429)
Exit	-	-1.407*** (0.521)	-1.255* (0.758)	-2.047*** (0.534)	-1.453*** (0.415)	-1.291*** (0.399)	-1.454*** (0.394)	-0.674 (0.832)	-0.493 (0.362)
Reentry	-	-1.743* (0.906)	-8.175*** (2.781)	-1.341 (0.901)	-1.294* (0.784)	-0.828 (2.549)	-0.466 (2.061)	-0.314 (0.440)	-0.826** (0.351)
Number of Programs	+	0.098* (0.051)	0.009 (0.096)	0.102** (0.041)	0.094*** (0.033)	0.072** (0.034)	0.051 (0.042)	0.012 (0.033)	0.039 (0.037)
Sampling1	+/-	0.836 (1.750)	-0.528 (0.963)	0.409 (0.978)	-0.759 (1.872)	-0.158 (1.018)	-0.165 (0.788)	0.327 (0.579)	0.597 (0.764)
Sampling2	+/-	2.944 (1.810)	3.185** (1.359)	1.937 (1.302)	-0.686 (1.933)	1.281 (1.075)	0.936 (0.853)	0.975 (0.645)	1.000 (0.783)

Sampling3	+/-	1.015 (1.760)	-0.173 (1.120)	0.473 (1.170)	-0.442 (1.901)	0.110 (1.049)	0.087 (0.843)	0.193 (0.611)	0.404 (0.783)
Sampling4	+/-	1.159 (2.004)	0.880 (1.560)	0.171 (1.303)	-2.016 (1.902)	-1.599 (1.106)	-1.882 (1.405)	-0.603 (0.910)	-0.613 (0.902)
IPIA	+/-	-0.610 (0.396)	-0.721 (0.790)	-0.357 (0.355)	-0.596** (0.285)	-0.564* (0.292)	-0.386 (0.316)	-0.110 (0.243)	-0.104 (0.143)
IPERA	+/-	-0.877* (0.480)	-8.673 (1.807)	-0.058 (0.363)	-0.333 (0.282)	-0.591* (0.301)	-0.552 (0.365)	-0.210 (0.303)	-0.349 (0.219)
R^2		0.2270	0.1368	0.1932	0.1945	0.1941	0.1851	0.1895	0.1313
<i>Parente-Santos Silva test</i> <i>for intra-cluster</i> <i>correlation</i>			T=10.494 P > T = 0.000	T=15.526 P > T = 0.000	T=20.704 P > T = 0.000	T=22.302 P > T = 0.000	T=20.00 P > T = 0.000	T=17.830 P > T = 0.000	T=15.970 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment rates (%). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.5 Reduced Model for Alternative Dependent Variable, IP Rate (%)

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$ IP: 0%	Quantile $\tau = 0.2919$ IP: 0.25%	Quantile $\tau = 0.5671$ IP: 1.53%	Quantile $\tau = 0.6459$ IP: 2.50%	Quantile $\tau = 0.7580$ IP: 5.01%	Quantile $\tau = 0.8801$ IP: 10%	Quantile $\tau = 0.9612$ IP: 20.26%
Career Executive	–	–2.269** (1.018)	–8.181*** (3.107)	–1.509*** (0.519)	–1.938*** (0.440)	–1.736** (0.812)	–0.823 (0.759)	–0.350 (0.353)	–1.031*** (0.237)
Non-Acting Executive	–	–1.799** (0.787)	–3.329 (3.555)	–0.608 (0.525)	–1.545*** (0.418)	–1.419*** (0.445)	–1.045** (0.403)	–0.767** (0.326)	–1.154*** (0.316)
R^2		0.0285	0.0246	0.0246	0.0285	0.0285	0.0255	0.0179	0.0268
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T=14.967 P > T = 0.000	T=21.948 P > T = 0.000	T=26.735 P > T = 0.000	T=26.065 P > T = 0.000	T=25.249 P > T = 0.000	T=21.187 P > T = 0.000	T=21.057 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment rates (%). Robust standard errors clustered by program appear inside parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.6 Model Fit Comparisons

	OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805	
Baseline Model									
R^2	Ln (DV)	0.4829	0.3548	0.4602	0.4731	0.4776	0.4784	0.4684	0.4495
	Unlogged DV	0.2159	0.1443	0.1733	0.1846	0.1889	0.2015	0.1530	0.1458
Model Controlling for Executive Tenure and Previous Department Experience									
R^2	Ln (DV)	0.4833	0.3515	0.4592	0.4724	0.4781	0.4762	0.4655	0.4483
	Unlogged DV	0.2223	0.1381	0.1708	0.1868	0.1916	0.2042	0.1692	0.1527

Above table shows the comparisons of R^2 values between the model using the natural logged dependent variable and the model using the unlogged dependent variable. At all estimated quantiles, higher R^2 values are presented in models using a natural logged measure of the dependent variable than models using an unlogged dependent variable, illustrating the better fit to the data. In the models using an unlogged dependent variable, the results also reveal the same expected sign for key covariates, but are estimated with less precision in some moderate improper payment sizes while estimated with more precision in some small and large improper payment sizes.

B.7 Subsample Analysis of Executive Bureaus (N*T=642)

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805
Career Executive	-	-1.718 (1.066)	-1.281 (4.107)	-1.912* (1.113)	-2.365*** (0.529)	-2.191*** (0.652)	-1.895** (0.889)	-0.933 (1.073)	-0.560 (0.373)
Non-Acting Executive	-	-2.264*** (0.737)	-2.549 (2.997)	-1.950*** (0.620)	-1.779*** (0.439)	-1.665*** (0.462)	-1.752*** (0.568)	-1.302* (0.759)	-0.975* (0.522)
Executive Turnover	+	-0.087 (0.320)	-0.911 (1.059)	-0.019 (0.418)	0.095 (0.205)	0.099 (0.262)	0.093 (0.248)	-0.006 (0.201)	0.118 (0.144)
Disaster Fund	-	-3.090** (1.196)	-8.214*** (2.738)	-6.554*** (2.271)	-4.193*** (1.223)	-2.155 (1.418)	-1.505* (0.804)	-1.127 (0.894)	-0.648 (0.630)
Section 57	+	0.805 (0.928)	1.878 (2.446)	-0.289 (1.103)	-0.447 (0.517)	0.082 (0.509)	0.517 (0.586)	-0.087 (0.688)	-0.555 (0.509)
Exit	-	-1.855*** (0.559)	-3.002 (2.357)	-2.435*** (0.794)	-2.198*** (0.709)	-1.649** (0.703)	-1.323*** (0.454)	-1.562*** (0.508)	-0.370 (0.352)
Reentry	-	-1.305 (0.927)	-3.808 (2.787)	-2.954* (1.683)	-1.239 (1.944)	-1.086 (1.478)	0.215 (2.232)	0.019 (1.267)	-0.961*** (0.369)
Number of Programs	+	0.187*** (0.064)	0.222 (0.176)	0.106 (0.072)	0.102* (0.054)	0.126*** (0.041)	0.117*** (0.035)	0.091 (0.050)	0.055 (0.038)
Sampling1	+/-	-0.993 (1.657)	-0.481 (3.845)	-2.117 (2.442)	-1.619 (1.833)	-0.612 (6.014)	-0.084 (1.254)	-0.827 (1.306)	-1.398* (0.712)
Sampling2	+/-	0.592 (1.775)	2.257 (3.734)	-0.811 (2.727)	-0.530 (2.007)	0.261 (6.021)	0.736 (1.299)	-0.014 (1.419)	-0.854 (0.913)
Sampling3	+/-	0.157 (1.734)	1.418 (4.104)	-1.159 (2.638)	-0.669 (1.886)	-0.054 (6.076)	0.368 (1.303)	-0.316 (1.290)	-1.018 (0.664)
Sampling4	+/-	-0.941 (1.794)	1.095 (4.078)	-2.144 (2.475)	-2.901 (1.895)	-2.267 (6.112)	-1.545 (1.795)	-2.523 (1.936)	-2.584*** (0.914)

IPIA	+/-	-0.402 (0.438)	-0.378 (1.241)	-0.486 (0.612)	-0.242 (0.333)	-0.387 (0.281)	-0.466 (0.285)	-0.539** (0.273)	-0.962** (0.481)
IPERA	+/-	-1.473** (0.715)	-9.597*** (3.377)	-1.001 (0.796)	-0.267 (0.481)	-0.308 (0.363)	-0.462 (0.333)	-0.425 (0.374)	-0.974*** (0.403)
Ln (Outlay)	+	1.108*** (0.204)	0.861* (0.477)	1.205*** (0.311)	1.297*** (0.119)	1.123*** (0.152)	1.042*** (0.139)	1.123*** (0.158)	1.154*** (0.119)
R^2		0.5631	0.4607	0.5427	0.5476	0.5526	0.5522	0.5412	0.5083
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T = 9.018 P > T = 0.000	T=12.289 P > T = 0.000	T=13.143 P > T = 0.000	T=13.512 P > T = 0.000	T=12.840 P > T = 0.000	T=15.182 P > T = 0.000	T=16.596 P > T = 0.000
N		642	642	642	642	642	642	642	642

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.8 Subsample Analysis of Independent Agencies (N*T=259)

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805
Career Executive	-	-0.679 (0.889)	2.591 (5.992)	-0.133 (0.896)	0.141 (0.535)	0.450 (0.460)	0.081 (0.466)	-0.791 (0.668)	0.017 (0.429)
Non-Acting Executive	-	-0.981 (0.959)	1.902 (5.561)	0.009 (0.957)	-0.309 (0.533)	-0.251 (0.708)	-0.493 (0.907)	-0.975 (0.716)	-0.152 (0.228)
Executive Turnover	+	-0.094 (0.576)	1.317 (2.343)	0.120 (0.441)	-0.377 (0.399)	-0.601 (0.413)	-0.533 (0.430)	-0.584 (0.436)	-0.132 (0.256)
Disaster Fund	-	-2.843 (2.451)	-4.130 (9.990)	-2.702 (4.656)	-3.367* (1.826)	-3.160*** (1.080)	-3.203** (1.323)	-2.862*** (0.940)	-2.668*** (0.416)
Section 57	+	1.436 (1.857)	3.649 (11.103)	0.563 (2.986)	0.010 (1.514)	-0.563 (1.113)	-0.389 (0.689)	-0.150 (0.873)	0.453 (0.287)
Exit	-	-1.110 (1.063)	-2.278 (6.526)	-0.382 (1.550)	-1.136 (0.876)	-1.630* (0.880)	-1.941** (0.769)	-1.093 (0.888)	-0.837*** (0.262)
Reentry	-	-2.527 (2.707)	-13.160** (5.342)	-0.119 (3.211)	-1.758* (0.945)	-1.542 (2.284)	-0.735 (0.925)	-0.711 (0.886)	-1.018*** (0.272)
Number of Programs	+	-0.130 (0.149)	-1.503*** (0.277)	0.177 (0.123)	0.130 (0.104)	0.104 (0.088)	0.104** (0.049)	0.096* (0.057)	0.101*** (0.032)
Sampling1	+/-	0.040 (1.848)	2.489 (8.650)	-1.876 (3.014)	-0.346 (2.085)	-0.676 (1.120)	-0.729 (1.296)	1.982 (1.478)	1.596*** (0.383)
Sampling2	+/-	3.094 (2.658)	7.535 (17.155)	2.970 (4.920)	2.530 (2.929)	1.041 (1.223)	0.472 (1.543)	2.521** (1.122)	2.571*** (0.547)
Sampling3	+/-	-0.965 (1.667)	0.862 (7.863)	-2.622 (3.195)	-1.105 (2.023)	-0.940 (0.988)	-1.160 (1.380)	0.780 (0.911)	0.140 (0.343)
Sampling4	+/-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
IPIA	+/-	-1.786	-1.574	-0.731	-0.450	-0.273	0.091	0.411	0.364

		(1.151)	(1.648)	(0.612)	(0.564)	(0.472)	(0.414)	(0.795)	(0.252)
IPERA	+/-	-0.297 (1.025)	-0.935 (1.597)	0.353 (1.013)	0.485 (0.551)	-0.029 (0.340)	-0.167 (0.427)	-0.257 (0.525)	0.118 (0.266)
Ln (Outlay)	+	0.686*** (0.164)	0.640*** (0.182)	1.125*** (0.202)	0.884*** (0.196)	0.852*** (0.221)	0.754*** (0.093)	0.676*** (0.169)	0.713*** (0.097)
R^2		0.3523	0.2340	0.3010	0.3063	0.2965	0.2863	0.2791	0.2824
<i>Parente-Santos Silva test</i>			T = 3.569	T=6.909	T=7.578	T=11.196	T=10.809	T=12.677	T=11.128
<i>for intra-cluster correlation</i>			P > T = 0.000	P > T = 0.000	P > T = 0.000	P > T = 0.000	P > T = 0.000	P > T = 0.000	P > T = 0.000
N		259	259	259	259	259	259	259	259

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). N/A: Results for Sampling4 variable are omitted since there is no variance in this subsample. * p < 0.10, ** p < 0.05, *** p < 0.01.

B.9. Robustness Check for Analyzing the Moderating Effect of Policy Complexity

IP (\$) in Millions		OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805
The Model Controlling for Executive Tenure and Previous Department Experience									
Interaction 1: Career Executive * Policy Complexity	Marginal Effects Based on Interdecile Change	-0.084 [-5.18, 5.01]	-1.348 [-22.56, 19.87]	-0.600 [-4.20, 3.00]	0.151 [-3.47, 3.77]	-0.226 [-3.18, 2.73]	-1.052 [-4.44, 2.33]	-0.320 [-4.94, 4.30]	0.648 [-0.74, 2.04]
Interaction 2: Non-Acting Executive * Policy Complexity	Marginal Effects Based on Interdecile Change	-2.783 [-6.59, 1.03]	-8.454 [-21.81, 4.90]	-3.927** [-7.75, -0.11]	-2.153 [-5.31, 1.01]	-0.599 [-2.80, 1.61]	-0.275 [-3.27, 2.72]	-0.273 [-2.20, 1.66]	-0.755 [-1.97, 0.47]
Joint Hypotheses Test of Interaction Terms		F=1.49 Prob>F= 0.23	F=1.87 Prob>F= 0.16	F=5.04 Prob>F= 0.01	F=2.12 Prob>F= 0.12	F=0.17 Prob>F= 0.85	F=0.45 Prob>F= 0.64	F=0.04 Prob>F= 0.96	F=3.50 Prob>F= 0.03

Note: 95% confidence intervals appear inside brackets. The full set of regression results for the models including interaction terms is available upon request. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.10 Robustness Check for Analyzing the Moderating Effect of Executive Tenure

IP (\$) in Millions		OLS (Mean)	Quantile $\tau = 0.0921$ IP: \$0	Quantile $\tau = 0.2109$ IP: \$1	Quantile $\tau = 0.3818$ IP: \$10	Quantile $\tau = 0.5483$ IP: \$40	Quantile $\tau = 0.6759$ IP: \$100	Quantile $\tau = 0.80$ IP: \$300	Quantile $\tau = 0.96$ IP: \$11,805
The Model Controlling for Executive Tenure and Previous Department Experience									
Interaction 1: Career Executive * Executive Tenure	Marginal Effects Based on Interdecile Change	0.346 [-1.59, 2.28]	1.307 [-3.80, 6.41]	0.002 [-1.50, 1.50]	0.149 [-1.16, 1.46]	-0.495 [-1.58, 0.59]	-1.059 [-2.76, 0.64]	-1.256*** [-2.04, -0.47]	-0.148 [-1.12, 0.82]
Interaction 2: Non-Acting Executive * Executive Tenure	Marginal Effects Based on Interdecile Change	2.352 [-1.85, 6.55]	7.379* [-1.07, 15.82]	2.281 [-6.89, 11.45]	0.552 [-3.41, 4.52]	1.994 [-1.10, 5.09]	1.514 [-1.71, 4.74]	1.261 [-0.46, 2.99]	1.478** [0.32, 2.64]
Joint Hypotheses Test of Interaction Terms		F=0.66 Prob>F=0.52	F=1.53 Prob>F=0.22	F=0.12 Prob>F=0.88	F=0.05 Prob>F=0.96	F=1.15 Prob>F=0.32	F=2.50 Prob>F=0.08	F=7.13 Prob>F=0.00	F=4.66 Prob>F=0.01

Note: 95% confidence intervals appear inside brackets. The full set of regression results for the models including interactions terms is available upon request. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.11 Robustness Check Controlling Different Presidential Administrations

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.0921$	Quantile $\tau = 0.2109$	Quantile $\tau = 0.3818$	Quantile $\tau = 0.5483$	Quantile $\tau = 0.6759$	Quantile $\tau = 0.80$	Quantile $\tau = 0.96$
IP (\$) in Millions			IP: \$0	IP: \$1	IP: \$10	IP: \$40	IP: \$100	IP: \$300	IP: \$11,805
Career Executive	–	–1.633* (0.959)	–1.082 (2.308)	–0.668 (0.847)	–1.683** (0.717)	–1.497*** (0.473)	–1.373 (1.040)	–0.747 (0.676)	–0.270 (0.230)
Non-Acting Executive	–	–2.030*** (0.677)	–1.049 (1.304)	–1.157 (0.936)	–1.864*** (0.638)	–1.743*** (0.417)	–1.434*** (0.516)	–0.954** (0.467)	–0.105 (0.220)
Republican	+/-	0.974** (0.413)	0.690 (1.012)	0.324 (0.633)	0.500 (0.335)	0.367 (0.253)	0.377 (0.325)	0.393 (0.285)	–0.131 (0.161)
Ln (Outlay)	+	1.054*** (0.139)	0.885 (0.538)	1.034*** (0.169)	1.067*** (0.157)	0.922*** (0.127)	0.950*** (0.129)	1.041*** (0.092)	1.076*** (0.077)
R^2		0.4851	0.3485	0.4599	0.4746	0.4788	0.4799	0.4697	0.4474
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T = 10.64 P > T = 0.000	T=13.291 P > T = 0.000	T=18.447 P > T = 0.000	T=20.483 P > T = 0.000	T=22.986 P > T = 0.000	T=19.745 P > T = 0.000	T=19.626 P > T = 0.000
N		901	901	901	901	901	901	901	901

Note: The dependent variable is the natural log of improper payment amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. IP indicates Improper Payment Amount in Millions (\$). The estimates of other control variables besides program outlay and presidential administrations are not presented in this table. The full set of regression results for the models including interactions terms is available upon request. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Chapter 4 Supplementary Appendix

C.1 Descriptive Statistics of Variables in Analysis

(N=807)	Mean	SD	min	max
<u>Study Variables</u>				
Ln (Overpayment)	15.10	6.38	0	24.52
Ln (Underpayment)	9.47	7.99	0	22.59
Career Executive	0.16	0.33	0	1
Non-Acting Executive	0.82	0.32	0	1
<u>Control Variables</u>				
Ideologically Aligned Agency	0.32	0.47	0	1
Ideologically Opposed Agency	0.23	0.42	0	1
Agency Insulation	-0.25	0.68	-0.85	1.84
Executive Turnover	0.32	0.47	0	1
Disaster Fund	0.12	0.32	0	1
Section 57	0.41	0.49	0	1
Exit	0.11	0.32	0	1
Reentry	0.02	0.13	0	1
Number of Programs	6.4	4.69	1	14
Sampling1	0.62	0.49	0	1
Sampling2	0.09	0.28	0	1
Sampling3	0.26	0.44	0	1
Sampling4	0.00	0.06	0	1
IPIA	0.23	0.42	0	1
IPERA	0.12	0.33	0	1
Ln (Outlay Amount)	21.34	3.13	10.59	27.51

C.2 Bivariate Correlations of Variables in Analysis

	Correlations									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Ln (Overpayment)	1									
(2) Ln (Underpayment)	0.54*	1								
(3) Career Executive	-0.04*	-0.01*	1							
(4) Non-Acting Executive	-0.11*	-0.08*	-0.32*	1						
(5) Aligned Agency	-0.10*	-0.23*	0.00	0.05*	1					
(6) Opposed Agency	0.07*	0.06*	-0.03*	-0.14*	-0.38*	1				
(7) Agency Insulation	-0.02*	-0.08*	-0.16*	0.05*	0.19*	0.05*	1			
(8) Executive Turnover	0.03*	0.01*	0.25*	-0.53*	-0.04*	-0.01*	-0.02*	1		
(9) Disaster Fund	-0.50*	-0.36*	-0.00*	-0.08*	0.06*	0.06*	-0.10*	0.03*	1	
(10) Section 57	0.42*	0.40*	-0.08*	-0.09*	-0.07*	0.14*	0.05*	-0.04*	-0.23*	1
(11) Exit	-0.32*	-0.22*	-0.02*	0.04*	0.01*	-0.01*	-0.04*	-0.03*	0.27*	-0.19*
(12) Reentry	-0.04*	-0.05*	0.01*	0.00*	-0.03*	-0.05*	0.03*	0.02*	-0.05*	-0.05*
(13) Number of Programs	-0.02*	-0.11*	-0.18*	0.13*	0.07*	-0.20*	-0.07*	0.01*	-0.08*	-0.19*
(14) Sampling1	-0.01*	0.07*	0.20*	-0.10*	0.02*	0.17*	-0.16*	-0.00*	-0.11*	-0.13*
(15) Sampling2	0.08*	-0.00*	-0.06*	0.01*	-0.13*	-0.07*	0.01*	0.12*	-0.09*	-0.17*
(16) Sampling3	0.10*	0.02*	-0.22*	0.09*	0.05*	-0.13*	0.20*	-0.08*	-0.03*	0.33*
(17) Sampling4	0.00	-0.01*	-0.03*	0.04*	-0.04*	-0.03*	-0.05*	0.00	-0.02*	-0.05*
(18) IPIA	0.12*	0.17*	-0.10*	0.14*	-0.13*	0.02*	-0.05*	-0.01*	-0.20*	0.18*
(19) IPERA	0.03*	-0.03*	0.00*	0.03*	0.06*	-0.04*	0.02*	-0.04*	-0.14*	0.10*
(20) Ln (Outlay Amount)	0.63*	0.53*	0.03*	-0.08*	-0.08*	0.11*	0.04*	-0.02*	-0.66*	0.62*

Note: * $p < .05$

C.2 Bivariate Correlations of Variables in Analysis (Cont.)

	Correlations									
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(11) Exit	1									
(12) Reentry	0.05*	1								
(13) Number of Programs	0.05*	0.08*	1							
(14) Sampling1	-0.01*	-0.10*	-0.07*	1						
(15) Sampling2	-0.06*	0.10*	0.16*	-0.39*	1					
(16) Sampling3	-0.05*	0.02*	-0.04*	-0.74*	-0.18*	1				
(17) Sampling4	0.04*	0.15*	0.10*	-0.08*	-0.02*	-0.04*	1			
(18) IPIA	-0.10*	0.05*	0.03*	-0.14*	0.18*	0.09*	0.02*	1		
(19) IPERA	-0.04*	0.04*	0.00	-0.04*	-0.06*	0.12*	-0.02*	-0.20*	1	
(20) Ln (Outlay Amount)	-0.33*	-0.06*	-0.23*	0.14*	-0.10*	0.12*	-0.02*	0.20*	0.10*	1

Note: * $p < .05$

C.3 Alternative Censored Quantile Regression (CQR) Estimates for U.S. Federal Program Overpayment Errors

Covariates	Sign	Quantile $\tau = 0.10$ OP=\$0	Quantile $\tau = 0.25$ OP=\$0.88 million	Quantile $\tau = 0.40$ OP=\$8 million	Quantile $\tau = 0.55$ OP=\$31 million	Quantile $\tau = 0.70$ OP=\$112 million	Quantile $\tau = 0.85$ OP=\$512 million	Quantile $\tau = 0.97$ OP=\$10.27 billion
Career Executive	-	-1.041 (2.675) [-6.283, 4.201]	-1.768 (1.582) [-4.870, 1.333]	-1.687** (0.791) [-3.238, -0.136]	-1.400** (0.549) [-2.475, -0.324]	-1.242** (0.519) [-2.260, -0.224]	-0.542 (0.557) [-1.634, 0.550]	0.069 (0.527) [-0.965, 1.102]
Non-Acting Executive	-	-3.373* (2.799) [-8.860, 2.113]	-1.046 (1.351) [-3.693, 1.601]	-1.245* (0.816) [-2.845, 0.354]	-1.478** (0.627) [-2.707, -0.249]	-1.182** (0.518) [-2.198, -0.167]	-0.659 (0.552) [-1.740, 0.423]	-0.113 (0.526) [-1.145, 0.918]
Ln (Outlay)	+	0.677 (0.496) [-0.295, 1.650]	1.102*** (0.222) [0.666, 1.538]	1.034*** (0.167) [0.707, 1.362]	0.896*** (0.103) [0.695, 1.097]	0.905*** (0.076) [0.757, 1.054]	0.907*** (0.071) [0.767, 1.046]	1.106*** (0.090) [0.090, 1.283]

Note: Only primary covariates and program outlays accounting for scaling effects are reported due to purposes of brevity. The dependent variable is the natural log of overpayment error amounts (adjusted for 2015 constant dollars). OP indicates overpayment error amounts (\$). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**C.4 Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors
(Estimated with Bootstrapped Standard Errors)**

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.10$ OP=\$0	Quantile $\tau = 0.25$ OP=\$0.88 million	Quantile $\tau = 0.40$ OP=\$8 million	Quantile $\tau = 0.55$ OP=\$31 million	Quantile $\tau = 0.70$ OP=\$112 million	Quantile $\tau = 0.85$ OP=\$512 million	Quantile $\tau = 0.97$ OP=\$10.27 billion
Career Executive	-	-1.263** (0.599)	-0.596 (1.711)	-1.063* (0.638)	-1.777*** (0.415)	-1.717*** (0.374)	-1.470*** (0.470)	-0.610 (0.431)	-0.242 (0.567)
Non-Acting Executive	-	-1.937*** (0.502)	-2.142 (1.320)	-1.085** (0.480)	-1.304** (0.437)	-1.568*** (0.399)	-1.294*** (0.347)	-0.661* (0.378)	-0.096 (0.578)
Aligned Agency	+	-0.285 (0.394)	-0.258 (1.551)	0.005 (0.437)	0.041 (0.261)	0.265 (0.220)	0.153 (0.253)	0.103 (0.249)	0.426 (0.410)
Opposed Agency	-	0.419 (0.436)	0.671 (1.292)	0.586 (0.469)	0.750** (0.304)	0.808*** (0.286)	1.028*** (0.277)	0.446* (0.247)	0.312 (0.475)
Agency Insulation	-	-0.719*** (0.270)	-3.564*** (1.196)	-1.341*** (0.472)	-0.525* (0.284)	-0.335* (0.196)	-0.260 (0.185)	0.020 (0.136)	-0.110 (0.181)
Executive Turnover	+	-0.045 (0.379)	0.212 (0.898)	-0.021 (0.378)	-0.037 (0.266)	0.107 (0.253)	0.213 (0.230)	0.075 (0.211)	0.876** (0.345)
Disaster Fund	-	-3.147*** (0.831)	-5.464* (3.041)	-7.505*** (1.608)	-3.755** (1.648)	-3.459*** (1.022)	-2.151*** (0.675)	-1.419** (0.585)	-1.376** (0.662)
Section 57	+	0.986* (0.571)	3.472 (2.233)	0.522 (0.591)	-0.049 (0.369)	-0.078 (0.316)	0.078 (0.310)	0.409 (0.320)	-0.691 (0.538)
Exit	-	-2.132*** (0.662)	-5.265** (2.193)	-2.435*** (0.786)	-2.643*** (0.524)	-1.866*** (0.628)	-1.443*** (0.338)	-1.379*** (0.457)	-0.561 (0.498)
Reentry	-	-1.306 (1.215)	-2.321 (2.330)	-2.419 (2.199)	-0.987 (1.371)	-1.692 (1.203)	-0.800 (1.082)	-1.196 (0.961)	-0.746 (0.789)
Number of Programs	+	0.126*** (0.045)	0.186 (0.151)	0.113*** (0.041)	0.123*** (0.027)	0.109*** (0.025)	0.086*** (0.026)	0.034 (0.030)	0.056 (0.042)
Sampling1	+/-	-0.763 (1.137)	-3.795* (2.163)	-1.976 (1.684)	0.389 (2.256)	-0.638 (2.465)	0.228 (1.307)	0.698 (0.974)	-1.379 (1.224)
Sampling2	+/-	2.251* (1.184)	1.418 (2.992)	-0.720 (1.842)	1.926 (2.279)	0.996 (2.403)	1.484 (1.324)	1.703* (1.003)	-1.108 (1.349)

Sampling3	+/-	0.262 (1.136)	0.140 (2.382)	-1.135 (1.744)	1.132 (2.213)	0.005 (2.405)	0.927 (1.290)	0.811 (0.936)	-1.118 (1.174)
Sampling4	+/-	0.796 (1.432)	2.049 (3.444)	-1.764 (2.630)	1.353 (2.784)	-1.054 (2.891)	0.563 (1.933)	0.898 (1.628)	-2.635 (2.150)
IPIA	+/-	-1.192*** (0.429)	-2.105 (1.920)	-0.817* (0.448)	-0.417 (0.259)	-0.487** (0.211)	-0.426* (0.245)	-0.193 (0.224)	-0.219 (0.371)
IPERA	+/-	-1.154* (0.610)	-7.193** (3.604)	-0.557 (0.700)	-0.270 (0.313)	-0.321 (0.273)	-0.139 (0.262)	-0.055 (0.240)	-0.137 (0.381)
Ln (Outlay)	+	1.003*** (0.103)	0.906*** (0.341)	1.006*** (0.141)	1.079*** (0.101)	0.976*** (0.085)	0.914*** (0.067)	0.905*** (0.073)	1.041*** (0.107)
Adjusted (Pseudo) R^2		0.4600	0.2932	0.3730	0.3700	0.3700	0.3812	0.4219	0.4306
N		807	807	807	807	807	807	807	807

Note: The dependent variable is the natural log of overpayment error amounts (adjusted for 2015 constant dollars). OP indicates overpayment error amounts (\$). Key covariates of interest appear in shaded regions. Bootstrap standard errors (500 replications) are in parentheses. Adjusted R-squared values are reported for OLS regression while Pseudo R-squared values are reported for quantile regression.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**C.5 Alternative Regression Model Estimates for U.S. Federal Program Underpayment Errors
(Estimated with Bootstrapped Standard Errors)**

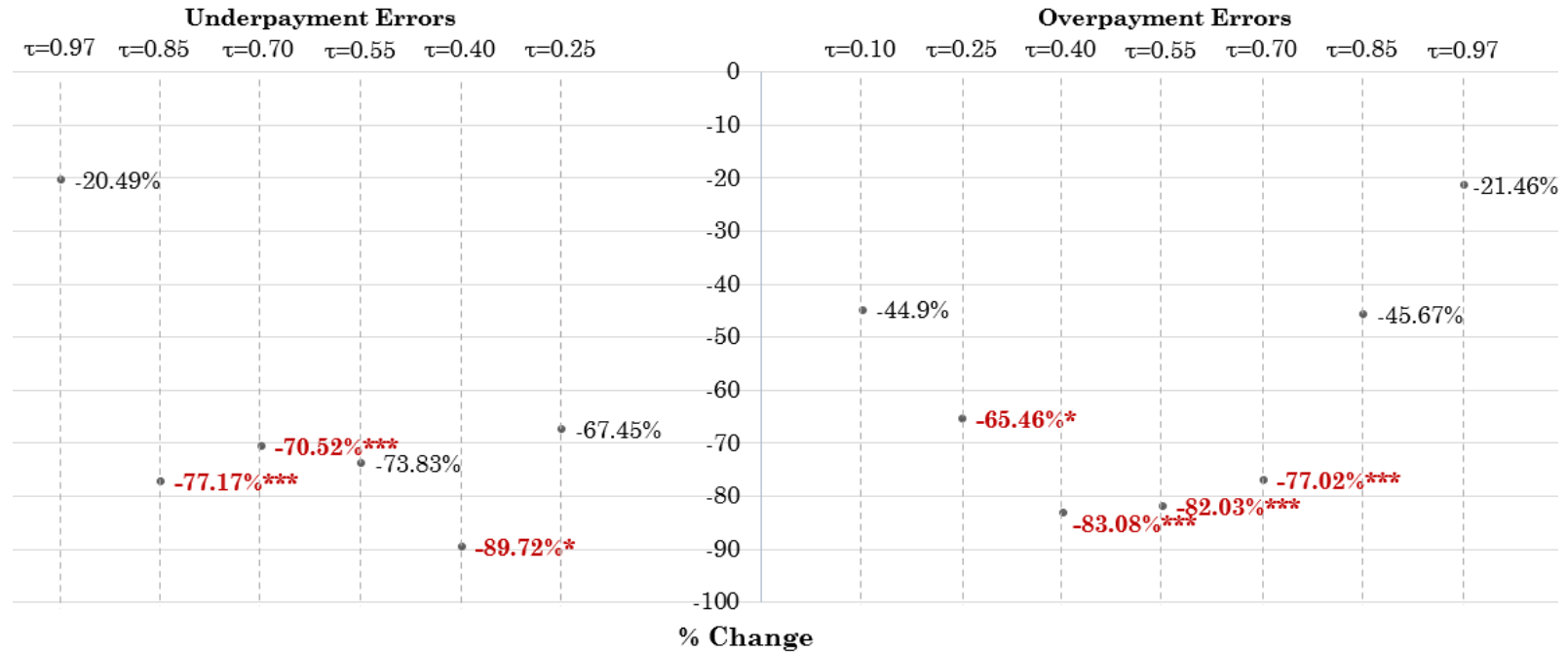
Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.25$ UP=\$0	Quantile $\tau = 0.40$ UP=\$3,000	Quantile $\tau = 0.55$ UP=0.55 million	Quantile $\tau = 0.70$ UP=4.6 million	Quantile $\tau = 0.85$ UP=\$77 million	Quantile $\tau = 0.97$ UP=1.083 billion
Career Executive	-	-1.319 (0.839)	-1.122 (1.643)	-2.275* (1.365)	-1.341 (0.819)	-1.221*** (0.440)	-1.477*** (0.515)	-0.229 (0.651)
Non-Acting Executive	-	-1.145 (0.962)	-1.930 (1.972)	-1.141 (1.113)	-0.378 (0.612)	-0.588 (0.425)	-0.575 (0.467)	-0.186 (0.519)
Aligned Agency	-	-3.384*** (0.548)	-2.172 (2.233)	-5.276*** (1.316)	-2.319** (0.923)	-1.200** (0.471)	-1.184*** (0.311)	-0.844*** (0.315)
Opposed Agency	+	-1.683*** (0.653)	-1.629 (1.822)	-1.508 (0.995)	-0.861 (0.539)	-0.437 (0.336)	-0.556* (0.330)	-0.387 (0.374)
Agency Insulation	-	-0.634* (0.361)	0.171 (0.620)	-2.603*** (0.550)	-2.071*** (0.665)	-0.691*** (0.226)	-0.177 (0.215)	-0.104 (0.169)
Executive Turnover	+	-0.112 (0.610)	-0.731 (1.241)	-0.278 (0.729)	0.255 (0.421)	0.302 (0.272)	0.283 (0.274)	-0.135 (0.273)
Disaster Fund	-	-1.405 (1.008)	0.213 (1.363)	-2.089 (1.668)	-5.196*** (1.393)	-6.487*** (1.370)	-2.588 (1.946)	-0.588 (0.834)
Section 57	+	2.377*** (0.718)	8.679*** (3.278)	2.593 (1.661)	2.006*** (0.654)	1.476*** (0.461)	1.508*** (0.358)	0.784* (0.453)
Exit	-	-1.265* (0.716)	-0.275 (0.907)	-2.772** (1.201)	-1.102 (0.865)	-1.127* (0.607)	-0.641 (0.615)	-1.100** (0.551)
Reentry	-	-1.304 (2.016)	-0.482 (2.876)	-1.668 (3.204)	-0.508 (3.639)	-0.567 (3.645)	-0.029 (1.708)	-0.678 (0.713)
Number of Programs	+	-0.000 (0.062)	0.002 (0.097)	-0.064 (0.113)	0.014 (0.057)	0.025 (0.038)	0.050* (0.029)	0.034 (0.045)
Sampling1	+/-	-2.123** (0.988)	-1.178 (1.642)	-3.674** (1.569)	-2.935** (1.440)	-2.565** (1.233)	1.342 (2.345)	-0.043 (1.965)
Sampling2	+/-	-1.861 (1.303)	-1.459 (2.465)	-3.513* (2.092)	-2.748* (1.570)	-1.893* (1.353)	1.596 (2.356)	-0.603 (2.005)

Sampling3	+/-	-3.318*** (1.057)	-2.598 (2.646)	-4.288** (1.680)	-3.531*** (1.357)	-2.747** (1.227)	0.812 (2.379)	-0.966 (1.906)
Sampling4	+/-	-3.056 (5.562)	-3.002 (7.923)	0.810 (8.578)	-2.160 (7.337)	-2.619** (6.080)	-0.288 (5.357)	-2.615 (5.002)
IPIA	+/-	0.170 (0.622)	-0.565 (2.219)	0.611 (0.816)	0.051 (0.439)	0.308 (0.267)	0.280 (0.266)	0.125 (0.351)
IPERA	+/-	-1.636** (0.788)	-1.317 (1.840)	-3.563** (1.698)	-0.864 (1.756)	-0.456 (0.374)	-0.409 (0.407)	0.241 (0.422)
Ln (Outlay)	+	1.070*** (0.135)	0.392 (0.410)	1.289*** (0.227)	1.149*** (0.136)	0.987*** (0.109)	0.994*** (0.088)	0.992*** (0.100)
<i>R</i> ²		0.3390	0.0320	0.2922	0.2967	0.3252	0.3505	0.3510
<i>N</i>		807	807	807	807	807	807	807

Note: The dependent variable is the natural log of underpayment error amounts (adjusted for 2015 constant dollars). UP indicates underpayment error amounts (\$). Since programs with an absence of underpayment errors account for 38.29% of the sample of data, the 10% quantile was not calculated to avoid redundancy. Key covariates of interest appear in shaded regions. Bootstrap standard errors (500 replications) are in parentheses. Adjusted R-squared values are reported for OLS regression while Pseudo R-squared values are reported for quantile regression.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

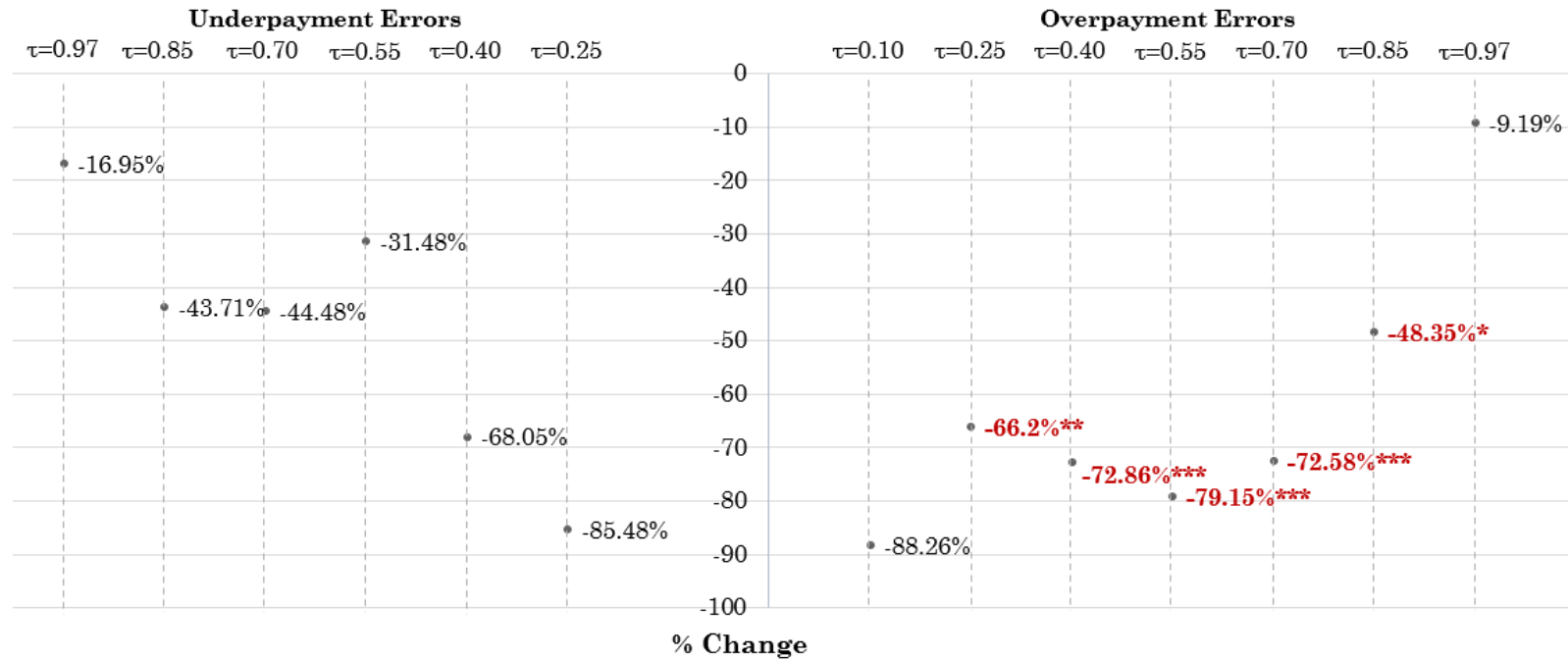
C.6 Average Differential Marginal Effects of Career Executives Versus Political Executives in Underpayment Errors and Overpayment Errors Using Alternative Regression Model with Bootstrapped Standard Errors



Note: Point estimates represent the average differential marginal effects (% factor change) between career executive-managed programs versus political executive-managed programs with respect to overpayment errors and underpayment errors at each quantile. Since programs with an absence of underpayment errors account for 38.29% of the sample, the 10th quantile was not calculated for underpayment errors to avoid redundancy.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.7 Average Differential Marginal Effects of Non-Acting Executives Versus Acting Executives in Underpayment Errors and Overpayment Errors Using Alternative Regression Model with Bootstrapped Standard Errors



Note: Point estimates represent the average differential marginal effects (% factor change) between acting executive-managed programs versus non-acting executive-managed programs with respect to overpayment errors and underpayment errors at each quantile. Since programs with an absence of underpayment errors account for 38.29% of the sample, the 10th quantile was not calculated for underpayment errors to avoid redundancy.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.8 Alternative Regression Model Estimates for U.S. Federal Program Overpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables)

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.10$ OP=\$0	Quantile $\tau = 0.25$ OP=\$0.88 million	Quantile $\tau = 0.40$ OP=\$8 million	Quantile $\tau = 0.55$ OP=\$31 million	Quantile $\tau = 0.70$ OP=\$112 million	Quantile $\tau = 0.85$ OP=\$512 million	Quantile $\tau = 0.97$ OP=\$10.27 billion
Career Executive	-	-1.395 (1.047)	-0.832 (2.252)	-0.892 (0.907)	-1.747** (0.716)	-1.744*** (0.614)	-1.617*** (0.542)	-0.839 (0.530)	-0.394 (0.310)
Non-Acting Executive	-	-1.981*** (0.720)	-2.371 (1.774)	-0.974* (0.560)	-1.519** (0.632)	-1.691*** (0.400)	-1.612*** (0.373)	-0.589 (0.369)	-0.402* (0.238)
President_Republican	-	0.691 (0.527)	1.019 (2.193)	0.060 (0.472)	0.268 (0.338)	0.345 (0.307)	0.220 (0.273)	-0.015 (0.299)	-0.519*** (0.170)
Agency Insulation	-	-0.785 (0.615)	-3.764** (1.468)	-1.324 (0.823)	-0.504 (0.572)	-0.500 (0.476)	-0.149 (0.379)	0.014 (0.208)	0.194** (0.097)
Executive Turnover	+	-0.048 (0.352)	-0.044 (0.841)	0.087 (0.413)	-0.255 (0.266)	-0.146 (0.303)	0.191 (0.262)	0.189 (0.182)	0.792*** (0.163)
Disaster Fund	-	-3.097*** (1.054)	-5.190 (3.716)	-6.986*** (1.830)	-3.678** (1.562)	-3.418*** (0.988)	-2.094*** (0.692)	-1.228** (0.579)	-0.838* (0.482)
Section 57	+	1.059 (0.964)	3.717 (3.331)	0.464 (0.759)	0.166 (0.530)	0.133 (0.498)	0.219 (0.444)	0.373 (0.519)	-0.432 (0.470)
Exit	-	-2.127*** (0.553)	-5.208** (2.409)	-2.592*** (0.553)	-2.443*** (0.587)	-2.010*** (0.714)	-1.434*** (0.315)	-1.544*** (0.358)	-0.491 (0.361)
Reentry	-	-1.264 (1.174)	-2.031 (2.111)	-2.417** (1.101)	-0.960 (1.265)	-1.520 (1.081)	-0.274 (0.866)	-1.388** (0.612)	-1.714*** (0.204)
Number of Programs	+	0.117 (0.076)	0.166 (0.143)	0.120* (0.064)	0.113*** (0.043)	0.103** (0.044)	0.079** (0.039)	0.016 (0.038)	0.070** (0.029)
Sampling1	+/-	-0.753 (1.363)	-3.048 (2.457)	-1.337 (1.825)	0.983 (2.475)	-0.362 (3.543)	0.061 (1.064)	0.632 (0.854)	-1.438*** (0.536)
Sampling2	+/-	2.218 (1.476)	2.035 (4.346)	-0.303 (2.334)	2.497 (2.565)	1.070 (3.498)	1.285 (1.118)	1.626* (0.942)	-1.197* (0.712)
Sampling3	+/-	0.228 (1.367)	0.755 (3.235)	-0.602 (1.981)	1.512 (2.342)	0.146 (3.473)	0.527 (1.128)	0.724 (0.832)	-0.852* (0.496)

Sampling4	+/-	0.864 (1.557)	3.507 (4.635)	-1.488 (2.636)	1.596 (2.687)	-0.940 (3.254)	-0.174 (1.435)	0.983 (1.082)	-1.736** (0.723)
IPIA	+/-	-1.705* (0.941)	-3.640 (2.950)	-0.751 (0.882)	-0.686 (0.478)	-0.743** (0.339)	-0.686* (0.352)	-0.166 (0.387)	0.026 (0.210)
IPERA	+/-	-1.099* (0.646)	-7.166* (3.923)	-0.677 (0.838)	-0.218 (0.366)	-0.251 (0.285)	-0.191 (0.262)	0.013 (0.322)	-0.192 (0.281)
Ln (Outlay)	+	0.999*** (0.156)	0.872** (0.420)	1.013*** (0.213)	1.047*** (0.150)	0.964*** (0.144)	0.929*** (0.117)	0.922*** (0.096)	1.065*** (0.078)
R^2		0.4716	0.3722	0.4437	0.4557	0.4602	0.4576	0.4401	0.4160
<i>Parente-Santos Silva test for intra-cluster correlation</i>			T = 11.20 P > T = 0.000	T = 12.76 P > T = 0.000	T = 14.80 P > T = 0.000	T = 18.88 P > T = 0.000	T = 19.06 P > T = 0.000	T = 16.98 P > T = 0.000	T = 16.14 P > T = 0.000
N		807	807	807	807	807	807	807	807

Note: The dependent variable is the natural log of overpayment error amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Key covariates of interest appear in shaded regions. OP indicates overpayment error amounts (\$).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.9 Alternative Regression Model Estimates for U.S. Federal Program Underpayment Errors (Including Dummy Indicator for Republican President Instead of President-Agency Alignment Variables)

Covariates	Expected Sign	OLS (Mean)	Quantile $\tau = 0.25$ UP=\$0	Quantile $\tau = 0.40$ UP=\$3,000	Quantile $\tau = 0.55$ UP=0.55 million	Quantile $\tau = 0.70$ UP=4.6 million	Quantile $\tau = 0.85$ UP=\$77 million	Quantile $\tau = 0.97$ UP=1.083 billion
Career Executive	-	-1.433 (1.362)	NA	-1.953 (2.299)	-0.908 (1.159)	-1.110 (0.693)	-1.284* (0.714)	-0.409 (0.493)
Non-Acting Executive	-	-1.004 (1.275)	NA	-0.603 (1.315)	-0.310 (0.710)	-0.331 (0.504)	-0.672 (0.524)	-0.362 (0.335)
President_Republican	+	1.068 (0.781)	NA	1.055 (1.150)	0.477 (0.614)	0.121 (0.471)	0.048 (0.466)	0.371 (0.307)
Agency Insulation	-	-1.276 (0.807)	NA	-2.735** (1.158)	-1.360 (1.095)	-0.734* (0.402)	-0.432 (0.385)	-0.316** (0.129)
Executive Turnover	+	0.051 (0.572)	NA	-0.255 (0.827)	0.088 (0.446)	0.310 (0.298)	0.313 (0.323)	-0.138 (0.144)
Disaster Fund	-	-2.097 (1.553)	NA	-2.505 (2.574)	-6.385*** (1.689)	-7.586*** (1.306)	-4.186 (3.368)	-1.567*** (0.575)
Section 57	+	2.392 (1.551)	NA	5.643* (3.289)	2.362** (1.186)	1.701** (0.760)	1.213* (0.677)	0.598 (0.365)
Exit	-	-1.191 (0.792)	NA	-1.876 (1.237)	-0.990 (0.681)	-1.196** (0.538)	0.312 (0.700)	-1.586*** (0.342)
Reentry	-	-0.970 (1.894)	NA	-1.405 (3.238)	-0.545 (2.160)	-0.468 (1.190)	1.012 (1.218)	-0.576*** (0.207)
Number of Programs	+	-0.020 (0.138)	NA	0.027 (0.186)	0.000 (0.081)	0.027 (0.051)	0.061 (0.047)	0.031 (0.025)
Sampling1	+/-	-2.400* (1.268)	NA	-3.373** (1.518)	-2.199** (0.927)	-2.742** (1.088)	-0.836 (3.695)	-0.154 (0.816)
Sampling2	+/-	-1.380 (1.702)	NA	-1.881 (2.443)	-0.942 (1.452)	-1.731 (1.147)	-0.188 (3.781)	-0.797 (0.795)
Sampling3	+/-	-3.182** (1.435)	NA	-4.388** (1.858)	-2.143* (1.135)	-2.698** (1.084)	-1.285 (3.941)	-0.812 (0.774)

Sampling4	+/-	-2.150 (3.235)	NA	2.894 (3.784)	-0.500 (2.242)	-2.556** (1.297)	-2.999 (3.959)	-2.382** (0.953)
IPIA	+/-	-0.493 (1.228)	NA	0.978 (1.438)	-0.175 (0.764)	-0.061 (0.617)	0.257 (0.640)	-0.393 (0.339)
IPERA	+/-	-1.696** (0.750)	NA	-4.549** (2.020)	-1.661 (2.762)	-0.317 (0.396)	-0.058 (0.468)	0.317 (0.334)
Ln (Outlay)	+	1.036*** (0.257)	NA	1.003*** (0.374)	0.972*** (0.195)	0.893*** (0.138)	1.059*** (0.162)	0.926*** (0.072)
R^2		0.3253		0.2986	0.3075	0.2855	0.2932	0.2931
<i>Parente-Santos Silva test for intra-cluster correlation</i>				T=17.81 P > T = 0.000	T=18.65 P > T = 0.000	T=18.60 P > T = 0.000	T=20.12 P > T = 0.000	T=13.78 P > T = 0.000
N		807	807	807	807	807	807	807

Note: The dependent variable is the natural log of underpayment error amounts (adjusted for 2015 constant dollars). Robust standard errors clustered by program appear inside parentheses. Since programs with an absence of underpayment errors account for 38.29% of the sample of data, the 10th quantile was not calculated to avoid redundancy. Due to model convergence problem, 25th quantile model was not obtained. Key covariates of interest appear in shaded regions. UP indicates underpayment error amounts (\$).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.