Escaping Pay-for-Performance*

Jason Chen
Auburn University

Jakub Hajda
HEC Montréal

Joseph Kalmenovitz
University of Rochester

November 11, 2025

How do regulators respond to performance pay? We study this question by exploiting the staggered adoption of performance pay by U.S. financial regulatory agencies between 1981 and 2006. We manually assemble employee-level data on 30,000 financial regulators from public and private payroll records, rulemaking documents, and employee surveys, allowing us to track their careers, incentives, and productivity inside and outside government. Using a stacked difference-in-differences design, we find that performance pay increased voluntary exits from the government by 43-57%. A separate reform affecting top federal executives produced similar effects, reinforcing external validity. Additional analysis highlights two broad channels underlying this response. Performance pay increases effort and thus potential private sector salaries. It simultaneously weakens public sector attachment via greater income volatility, a more competitive work environment, and distrust toward performance evaluations. Consequently, regulators with better outside options and a weaker public sector attachment – leave. Taken together, we highlight a fundamental trade-off in incentive design: performance pay raises productivity but reshapes the composition of the public workforce. We develop and estimate a structural model that quantifies this trade-off and evaluates counterfactual pay policies balancing incentives and retention.

*Kalmenovitz is with the University of Rochester (jkalmeno@simon.rochester.edu, corresponding author); Hajda is with HEC Montréal (jakub.hajda@hec.ca); Chen is with Auburn University (jason.chen@auburn.edu). We received excellent comments from our discussants Harold Cuffe, Alessia Damonte, Ariadna Dumitrescu, Yaniv Grinstein, Tomislav Ladika, Kevin Murphy, Alexander Nitschke, Thomas Ruchti, Tesfaye Sebastiano Salarin, Sophie Shive, John Thanassoulis, and Chen Zhao; Chris Carrigan, David Yermack, and Ján Zábojník; and seminar and conference participants at Auburn University, Hebrew University, University of Pennsylvania, National Taiwan University, National Chengchi University, Queen's University, Finance Organizations & Markets Conference, Drexel Corporate Governance Conference, Wellington Finance Summit, Swiss Society for Financial Market Research, Erasmus Corporate Governance Conference, Financial Management Association, Midwest Finance Association, Northern Finance Association, Eastern Finance Association, Southern Finance Association, Boca Corporate Finance & Governance Conference, the Annual International Corporate Governance Society Conference, and Global Finance Conference for excellent comments.

Introduction

Since the 2024 Presidential elections, federal regulators have come under a harsh spotlight. As a warning shot, in February 2025 the Department of Government Efficiency (DOGE) ordered all employees to report their top five accomplishments. More fundamentally, the new administration seeks to monitor the performance of federal employees and to tighten the link between pay and performance.[1] This would presumably put the public sector on par with the private sector, where employees must often demonstrate satisfactory performance to justify their pay. Although several regulatory agencies have already experimented with performance-based pay, these reforms remain poorly understood by economists. In particular, little is known about their unintended consequences: while performance pay may encourage greater effort, it can also create dissatisfaction or distort incentives in ways that limit overall efficiency gains.

This backdrop motivates our central research question: how do regulators respond to performance pay? We address this question by exploiting the staggered adoption of performance pay by U.S. financial regulatory agencies over several decades. Using newly assembled employee-level data that track careers both within and beyond government, we find that the introduction of performance pay led to a substantial rise in voluntary exits from the public sector.[2] To interpret this response, we further show how performance pay reshapes the landscape via two broad economic channels: it improves the outside pay and weakens the appeal of the public sector. With stronger incentives, regulators exert more effort. This increases potential salary gains in the private sector through accumulation of human capital and better signals to potential employers. At the same time, performance pay weakens the appeal of a public sector job due to greater income risk, competitive work environment, and aversion toward the government's performance evaluation process. The result is selective attrition: regulators with more lucrative private sector opportunities exit, leaving behind regulators

[1] See, for instance, a Presidential Memorandum from January 2025 (90 FR 8481).

[2] We find a comparable response to an independent performance pay reform for federal executives.

with stronger intrinsic attachment to the public sector. To quantify this trade-off, we develop and estimate a structural model linking performance incentives, effort, and turnover. The model allows us to evaluate counterfactual pay policies that balance incentives against retention, providing a quantitative framework for designing optimal public-sector compensation.

Taken together, our findings reveal a central trade-off in the design of incentives within the public sector. Performance pay enhances productivity among regulators, yet simultaneously raises the value of the outside option for those most responsive to incentives. Thus, reforms intended to reward excellence can unintentionally accelerate turnover and alter the composition of the regulatory workforce. By quantifying this trade-off, our analysis highlights how incentive design shapes both the efficiency and the stability of public institutions. More broadly, we contribute to the literatures on the revolving door and performance incentives, by documenting how incentive pay affects effort and turnover in a unique setting where individual performance and regulatory outcomes are directly observable.

In the first part of the paper, we examine the staggered adoption of pay-for-performance (P4P) systems by nine U.S. financial regulatory agencies between 1981 and 2006.[3] The reforms aimed to link individual compensation more closely to performance evaluations, reduce pay compression, and retain high-performing employees. In most cases, this involved replacing nearly automatic step increases with merit-based raises and introducing wider pay bands that allowed managers to reward top performers and differentiate pay more sharply across staff. Although these agencies share broadly similar statutory mandates and labor market environments, they introduced P4P at different points in time, generating a quasi-experimental setting to study how incentive reforms reshape regulator behavior. We source a new dataset with comprehensive federal payroll records, covering over three decades of em-

[3]Those are the Office of the Comptroller of the Currency (OCC, 1981); National Credit Union Administration (NCUA, 1991); Office of Thrift Supervision (OTS, 1991); Office of Federal Housing Enterprise Oversight (OFHEO, 1992); Farm Credit Administration (FCA, 1993); Internal Revenue Service (IRS, 2000); Securities and Exchange Commission (SEC, 2002); Federal Deposit Insurance Corporation (FDIC, 2003); and Commodity Futures Trading Commission (CFTC, 2006).

ployment histories for 30,000 individual financial regulators. We further link these records to two complementary data sources at the individual level: newly constructed measures of rulemaking activity from the Federal Register, which capture each regulator's participation in rulemaking, and career histories from Revelio Labs, which document pre- and post-government employment. Together, these data provide an unusually detailed view of the internal dynamics of major financial regulatory agencies and the broader consequences of incentive reforms in those agencies.

To estimate the causal effect of P4P on turnover, we employ a stacked difference-in-differences design. Our specification compares agencies adopting performance pay (treated) to those that had not yet implemented it (control), within a seven-year event window around each reform. Reassuringly, treated and control agencies are balanced along observable characteristics prior to reform, and in particular show no distinct turnover trajectories before the adoption of performance pay. We further augment the specification with rich fixed effects, especially agency×cohort and occupation×year×cohort, to absorb persistent differences across agencies and time-varying shocks within occupations. Together, these features ensure that the estimated effect is identified from within-cohort variation in exit rates across agencies over time. Across all specifications, we find a substantial and persistent rise in exits following the adoption of performance pay, ranging from 43% to 57% relative to the sample mean. Dynamic estimates show no evidence of pre-trends, and the effect is robust to alternative sample restrictions and additional controls and fixed effects.

In this setting, each reform occurred in a different year and agency. It is therefore unlikely that non-causal factors could systematically generate the repeated pattern of post-reform turnover across cohorts, supporting a causal interpretation. To corroborate these findings and assess the external validity of our framework, we examine a separate reform that introduced performance-based pay for top federal executives. This group, known in government parlance as the Senior Executive Service (SES), comprises 24,000 officials who manage key regulatory and policy functions across the entire government. Using a similar

payroll dataset, we track the executives' careers before and after the shift to performance pay in 2004. In a difference-in-differences design, we find that the reform increased exit rates among treated executives by 39–41%. The magnitude of the response is strikingly similar to that observed among financial regulators, reinforcing the causal interpretation of our results. Because the SES reform applied simultaneously across hundreds of agencies, the rise in turnover cannot be attributed to idiosyncratic events or conditions in any single organization. Instead, it suggests that the response to performance pay generalizes well beyond the financial regulatory domain.

The consistent rise in turnover across both settings, financial agencies and executives, supports a causal link between performance pay and exits. At first blush, it is puzzling: the pay reforms intended to retain talented regulators, not to drive them out. Moreover, pay in the private sector is even more sensitive to performance and poor performance can in fact lead to bigger pay cuts. Thus, government wage contracts with a performance component should resemble a call option, where performance pay increases the volatility and hence the contract value. Nevertheless, our results clearly suggest that performance pay in the public sector raises the value of the outside option even more, resulting in exits.

To understand this dynamic, we outline two broad channels triggered by performance pay reforms: the reforms could increase potential private sector salaries, as well as reducing the intrinsic appeal of government service. Together, these channels can jointly explain our main finding: the aggregate rise in exits. Moreover, they can further explain the heterogeneous effect on different subsamples, and the reforms' impact on other outcomes beyond turnover. Concretely, we present five key results consistent with those two channels.

First, performance pay increases rulemaking activity. Using our new regulator-level measures of rulemaking from the *Federal Register*, we find that treated regulators increase the number of new regulatory documents by 75% and the likelihood of contributing to at least one new regulatory document by 63%, relative to the sample means.[4] By stimulating greater

[4]Interestingly, the total documents produced by treated agencies increases, while the number of coauthors per document remains unchanged.

effort, the reforms achieved one of their intended goals – enhancing regulatory productivity, highlighting the bright side of performance pay. At the same time, the increased effort can improve the potential private sector pay and weaken the preference for a public sector job, mechanisms which we explicitly examine below.

Second, performance pay provided better outside opportunities for regulators. Using Revelio data, we find that treated regulators who exited after the reform earned 53%-69% higher starting salaries than comparable leavers from control agencies. By linking compensation more tightly to individual results, the reform encouraged greater effort (human capital) and made productivity more visible to external employers (signal). As the external reward for regulators increased, the incentive to exit the government increased as well.

Third, performance pay increased income risk inside the government. Pay dispersion across regulators widened by 25.8%. Pay volatility per individual rose as well, with the probability of promotions and pay cuts increasing by more than 50% compared to the pre-reform corresponding probabilities. These shifts indicate that performance pay raised the stakes of annual evaluations, replacing the traditional stability of government pay with greater income risk (Burgess and Ratto, 2003; Friebel, Kosfeld, and Thielmann, 2019). For risk-averse employees, this volatility undermined a key attraction of public employment and contributed to the desire to exit.[5]

Fourth, performance pay is susceptible to distrust among employees, as its effectiveness depends on credible and transparent evaluations. Because performance in the government is difficult to measure and tasks are multi-dimensional, many regulators perceived the new appraisal process as opaque or unfair. Consistent with this possibility, we find that exits were concentrated in agencies where evaluations were perceived as less transparent, among employees with few comparable peers (noisier assessments) , and among those who had underperformed under the prior pay system. Together, these results suggest that performance

[5]Income volatility may also push out more risk-averse regulators. While this may contribute to turnover, the observed rise in private sector pay, and the fact that many leavers likely move to even more risky positions, suggest that risk aversion alone cannot explain the overall pattern.

pay weakened public sector attachment by exposing regulators to evaluation systems they viewed as unreliable, prompting those most averse to such systems to seek employment in environments where benchmarks are perceived as more credible.

Fifth, performance pay triggered selective exits based on ex-ante intrinsic motivation to serve in the public sector. Employees with weaker attachment to public service – those hired during recessions or who gave up little pay to join the government – were significantly more likely to leave after the reform. In contrast, intrinsically motivated regulators – those who entered during strong labor markets or accepted large pay cuts to serve – were more likely to remain. These results indicate that performance pay acted as a sorting mechanism, amplifying turnover among those less mission-driven while retaining those with stronger public sector commitment.

Taken together, the evidence demonstrates that performance pay reshapes the landscape via two broad economic channels: it improves the outside pay and weakens the appeal of the public sector. In particular, by providing stronger incentives, the pay reforms induce more regulatory effort. This increases potential salary gains in the private sector through human capital accumulation and signal precision. At the same time, performance pay weakens the preference for public service for multiple reasons. It increases income risk by introducing pay volatility, creates a more competitive and productivity-driven work environment, and sows distrust by installing performance evaluation systems perceived as opaque and unfair. The result is selective attrition: regulators with more lucrative private sector opportunities are more likely to leave, while those who remain have stronger intrinsic motivation and lower aversion to the government's performance evaluation process.

The patterns we uncover point to a fundamental trade-off: performance pay enhances productivity among regulators, yet simultaneously raises the value of the outside option for the very employees most responsive to incentives. This dual effect underscores the broader implications of incentive reforms in the public sector: policies that succeed in boosting effort may inadvertently erode the incentive to stay, triggering a revolving door effect that shifts

the composition of the regulatory workforce. Effective policy design must account for this trade-off. In an extension of our paper, we develop and estimate a structural model to better quantify this trade-off. For instance, we show that regulators who remain in government after the adoption of performance pay have stronger preferences for public service. We further evaluate counterfactual policies with alternative balances between incentives, effort, and turnover, to inform policy discussions on regulatory pay and performance.

Three final observations are in order. First, although our empirical analysis focuses on financial regulators, the mechanisms we uncover are more general. We provide reduced-form evidence from an independent setting – federal executives – and the channels we highlight are not specific to any agency, occupation, or pay scale. We therefore view our conclusions as relevant to a broad range of public sector employees operating under similar incentive structures. Second, several mechanisms we identify may extend beyond the public sector. The interplay between performance pay, intrinsic motivation, and risk preferences is common to many mission-driven organizations, including nonprofits and regulated industries. Likewise, the role of performance evaluation noise, peer comparisons, and outside option dynamics are central features of incentive design in the private sector. While the specific institutional context differs, the economic forces we document may inform broader debates about the design and unintended consequences of performance-based pay. Third, our analysis stops short of evaluating the welfare consequences of performance pay. We do not observe all government outputs and therefore cannot quantify changes in the quality or quantity of public good provision. Moreover, our model is partial equilibrium, focusing on the optimization problem of a representative regulator rather than system-wide productivity. Assessing these general equilibrium and welfare effects remains an important direction for future research.

Our work relates primarily to the literature on the revolving door, regulatory incentives, and regulatory performance. Concretely, we make three contributions. *First*, studies tend to link performance to the level of pay (Dal Bó, Finan, and Rossi, 2013; Kalmenovitz, 2021) and to organizational features such as field offices (Gopalan, Kalda, and Manela, 2021), su-

pervision hours (Hirtle, Kovner, and Plosser, 2020), and jurisdictional overlap (Kalmenovitz, Lowry, and Volkova, 2025). Our paper highlights the importance of the pay structure in both the public and private sectors, and specifically the role of performance-based pay.

Second, we link the revolving door to the pay structure. Studies typically focus on the prevalence of the revolving door and whether it induces regulatory leniency, without considering how it is affected by the design of regulatory pay (deHaan et al., 2015; Lucca, Seru, and Trebbi, 2014; Shive and Forster, 2017; Cornaggia, Cornaggia, and Xia, 2016; Kempf, 2020; Lambert, 2019; Kalmenovitz, Vij, and Xiao, 2025). Our work shows that pay design directly affects the revolving door, and in particular the incentive to switch sectors in fact increases when regulators receive performance pay. More broadly, the pay structure affects the joint decision on exit and effort: a pay policy which increases effort will raise the value of the outside option, motivating regulators to quit.

Third, we highlight the self selection induced by the pay structure. In related papers, performance pay serves as a screening mechanism for high-ability workers (Lazear and Rosen, 1981; Lazear, 2000; Goldmanis and Ray, 2015). Our paper finds that performance pay can motivate productive regulators to leave, a dynamic that arises given the broader context of public and private sector. A related theoretical literature studies how workers sort into private and public sector jobs based on intrinsic motivation, skills, and performance pay (Macchiavello, 2008; Bond and Glode, 2014; Delfgaauw and Dur, 2010). Our empirical approach combines reduced-form evidence with a structural model, shedding more light on how changes to performance pay affect incentives, effort, and the revolving door between the public and private sectors.

Our work also contributes to the literature on performance pay and compensation design. Most studies of performance pay focus on the private sector, particularly managers (Yermack, 2004; Frydman and Jenter, 2010a; Edmans, Gosling, and Jenter, 2023a; Maksimovic and Yang, 2023), leaving regulatory positions relatively understudied. We extend this literature by examining how performance pay affects both effort and turnover among federal regulators

who perform economically consequential functions. We document how changes to incentive pay can induce significant behavioral responses. While our empirical setting centers on financial regulators and federal executives, the mechanisms we uncover – joint determination of effort and exit under performance pay – apply broadly to regulators operating in similar institutional settings. In doing so, our findings shed light on a large, important, and yet overlooked segment of the workforce, working on complex projects with major economic implications.

1 Background and Data

1.1 Institutional setting

Our empirical setting is the staggered adoption of pay-for-performance (P4P) schemes by nine U.S. financial agencies. The reforms were implemented gradually over several decades, starting with the Office of the Comptroller of the Currency (OCC, 1981) and followed by the National Credit Union Administration (NCUA, 1991); the Office of Thrift Supervision (OTS, 1991); the Office of Federal Housing Enterprise Oversight (OFHEO, 1992); the Farm Credit Administration (FCA, 1993); the Internal Revenue Service (IRS, 2000); the Securities and Exchange Commission (SEC, 2002); the Federal Deposit Insurance Corporation (FDIC, 2003); and the Commodity Futures Trading Commission (CFTC, 2006).

The performance-pay reforms aimed to accomplish three goals. First, to align individual performance more closely with organizational objectives. By embedding those objectives directly into individual performance plans, regulators aimed to create a clearer link between employees' day-to-day responsibilities and the agency's mission, thereby fostering a more motivated and productive workforce (Ginsberg, 2008). Second, to reduce pay compression and financially motivate employees to exert more effort. For example, investigating the SEC's pre-reform pay structure, a commission found that "there are very limited incentives to become a manager. A 'promotion' to supervisory status often means more responsibility

with no increase in pay" (SEC, 2002). Performance-based compensation was intended to widen the spread between high- and average-performing staff, and to ensure that career advancement came with meaningful financial rewards. Finally, the reforms aimed to allow financial agencies to "recruit and retain employees critical to meeting their organizational missions" (GAO, 2007). Unlike most federal agencies, financial regulators compete directly with banks and other employers offering substantially higher pay. By restructuring their compensation systems, agencies were able to reward productive staff with salaries closer to private sector benchmarks.

While the goals and broad features of all pay-for-performance reforms were similar, they differ along several aspects. To conserve space, we defer the detailed description of each reform to Appendix A.1. We exploit some of the underlying heterogeneity in Section 4, as we dig deeper into the economic mechanisms.

1.2 Data

In this section, we describe the primary datasets we have constructed. Our focus is on ten financial agencies: nine agencies adopting independent performance pay systems at different points (listed above in Section 1.1), and the Financial Crimes Enforcement Network, which serves in the control group (see Section 3.1).

1.2.1 Payroll

We source a comprehensive dataset covering all employees who worked at any of the ten financial agencies between 1973 and 2013. The dataset, released by BuzzFeed News following a Freedom of Information Act request, provides information on each employee's agency, occupation, hiring date, location, and compensation. To the best of our knowledge, it captures the universe of financial regulators during this period. Our empirical strategy, detailed in Section 3.1, focuses on regulators employed at financial agencies prior to the introduction of performance pay and follows them over a seven-year event window centered

on the reform year. We begin with 62,559 unique financial regulators who were employed at either treated or control agencies three years before their respective reforms and track their payroll records over the seven-year window, yielding 340,109 regulator×year observations.

To execute our identification strategy, we restrict the sample in several ways. Because our empirical design exploits within-individual variation over time and compares employees across locations and agencies, accurate information on pay, location, education, and rank is essential. We therefore exclude regulator×year observations with missing or inconsistent entries in these fields to ensure clean identification of the treatment effect. For instance, missing location data would compromise the ability to absorb regional labor market shocks. After these filters, the final – and main – sample consists of 31,871 unique financial regulators across ten agencies, totaling 189,032 regulator×year observations.[6]

1.2.2 Rulemaking

Part of the analysis requires a reasonable measure of regulatory activity. To that end, we collect new data on rulemaking from the Federal Register (FR), the official daily publication of the U.S. government. The FR is the primary channel through which federal agencies announce proposed and final rules, making it an indispensable source for tracking regulatory output over time (Chen and Kalmenovitz, 2024; Kalmenovitz, Lowry, and Volkova, 2025). Starting in 2000, the FR provides detailed metadata on each rulemaking document, including the names of the individual regulators involved in drafting the rule. These features allow us to link individual regulators to specific regulatory events, providing a direct measure of regulatory effort.

Since the FR data begin in 2000, we focus on agencies in the 2002, 2003, and 2006 cohorts, for which both pre- and post-reform data are available. We restrict the sample to 2000-2009, which covers the post-reform event window of the last cohort, resulting in 2,980 regulatory

[6]Note that this number is smaller than the counts reported in the regression tables, because the same regulator×year observation may appear in overlapping cohorts.

documents associated with up to 1,143 unique regulators.[7] We match those to identifiable payroll records in the corresponding period, using agency affiliation along with first and last names. For individuals with complete first names, we perform a fuzzy match on the first name combined with an exact match on agency and last name. We then use the ChatGPT API to evaluate each potential match, taking into account that nicknames are sometimes used in the Federal Register (e.g., "Tim" instead of "Timothy"). Each potential match is subsequently reviewed manually to exclude false positives. For individuals with incomplete first names, we proceed more cautiously given the higher risk of matching error. We begin by matching the first name initial in the payroll data to the first letter of the first name in the Federal Register, requiring exact matches on last name and agency. We then manually inspect each candidate match, retaining those where the middle name is an exact match and those with relatively uncommon last names. For example, an individual with the last name "Seifert" would be retained whereas those with "Smith" would not. When multiple matches remain for the same name, we prioritize cases with an exact first-name match. If ambiguity persists, we retain the individual identified as an attorney in the payroll data, since the vast majority of rulemakers are attorneys (Chen, Kalmenovitz, and Sosyura, 2025).

This process yields 724 unique financial regulators from six unique agencies (SEC, FDIC, OTS, FinCEN, CFTC, and FHEO) who are also in our main sample.

1.2.3 Outside option

Parts of the analysis require information on the outside option. To that end, we combine the federal payroll dataset with detailed career histories from Revelio Labs. The Revelio data allow us to examine the outside option facing financial regulators, both at the point of entry to the public sector and at the point of exit. In addition to compensation information,[8]

[7] The true number of unique regulators is somewhat smaller because names are not always recorded consistently across documents. For instance, a regulator may include a middle name in some filings but omit it in others.

[8] Revelio imputes compensation using job characteristics such as title, company, and location (Dorn et al., 2025). While this approach could inevitably include measurement error, we are not aware of any systematic bias correlated with the timing of the performance pay treatment.

Revelio provides data on organizational type, seniority, and industry affiliation, enabling a richer assessment of how pay-for-performance reforms shaped regulators' labor market opportunities and mobility patterns.

We match payroll records to Revelio profiles in the following manner. We begin with 25,135 unique financial regulators with unmasked names from our main sample. Then, from Revelio, we extract individual profiles sharing the same last name, first initial, and employer. For each potential match, we require that the difference in the year of entry into and exit from a financial agency between the two datasets does not exceed two years. In addition, we require the predicted age in Revelio, calculated by assuming college graduation at age 22, to be within four years of the reported age in the payroll data. We employ different procedures depending on whether the payroll record contains a complete first name. For regulators with complete first names, we use the ChatGPT API to evaluate potential matches, accounting for nickname variations. For those with incomplete first names, we retain only Revelio profiles that have no duplicates in the payroll dataset. We then combine the two sets of potential matches. When multiple payroll records remain linked to the same Revelio profile, we resolve duplicates by retaining (i) the match with the smallest difference in year of entry into the agency, (ii) if necessary, the smallest difference in year of exit, and (iii) if ambiguity persists, the smallest difference in predicted age across the two datasets.

Using this procedure, we identify LinkedIn profiles for 4,011 unique financial regulators across nine financial agencies (the IRS masks the names of all its employees). Consistent with typical limitations of online labor market data, coverage is restricted to individuals with publicly available LinkedIn profiles, who are disproportionately younger and more recent employees.

2 Descriptive Statistics

2.1 Career trajectories

Table 1 reports descriptive for our main sample. Compensation variables are adjusted for inflation and expressed in constant 2023 USD. In Panels A and B, we focus on the government portion of their careers based on payroll records. The average regulator is 42 years old and has 11 years of government experience. The unconditional turnover rate among regulators is 8.2% and retirement rate, defined as exiting the government at age 60 or older, is 1%. 84.3% and 32.6% of regulators have college and postgraduate education, respectively. The average regulator earns $123,588 and the unconditional probability of experiencing a large pay cut is 4.4%.[9] On average, 0.6% of employees are promoted to senior rank each year.[10] In the subsample of 724 rulemakers, 18.1% initiate at least one new rulemaking draft – publishing a proposed or final rule in the *Federal Register* – in a given year. Conditional on publishing, the average annual number of new drafts started is 1.9.

In Panel C, we turn to the post-government portion of their careers. We focus on 904 unique financial regulators with available Revelio records who left the government,[11] and describe their first job upon leaving government service. The overwhelming majority of ex-regulators (85%) moved to for-profit firms, while a smaller share joined nonprofit organizations (8%), universities (1%), or state and local government (6%). Roughly half of those entering the private sector joined publicly listed companies. Compensation differences between the public and private sectors are stark: the average government salary prior to exit was about $135,000, compared to an average starting private sector salary of over $205,000. When including bonuses and other forms of remuneration, total private compensation averaged nearly $300,000. On average, performance pay constituted nearly one-quarter of

[9] A large pay cut is defined at a real pay cut of at least 2.95%, which is the 95th percentile of real pay cuts during the pre-period of our main sample.

[10] According to the OPM, senior rank are defined as positions within the EX, ES, EO, EM, CM, OE, SL, SO, SS, VE, WL, WS, XL, or XS pay plans.

[11] This is a subset of the 4,011 regulators with available Revelio records.

this compensation, highlighting how different the structure of pay packages is outside of government.[12]

In Figure 1 we report information on the first post-government job held by revolving door financial regulators. A substantial fraction moved directly into high-level positions, with over 40% starting at the director level or above (including 3.6% at the very top tier such as CFO or CEO). At the same time, almost one-third of exits were into junior or associate roles, indicating that not all regulators were able to command seniority immediately upon transition. This distribution reflects the heterogeneity of the regulators in the sample and the broad range of career paths available after public service; we will exploit this heterogeneity in Section 4.2. In Panel B, we report the industries in which these individuals landed. financial services (35.8%) and legal services (35.1%) dominate, consistent with the professional expertise of regulators and the demand for their skills in compliance-heavy sectors. Consulting (9.2%) also features prominently, reflecting the portability of regulatory knowledge into advisory roles. Other industries – ranging from professional associations and public sector management to technology, energy, and healthcare – are present, but in much smaller shares.

2.2 Pay-for-performance reforms

While the performance-pay reforms were implemented at different points in time, their unified goal was to create meaningful distinctions in pay between higher- and lower-performing employees. We illustrate this in Figure 2, which plots the inflation-adjusted salary distribution of affected agencies three years before (red dotted line) and three years after (blue solid line) the reform. Because the reforms occurred at different times, we use inflation-adjusted pay rather than raw pay. Prior to adoption, salary distributions are relatively compressed; afterward, they widen substantially. The standard deviation of pay rose from $46,785 to

[12]Because Revelio reports total compensation only for the final year of each position, we approximate starting total compensation by scaling the reported final value using the observed growth in base salary between the first and last paychecks.

$58,844, a 25.8% increase. Similarly, the pay gap between the 90th and 10th percentiles expanded from $121,719 to $160,112, a 31.5% increase. Thus, the implementation of performance pay led to a marked widening of pay differentials among financial regulators. This provides an indirect evidence that the new pay-for-performance systems were not merely "cheap talk," but rather had a demonstrable impact on regulator compensation wherever implemented.[13] Note that expectation of a greater dispersion could also generate an income risk, and increase regulators' aversion to the new pay system ex-ante. We return to that in Section 4.3.

3 Impact on Exits

We begin our analysis by examining whether performance pay affected regulators' propensity to leave government service. We describe our empirical strategy, present the main results, and conduct robustness tests. These baseline findings set the stage for the analysis in Section 4, which clarifies the mechanisms through which performance pay influences both effort and attrition.

3.1 Empirical strategy

To examine how the staggered adoption of pay-for-performance (P4P) by financial agencies affected employees' propensity to exit government, we estimate the following stacked difference-in-differences specification:

$$Exit_{c,i,a,t} = \beta \cdot Post_{c,t} \times Treated_{c,a} + \gamma' X_{c,pre,i} + \alpha_{c,o,t} + \alpha_{c,a} + \varepsilon_{c,i,a,t}. \quad (1)$$

Here, $Exit_{c,i,a,t} = 1$ if individual i in agency a and cohort c exits government employment in year $t+1$, and 0 otherwise. $Treated_{c,a} = 1$ if agency a adopted P4P in cohort c, and

[13] Through the lens of our model, wages are deterministic (tenure-based) in the absence of performance pay, and performance pay adds a stochastic component which increases the variance of pay.

$Post_{c,t} = 1$ in the post-adoption period of cohort c.

We include a rich set of fixed effects to ensure credible identification. Specifically, we include agency×cohort fixed effects ($\alpha_{c,a}$) and occupation×year×cohort fixed effects ($\alpha_{c,o,t}$). The former absorb all time-invariant differences across agencies within a given cohort, for example, persistent differences in workplace culture or baseline turnover rates between the SEC and the FDIC that predate the reform. The latter control for time-varying occupational shocks, such as changes in labor market conditions for lawyers, that could otherwise confound exit behavior. This approach mirrors that of the GAO, which compares exit rates and salaries across financial agencies within the same occupation.[14] These fixed effects also subsume the main effects of $Treated_{c,a}$ and $Post_{c,t}$. We further control for pre-reform pay and tenure, using pre-treatment averages to avoid "bad controls." Standard errors are clustered at the agency level. The coefficient of interest, β, captures the differential change in exit rates between treated and control agencies within the same occupation, cohort, and year, providing a plausibly clean estimate of the reform's effect on exits.

We construct the sample using a seven-year event window (±3 years) around each agency's adoption of P4P. The control group consists of agencies that had not yet switched to P4P at the time of treatment (eventual adopters), the Financial Crimes Enforcement Network (never adopter), and the Office of Thrift Supervision (always adopter).[15] Nearly all control agencies adopted P4P eventually but at different times, providing a natural comparison across cohorts. To avoid confounding reforms, we exclude agencies that made substantial changes to their pay structure within five years before the event window.[16]

Finally, Table 2 confirms that treated and control employees are highly comparable in observable characteristics prior to the reforms. Average real pay is $110,090 among con-

[14] See Figure 3 and Table 10 of GAO-07-678.

[15] The Office of Thrift Supervision adopted P4P in 1991, but our data do not separately identify the agency before then, and we therefore refer to it as an "always adopter" (see Appendix A.1).

[16] There are three such episodes in our sample. The FDIC underwent a 1989 reform that aligned its pay with other financial regulators, so we exclude it from the 1991 and 1993 cohorts. Similarly, we exclude the Farm Credit Administration from the 1991 cohort (due to the 1991 VE pay plan reform) and the CFTC from the 2002-2003 cohorts (due to its 2003 switch from GS to CT pay).

trol employees and $127,635 among treated employees, a difference that is not statistically significant. Promotions and large pay cuts are also virtually identical across groups, and managerial representation is 21.3% versus 24.8%, neither of which differs significantly across groups. Exit and retirement rates are nearly identical (8.5% versus 7.9% and 1.1% versus 0.8%, respectively), as are age (41.8 versus 41.0 years) and tenure (11.9 versus 12.5 years). Treated employees have more college education (87.0% versus 80.6%) and less postgraduate education (28.3% versus 33.3%), but the gaps are modest and statistically insignificant.[17] Although balance is not strictly required for identification, the absence of systematic pre-reform differences is important because it rules out the possibility that treated agencies were already on distinct turnover trajectories before the adoption of performance pay. This confirms the conclusion in Appendix A.1, whereby the staggered timing of the reforms was driven by agency-specific statutory authority, rather than by shocks to the demand for regulators specific to treated agencies.

3.2 Results

Table 3 presents the main results. In column 1, we report the baseline specification with agency×cohort and occupation×year×cohort FE. Column 2 adds education×year×cohort fixed effects, allowing comparisons among employees with similar educational backgrounds in the same year.[18] Column 3 introduces manager×year×cohort fixed effects, controlling for time-varying shocks that differentially affect managers and non-managers. Column 4 adds age-bin×year×cohort fixed effects, and column 5 adds city×year×cohort fixed effects to absorb local economic conditions. Across all specifications, the coefficient on *Treated*×*Post* is positive and statistically significant, indicating that treated agencies experienced higher turnover after adopting P4P. The magnitudes are economically meaningful, with exit propensities rising by 43-57% relative to the sample mean. The negative coefficients on pre-reform

[17] We address concerns that pre-reform pay and educational differences explain our findings: we control for pre-reform pay and demonstrate that adding education-by-year fixed effects leaves the results unchanged.

[18] Education is grouped into high school, college, and graduate degree.

pay and tenure indicate that regulators with higher pay and longer tenure are less likely to exit, consistent with more successful and well-established regulators having stronger attachment to government service and higher thresholds for pursuing outside opportunities.

The key identifying assumption is that, absent the reform, exit rates in treated and control groups would have followed parallel trends. While inherently untestable, the staggered adoption of P4P across multiple cohorts makes alternative explanations less plausible, since any confounder would have to account for sharp increases in exits at different agencies and times. To further probe this assumption, we estimate a dynamic version of Equation 1:

$$Exit_{c,i,t+1} = \sum_{k=-3}^{+3} \eta_k \cdot Treated_{c,a} \cdot [\mathbb{1}(Event\,Time_{c,t} = k)] + \alpha_{c,a} + \alpha_{c,o,t} + \epsilon_{c,i,o,a,t}, \quad (2)$$

Here, $\mathbb{1}(EventTime_{c,t} = k)$ indicates that year t in cohort c is k years from P4P adoption. As in our baseline specification, we include agency×cohort fixed effects and occupation×year×cohort fixed effects. Figure 3 plots the estimated coefficients η_k. The figure shows a sharp increase in exits immediately after reform, followed by a peak within two years and a decline toward the end of the event window. The absence of pre-trends and the clear post-reform spike support a causal interpretation.

In Table 4, we assess the robustness of our results. We first separate exits by age, distinguishing departures potentially due to early retirement (age 60+) from those more likely associated with transitions to private sector jobs (age 60 or below). The increase in exits is concentrated among younger regulators, with estimated effect of 0.032, while the estimate for older regulators is small and insignificant. Relative to the baseline estimate in column 1 of Table 3 (0.034), roughly 94% of the increase in exits comes from younger employees, consistent with a private sector pull rather than retirement. Column 3 implements the Bertrand, Duflo, and Mullainathan (2004) correction for serial correlation by collapsing the data into pre- and post-periods. Each individual×cohort contributes at most two observations, with the outcome equal to one if the employee exited in that period. Even in this collapsed speci-

fication, the coefficient on *Treated*×*Post* remains positive and significant, ruling out inflated t-statistics from serial correlation. Finally, we verify that our results are not driven by the inclusion of the Internal Revenue Service (IRS). As discussed in Appendix A.1, the IRS's pay reform was implemented unevenly across occupational groups, as many non-managers did not transition to the new pay structure. We therefore now entirely exclude the IRS cohort (year 2000) and re-estimate our baseline specification. The results, reported in column 4, show that the estimated effect remains virtually unchanged, confirming that our main findings are not driven by partial coverage of the IRS reform.

Taken together, we document a consistent and economically meaningful relationship between pay-for-performance and turnover. Across specifications, the estimated coefficients are stable in magnitude and significance, invariant to the inclusion of increasingly stringent fixed effects. The dynamic event study reveals a sharp increase in exits immediately after reform, with no evidence of pre-trends, reinforcing the interpretation of a causal response. The concentration of effects among younger employees, who are more likely to transition to private sector positions, suggests that the increase reflects behavioral responses to incentives rather than demographic attrition. Thus, the results provide evidence that performance pay in the government motivates regulators to pursue opportunities in the private sector.

3.3 Evidence from executive pay

In this section, we turn to a separate reform affecting a different population: federal executives, known in government parlance as the Senior Executive Service (SES). While our main analysis focuses on financial agencies, this section offers an independent setting in which performance pay was introduced at scale, and shows that the link between incentives and exits holds beyond our core sample. Thus, it provides out-of-sample evidence on the impact of performance pay on turnover. Moreover, this is an extension of economic interest, given the pivotal role these officials play in federal governance.

Created in 1979 under the Civil Service Reform Act,[19] the SES was designed to attract and retain senior managers just below Presidential appointees. These executives link political leadership to the federal workforce: for instance, while the EPA Administrator is a political appointee, more than 200 career executives manage its 15,000 employees. At inception, SES pay was bounded between 120% of GS-15 and 100% of Executive Level IV (EX-IV).[20] Raises were nearly automatic and uniform. By the late 1990s, two problems emerged: the absence of genuine pay-for-performance (most executives were rated "outstanding", Office of Personnel Management, 2001), and pay compression, as the SES ceiling did not keep pace with GS-15 increases. A reform in 2004 addressed both issues. Agencies were required to implement OPM-certified performance appraisal systems, after which their executive pay cap rose substantially to EX-III or EX-II.[21] Automatic raises were replaced with performance-based adjustments, creating scope for meaningful distinctions across executives.

The 2004 executive pay reform provides a sharp shock to executive compensation. Before the reform, SES pay was compressed and performance distinctions were muted. The reform simultaneously raised the pay ceiling and tied raises explicitly to performance evaluations, thereby creating meaningful dispersion in realized compensation. We use this reform to study the impact of performance pay on exits. Based on the same comprehensive payroll records, we now focus on 23,763 unique executives across 397 agencies (156,634 executive×year observations). Executives are on average 51 years old, have 16.5 years of tenure, and earn $212,000, about 14.2% below the cap, with annual growth of 1.3%. Turnover is 10.8%, and education levels are high (96.6% college, 75.4% postgraduate).[22] Similar to the pay reforms among financial agencies, the executive pay reform raised the pay range substantially: the upper bound was 12% above the lower bound before the reform, but 40% afterwards. Vari-

[19] P.L. 95-454, Title IV, amending various sections of the U.S. Code, Title 5, Chapters 31, 33, and 35. This description is based primarily on Government Accountability Office (1980); Congressional Research Service. (2007); Congressional Research Service (2012); and Congressional Research Service (2021).

[20] GS is the General Schedule, the most common pay system in the federal government. EX is the Executive Schedule, reserved for political appointees and organized in reverse order, such that EX-IV is the lowest rank and EX-I is the highest.

[21] §1322 of the Homeland Security Act of 2002 (P.L. 107-296).

[22] See Table A.1.

ation in pay also increased, with the standard deviation of pay more than doubling. These patterns confirm that pay-for-performance was implemented at scale, raising ceilings and generating heterogeneity in realized pay.

Finally, we examine how the pay reform affected executive exits. Our empirical strategy closely follows the framework used for financial agencies (Section 3.1), but with one important distinction: whereas the previous analysis exploits staggered adoption of pay-for-performance across cohorts, here we study a single system-wide policy shock. Accordingly, we estimate the following difference-in-differences specification:

$$Exit_{i,t+1} = \beta \cdot Post_t \cdot Treated_i + X_{i,t}^{'} + \alpha_t + \alpha_a + \varepsilon_{i,t}. \tag{3}$$

Here, $Exit_{i,t+1} = 1$ if executive i leaves the government at time $t+1$, and $Post_t = 1$ for years $t \geq 2004$. We focus on the period 2001-2007, forming a symmetric $\pm$3-year window around the reform. $Treated_i = 1$ for executives who were part of the Senior Executive Service (SES) pay scale prior to the reform.[23] Because the SES pay reform applied only to agencies within the SES system, our control group consists of agencies that have never been part of SES and instead operated under independent pay systems for their managerial staff. We further exclude agencies that were created or dissolved within $\pm$5 years of the reform, since openings and closures can induce abnormal turnover. This restriction yields a balanced sample of 26 agencies (see Appendix A.2). Within those agencies, we limit the control group to employees in managerial roles prior to the SES reform, to ensure comparability with the treated executives.[24] We control for pre-reform pay and tenure (both in logs) and include agency fixed effects (α_a) and occupation$\times$year fixed effects (α_t), which absorb the main effects of $Post$ and $Treated$, respectively. Standard errors are clustered at the agency level. The coefficient of interest, β, captures the differential change in exit rates among treated executives after the reform, relative to a similarly high-ranked control group.

[23] We exclude executives entering after the reform, who have self-selected into pay-for-performance.

[24] Those are defined as supervisory levels 2 through 7, according to the Office of Personnel Management.

Table 5 presents the results. Column 1 reports estimates from the baseline specification, showing a substantial rise in exit rates among treated executives following the reform. Column 2 adds education×year fixed effects, column 3 introduces age-bin×year fixed effects, and column 4 includes city×year fixed effects to account for spatial variation in employment opportunities. Across all specifications, treated executives exhibit a 4.1-4.5 percentage point increase in their probability of leaving the government relative to comparable executives in control agencies. This effect is economically large, corresponding to roughly a 40% increase relative to the sample mean, and remains highly stable across progressively stringent specifications. The magnitude and direction of the estimates are closely aligned with those obtained for financial regulators in Section 3.2, underscoring the consistency of behavioral responses to pay-for-performance reforms across institutional settings.

The dynamic analysis in Figure A.1, similar to the one in Figure 3, further supports a causal interpretation. Exit rates for treated and control groups evolve in parallel prior to the reform, with no evidence of pre-trends, and diverge sharply immediately afterward. The sharp and sustained post-reform increase in exits among treated executives, coupled with the stability of pre-reform trends, reinforces the validity of the parallel-trends assumption and the credibility of the estimated treatment effect.

Taken together, these results indicate that the 2004 executive pay reform substantially increased turnover among senior government officials, mirroring the pattern observed in financial agencies. The consistency of the findings across two distinct institutional contexts strengthens the causal front: the SES reform applied simultaneously across dozens of agencies with diverse missions, making it unlikely that the effect is driven by idiosyncratic factors at any single organization. Moreover, the fact that a different reform produced nearly identical responses suggests that our findings generalize beyond financial regulators. This broader applicability motivates the more generalized framework developed in the next section, which formalizes the mechanisms underlying effort, exit, and selection under performance pay.

4 Channels

In Section 3, we show that the adoption of performance pay led to a sharp increase in exits. This suggests that the relative value of private sector employment rose compared to that of government service. In this section we investigate this shift and highlight two interrelated channels: performance pay improves the potential salary gains in the private sector while weakening the appeal of government service. Together, these channels jointly explain the aggregate rise in exits documented earlier, as well as the heterogeneous effect on different subsamples and the impact on other outcomes which we study below.

4.1 Effort

A natural first step is to ask whether the reforms changed regulators' effort within the government. The performance pay systems were designed to link compensation more directly to measurable outcomes. If they were effective, we should see an increase in observable productivity among employees who remained in government service. To test this, we exploit the new regulator–level dataset on rulemaking activity, which records both the number of regulatory documents to which each individual contributed and an indicator for whether the regulator participated in at least one new rulemaking draft in a given year. These data allow us to quantify the behavioral response of regulators to the adoption of performance pay in real time. It fits within the broader literature on incentives and multitasking, which predicts that agents reallocate effort toward observable and contractible tasks once rewards become performance-contingent (Holmstrom and Milgrom, 1991; Prendergast, 1999).

Table 6 reports the results. Columns 1–3 use the number of regulatory documents ($\#RegDoc$) as the dependent variable, while Columns 4–6 use an indicator for working on at least one rule ($\mathbb{1}(RegDoc)$). We use a Poisson model for the former and OLS for the latter, because the number of rules is a highly skewed count variable left-censored at zero. Across specifications, we find that treated regulators became significantly more productive

after the adoption of P4P. The estimated coefficients on *Treated*×*Post* range from 0.56 to 0.59 in the Poisson regressions and from 0.15 to 0.16 in the OLS specifications, implying that treated regulators increase their rulemaking activity relative to control regulators, following the performance pay reforms. The magnitude of the effect is economically significant. Specifically, the propensity to produce a new regulatory document rises by 62-68% relative to the sample mean, consistent with a substantial increase in effort intensity among employees who remained in government service following the implementation of performance pay reforms.

Importantly, the inclusion of rich fixed effects rules out several alternative explanations. The Cohort×Occupation×Year and Cohort×Education×Year fixed effects ensure that productivity gains are not driven by changes in job composition or workforce quality within cohorts. The Cohort×Manager×Year and Cohort×City×Year fixed effects account for differences in managerial practices or local economic conditions that might otherwise correlate with both pay reform and productivity. The treatment effect remains positive and significant, reinforcing the interpretation that the reform causally raised individual effort.

In sum, the evidence points to a clear increase in output following the introduction of performance pay. Regulators exposed to the reforms produced substantially more rules, consistent with stronger incentives to deliver observable results.[25] These productivity gains suggest that the reforms achieved one of their intended objectives: stimulating greater effort among those who stayed. At the same time, the increase in effort provides an important backdrop for understanding why exits rose. Higher effort can erode the intrinsic appeal of public service, increase income risk, and raise the potential salary in the private sector – factors that together contribute to greater turnover. We will examine those ripple effects in subsequent sections.

[25]In untabulated results, we find that P4P adoption increases the total regulatory documents produced by the agency, but did not have a meaningful impact on the number of "coauthors" per document.

4.2 Private sector pay

Building on our prior findings, it is possible that performance pay improves potential private sector earnings for regulators. As they exert more effort, they accumulate transferable human capital. For example, attorneys who draft more rules acquire specialized legal skills that are highly valued by regulated firms. Moreover, by placing greater weight on individual performance, the reform makes employees' productivity more visible and improves the signal, enabling outside employers to better distinguish between high- and low-performing regulators. For instance, when promotions and pay adjustments become more tightly linked to measured performance, they serve as credible indicators of ability for potential private-sector recruiters. Either way, regulators exposed to performance pay should enjoy stronger post-government career opportunities, reflecting both higher human capital accumulation and clearer signals to potential employers.

Note that this channel is motivated by two ideas from the revolving door literature (Kalmenovitz, Vij, and Xiao (2025)): *schooling*, whereby regulators acquire skills that enhance their private sector value, and *regulatory capture*, whereby regulators take actions that may benefit prospective employers. Both ideas explain why performance pay may strengthen the outside option, and distinguishing between the two is beyond the scope of our paper. This channel is also related to Holmström (1999), who shows that workers exert effort not only to raise current output but also to influence how the labor market learns about their ability over time. In his framework, effort increases the precision of the market's signal about ability, thereby improving future wage prospects. Similarly, by making individual performance more observable, performance pay reforms amplify these career-based incentives and, as a result, strengthen regulators' outside options.

To examine how performance pay shapes private sector opportunities, we use Revelio Labs data. The sample includes individual regulators who left government service and entered a private sector position. We focus on the compensation of ex-regulators in their first post-government job: the starting base salary; the total starting salary, which includes base

pay plus any performance-related bonuses; and performance pay, defined as the fraction of total pay aside from base pay. All outcomes are in logs. Similar to Equation 1, the main independent variable is the interaction of treatment status with the post-reform period (*Treated* × *Post*). In this test, all columns include cohort-by-agency and cohort-by-exit-year fixed effects. This specification absorbs persistent differences across agencies and exit cohorts, while focusing on differential changes following the reform.

The results in Table 7 indicate a sizable increase in outside pay for treated regulators after the reform. Relative to a control group of leavers, treated regulators earned 53% ($e^{0.43} - 1$) higher starting salaries (column 1) and 69% ($e^{0.525} - 1$) higher total compensation (column 2). Even when isolating performance-related pay, they earned about 35% ($e^{0.304} - 1$) more (column 3), though this estimate is less precise. These findings suggest that stronger incentives inside government not only raised effort but also enhanced the market value of regulators' skills, consistent with Bond and Glode (2014), who show that higher-ability workers self-select into jobs with greater performance pay. By saturating the model with cohort-by-agency and cohort-by-exit-year fixed effects, we rule out alternative explanations such as agency-specific pay premia, exit-year shocks (e.g., macroeconomic conditions), or stable differences in the types of employees leaving particular agencies. Identification therefore comes from within-agency, within-cohort comparisons of treated and control employees before and after the reform.

Taken together, the evidence indicates that performance pay substantially strengthened the outside option for treated regulators. Former regulators secured more generous compensation packages in the private sector, consistent with the idea that the reform increased their human capital and improved the signal to potential employers. These effects help explain the rise in exits observed in Section 3. By raising effort, performance pay created more profitable private sector job opportunities, resulting in selective departures from public service.

4.3 Income risk

Having established that effort increased, we next examine how the reform affected the distribution of pay inside the government. As shown in Figure 2, the variation in salaries *across* regulators rose sharply after the reform, consistent with a stronger link between compensation and performance. We now investigate whether the reform increased pay volatility *within* regulator over time. This pattern has two important implications. First, as mentioned in Section 2.2, greater pay dispersion provides indirect evidence of a stronger pay-for-performance mechanism. Second, the expected volatility in earnings may weaken the public sector preference ex-ante, and thus contribute to rising exit rates.

Empirically, we compute the annual pay adjustments for each regulator, and estimate a version of Equation 1 with a set of pay adjustment variables.[26] Table 8 reports the results. Columns 1–3 use log pay as the dependent variable. In Columns 4–6, the dependent variable is an indicator equal to one if an employee is promoted to senior rank. On average, employees promoted to senior rank experience a 10.4% increase in real salary. Conversely, Columns 7–9 use an indicator for large real pay cuts, defined as a decline in inflation-adjusted pay of at least -2.95% (the 95th percentile of the pre-reform distribution). These outcomes capture extreme realizations of performance-based pay.

The results are consistent with stronger performance incentives. The probability of receiving a promotion rises by 0.8-0.9 percentage points, and the likelihood of experiencing a large pay cut increases by 2.1-3.8 percentage points. Those estimates are more than half of the pre-reform corresponding probabilities. At the same time, we find no significant change in mean pay. Combined, this confirms that the reform primarily affected the dispersion rather than the level of compensation. In economic terms, the introduction of performance pay heightened the stakes of annual evaluations: high-performing regulators were more likely to be rewarded, while those with weaker evaluations faced tangible financial penalties.

In sum, we document greater pay volatility following the performance pay reform, in-

[26] Since here pay is an outcome, we do not control for pre-reform pay.

creasing both the upside and downside risk of pay. For some regulators, this heightened volatility undermines the stability traditionally valued in government employment. Viewed in conjunction with our previous findings, this set of results sharpens the fundamental trade-off triggered by performance pay: it can boost effort among stayers, while encouraging exits among those unwilling to bear higher income risk, as it weakens their public sector preference.[27]

4.4 Aversion to government evaluation

In the public sector, performance is difficult to measure (Khan et al., 2022) and tasks are often multi-dimensional and non-verifiable (Holmstrom and Milgrom, 1991). Because of that, performance pay can backfire. If regulators distrust the new evaluation process and perceive it as opaque or poorly aligned with effort, their preference for the public sector is weakened. As they feel disadvantaged under the new system, they pursue for private sector opportunities, where evaluation mechanisms may be often viewed as more transparent and merit-based.

Empirically, we develop three measures capturing distinct sources of evaluation aversion. For each measure, we divide the sample into high- and low-aversion groups and re-estimate our baseline specification (Equation 1) separately for each group. The results, reported in Table 9, show that the increase in exits is concentrated among employees more averse to the government's performance evaluation process. Our first measure is based on the Federal Employee Viewpoint Survey (FEVS), a large-scale study of federal employees. We focus on Question 32:[28]

> *"In my most recent performance appraisal, I understood what I had to do to be rated at different performance levels (for example, Fully Successful, Outstanding)."*

[27] Income volatility can also push out more risk-averse regulators. While this mechanism may contribute to turnover, the observed rise in private-sector wages and the fact that many leavers likely move to higher-risk private positions suggest that risk aversion alone cannot explain the overall pattern.

[28] Survey results are available online (source).

This item captures whether employees perceive evaluation standards as clear. Optional responses were Strongly Disagree; Disagree; Agree; Neither Agree nor Disagree; Agree; and Strongly Agree. Because it was first included in 2008 (after the last cohort adopted P4P), we use that year to build a cross-sectional measure of post-reforms transparency. For each agency, we calculate the share of employees who disagreed or strongly disagreed with the statement. Greater share indicates that employees view the appraisal criteria as uncertain. We classify agencies with the highest (lowest) disagreement rates as opaque (transparent) appraisal systems and re-estimate Equation 1 separately for each group.[29] Columns 1–2 report the results. In opaque agencies, the treatment effect is positive and significant at the 10 percent level, while in transparent agencies it is negative and indistinguishable from zero. The difference between the two estimates is significant at the 5 percent level. These findings indicate that exits were concentrated in agencies where employees are dissatisfied with the performance evaluation process. Note that this is an ex-post measure based on the perceptions of those who remained, which likely attenuates the effect if the most dissatisfied employees already have exited.

Our second test, motivated by the literature on relative performance evaluation (e.g., Jayaraman et al., 2021), is based on the number of comparable peers. Intuitively, when many peers perform similar tasks, the supervisor can better filter out idiosyncratic shocks and conduct more precise evaluations (Hölmstrom, 1979; Holmstrom, 1982). Thus, evaluations are expected to be noisier when an employee has few peers and more precise when many comparable colleagues exist. We define peers as employees working in the same office and occupation one year before reform, for example, attorneys in the SEC's Philadelphia office. We split the sample at the cohort-agency median number of peers and re-estimate our baseline specification separately for each group. As shown in columns 3-4 of Table 9, the effect of performance pay on exits is substantially larger among employees with fewer peers (0.041) than among those with many peers (0.026), with the difference statistically significant at

[29]The FDIC did not participate in the survey, eliminating 69,780 observations from the primary sample.

the 1% level. These results suggest that noisy performance assessments amplify the effects of incentive pay on exits.

Our final test is based on employee-specific ability to navigate the government's compensation system. Employees who earned higher pay relative to observables *before* the reform know how to navigate the system effectively. They expect to master the new evaluation system and thus have weaker incentive to leave. In contrast, employees who fared poorly before the reform would view the new evaluation system as continuation of their disadvantage, providing another incentive to exit. To measure these differences, we use pre-reform data to estimate the following model in the spirit of Abowd, Kramarz, and Margolis (1999):[30]

$$\log(Pay)_{c,i,t} = \log(Tenure)_{c,i,t} + \alpha_{c,i} + \alpha_{c,o,t} + \alpha_{c,e,t} + \epsilon_{c,i,o,t} \tag{4}$$

Here, $\log(Pay)_{c,i,t}$ and $\log(Tenure)_{c,i,t}$ denote the log of pay and tenure for employee i in cohort c and year t. We include three sets of fixed effects: employee ($\alpha_{c,i}$), office×year ($\alpha_{c,o,t}$), and education×year ($\alpha_{c,e,t}$).[31] We estimate Equation 4 over the decade preceding each reform,[32] and interpret the estimated employee fixed effects as proxies for abnormal pay: persistent deviations from predicted salary after accounting for office-wide pay trends, occupation, educational background, and tenure. We divide the sample into low- and high-abnormal pay groups based on the yearly median and re-estimate Equation 1 for each group. As shown in columns 5-6 of Table 9, the post-reform increase in exits is concentrated among employees with low abnormal salaries, whose estimated coefficient (0.039) is more than twice that of high-abnormal-salary employees (0.015), with the difference statistically significant. It is consistent with the idea that the evaluation process was less attractive for regulators who had struggled to secure rewards under the previous system.

Taken together, the analyses suggest that aversion toward the government's evaluation

[30] This approach is analogous to estimating judicial leniency using judge fixed effects (e.g., Bernstein, Colonnelli, and Iverson, 2019) or CEO styles using CEO fixed effects (e.g., Bertrand and Schoar, 2003).

[31] An office is defined as a unique agency×city combination, such as the Philadelphia office of the Securities and Exchange Commission.

[32] Regression results are reported in Table A.2.

process amplified exits. Regulators were most likely to leave when performance appraisals within the government were perceived as opaque, when a small number of peers made evaluations noisier, or when they had persistently underperformed relative to peers under the prior system. These patterns point to a common mechanism: distrust toward the government's evaluation system weakened the intrinsic attachment to public service by lowering the perceived fairness and credibility of performance assessments. As a result, regulators sought environments where performance metrics were perceived as clearer and more credible. This dynamic highlights how performance pay, by exposing regulators to ambiguous evaluations, eroded the intrinsic preference for public service and thus lowered the barrier to exit.

4.5 Weak ex-ante attachment

In this section, we examine the relation between performance pay and public sector preference from a different perspective. Our earlier findings suggest that performance pay weakened the intrinsic motivation to serve in the public sector through multiple channels. Building on that, we hypothesize that employees with weak motivation to serve in the public sector ex-ante – *prior* to the reform – will respond more strongly to performance pay. By contrast, employees with strong ex-ante attachment to public service are less likely to exit, even if performance pay heightens their disutility.

To test this mechanism empirically, we develop two observable proxies for intrinsic motivation that reflect differences in employees' attachment to public service. For each measure, we classify employees into two groups, those with relatively *high* and *low* intrinsic motivation, and re-estimate our baseline difference-in-differences specification (Equation 1) separately for each group. By comparing the estimated coefficients across subsamples, we test whether the effect on exit is concentrated among employees with weaker intrinsic attachment to government service. The results are reported in Table 10.

Our first measure classifies employees based on whether they entered during an economic downturn (recession). The rationale is that recession-hires are more likely to have joined the

government for job security in the face of weak private sector prospects, rather than out of a strong belief in the organization's mission. We estimate our baseline model separately for employees who joined during recession and non-recession years. The point estimate is positive and statistically significant for recession hires (column 1), but statistically insignificant for non-recession hires (column 2), and the difference between the coefficients is statistically significant at the 10% level. These results indicate that the increase in exit rates is driven by employees who entered during recessions, consistent with the notion that such employees have weaker preferences for government work and are therefore more likely to exit when performance pay heightens disutility. By contrast, those who entered government service during strong labor markets likely have a stronger intrinsic preference for public sector work, which dampens the exit response to dissatisfaction from performance pay.

As an alternative proxy for intrinsic motivation, we use the salary forgone upon entering government service. This measure captures the extent to which individuals sacrifice private sector pay to work in regulation. We compute the forgone salary as the difference between the employee's most recent pre-entry compensation and their initial salary at a financial agency. Because this measure requires prior private-sector experience, the sample is limited to employees with available pre-government pay data. We then split the sample at the median into groups with small and high forgone salaries and re-estimate Equation 1 separately for each group. Columns (3)-(4) report the results. The increase in exits following the reform is concentrated among employees with small forgone salaries, who gave up little to join the public sector, consistent with weaker intrinsic motivation for government work. The *Treated*×*Post* coefficient is positive and significant in column (3) but negative (and imprecisely estimated) in column (4); the difference between the two is statistically significant at the 10 percent level.

In sum, we find that employees with weaker intrinsic attachment to public service were more likely to exit under P4P. Regulators hired during recessions and those who sacrificed little salary to join the government displayed stronger exit responses, consistent with lower

mission orientation. By contrast, employees with stronger intrinsic motivation were less likely to leave. Thus, performance pay serves as a sorting mechanism, motivating employees with weaker public sector attachment to exit and those with stronger attachment to stay.

4.6 Joint interpretation

To summarize, performance pay motivated regulators to exert more effort (Section 4.1). Regulators who chose to exit received higher salaries in the private sector (Section 4.2). Inside the government, income risk has increased (Section 4.3). In the cross-section, regulators more averse to the government's evaluation system (Section 4.4) and with weaker ex-ante motivation for public service (Section 4.5) were particularly eager to exit. Combined, these findings highlight the broader implications of introducing performance pay for regulators. Performance pay strengthens incentives and thus raises effort among regulators. This increases potential salary gains in the private sector through human capital accumulation and signal precision. At the same time, performance pay weakens the preference for public service through multiple channels. It increases income risk by introducing pay volatility, creates a more competitive and productivity-driven work environment, and sows distrust by installing performance evaluation systems perceived as opaque and unfair. The result is selective attrition, as the reforms stimulate effort while simultaneously increasing turnover and altering the workforce composition. In particular, regulators with more lucrative private sector opportunities are more likely to leave, while those who remain have stronger intrinsic motivation and lower aversion to the government's performance evaluation process.

5 Extension: Model

In this section, we introduce a structural model of how performance pay reshapes incentives through several interrelated channels. The model is closely connected to the empirical results and provides a unifying framework to interpret them. Moreover, it quantifies unobserved

parameters such as public sector preferences, and it allows us to study counterfactual pay policies with different impact on effort and turnover. For clarity and brevity, we provide only an intuitive overview of the model here, and we discuss the formal setup and estimation procedure in Appendix A.3.

The model builds on three institutional facts about compensation in the public and private sectors. First, government pay features both a floor and a ceiling: pay does not decline after poor performance, but it cannot exceed a system-wide cap even after exceptional performance. Second, private sector pay is more sensitive to performance, with both downside risk and uncapped upside potential. Third, regulators place positive value on working in the public sector beyond wages, reflecting mission, stability, and culture (Preston, 1989; Besley and Ghatak, 2005; Macchiavello, 2008; Pitts, Marvel, and Fernandez, 2011; Jaimovich and Rud, 2014).

Performance pay reshapes incentives along several margins, consistent with the empirical findings in Section 3 and Section 4. We illustrate the dynamic in Figure 4. By linking compensation more closely to performance, the reforms directly raise the return to effort. Regulators respond by working harder and becoming more productive. Higher observed productivity improves private sector job opportunities, consistent with greater human capital or clearer signal to potential employers. At the same time, performance pay could reduce the appeal of public sector job through several channels. For instance, a performance-driven culture can erode the "quiet life" aspect of public employment (Dixit, 2002), and the greater pay volatility can undermine a key attraction of government work for risk-averse regulators (Burgess and Ratto, 2003; Friebel, Kosfeld, and Thielmann, 2019). The combined effect is that regulators' preference for government work weakens just as their outside opportunities become more rewarding, raising the value of exit. The result is selective attrition, as the reforms stimulate effort while simultaneously increasing turnover and altering the workforce composition. In particular, regulators with more lucrative private sector opportunities are more likely to leave, while those who remain have stronger intrinsic motivation and lower

aversion to the government's performance evaluation process.

Finally, we extend the analysis by estimating key parameters and examining counterfactual policies. We recover the performance sensitivity in the government and the implied public sector preference, and use the estimated parameters to evaluate alternative policies. The counterfactual exercises reinforce the central insight of the model, namely, that incentive design in the public sector jointly determines effort and self-selection. They further demonstrate how even modest adjustments to pay formulas can produce large shifts in both effort and turnover decisions.

6 Conclusions

This paper examines how performance pay affects incentives and retention in the public sector, using the staggered adoption of pay-for-performance systems across U.S. financial regulatory agencies. We find that performance pay increases effort and productivity among regulators but also raises exits, particularly among those most responsive to incentives. We highlight two primary channels underlying this dynamic: introducing stronger performance incentives improves potential pay in the private sector while weakening the preference for government service. At the same time, it increases regulatory effort. The resulting trade-off underscores the challenge of incentive design in the public sector: mechanisms that successfully stimulate productivity may inadvertently encourage resignations and change the composition of the regulatory workforce.

Our evidence also reveals meaningful heterogeneity in responses. Exits are concentrated among regulators who are more averse to government performance evaluation and those with weaker ex-ante attachment to public service. These patterns highlight that employees differ not only in their productivity but also in their tolerance for evaluation and their intrinsic motivation to serve. Performance pay therefore reshapes both behavior and self-selection, as regulators cross the boundaries between the private and public sectors.

Several avenues remain open for future work. First, our analysis focuses on individual behavior and partial equilibrium effects; assessing the general equilibrium and welfare implications of performance pay remains an important next step. Second, richer data on evaluation processes and peer networks could clarify how bureaucratic assessment systems shape morale and effort allocation. Finally, extending this framework to other segments of the public workforce – such as educators, law enforcement, or healthcare administrators – would help determine the scope of the trade-off between incentives and retention.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67:251–333.

Andrews, I., M. Gentzkow, and J. M. Shapiro. 2017. Measuring the Sensitivity of Parameter Estimates to Estimation Moments. *Quarterly Journal of Economics* 132:1553–92.

Bazdresch, S., R. J. Kahn, and T. M. Whited. 2017. Estimating and Testing Dynamic Corporate Finance Models. *Review of Financial Studies* 31:322–61.

Bernstein, S., E. Colonnelli, and B. Iverson. 2019. Asset allocation in bankruptcy. *Journal of Finance* 74:5–53.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119:249–75.

Bertrand, M., and A. Schoar. 2003. Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics* 118:1169–208.

Besley, T., and M. Ghatak. 2005. Competition and incentives with motivated agents. *American Economic Review* 95:616–36.

Bond, P., and V. Glode. 2014. The labor market for bankers and regulators. *Review of Financial Studies* 27:2539–79.

Briggs, J., D. Cesarini, E. Lindqvist, and R. Östling. 2021. Windfall gains and stock market participation. *Journal of Financial Economics* 139:57–83.

Burgess, S., and M. Ratto. 2003. The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy* 19:285–300.

Chen, J., J. Kalmenovitz, and D. Sosyura. 2025. The deep state and the liberal drift of financial regulators. *Working paper* .

Chen, P.-L., and J. Kalmenovitz. 2024. Regulatory similarity. *Journal of Law and Economics* 67:691–730.

Congressional Research Service. 2007. Senior Executive Service (SES) Pay for Performance System. *Available online* .

Congressional Research Service. 2012. The senior executive service: Background and options for reform. *Available online* .

———. 2021. The senior executive service: An overview. *Available online* .

Cornaggia, J., K. J. Cornaggia, and H. Xia. 2016. Revolving doors on wall street. *Journal of Financial Economics* 120:400–19.

Dal Bó, E., F. Finan, and M. A. Rossi. 2013. Strengthening state capabilities: The role of financial incentives in the call to public service. *Quarterly Journal of Economics* 128:1169–218.

deHaan, E., S. Kedia, K. Koh, and S. Rajgopal. 2015. The revolving door and the SEC's enforcement outcomes: Initial evidence from civil litigation. *Journal of Accounting and Economics* 60:65–96.

Delfgaauw, J., and R. Dur. 2010. Managerial talent, motivation, and self-selection into public management. *Journal of Public Economics* 94:654–60.

Dixit, A. 2002. Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources* 696–727.

Dorn, D., F. Schoner, M. Seebacher, L. Simon, and L. Woessmann. 2025. Multidimensional skills on linkedin profiles: Measuring human capital and the gender skill gap. *Working paper* .

Edmans, A., T. Gosling, and D. Jenter. 2023a. CEO compensation: Evidence from the field. *Journal of Financial Economics* 150:103718–.

———. 2023b. CEO compensation: Evidence from the field. *Journal of Financial Economics* 150:103718–.

Friebel, G., M. Kosfeld, and G. Thielmann. 2019. Trust the police? Self-selection of motivated agents into the German police force. *American Economic Journal: Microeconomics* 11:59–78.

Frydman, C., and D. Jenter. 2010a. CEO compensation. *Annual Review of Financial Economics* 2:75–102.

———. 2010b. CEO Compensation. *Annual Review of Financial Economics* 2:75–102. doi:10.1146/annurev-financial-120209-133958.

Ginsberg, W. R. 2008. Pay-for-performance: Linking employee pay to performance appraisal. Congressional Research Service, the Library of Congress.

Goldmanis, M., and K. Ray. 2015. Sorting effects of performance pay. *Management Science* 61:335–53.

Gopalan, Y., A. Kalda, and A. Manela. 2021. Hub-and-spoke regulation and bank leverage. *Review of Finance* 25:1499–545.

Government Accountability Office. 1980. Civil service reform: Where it stands today. *Available online* .

Han, C., and P. Phillips. 2010. GMM estimation for dynamic panels with fixed effects and strong instruments at unity. *Econometric Theory* 26:119–51.

Hennessy, C. A., and T. M. Whited. 2007. How Costly Is External Financing? Evidence from a Structural Estimation. *Journal of Finance* 62:1705–45.

Hirshleifer, D., and Y. Suh. 1992. Risk, managerial effort, and project choice. *Journal of Financial Intermediation* 2:308–45.

Hirtle, B., A. Kovner, and M. Plosser. 2020. The impact of supervision on bank performance. *Journal of Finance* 75:2765–808.

Hölmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 74–91.

Holmstrom, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 324–40.

Holmström, B. 1999. Managerial incentive problems: A dynamic perspective. *The Review*

of Economic studies 66:169–82.

Holmstrom, B., and P. Milgrom. 1991. Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization* 7:24–52.

Jaimovich, E., and J. P. Rud. 2014. Excessive public employment and rent-seeking traps. *Journal of Development Economics* 106:144–55.

Jayaraman, S., T. Milbourn, F. Peters, and H. Seo. 2021. Product market peers and relative performance evaluation. *The Accounting Review* 96:341–66.

Kalmenovitz, J. 2021. Incentivizing financial regulators. *Review of Financial Studies* 34:4745–84.

Kalmenovitz, J., M. Lowry, and E. Volkova. 2025. Regulatory fragmentation. *Journal of Finance* 80:1081–126.

Kalmenovitz, J., S. Vij, and K. Xiao. 2025. Closing the revolving door. *Journal of Finance* .

Kempf, E. 2020. The job rating game: Revolving doors and analyst incentives. *Journal of Financial Economics* 135:41–67.

Khan, A., G. Xu, R. Burgess, and T. Besley. 2022. Bureaucracy and development. *Annual Review of Economics* 14:397–424.

Lambert, T. 2019. Lobbying on regulatory enforcement actions: Evidence from US commercial and savings banks. *Management Science* 65:2545–72.

Lazear, E. P. 2000. Performance pay and productivity. *American Economic Review* 90:1346–61.

Lazear, E. P., and S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89:841–64.

Lucca, D., A. Seru, and F. Trebbi. 2014. The revolving door and worker flows in banking regulation. *Journal of Monetary Economics* 65:17–32.

Macchiavello, R. 2008. Public sector motivation and development failures. *Journal of Development Economics* 86:201–13.

Maksimovic, V., and L. Yang. 2023. What do unions do? incentives and investments. *Working paper* .

Michaelides, A., and S. Ng. 2000. Estimating the rational expectations model of speculative storage: A monte carlo comparison of three simulation estimators. *Journal of Econometrics* 96:231–66.

Murphy, K. J. 1985. Corporate performance and managerial remuneration: An empirical analysis. *Journal of Accounting and Economics* 7:11–42.

Newey, W. K., and D. McFadden. 1994. Chapter 36: Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, vol. 4 of *Handbook of Econometrics*, 2111–245. Elsevier.

Office of Personnel Management. 2001. Senior executive excellence and accountability. *Available online* .

———. 2019. Retirement Statistics & Trend Analysis, Fiscal Year 2015 to Fiscal Year 2019. *Available online* .

Page, T. B. 2018. CEO attributes, compensation, and firm value: Evidence from a structural estimation. *Journal of Financial Economics* 128:378–401.

Pitts, D., J. Marvel, and S. Fernandez. 2011. So hard to say goodbye? Turnover intention

among US federal employees. *Public Administration Review* 71:751–60.
Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37:7–63.
Preston, A. E. 1989. The nonprofit worker in a for-profit world. *Journal of Labor Economics* 7:438–63.
Shive, S. A., and M. M. Forster. 2017. The revolving door for financial regulators. *Review of Finance* 21:1445–84.
Yermack, D. 2004. Remuneration, retention, and reputation incentives for outside directors. *Journal of Finance* 59:2281–308.

Figure 1: **The Outside Option of Financial Regulators**

The figure summarizes the first private sector positions of 904 former financial regulators, based on manually linked information derived from payroll records and Revelio Labs career data. Panel A shows the distribution of seniority levels in these initial private sector roles: over 40% transition directly into director-level or higher positions, while nearly one-third begin in junior or associate roles. Panel B reports the industries that absorb former regulators. Financial services (35.8%) and legal services (35.1%) dominate, followed by consulting (9.2%). See Section 2.1.

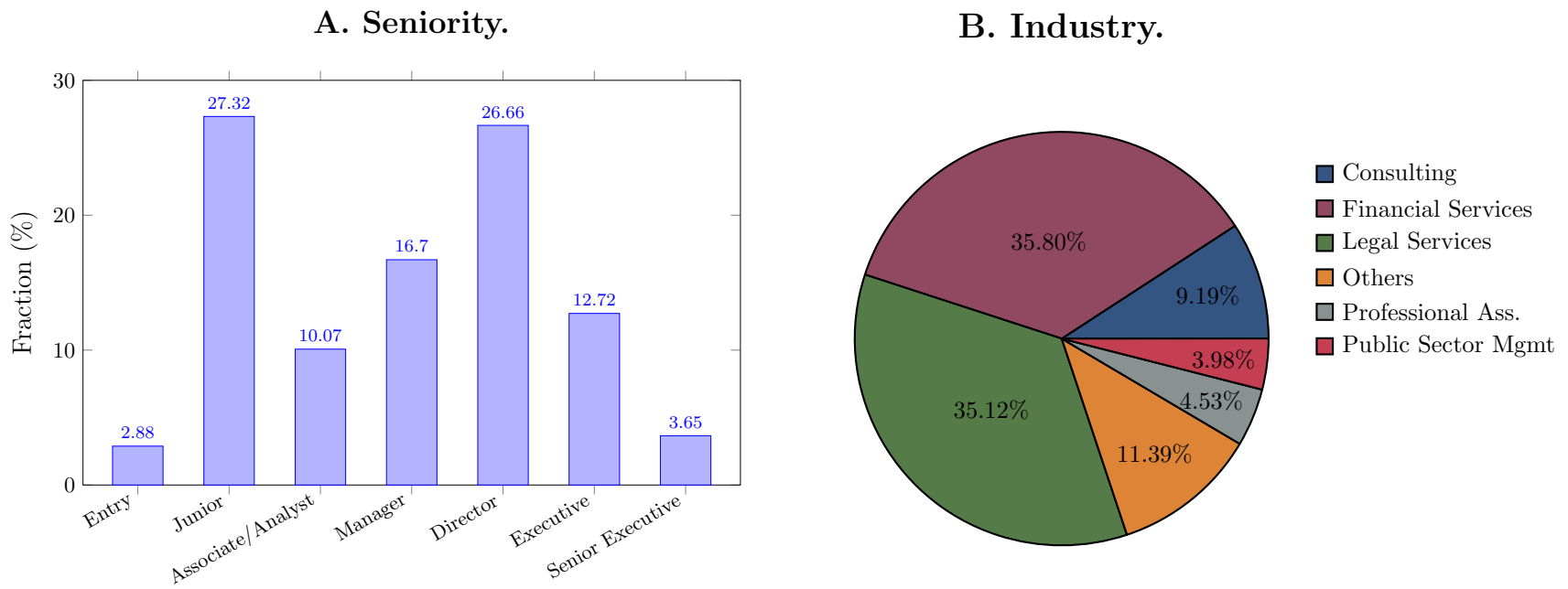

Figure 2: **Impact of Pay-for-Performance on Salary Dispersion**

The figure plots the inflation-adjusted salary distribution of financial regulatory agencies, three years before (red dotted line) and three years after (blue solid line) the adoption of pay-for-performance reforms. The data are drawn from federal payroll records covering all employees in affected agencies, adjusted to constant 2023 dollars. Prior to reform, salary distributions were relatively compressed; afterward, they widened substantially. The standard deviation of pay rose by 26%, and the gap between the 90th and 10th percentiles expanded by more than 30%. The changes illustrate the shift toward merit-based pay differentiation. See Section 2.2.

Figure 3: **Dynamic Response of Financial Regulators to Performance Pay**

The figure reports the coefficients (η_k) from the dynamic stacked difference-in-differences specification, $Exit_{c,i,t+1} = \sum_{k=-3}^{+3} \eta_k \cdot Treated_{c,a} \times \mathbb{1}(EventTime_{c,t} = k) + \alpha_{c,a} + \alpha_{c,o,t} + \varepsilon_{c,i,a,t}$, which traces exit rates around the adoption of pay-for-performance reforms by U.S. financial regulatory agencies. Each η_k captures the difference in exit rates between treated and control agencies k years relative to the reform year. The sample includes all regulators employed at least three years before each agency's reform, with fixed effects absorbing persistent and time-varying differences across agencies and occupations. The plotted coefficients show no evidence of pre-trends prior to reform and a sharp, statistically significant increase in exits ($k > 0$) immediately afterward. Vertical bars denote 90% confidence intervals. See Section 3.2.

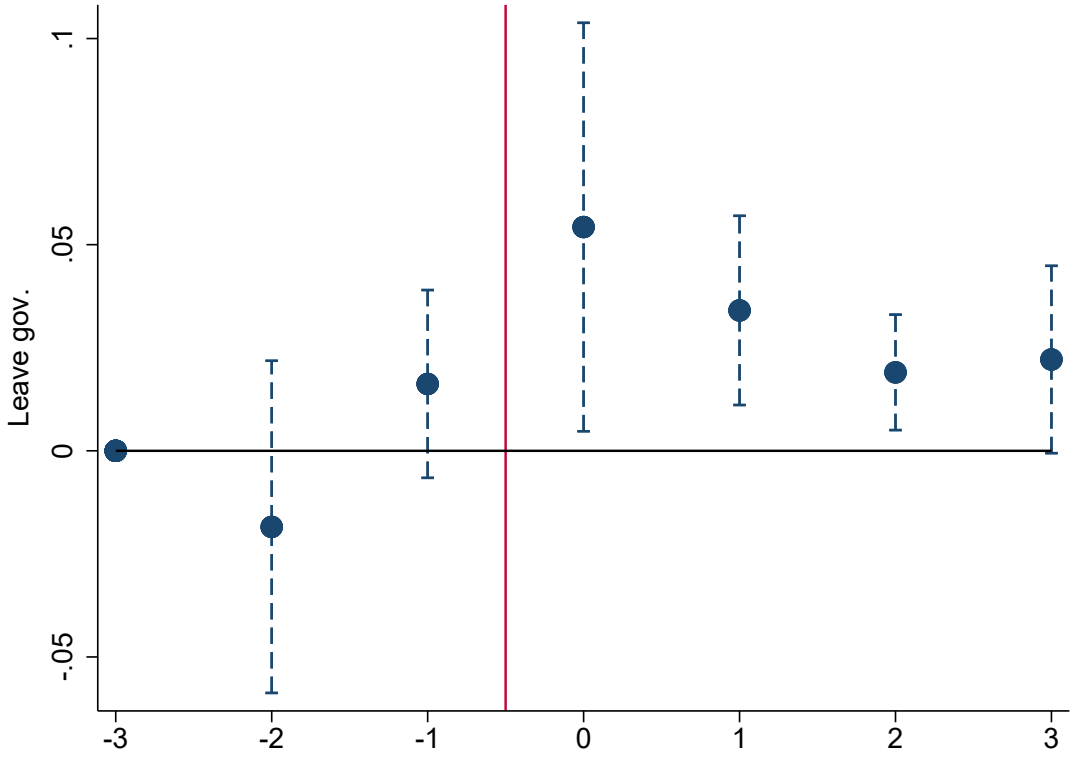

Figure 4: **Mechanisms Linking Performance Pay to Effort and Exit**

The figure illustrates the unified model, outlined intuitively in Section 5 and formally in Appendix A.3. Performance pay prompts regulators to exert greater effort and thereby increase productivity. Higher effort, in turn, amplifies pay volatility within the government and raises potential pay in the private sector, as human capital accumulates and the signal quality improves. At the same time, the shift to performance pay weakens the public sector preference for multiple reasons, such as aversion to government evaluation and disutility from effort and income risk. The interaction of these forces explains the observed pattern of higher effort and more exits (average effect), and the greater impact on regulators with weaker attachment to public service and greater aversion to being evaluated inside the government (heterogeneous effect).

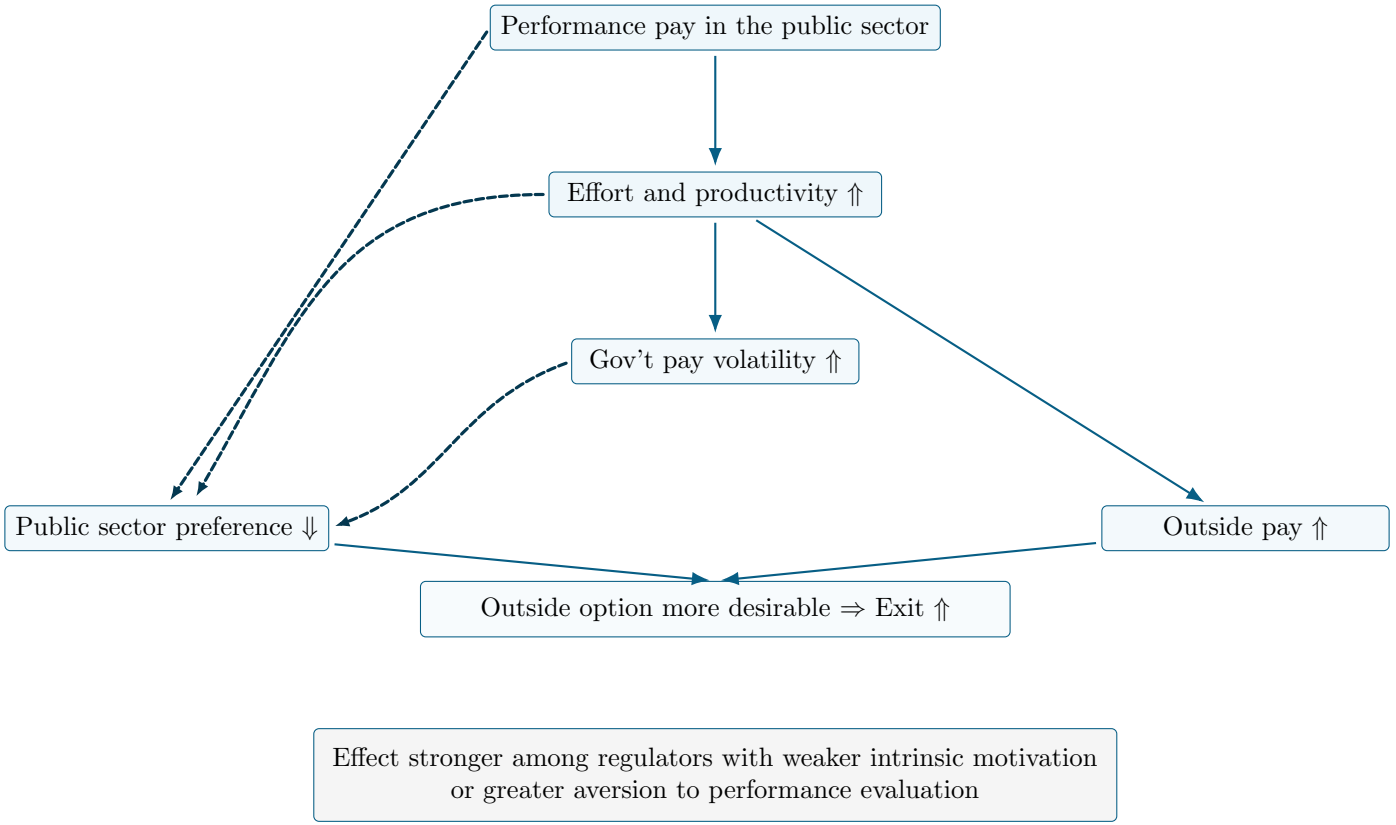

Table 1: **Career Trajectories of Financial Regulators**

In Panel A, we use the full sample of 31,871 financial regulators employed at ten U.S. financial regulatory agencies between 1973 and 2013. *Pay* denotes the employee's annual salary in constant 2023 USD. $Promotion = 1$ if the employee receives promotion to a senior rank. $Large\ Pay\ Cut = 1$ for real pay cuts above the 95th percentile (at least 2.95%). *Tenure* is the number of years since joining federal service. *Age* is the employee's age. $Manager = 1$ if the employee holds a managerial position in year t. $Exit = 1$ if the employee leaves the government in year $t+1$. $Retire = 1$ if the employee leaves the government in year $t+1$ at age 60 or older. $College = 1$ ($Postgrad = 1$) if the employee has completed a college (postgraduate) degree. In Panel B, we focus on 724 financial regulators engaged in rulemaking between 2000 and 2009. $\mathbb{1}(RegDoc) = 1$ if the regulator authored a proposed or final rule in year t, and $\#RegDoc$ is the number of new regulatory documents drafted conditional on non-zero. In Panel C, we focus on 904 financial regulators with matched Revelio Labs data who left the government, and report information on their first post-government position: $ForProfit$ for for-profit companies; $Nonprofit$ for nonprofit organizations, excluding universities; *Academic* for academic institutions; *PublicSector* for employment in non-financial agencies; and *PublicFirm* for publicly traded companies. We also report the ex-regulator's first post-government salary: base pay ($Pay^{private,base}$), total pay ($Pay^{private,total}$), and performance pay ($Pay^{private,P4P}$) (defined as $1 - \frac{Pay^{private,base}}{Pay^{private,total}}$). See Section 2.1.

Statistic:	Mean	S.D.	Min	Max
Panel A. Government careers				
Pay ($)	$123,588	$52,561	$38,186	$265,254
Promotion (0/1)	0.6	7.9	0.0	100.0
Large Pay Cut (0/1)	4.4	20.4	0.0	100.0
Tenure (years)	11.4	8.1	0.0	36.0
Age (years)	41.9	11.0	17.0	75.0
Manager (0/1)	23.9	42.6	0.0	100.0
Exit (0/1)	8.2	27.4	0.0	100.0
Retire (0/1)	1.0	10.0	0.0	100.0
College (0/1)	84.3	36.4	0.0	100.0
Postgrad (0/1)	32.6	46.9	0.0	100.0
Panel B. Rulemaking activity				
$\mathbb{1}(RegDoc)$ (0/1)	18.1	38.5	0.0	100.0
$\#RegDoc$	1.9	1.4	1.0	14.0

Statistic:	Mean	S.D.	Min	Max
Panel C. Post-government careers				
$ForProfit$	0.85	0.36	0.00	1.00
$Nonprofit$	0.08	0.27	0.00	1.00
$Academic$	0.01	0.10	0.00	1.00
$PublicSector$	0.06	0.24	0.00	1.00
$PublicFirm$	0.51	0.50	0.00	1.00
$Pay^{private,base}$	\$205,868	\$116,357	\$25,545	\$926,978
$Pay^{private,total}$	\$298,646	\$222,566	\$29,500	\$1,895,206
$Pay^{private,P4P}$	0.23	0.15	0.04	0.72
$Pay^{gov't}$	\$135,333	\$43,436	\$33,816	\$300,108

Table 2: **Balancing Test**

This table compares pre-reform characteristics between treated and control groups. The treated group consists of employees in agencies that adopted pay-for-performance (P4P) during the sample period, while the control group includes agencies that had not yet implemented P4P in the same event window. All variables are measured in the three-year period preceding each agency's reform. *Pay* denotes the employee's annual salary in constant 2023 USD. $Promotion = 1$ if the employee receives promotion to a senior rank. $Large\ Pay\ Cut = 1$ for real pay cuts above the 95th percentile (at least 2.95%). *Tenure* is the number of years since joining federal service. *Age* is the employee's age. $Manager = 1$ if the employee holds a managerial position in year t. $Exit = 1$ if the employee leaves the government in year $t+1$. $Retire = 1$ if the employee leaves the government in year $t+1$ at age 60 or older. $College = 1$ ($Postgrad = 1$) if the employee has completed a college (postgraduate) degree. The table shows that treated and control groups are well balanced along all observable dimensions prior to reform, supporting the validity of the identification strategy. See Section 3.1.

	Control	Treated	Dif	P-value
Pay ($)	$110,090	$127,635	-$17,545	0.241
Promotion (0/1)	0.003	0.002	0.001	0.465
Large Pay Cut (0/1)	0.053	0.042	0.012	0.812
Tenure (years)	11.874	12.538	-0.665	0.845
Age (years)	41.826	40.962	0.864	0.796
Manager (0/1)	0.213	0.248	-0.034	0.431
Exit (0/1)	0.085	0.079	0.006	0.674
Retire (0/1)	0.011	0.008	0.003	0.333
College (0/1)	0.806	0.870	-0.064	0.137
Postgrad (0/1)	0.333	0.283	0.050	0.535

Table 3: **Performance Pay and Exits**

Results from estimating Equation 1, which implements a stacked difference-in-differences design comparing exit behavior across agencies that adopted pay-for-performance (P4P) at different times. The sample includes ten financial regulatory agencies over a seven-year event window (three years before and after each reform). A cohort is defined by the agency and the year in which it introduced P4P. The dependent variable, $Exit = 1$, indicates that a regulator leaves government service in year $t + 1$. $Treated = 1$ for agencies adopting performance pay in that cohort, and $Post = 1$ in years following adoption. $Log(Pay_{pre})$ and $Log(Tenure_{pre})$ are the logarithms of average pay and tenure prior to the reform, respectively. All specifications include agency×cohort and occupation×year×cohort fixed effects, ensuring that identification derives from within-cohort, within-occupation variation over time. Effect (Mean%) is the coefficient on Treated×Post divided by the sample mean, capturing the proportional change in exit rates among treated agencies following the reform. Standard errors, clustered at the agency level, are in parentheses. See Section 3.2.

Outcome:	*Exit*				
	(1)	(2)	(3)	(4)	(5)
Treated×Post	0.034***	0.034***	0.037***	0.042***	0.045***
	(0.008)	(0.009)	(0.010)	(0.011)	(0.009)
$Log(Pay_{pre})$	-0.121***	-0.128***	-0.161***	-0.174***	-0.190***
	(0.028)	(0.028)	(0.037)	(0.039)	(0.040)
$Log(Tenure_{pre})$	-0.040***	-0.039***	-0.039***	-0.061***	-0.061***
	(0.011)	(0.010)	(0.011)	(0.013)	(0.014)
Cohort×Agency FE	Yes	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	Yes	Yes	Yes
Cohort×Education×Year FE	No	Yes	Yes	Yes	Yes
Cohort×Manager×Year	No	No	Yes	Yes	Yes
Cohort×AgeBin×Year FE	No	No	No	Yes	Yes
Cohort×City×Year FE	No	No	No	No	Yes
R^2	0.058	0.059	0.063	0.096	0.104
N	246,311	246,311	246,311	246,287	244,294
Effect (%Mean)	43.14	44.10	47.25	53.92	57.35

Table 4: **Performance Pay and Exits: robustness**

This table examines the robustness of our main results (Table 3). In column (1) (column (2)), the dependent variable equals 1 if the regulator exits the government in year $t+1$ and is younger (older) than 60. The estimated effect is concentrated among younger regulators, indicating that the rise in exits reflects transitions to private sector jobs rather than retirements. Column (3) implements the collapse test of Bertrand, Duflo, and Mullainathan (2004) to address concerns about serial correlation inflating t-statistics; the coefficient remains positive and significant, confirming that our findings are not driven by serial dependence. For each cohort, we collapse the data into at most two observations-one for the pre-treatment period and one for the post-treatment period-with the outcome equal to 1 if the individual exited the government in any year of that period. Column (4) excludes the 2000 cohort, in which the treated agency is the IRS; the results are virtually unchanged, indicating that the main effect is not driven by the partial coverage of the IRS reform. $Post = 1$ denotes the period after P4P adoption, and $Treated = 1$ for agencies that adopted the P4P system in a given cohort. $Log(Pay_{pre})$ and $Log(Tenure_{pre})$ are the logarithms of average pre-reform pay and tenure, respectively. Standard errors, clustered at the agency level, are reported in parentheses. See Section 3.2.

Sample:	Full	Full	Collapsed	Exclude 2000
Outcome:	$Exit_{age\leq 60}$	$Exit_{age>60}$	$Exit$	$Exit$
	(1)	(2)	(3)	(4)
Treated×Post	0.032***	0.002	0.053**	0.033***
	(0.008)	(0.001)	(0.018)	(0.009)
$Log(Pay_{pre})$	-0.117***	-0.004*	-0.207***	-0.121***
	(0.028)	(0.002)	(0.055)	(0.024)
$Log(Tenure_{pre})$	-0.047***	0.008***	-0.037**	-0.042***
	(0.009)	(0.002)	(0.015)	(0.010)
Cohort×Agency FE	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	-	Yes
Cohort×Occupation×Post FE	-	-	Yes	-
R^2	0.068	0.018	0.080	0.060
N	246,311	246,311	102,733	200,358

Table 5: **Performance Pay and Exits: Out-of-Sample Evidence**

Results from estimating Equation 3, which examines the 2004 executive pay reform affecting members of the Senior Executive Service (SES). The sample includes executives across 397 federal agencies between 2001 and 2007. $Exit = 1$ if the executive leaves the government in year $t+1$, $Post = 1$ for years $t \geq 2004$, and $Treated = 1$ for executives employed under the SES pay system prior to the reform. The control group consists of comparable managerial employees in agencies that were not part of the SES system. $Log(Pay_{pre})$ and $Log(Tenure_{pre})$ denote the logarithms of each executive's average pre-reform pay and tenure, respectively. All regressions include agency and occupation×year fixed effects. Effect (%Mean) reports the coefficient on *Treated*×*Post* as a percentage of the sample mean, indicating the proportional increase in exit rates among treated executives after the reform. Standard errors, clustered at the agency level, are shown in parentheses. See Section 3.3.

Outcome:	*Exit*			
	(1)	(2)	(3)	(4)
Treated×Post	0.041**	0.042**	0.040**	0.042***
	(0.017)	(0.017)	(0.016)	(0.016)
$Log(Pay_{pre})$	-0.223***	-0.228***	-0.246***	-0.235***
	(0.085)	(0.084)	(0.076)	(0.075)
$Log(Tenure_{pre})$	-0.017***	-0.017***	-0.038***	-0.039***
	(0.005)	(0.005)	(0.004)	(0.005)
Agency FE	Yes	Yes	Yes	Yes
Occupation×Year FE	Yes	Yes	Yes	Yes
Education×Year FE	No	Yes	Yes	Yes
AgeBin×Year FE	No	No	Yes	Yes
City×Year FE	No	No	No	Yes
R^2	0.027	0.027	0.058	0.057
N	34,924	34,924	34,922	33,972
Effect (%Mean)	39.63	40.80	39.18	40.71

Table 6: **Effect of Performance Pay on Rulemaking Activity**

This table examines how the staggered adoption of performance pay (P4P) affects rulemaking activity among financial regulators. The sample is restricted to individuals who contributed to at least one rule and to the 2002, 2003, and 2006 reform cohorts, for which detailed rulemaking data from the Federal Register are available. $\#RegDoc$ is the number of new regulatory documents an individual worked on in a given year, and $\mathbb{1}(RegDoc) = 1$ if the individual contributed to any new draft of proposed or final rule in that year. $Treated = 1$ for agencies adopting performance pay in that cohort, and $Post = 1$ in years following adoption. $Log(Pay_{pre})$ and $Log(Tenure_{pre})$ denote the logarithms of each individual's average pay and tenure prior to reform. Effect (%Mean) reports the coefficient on $Treated \times Post$ as a percentage of the corresponding sample mean, quantifying the proportional change in rulemaking among treated agencies following the reform. The results show that performance pay substantially increased regulators' rulemaking activity, consistent with higher effort under stronger incentives. Standard errors, clustered at the agency level, are shown in parentheses. See Section 4.1.

Outcome:	$\#RegDoc$			$\mathbb{1}(RegDoc)$		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated×Post	0.590**	0.561**	0.576***	0.147**	0.151**	0.160**
	(0.246)	(0.226)	(0.209)	(0.046)	(0.049)	(0.041)
$Log(Pay_{pre})$	0.479	0.285	0.083	0.066	-0.000	-0.118
	(0.314)	(0.207)	(0.230)	(0.049)	(0.084)	(0.089)
$Log(Tenure_{pre})$	0.099	0.050	0.031	-0.001	-0.013	-0.042*
	(0.077)	(0.084)	(0.088)	(0.016)	(0.020)	(0.019)
Cohort×Agency FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Education×Year FE	No	Yes	Yes	No	Yes	Yes
Cohort×Manager×Year	No	Yes	Yes	No	Yes	Yes
Cohort×AgeBin×Year FE	No	No	Yes	No	No	Yes
Cohort×City×Year FE	No	No	Yes	No	No	Yes
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS
R^2				0.061	0.064	0.042
N	3,833	3,833	3,830	3,833	3,833	3,830
Effect (%Mean)				62.52	64.21	68.39

Table 7: **Performance Pay and the Value of the Outside Option**

This table examines how the adoption of pay-for-performance (P4P) affects the value of the outside option among financial regulators who left government service. The sample includes regulators with matched Revelio Labs career data, for whom both government and first post-government salaries are observed. The dependent variable captures the log of the individual's first private-sector compensation: base pay ($Pay^{private,base}$), total pay ($Pay^{private,total}$), and performance pay ($Pay^{private,P4P}$), defined as $1 - \frac{Pay^{private,base}}{Pay^{private,total}}$. These measures reflect the fixed and variable components of compensation in the private sector. By comparing across pre- and post-reform cohorts, we show that stronger incentives inside the government are associated with higher private sector pay upon exit, consistent with the reform increasing the attractiveness of outside opportunities. Standard errors are clustered at the agency level. See Section 4.2.

Outcome:	$Log(Pay^{private,base})$	$Log(Pay^{private,total})$	$Log(Pay^{private,P4P})$
	(1)	(2)	(3)
Treated×Post	0.430***	0.525***	0.304*
	(0.098)	(0.113)	(0.139)
Cohort×Agency	Yes	Yes	Yes
Cohort×Exit Year	Yes	Yes	Yes
R^2	0.210	0.162	0.032
N	1,198	1,198	1,191

Table 8: **Effect of Performance Pay on Pay Volatility**

This table examines how the adoption of pay-for-performance (P4P) reforms affected the level and dispersion of government pay. The dependent variables are the log of annual real salary (columns 1-3); an indicator for promotion to a senior position (columns 4–6); and an indicator for a real pay cut above the 95th percentile (at least 2.95%; columns 7–9). Effect (%Mean) reports the coefficient on *Treated*×*Post* as a percentage of the corresponding sample mean, quantifying the proportional change in each outcome for treated agencies following the reform. The results indicate that the introduction of P4P increases the probability of either a promotion or a pay cut, while average pay remains stable. Standard errors, clustered at the agency level, are shown in parentheses. See Section 4.3.

	Log(Pay)			*Promotion*			*Large Pay Cut*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated×Post	0.028	0.050	0.026	0.009*	0.009*	0.008*	0.034**	0.038**	0.021**
	(0.029)	(0.037)	(0.029)	(0.004)	(0.004)	(0.004)	(0.013)	(0.015)	(0.009)
$Log(Tenure_{pre})$	0.210***	0.175***	0.120***	0.002**	-0.000	0.000	0.009**	0.007**	0.003*
	(0.030)	(0.032)	(0.024)	(0.001)	(0.000)	(0.000)	(0.003)	(0.003)	(0.001)
Cohort×Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Education×Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Cohort×Manager×Year	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Cohort×AgeBin×Year FE	No	No	Yes	No	No	Yes	No	No	Yes
Cohort×City×Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R^2	0.862	0.895	0.913	0.056	0.203	0.203	0.272	0.281	0.345
N	246,311	246,311	244,294	246,047	246,047	244,030	229,964	229,964	228,008
Effect (%Mean)				190.08	200.54	180.63	106.09	118.28	65.87

Table 9: **Aversion to Government Evaluation and Exit Behavior**

This table tests whether dissatisfaction with government performance evaluation amplifies the response to performance pay (P4P). Columns (1)-(2) use the 2008 Federal Employee Viewpoint Survey to classify agencies as opaque (transparent) when a high (low) share of employees report that performance appraisal standards are unclear, capturing perceived opacity of government evaluations. Columns (3)-(4) use the number of peers – employees working in the same office×occupation one year before reform – as a proxy for appraisal precision. Employees with few (more) peers face noisier (more precise) evaluations. Columns (5)-(6) split the sample by abnormal pay, estimated from individual fixed effects in Equation 4. Employees with below-median fixed effects historically received lower pay than predicted and likely view government evaluation processes as unfavorable. $Exit = 1$ if a regulator leaves government service in year $t+1$, $Post = 1$ in post-reform years, and $Treated = 1$ for agencies adopting P4P. Controls include $Log(Pay_{pre})$ and $Log(Tenure_{pre})$. p-values test equality with the preceding column. Standard errors are clustered at the agency level. See Section 4.4.

Outcome:	*Exit*					
Measure:	Employee Survey		#Peers		Abnormal Pay	
Sample:	Opaque	Transparent	Few	Many	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Treated×Post	0.038*	-0.008	0.041***	0.026***	0.039**	0.015**
	(0.015)	(0.014)	(0.010)	(0.008)	(0.017)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Agency FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.069	0.047	0.055	0.062	0.034	0.024
N	143,123	33,408	122,639	115,760	112,737	113,017
P-value (vs. previous column)		0.05		0.00		0.01

Table 10: **Intrinsic Motivation and Exit Response to Performance Pay**

This table examines how weaker intrinsic motivation toward public service amplifies the response to performance pay (P4P). Columns (1)-(2) split the sample by whether employees entered government service during a recession, defined using official NBER recession dates. Entering during a downturn proxies for weak intrinsic motivation, as private sector options were temporarily limited. Columns (3)-(4) classify employees by the salary forgone upon entering government service, measured as the difference between their last pre-entry salary and their initial federal salary. We split the sample at the median into large (above-median) and small (below-median) forgone-salary groups, where smaller forgone pay implies weaker attachment to the public sector. $Exit = 1$ if a financial regulator leaves government service in year $t + 1$, $Post = 1$ for post-reform years, and $Treated = 1$ for agencies adopting P4P in each cohort. All specifications include $Log(Pay_{pre})$ and $Log(Tenure_{pre})$ as controls. p-values (vs. column (1)) and (vs. column (2)) report tests of equality with the preceding specification. Regulators with weaker public sector orientation are more likely to exit, suggesting that drives out employees less intrinsically motivated by public service. Standard errors, clustered at the agency level, are reported in parentheses. See Section 4.5.

Outcome:	$Exit = 1$			
Measure:	Economic environment		Salary Forgone at Entry	
Subsample:	Recession	Non-Recession	Small	High
	(1)	(2)	(3)	(4)
Treated×Post	0.052***	0.029	0.041***	-0.081
	(0.007)	(0.016)	(0.011)	(0.063)
Controls	Yes	Yes	Yes	Yes
Cohort×Agency FE	Yes	Yes	Yes	Yes
Cohort×Occupation×Year FE	Yes	Yes	Yes	Yes
R^2	0.077	0.094	0.030	0.095
N	70,983	115,171	2,600	2,643
P-value (vs. previous column)		0.10		0.07

Internet Appendix

A.1 Pay-for-Performance Reforms

In Section 1.1, we described the pay-for-performance reforms among financial agencies in broad strokes. In this section, we describe each performance pay system of each agency. The timing of reforms across agencies was tied to agency-specific statutory authority. The earliest adopter (OCC) was uniquely able to experiment with performance-based adjustments due to a statutory exemption from the General Schedule dating back to 1974. The early-1990s wave of reforms followed the Financial Institutions Reform, Recovery, and Enforcement Act of 1989, which granted several regulators independent pay authority in response to the savings and loan crisis. The early-2000s wave was driven largely by pay parity concerns, as the SEC and CFTC remained on the GS system well after other banking regulators had moved to performance pay. It coincided with a broader government-wide push to introduce more business-like principles and practices from the private sector. Together, these staggered, statute-driven transitions provide the quasi-experimental variation we exploit in the paper.

Commodity Futures Trading Commission (CFTC) In October 2006, the CFTC began transitioning to a new performance pay system. In preparation for this transition, the agency moved in 2003 from the General Schedule (GS) pay plan to its own Commission-Track (CT) pay plan. Under the new pay-for-performance system, employees are grouped into five performance categories based on annual evaluations (OPM, 1998). Merit-based pay increases are awarded to employees in the higher performance categories, while those in the lowest category receive no increase.

Farm Credit Administration (FCA) The FCA adopted performance pay in 1993, by transitioning to the Variable Grade (VG) pay plan (FCA, 2016). The new pay plan eliminated the automatic within-grade step increases, introducing instead a merit pay matrix. Each year, supervisors conduct performance appraisals to evaluate employees against established objectives and competency standards. These ratings are reviewed by higher-level managers and the human resources office. Employees can be placed anywhere within their assigned salary range, with increases tied both to the performance rating and to the employee's position relative to market benchmarks. Employees with the same performance rating in the same grade who are paid below the market rate receive larger percentage increases than those already above the benchmark. Employees in the lowest performance category do not receive salary increases.

Federal Deposit Insurance Corporation (FDIC) In 2003, the FDIC implemented a new pay scheme tying all pay raises to performance based on measurable objectives. It also launched the Corporate Success Award, to further differentiate annual merit-based pay adjustments for the entire workforce on the basis of performance (FDIC, 2003). FDIC operates two separate performance systems. Non-managers must first earn a "meets expectations" rating to be eligible for performance-based increase, and are then placed into one of four pay groups based on their performance and "corporate contributions" relative to peers (GAO,

2007). The top-level pay ratings are limited to 25% of the employees (Source). Managers and executives are rated in three categories: Meets Expectations, Marginal, and Does Not Meet Expectations (GAO, 2007).

Internal Revenue Service (IRS) The IRS instituted a new performance management system, and the corresponding IR pay plan, in December 1999. We therefore treat 2000 as the first post-adoption year. Under the IR system, employees are rated on four levels: Outstanding, Exceeded, Met, and Less than Met. Annual performance-based increases scale with the rating. For example, in 2007, those with an Outstanding rating received a 6.5% pay increase, while those rated Less than Met" received no merit increase (OPM, 2007). Note that many non-managers, primarily unionized employees, did not transition to the IR system (source). We classify individuals as affected by the pay reform if they were originally on a managerial pay plan (pay plans ES, EX, SL) or in occupations with high transition rates to the IR: Program Management (60%), Distribution Facilities & Storage Management (50%), Financial Management (45%), and Administrative Officer (41%). We restrict the IRS sample to these likely affected individuals and define treatment status at the agency level to maintain consistency with our classification for other agencies. For robustness, we verify that our main results remain virtually unchanged when we exclude the IRS altogether (see Table 4). Additionally, in an untabulated test, we show that our result continue to hold if we focus on program managers and senior officials.

National Credit Union Administration (NCUA) The NCUA implemented performance pay in 1991, switching from the GS pay system to the Credit Union (CU) pay system. Under the CU regime, employees receive a performance rating score ranging from 0 to 300 to determine their pay increases. Employees with the same score receive the same percentage increase, while those rated "unsatisfactory"/"minimally successful" receive no increase (OPM, 2001).

Office of the Comptroller of the Currency (OCC). In 1981, the OCC became one of the first federal agencies to adopt a pay-for-performance. At that time, most federal agencies still relied on incremental, across-the-board salary increases.[1] To implement this change, the OCC switched its own pay plan, known as CP (OPM, 2007). Under the OCC's performance pay system, each employee's performance is evaluated annually by their manager. To promote fairness and consistency, managers often consult with peers and human resources when assigning ratings. Employees are then placed into one of four performance categories, and only those in the top two categories are eligible for merit-based pay increases (OPM, 2007).

Office of Thrift Supervision (OTS) The OTS implemented a performance pay system in 1991, shortly after its inception (GAO, 2007).[2] Under the OTS pay system, employees were evaluated on a five-level performance rating scale. Pay increases were tied to these ratings, with managers retaining discretion to determine the specific raise within the allowable range

[1] See *The Practice of Merit*.

[2] Although the agency was officially created in 1989, our data indicate that it first employed staff in 1990.

for each category. For example,in 2006, OTS employees in the lowest two rating category received no pay increase (OPM, 2007).

Office of Federal Housing Enterprise Oversight (OFHEO) The OFHEO has operated a performance-based pay system since its inception in 1992. Employees are appraised on a five-level scale for each element in their performance plans. The standards define the midpoint, "fully successful", and specify what raters should examine to judge performance against that benchmark. Each element is scored, producing a total performance score. Merit increases are tied directly to the overall rating, so employees can anticipate their pay adjustment from the rating received (OPM, 2007).

Securities and Exchange Commission (SEC) The SEC shifted to an independent, performance-based pay system in 2002 after Congress enacted the Pay Parity Act. In the process, the SEC switched from GS its own SK/SO pay plan (SEC, 2002). Similar to the FDIC, the SEC has separate system for managers and non-managers. Non-managers are first rated as acceptable or unacceptable by their supervisors. Those rated unacceptable do not receive annual pay adjustments. Employees rated acceptable then enter a second phase in which they are placed into four contribution categories ("highest quality," "high quality," "quality," "no significant contribution beyond acceptable"). Finally, a compensation committees review the results of each employee and recommend merit increases (OPM, 2007). Managers are appraised under a separate performance system aligned with organizational goals, and senior officers have written performance plans for each fiscal year and required progress reviews, with ratings used to inform merit pay and awards (SEC, 2007; GAO, 2007).

A.2 Executive Pay Reforms

In Section 3.3, we discuss the executive pay reform. The difference-in-differences specification (Equation 3) requires a control group consisting of agencies that were not affected by the reform. It includes the African Development Foundation; Christopher Columbus Fellowship Foundation; U.S. Tax Court; Department of Agriculture, National Appeals Board; Department of Education, Advisory Councils and Committees; Department of Education, National Assessment Governing Board; Department of Treasury, Office of Thrift Supervision; Export-Import Bank of the United States; Federal Election Commission; Federal Financial Institutions Examination Council; Government Publishing Office; International Boundary Commission (U.S. and Canada); International Joint Commission (U.S. and Canada); James Madison Memorial Fellowship Foundation; Medicare Payment Advisory Commission; Morris K. Udall Scholarship Foundation; National Security Council; Pension Benefit Guaranty Corporation; Office of Administration within the Executive Office of the President; Presidio Trust; Smithsonian Institution; Smithsonian Institution, John F. Kennedy Center for the Performing Arts; Smithsonian Institution, National Gallery of Art; Smithsonian Institution, Woodrow Wilson International Center for Scholars; U.S. Holocaust Memorial Council; and Utah Reclamation Mitigation and Conservation Commission.

A.3 Structural Model

We formalize the link between performance pay, effort, and exit in a dynamic structural model of public sector employee (regulator). The formal description below is the basis for the intuitive description in Section 5 in the main text.

A.3.1 Setup

A regulator earns a realized wage, which we model as:

$$\tilde{w}_t^g = \min\{\overline{w}_t,\ w_t^g\} \tag{A.1}$$

The left-hand side term, $\overline{w}_t$, is a set of government-wide pay caps which apply to all regulators. The right-hand side term, w_t^g, is a regulator-specific uncapped pay. This uncapped pay can exceed the pay cap, but the realized pay ($\tilde{w}_t^g$) cannot. This fundamental tension has important implications in our model.

The regulator's uncapped pay (w_t^g) is a function of tenure and performance:

$$\log(w_t^g) \;=\; \underbrace{\alpha_{Base} TenurePay_t}_{\text{deterministic}} \;+\; \underbrace{\alpha_{P4P} \max\{0,\ \log(z_t)\}}_{\text{stochastic}} \tag{A.2}$$

In words, the uncapped pay consists of two factors: a deterministic component which depends on tenure, $TenurePay_t$, and a stochastic pay-for-performance component which depends on productivity, z_t. The weights of the two parts sum up to one and are given by α_{Base} and α_{P4P}, respectively. The parameter α_{P4P} represents the exposure to pay-for-performance. If $\alpha_{P4P} = 0$, there is no pay-for-performance, meaning that wages are deterministic and based solely on tenure. As α_{P4P} increases, the regulator is more exposed to pay-for-performance. Crucially, bad performance can never reduce the regulator's pay, while good performance can increase his pay beyond the deterministic tenure-based component (up to the pay cap). This condition is captured by the right-hand side term in Equation A.2.

The regulator's productivity (z_t) evolves according to:

$$\log(z_{t+1}) \;=\; \underbrace{\rho \log(z_t)}_{\substack{\text{past}\\ \text{productivity}}} \;+\; \underbrace{\mu_t(1-\rho)}_{\text{drift}} \;+\; \underbrace{\sigma\varepsilon_{t+1}}_{\text{shock}}, \qquad \varepsilon_{t+1} \sim \mathcal{N}(0,1), \tag{A.3}$$

such that the next period productivity z_{t+1} is a function of last period productivity z_t, a drift μ_t, and a productivity shock ε_{t+1}. The drift is influenced by the regulator's tenure and their choice of effort, f_t:

$$\mu_t(f_t) = \log(f_t) + TenurePay_t. \tag{A.4}$$

Looking jointly at Equation A.2, Equation A.3, and Equation A.4, note the following. First, productivity is expected to increase with effort (f_t) and over time ($TenurePay_t$). Furthermore, the importance of the pay-for-performance component grows over time, through the influence of $Tenure_t$ on the drift μ_t.[3] When exerting a neutral level of effort (defined as

[3] This ensures that the deterministic component does not outgrow the stochastic one over time.

$\log(f_t) = 0$), the regulator expects to be paid as if there was no pay-for-performance, since the expected value of $\mathbb{E}(z_t) = TenurePay_t$. The parameters ρ and σ represent the persistence and volatility of the productivity process. By exerting more effort today, the regulator expects to increase productivity moving forward. However, the realization of productivity also depends on the shock ε. In other words, the regulator can be rewarded for performance that's either due to effort f_t or the random noise ε.

The regulator has an outside option, meaning a potential job in the private sector. We model the expected value of the outside option as:

$$o_t = \frac{1}{\theta} \cdot \phi_t \cdot w_t^o, \tag{A.5}$$

where w_t^o is the private sector wage, ϕ_t is the private sector pay differential, and θ is the public sector preference. The private sector wage w_t^o evolves differently than the public sector wage w_t^g. Concretely, we model it as:

$$\log(w_t^o) = \underbrace{\alpha_{Base}^o TenurePay_t^o}_{\text{deterministic}} + \underbrace{\alpha_{P4P}^o z_t}_{\text{stochastic}}. \tag{A.6}$$

There is some similarity between the private sector wage (Equation A.6) and the public sector wage (Equation A.1 and Equation A.2): they are both determined by tenure and productivity. However, there are three important differences. *First*, tenure-based pay ($TenurePay_t$) could grow at a different rate, and performance pay could have a different weight (α_{P4P}). *Second*, the realized public sector wage never decreases, even with dismal performance. Private sector employees, on the other hand, could experience cuts in the variable component of their salary due to poor performance, which is captured by the right-hand side term in Equation A.6. *Third*, the realized public sector wage is capped from above, meaning that the upside of good performance is limited. The realized private sector wage, on the other hand, has no such cap. Thus, unlike in the government pay, pay-for-performance in the private sector has no floor or ceiling.

To derive the value of the outside option in Equation A.5, we consider two additional adjustments. First, we include a private sector multiplier (ϕ_t), to account for the substantial wage differential between the public and private sector. Second, we include the parameter $\theta \in [1, \infty)$, representing a preference for the public sector. Equivalently, $\frac{1}{\theta}$ is a discount of the private sector. If $\theta = 1$, the regulator has no particular preference for the public sector. As θ increases, the regulator discounts the private sector wage even more. The addition of θ is motivated by earlier studies, both theoretical and empirical, arguing that employees in the public sector have a strong intrinsic motivation to serve the public interest (Macchiavello, 2008; Pitts, Marvel, and Fernandez, 2011; Bond and Glode, 2014).

A.3.2 Decision on effort and resignation

At time t, after observing the pay for this period, the regulator jointly makes two decisions: the level of effort and the employer for the next period. In other words, they choose whether to continue working in the public sector for one more period ($q_t = 0$), or exercise their outside option and quit ($q_t = 1$).When deciding whether to quit today, the regulator also

considers how much effort they will exert tomorrow, since that would affect the performance pay component and the evolution of productivity going forward. Let t_Q denote the time at which the regulator voluntarily quits to join the private sector. The regulator also faces a mandatory retirement age at t_R. Retirement income is non-stochastic and and defined by the replacement rate λ of the final wage $\lambda\widetilde{w}^g_{t_R}$. The regulator collects the retirement income for t_N more years upon retirement. Therefore, $\lambda\widetilde{w}_{t_R}$ determines the terminal value at the retirement age, V^R. Given this restriction, $t_Q < t_R$. Let $t_E \in \{t_Q, t_R\}$ indicate the time at which the regulator exits the public sector, either by choice (t_Q) or due to mandatory retirement age (t_R).

We assume that the regulator is risk averse and effort averse, with a constant relative risk aversion over their period t wage (similar to Hirshleifer and Suh, 1992; Page, 2018):

$$u(w, f) = \frac{(w-1)^{1-\gamma}}{1-\gamma} - \xi(\overline{f} - f)^{-1}, \tag{A.7}$$

where $\gamma > 0$ and $\xi > 0$ represent risk aversion and cost of effort, respectively, and $\overline{f}$ is the maximum possible effort. This specification implies that the regulator prefers receiving a certain wage w over a risky wage with the same expected value. Moreover, high levels of effort induce strong disutility for the regulator.

The regulator's salary is set at the beginning of period t and consumed during the period. The regulator's expected utility at time t is then:

$$U_t = \underbrace{\mathbb{E}_t \sum_{s=t}^{t_E} \beta^{s-t} u(\widetilde{w}^g_s, f_{s+1})}_{\text{gov't period}} + \underbrace{\beta^{t_E+1-t} q_{t_E} U^E}_{\text{post-gov't period}}, \tag{A.8}$$

where β is the discount factor. The left-hand side term captures the expected utility from government pay, until the exit point t_E (whether by choice or due to retirement).[4] The right-hand side term captures the next period's expected utility from receiving non-government pay. This includes the regulator's retirement benefits or the private sector income, if the regulator quits before the retirement age. We assume that, when deciding whether to exit the public sector, the regulator chooses how much effort to exert in the private sector ($f^q_{t_E}$) and expects to keep his effort at this level throughout the private sector career.[5]

We consider two cases of Equation A.8. In the first scenario, the regulator retires from the public sector at the mandatory age. In other words, $t_E = t_R$ and $q_{t_R} = 1$. In this case, the term U^E is:

$$U^E = U^R = \sum_{s=t_R}^{t_N} \beta^{s-t_R} u(\lambda\widetilde{w}^g_{t_R}).$$

which is the present value of retirement paychecks, based on the regulator's terminal

[4]Note that period s effort f_s was chosen at time $s-1$ and affects wages at time s. However, it is the decision about next period's effort f_{s+1} that affects the regulator's utility today.

[5]This is a simplifying assumption at the point of exit, when the regulator calculates the value of the outside option. We do not track the regulator's post-government career, and they may decide to change the level of effort at some point.

wage in the public sector $\widetilde{w}_{t_R}$ (which depends on the terminal level of effort t_R chosen at time t_{R-1}) and the replacement factor λ. In the second scenario, the regulator chooses to quit at time $t_Q = t_E < t_R$ and $q_{t_Q} = 1$. In other words, their government career is followed by a private sector career and then retirement. In this case, U^E is:

$$U^E = U^Q = \underbrace{\sum_{s=t_Q}^{t_R} \beta^{s-t_Q} \mathbb{E}_{t_Q}[u(o_s, f_{t_Q})]}_{\text{private sector career}} + \underbrace{\sum_{s=t_R}^{t_N} \beta^{s-(t_R-t_Q)} \mathbb{E}_{t_Q}[u(\lambda o_{t_R}, f_{t_R})]}_{\text{retirement period}},$$

where $\mathbb{E}_{t_Q}[u(o_s, f_{t_Q})]$ captures the expected utility of the private sector wage at time s, if the regulator quits the government job at time t_Q, and $\mathbb{E}_{t_Q}[u(\lambda o_{t_R}, f_{t_Q})]$ is the expected utility of the retirement benefits when quitting the government job at time t_Q.[6]

A.3.3 Model solution

Every period, the regulator faces three state variables: current level of effort f_t, current realization of private sector premium ϕ_t, and current productivity z_t. The regulator then chooses whether to quit or continue (q), and the level effort if staying (f) or quitting (f^q). Their goal is to maximize the present value of discounted future income. This decision is summarized in the following Bellman equation:

$$\begin{aligned} U(f_t, \phi_t, z_t) = \max_{f_{t+1}, f^q_{t+1}, q_{t+1}} \Big\{ & u(\widetilde{w}^g_t, f_{t+1}) + (1 - q_{t+1}) \beta \mathbb{E}\left[U(f_{t+1}, \phi_{t+1}, z_{t+1})\right] \\ & + q_{t+1} \beta \mathbb{E}\left[U^E(f^q_{t+1}, \phi_{t+1}, z_{t+1})\right] \Big\}. \end{aligned} \tag{A.9}$$

The problem in Equation A.9 is solved by discretizing the state variables f, ϕ and z and then applying backward induction, starting at $t = T$. The solution consists of two decision rules: whether to quit the public sector job and how much effort to exert, with respect to the next period. Each decision rule is a function of the three state variables: productivity z, effort f, and tenure t.

A.3.4 Optimal effort and exit policies

We illustrate the optimal effort and quitting choices in Figure A.2. For illustration purposes, we use the baseline parameters discussed in Appendix A.3.5 and in the Internet Appendix (Model Estimation).

Starting with effort, note that greater effort yields a short-term disutility due to cost of effort ξ. On the other hand, greater effort will increase productivity and hence the expected wages in the next period. This dynamic is captured in Panel A. First, effort decreases with tenure (the areas become darker), as the upside of effort is limited: the deterministic component of pay pushes the realized pay closer to the cap, which reduces the incentive to exert additional effort. Second, there is a non-linear relation between effort and

[6] For tractability, we assume that productivity is transferable and that effort in public sector increases the value of the outside option.

productivity. For low-productivity employees, the deterministic component and the pay floor ensures that they get sufficient compensation, resulting in no incentive to exert effort. For high-productivity employees, the pay-for-performance component pushes the realized pay closer to the cap, and the regulator will stay close to the cap due to the persistence of the productivity process. Therefore, they also have a weak incentive to exert effort. However, as they approach the area in which they optimally exit, they again try to exert more effort so as to maximize their uncapped pay in the private sector.

Exit policy involves a different trade-off. The private sector can offer higher pay, due to pay differentials ϕ and the lack of pay cap. On the other hand, the private sector pay is more exposed to performance shocks, resulting in lower pay during periods of bad performance. We capture this dynamic in Panel B. First, exits increase with productivity. Productive regulators prefer the uncapped pay of the private sector, since the upside from being a top performer in the public sector is capped and the value of their outside option is higher. Second, as the regulator is close to retirement age, two opposing factors emerge. On the one hand, the regulator is more likely to reach the maximum allowable pay in the public sector. Thus, the upside from staying in the government is even lower. On the other hand, adverse productivity shocks can lower his last wage, which would be the benchmark for his entire retirement income stream. On balance, as the regulator gets especially close to retirement, quitting rates decrease.

Panel B also captures the interdependence between exit and effort. We do so by plotting the exit policy in two scenarios, with high and low levels of effort (light and dark shades, respectively). We find that effort increases the incentive to quit: as we shift from the dark shade to the light one, the quitting region expands to younger and less productive employees. This is because the private sector rewards effort more extensively, without imposing an upper bound on salaries. Younger employees with higher effort are more likely to hit the pay cap, even though the deterministic component of their wages is relatively smaller, and thus their incentives to quit increase. Moreover, greater effort increases the expected productivity, and thus even medium-productivity employees prefer to exit.

A.3.5 Model-Based Interpretation of the P4P Reform

In this section we study the quantitative implications of our model. We provide the details regarding the data, variables used, estimation procedure, and identification of structural parameters in the Internet Appendix (Model Estimation).

We first estimate the structural parameters for the full sample of federal executives (Table A.5, Panel A). Productivity is highly persistent ($\rho = 0.72$) and smooth ($\sigma = 0.16$). The cost of effort ξ is small in level but meaningful: a 1% increase reduces log wages by 0.02% as executives supply less effort. Most strikingly, we find that only 9% of federal executive pay behaves as if tied to performance ($\alpha_{P4P} = 0.0918$), with the remaining 91% driven by deterministic components.[7] This is far below private sector norms, where performance pay accounts for roughly half of non-CEO executive pay and 75-90% for CEOs by the 2000s (Murphy, 1985; Frydman and Jenter, 2010b; Edmans, Gosling, and Jenter, 2023b). We also estimate a strong public-sector preference ($\theta = 2.11$): executives discount private wages by

[7] We can only say that executives behave *as if* this were the case.

53%. For instance, an average federal salary of $225k corresponds to an expected private salary of $282k, but is valued as if it were only $134k. This preference is large enough to offset higher market pay and helps explain the persistence of federal careers. The parameters are precisely estimated and the model replicates key features of the data: simulated pay gaps (5.85% vs. 5.53%) and turnover (14.0% vs. 12.2%) closely match the data. Overall, the estimates suggest that federal executives face weak performance incentives but derive substantial utility from public service.

We next estimate parameters for executives observed after the reform (2005-2012), reported in Table A.5, Panel B. The key change is that the performance component rises sharply to 24% (vs. 9% pre-reform), validating the reform's intent. Executives also exert 4.5% more effort, showing that incentives succeeded in raising productivity. At the same time, the estimated public sector preference increases from 2.11 to 2.39, implying that stayers are an increasingly self-selected group. This shift likely reflects sorting: the reform crowded out some of the intrinsic motivation for public service, raising the relative value of the outside option. Executives with weaker attachment exited, while those with stronger attachment remained.

Overall, the model suggests that performance pay induced more effort but simultaneously raised turnover pressure. Three mechanisms operate jointly: weaker public sector preference, stronger outside option value, and the higher chance of hitting the government pay cap. Together these explain the observed post-reform pattern of higher effort, higher exits, and more mission-oriented stayers.

A.3.6 Alternative pay policies

In this section, we use the model to explore how alternative pay policies would affect both effort and exit. We focus on two levers: the weight on performance pay (α_{P4P}) and the pay cap ($\overline{w}_0$). Figure A.3 shows that both raise effort, but they have opposite effects on exit: more performance pay increases turnover (by making the private sector more attractive and wages more volatile), while a higher cap reduces it. Quantitatively, a 1% increase in performance pay raises exits by 0.8% and effort by 0.11%, while a 1% higher cap lowers exits by 1.0% and raises effort by 0.02%. Thus, performance pay is the stronger lever for effort, while caps dominate in shaping retention.

We then simulate counterfactual policies. A "no extra exits" policy balances a moderate increase in the cap with a large increase in performance pay, holding θ fixed. This ensures that effort rises without higher turnover (Figure A.4). Other policies target different parameters (Table A.6). Reducing the cost of effort (ξ), for instance through remote work or AI tools, raises effort but also accelerates exits, since the private option becomes more valuable. Weakening public sector preference (θ), or widening public-private pay differentials (ϕ), has a similar effect: more effort but also a wave of exits. In all cases, the mechanism is the same: policies that make effort more rewarding also raise the value of leaving.

In sum, our counterfactual analyses reveal a fundamental trade-off. Any policy that successfully increases effort also raises turnover, especially among more productive executives. In the short run, this boosts average productivity in government. In the long run, however, the workforce tilts toward less productive but more mission-oriented stayers, highlighting the tension between effort, retention, and the composition of the public service.

Figure A.1: **Dynamic Response of Federal Executives to Performance Pay**

The figure reports the coefficients (η_k) from the dynamic difference-in-differences specification, $Exit_{i,t+1} = \sum_{k=-3}^{+3} \eta_k \cdot Post2004_t \times \mathbb{1}(EventTime_t = k) + \alpha_i + \alpha_t + \varepsilon_{i,t}$, which traces the evolution of exit rates among federal executives around the 2004 reform of the Senior Executive Service (SES) pay system. Each η_k measures the change in the probability of leaving government service k years relative to the reform year (vertical red line). The sample includes treated and control executives observed between 2001 and 2007, where control executives are those in comparable managerial positions outside the Senior Executive Service (SES) system. Pre-reform coefficients ($k < 0$) are close to zero, indicating no evidence of pre-trends, while exit rates rise sharply following the reform and remain elevated for several years thereafter, consistent with higher turnover under performance pay. Vertical bars represent 90% confidence intervals. See Section 3.3.

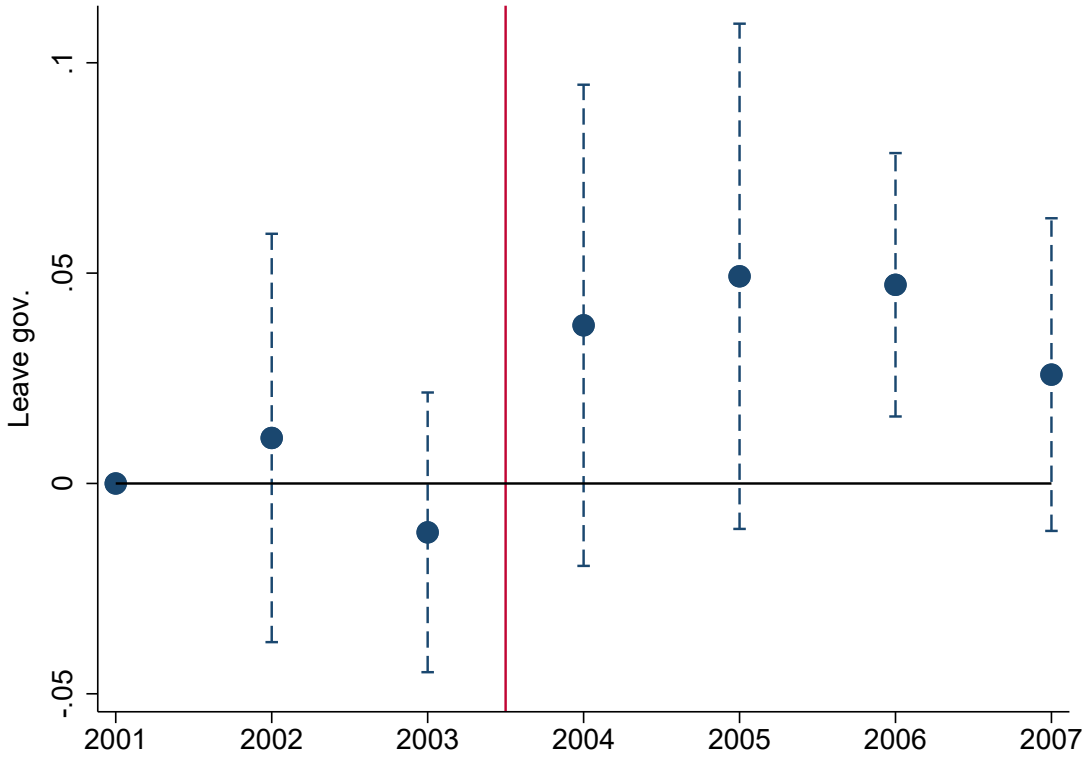

Figure A.2: **Optimal Effort and Exit Decisions under Performance Pay**

In Panel A, we plot the optimal effort policy implied by the structural model, using the baseline parameters discussed in Appendix A.3.4. Darker areas correspond to lower optimal effort, expressed as the percentage deviation from the minimum level of effort. The figure shows that effort increases with both productivity and tenure, as higher-ability regulators and those later in their careers exert more effort. In Panel B, we plot the optimal exit policy for two effort levels – high (lighter shade) and low (darker shade) – given the same baseline parameters discussed in Appendix A.3.4. Effort expands the optimal exit region across a broader range of productivity and tenure states, reflecting the higher opportunity cost of continued government service under stronger incentives. Together, the panels illustrate how performance pay simultaneously raises effort and induces selection out of the public sector, capturing the key trade-off in the model between incentives and retention.

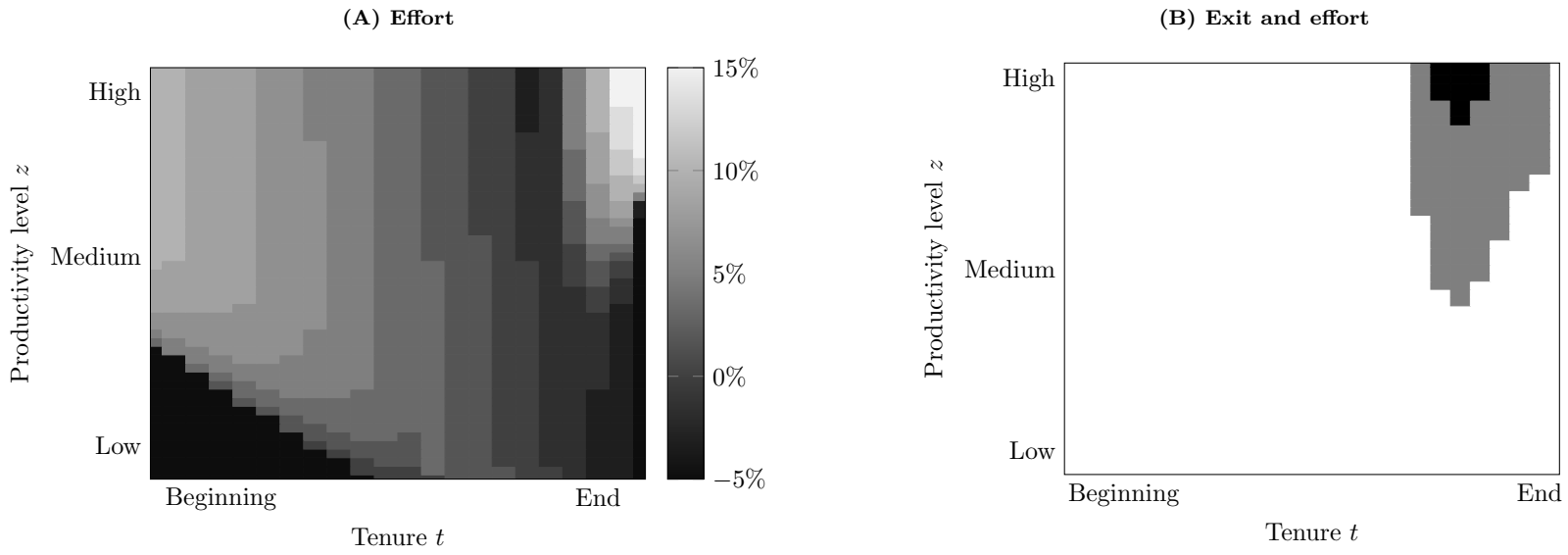

Figure A.3: **Comparative Statics of Effort and Exit with Respect to Pay Parameters**

The figure shows how effort (Panel A) and exit (Panel B) respond to changes in the two key parameters of pay packages: the intensity of performance sensitivity, α_{P4P}, and the government pay cap, $\overline{w}_0$. Both parameters are expressed as percentage deviations from their baseline estimated values reported in Table A.5, Panel A, while all other parameters are held constant. Both performance pay (α_{P4P}) and pay cap ($\overline{w}_0$) raise effort, but they have opposite effects on exit: more performance pay increases turnover (by making the private sector more attractive and wages more volatile), while a higher cap reduces it. See Appendix A.3.6.

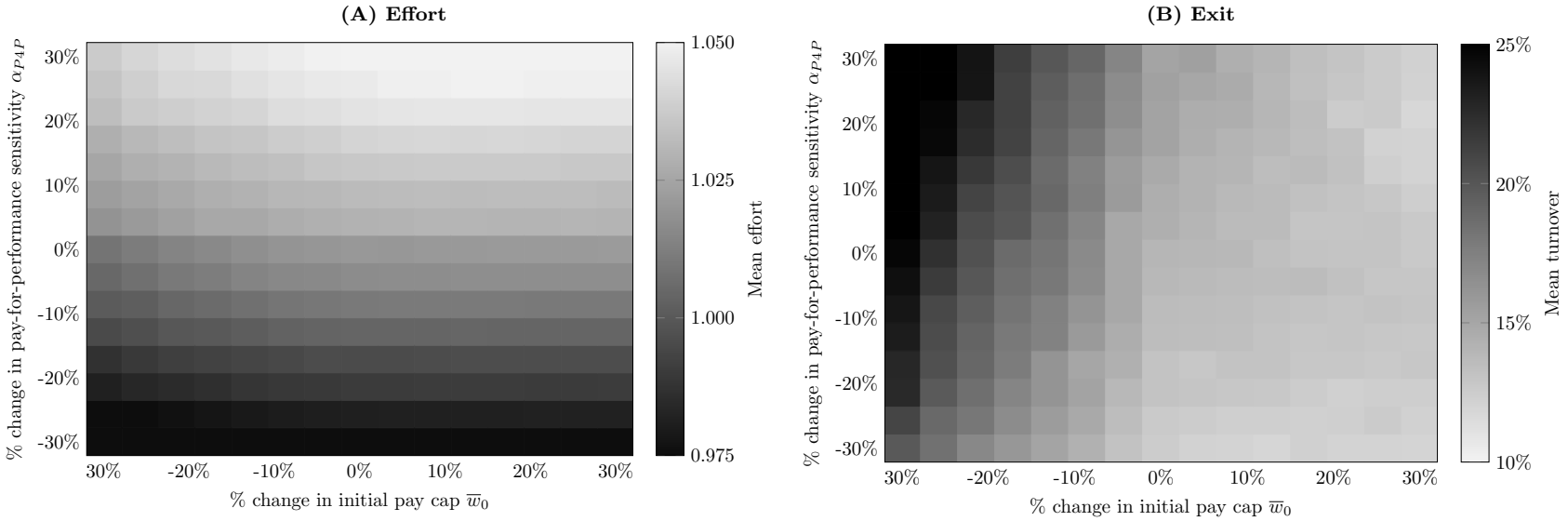

Figure A.4: **Trade-Off Between Incentive Intensity and Pay Caps**

The figure traces combinations of pay-for-performance sensitivity (α_{P4P}) and government pay cap ($\overline{w}_0$) that minimize the deviation in turnover from its baseline value, holding all other parameters fixed. Each point on the curve represents a simulated equilibrium outcome obtained from the structural model, where higher α_{P4P} raises exit risk, while a higher $\overline{w}_0$ offsets this effect by relaxing the constraint on potential government earnings. The curve captures combinations of the two policy parameters that maintain a constant turnover rate, illustrating the trade-off between incentive intensity and pay ceilings. The curve is a second-order polynomial interpolation based on discrete counterfactual simulations, and the square marker denotes the baseline parameter values reported in Table A.5. See Appendix A.3.6.

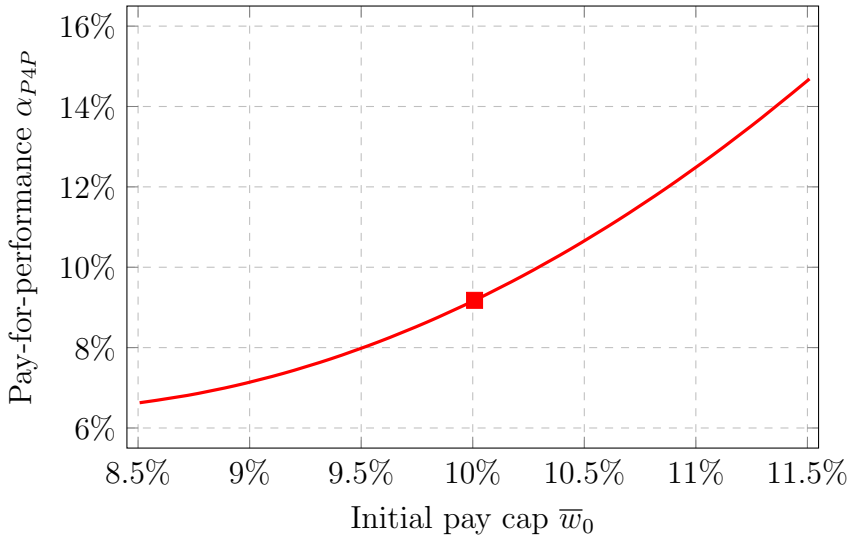

Table A.1: **Career Trajectories of Federal Executives**

The sample includes all federal executives who held an executive position between 1973 and 2013, focusing on the years in which they served in such positions. *Pay* denotes the executive's annual salary in constant 2023 USD, and ΔPay is the year-on-year change in salary, both winsorized at the 1st and 99th percentiles. *Tenure* is the number of years since entering the public sector, and *Age* is the executive's age. $Exit = 1$ when the executive leaves government service, regardless of rank at the time of departure. $College = 1$ and $Postgrad = 1$ indicate whether the executive holds a college or postgraduate degree, respectively. See Section 3.3.

Statistic:	Avg.	Median	S.D.	Min	Max	Obs.
$Pay(\$)$	212,929.5	217,500.0	21,909.4	167,916.0	251,080.7	156,634
ΔPay	1.3	0.2	5.2	-9.4	23.3	130,816
$Tenure$	16.5	16.0	9.2	0.0	40.0	156,634
Age	51.5	52.0	8.3	22.0	75.0	156,633
$Exit$	10.8	0.0	31.0	0.0	100.0	151,902
$College$	96.6	100.0	18.1	0.0	100.0	156,634
$Postgrad$	75.4	100.0	43.1	0.0	100.0	156,634

Table A.2: **Estimating Abnormal Pay**

Results from estimating Equation 4. The dependent variable is the logarithm of pay, $Log(Pay)$, and $Log(Tenure)$ denotes the logarithm of tenure. This specification decomposes individual-level pay into components attributable to observable characteristics versus persistent individual effects, following Abowd, Kramarz, and Margolis (1999). The resulting individual fixed effects provide a measure of abnormal pay, which we use in Section 4.4 to proxy for the ability to navigate the government pay system. Robust standard errors are reported in parentheses. See Section 4.4.

Outcome:	*Log(Pay)*		
	(1)	(2)	(3)
$Log(Tenure)$	0.191***	0.193***	0.182***
	(0.001)	(0.001)	(0.001)
Cohort×Individual FE	Yes	Yes	Yes
Cohort×Education×Year FE	Yes	Yes	Yes
Cohort×Office×Year FE	Yes	No	Yes
Cohort×Occupation×Year FE	No	No	Yes
Cohort×Agency×Year FE	No	Yes	No
Cohort×City×Year FE	No	Yes	No
R^2	0.989	0.989	0.991
N	309,329	310,985	307,753

Table A.3: **Aversion to Government Evaluation and Exit Behavior: Robustness**

This table is similar to columns 5-6 in Table 9, except for how we estimate abnormal salaries in the pre-reform period (Equation 4). In Panel A, we replace office×year fixed effects with agency×year and city×year fixed effects. In Panel B, we add occupation×year fixed effects. We continue to define high and low-abnormal pay as above and below cohort-median, respectively. See Section 4.4.

Outcome:	*Exit*	
Abnormal pay:	Low	High
	(1)	(2)
Panel A: Agency×year & City×year FE & Edu×Year FE		
Treated×Post	0.040**	0.015**
	(0.017)	(0.005)
R^2	0.034	0.023
N	112,753	113,011
P-value (vs. column(1))		0.01
Panel B: Office×year & Occupation×year FE & Edu×Year FE		
Treated×Post	0.040**	0.014**
	(0.017)	(0.005)
R^2	0.037	0.022
N	112,617	112,941
P-value (vs. column(1))		0.10

Table A.4: **Pre-Government Careers of Financial Regulators**

Panel A: Prior employer. This table reports the pre-government careers of financial regulators. The data are from Revelio Labs and include unique financial regulators who held positions outside of the government. $ForProfit$ Firm refers to private sector, for-profit companies. $Nonprofit$ includes nonprofit organizations, excluding universities. *Academic* refers to academic institutions. $PublicSector$ indicates employment in non-financial agencies. $PublicFirm$ refers to publicly traded companies. $Pay^{gov't}$ denotes the regulators first annual salary at financial agencies (in 2023 dollars). $Pay^{private,base}$ is the employee's final base salary (in 2023 dollars) before joining the government. *Total Private Comp.* is the employee's final total compensation (Base salary + Bonus in 2023 dollars) before joining financial agencies. *Performance pay* is the final performance pay $(1 - \frac{Private\ Base\ Salary}{Total\ Private\ Comp.})$ before financial agencies. See Section 4.5.

	Mean	**SD**	**Min**	**Max**
$ForProfit$	0.63	0.48	0.00	1.00
$Nonprofit$	0.07	0.25	0.00	1.00
$Academic$	0.02	0.14	0.00	1.00
$PublicSector$	0.28	0.45	0.00	1.00
$PublicFirm$	0.26	0.44	0.00	1.00
$Pay^{private,base}$	\$125,835	\$70,725	\$9,882	\$593,139
$Pay^{private,total}$	\$153,170	\$109,235	\$10,038	\$934,275
$Pay^{private,P4P}$	0.14	0.11	0.00	0.80
Pay^{gov}	\$88,569	\$47,002	\$10,543	\$238,495

Panel B: Seniority. This table reports the seniority level of financial regulators' last job before entering the financial agencies. Seniority is categorized into seven levels, as defined by Revelio Labs. See Section 4.5.

Seniority	**Number**	**Fraction** (%)
Senior Executive (e.g., CFO; COO; CEO)	6	0.80
Executive (e.g., Partner, Managing Director)	15	2.01
Director (e.g., Head of Legal)	73	9.77
Manager (e.g., Lead Lawyer)	73	9.77
Associate/Analyst (e.g., Attorney)	89	11.91
Junior (e.g., Legal Adviser)	384	51.41
Entry (e.g., Paralegal)	107	14.32

Panel C: Industries. This table reports the top 15 industries in which financial regulators held prior to entering financial agencies. Industry is classified using the *Revelio Industry Classification K=50* clusters. See Section 4.5.

Industry:	**Number**	**Fraction (%)**
Legal Services	256	45.55
Public Sector Management	80	14.23
Financial Services	67	11.92
Consulting and Advisory Services	53	9.43
Aerospace and Defense	26	4.63
Education Services	16	2.85
Professional and Trade Associations	10	1.78
Information Technology Services	7	1.25
Healthcare and Wellness Services	5	0.89
Miscellaneous	5	0.89
Energy and Resources	4	0.71
Research and Development	4	0.71
Industrial Manufacturing	3	0.53
Commercial Aviation	2	0.36
Culture and Entertainment	2	0.36

Table A.5: **Model-Based Interpretation of the P4P Reform**

Panel A. Full sample. In the top panel, we report the estimates of the structural parameters: α_{P4P} is the sensitivity to pay-for-performance; θ is the public sector preference; ρ is the persistence of productivity; σ is the volatility of productivity; ξ is the cost of effort. In the bottom panel, we report the data-implied and the model-implied moments. We use the full sample of federal executives between 1996-2012 and the Simulated Method of Moments, which chooses model parameters by minimizing the distance between the moments from a simulated panel of firms and their data counterparts. We report the estimated parameters and their standard errors, clustered at the executive level. See Appendix A.3.5.

Parameter estimates:

Parameter	Symbol	Estimate	Std. error
Pay-for-performance	α_{P4P}	0.0918	0.0151
Preference for public sector	θ	2.1126	0.1057
Productivity persistence	ρ	0.7245	0.1011
Productivity volatility	σ	0.1613	0.0091
Effort aversion	ξ	0.0029	0.0003

Model-implied and data-implied moments:

Moment	Simulated	Actual
Variance of residual log wage	0.0004	0.0009
Serial correlation of residual log wage	0.1445	0.1989
Mean pay gap	0.0585	0.0553
Variance of pay gap	0.0010	0.0013
Serial correlation of pay gap	0.1450	0.1229
Mean turnover	0.1399	0.1218
Covariance of wages and pay gaps	-0.0004	-0.0006

Panel B. After P4P. In the top panel, we report the estimates of the structural parameters: α_{P4P} is the sensitivity to pay-for-performance; θ is the public sector preference; ρ is the persistence of productivity; σ is the volatility of productivity; ξ is the cost of effort. In the bottom panel, we report the data-implied and the model-implied moments. We use the sample of federal executives after the reform, 2005-2012, and the Simulated Method of Moments, which chooses model parameters by minimizing the distance between the moments from a simulated panel of firms and their data counterparts. We report the estimated parameters and their standard errors, clustered at the executive level. See Appendix A.3.5.

Parameter estimates:

Parameter	Symbol	Estimate	Std. error
Pay-for-performance	α_{P4P}	0.2355	0.0036
Preference for public sector	θ	2.3982	0.0125
Productivity persistence	ρ	0.7102	0.0161
Productivity volatility	σ	0.2051	0.0041
Effort aversion	ξ	0.0015	0.0002

Model-implied and data-implied moments:

Moment	Simulated	Actual
Variance of residual log wage	0.0010	0.0006
Serial correlation of residual log wage	0.1640	0.2143
Mean pay gap	0.0295	0.0803
Variance of pay gap	0.0020	0.0009
Serial correlation of pay gap	0.1662	0.1164
Mean turnover	0.1222	0.1229
Covariance of wages and pay gaps	-0.0010	-0.0004

Table A.6: **Counterfactual Policies**

We consider three counterfactual policies that change the cost of effort (left), private sector discount (middle), and pay differential (right). In each case, we consider a 1% deviation from the baseline parameter estimate in Table IA.1 and Table A.5, while holding all other parameters constant. Note that a 0.1% change in private sector discount ($\frac{1}{\theta}$) corresponds to a 0.1% change in public sector preference (θ). For each policy, we report the resultant impact on turnover, effort, and wages. See Appendix A.3.6.

Parameter:	Cost of effort	Public sector preference	Pay differential
	ξ	θ	μ_ϕ
Magnitude:	-1%	-0.1%	$+1\%$
%Δ Mean turnover	2.42%	6.32%	14.84%
%Δ Mean effort	0.13%	0.08%	0.22%
%Δ Mean productivity	0.18%	0.09%	0.27%
%Δ Mean log wages	-0.01%	-0.09%	-0.15%
%Δ Mean log wages (non-quitters)	0.03%	0.01%	0.02%

Internet Appendix: Model Estimation

In this Internet Appendix (Model Estimation), we explain how we structurally estimate the model. We first outline the data used to construct variables used in the estimation procedure and estimate certain parameters outside of the model. We then describe the estimation procedure and the identification of structural parameters.

IA.1 Data and Variables

We focus on the sample of treated executives on the SES pay plan over the sample period of 1996–2012, that is, ±8 years around the reform. estimate the model using the sample of federal executives, rather than financial agencies, because the estimation technique is better suited for a single event rather than a sequence of staggered reforms.

We first estimate several parameters outside of the model, such as the deterministic components of wages. Those are reported in Table IA.1. We then estimate the remaining parameters of interest: pay-for-performance (α_{P4P}), public sector preference (θ), cost of effort (ξ), and the persistence (ρ) and volatility (σ) of productivity. We start by estimating the parameters describing the public sector wage. We model $TenurePay_t$, the deterministic component of public sector wages (Equation A.2), as a linear function of tenure:

$$TenurePay_t = l_0 + l_t \times t \tag{IA.1}$$

We scale each executive's wage by their first available salary in the public sector.[1] As a result of this scaling, the executive's wage at time 0 is 1, which implies that $l_0 = 0$. To estimate l_t, we regress the log of the relative salary on tenure while controlling for executive, agency, city, and occupation fixed effects to remove unobserved heterogeneity. Based on the results (untabulated), we set the deterministic trend (l_t) to be 0.0320. This means that the annual pay raise is 3.2%.

Next, we turn to the executive pay cap ($\overline{w}$). As with wages, we normalize it by the executive's first initial salary. The scaled variable represents the growth potential of the executive's salary. We model the growth potential as growing linearly with time t, starting at the initial level $\overline{w}_0$:

$$\log(\overline{w}_t) = \overline{w}_0 + \overline{w}_t \times t \tag{IA.2}$$

To estimate $\overline{w}_0$, we calculate the average initial log growth potential in the year of hiring (for executives whose initial tenure year is in our sample), and find that the starting salary is 10.5% lower than the maximum allowable salary in that year. We then regress the log growth potential on time trend while controlling for executive, agency, city, and occupation fixed effects. Based on the results (untabulated) we set the deterministic trend ($\overline{w}_t$) to be 0.0234. This estimate implies that, on average, growth potential is increasing by 2.37% with each year of tenure.

Next, we model the private sector pay differential as a lognormal random variable:

$$\log(\phi_t) \sim \mathcal{N}\left(\mu_\phi, \sigma_\phi^2\right) \tag{IA.3}$$

[1] We thus effectively focus on the evolution of salary within executive, rather than across executives.

To obtain μ_ϕ and σ_ϕ, we source data from the Federal Salary Council, which recommends annual adjustments the federal pay due to changes in the private sector labor markets.[2] We obtain all the reports for the years 2001-2012, and calculate the average and standard deviation of the pay differential across all years. The average pay differential is 21.97% with a standard deviation of 5.91%. We use these values for μ_ϕ and σ_ϕ, respectively.

To estimate the private sector wage w_t^o, we assume that the pay-for-performance in private sector α_{P4P}^o is 50%. This is based on the private sector compensation structure estimates of **?** for non-CEO executives. For the deterministic part of private sector wage, $TenurePay_t^o$, we assume that, upon exiting the government, the executive's new private sector salary grows at a rate of $l_t + p_t$, where l_t is the growth rate of the deterministic component in the government (Equation IA.1). This means that it is exceeds the public sector salary growth rate by p_t. The estimated incremental growth rate of pay p_t is 0.0230 and we calculate it as the difference between the average salary growth of private sector workers in management, business, and financial occupations (which plausibly correspond to the federal executives in our sample) from the BLS Employment Cost Index, and the average growth of log wages of sample executives over the sample period.

We set the executive's risk aversion coefficient at $\gamma = 3$.[3] The risk-free interest rate r equals 2.75%, which is the average 3-month Treasury bill rate over the sample period. We assume that the executive works for a maximum of $T = 25$ years[4] and lives for additional $t_N = 15$ years after retiring from the public sector. Similar to Briggs et al. (2021), we set the retirement replacement factor (λ) to 60%.

IA.2 Estimation

We now turn to the key parameters of the model: the pay-for-performance (α_{P4P}), the public sector preference (θ), the persistence and volatility of productivity (ρ and σ), and cost of effort (ξ). We estimate those using the Simulated Method of Moments (SMM).

We first solve the model numerically, given the parameters, and generate simulated data from the model. Then, we compute a set of moments from the simulated and the actual data. The SMM estimation procedure determines the parameter values that minimize the weighted distance between the model-implied moments and their empirical counterparts. An important issue is the unobserved heterogeneity in the data. In the model, the only source of heterogeneity is the draws of productivity shock (ε). In reality, productivity shocks could be correlated with unobserved factors such as talent or ability, or with factors not explicitly captured by the model such as geographical location or agency. To address this challenge, we follow Hennessy and Whited (2007) and extract as much of observed heterogeneity from data as possible to make the model- and data-implied moments comparable. In particular, when calculating the variance (and covariance) of wages using residual wage after controlling

[2] Its reports can be retrieved from the OPM website.

[3] It is close to, but higher than, the value of 2.05 reported in Page (2018) using CEOs in the private sector. This is because public sector employees are likely more risk averse. In untabulated results, we demonstrate that lower risk aversion results in greater effort and more turnover.

[4] According to Office of Personnel Management (2019), the average length of service at retirement was 24.9 years in 2019.

for the deterministic growth rate and executive-, agency-, city- and occupation fixed effects. Similarly, when calculating the variance and covariance of pay gaps we also control for executive-, agency-, city- and occupation fixed effects. We use the Han and Phillips (2010) estimator to calculate the serial correlation of wages and pay gaps given the panel structure of the data.

Let the pooled time series of all firms be $x_i = x_1, \ldots, x_N$, where $N = n \times T$ is the total number of firm-year observations. Using the transformed data, we compute a set of moments $h(x_i)$. We create the simulated moments by first solving the model given a vector of parameters $\Theta = (\alpha_{P4P}, \theta, \rho, \sigma, \xi)$ and then generating simulated data y from the model. We simulate $S = 10$ datasets of $N = 10,000$ executives over $T = 25$ years, following Michaelides and Ng (2000), who find that a simulation estimator behaves well in finite samples if the simulated sample is approximately ten times as large as the actual data sample. The resulting moments in a given simulated sample are given by the vector $h(y_s, \Theta)$.

The simulated methods of moments estimator $\widehat{\beta}$ is then the solution to:

$$\widehat{\Theta} = \arg\min_{\Theta} \left[g(x) - g(y, \Theta)\right]' W \left[g(x) - g(y, \Theta)\right], \tag{IA.4}$$

where $g(x) = \frac{1}{N}\sum_{i=1}^{N} h(x_i)$ and $g(y, \Theta) = \frac{1}{S}\sum_{s=1}^{S} h(y_s, \Theta)$ are the sample means of the actual and model-implied data, and W a positive definite weight matrix, which we calculate following Bazdresch, Kahn, and Whited (2017) as the optimal clustered weight matrix. We use simulated annealing to find the optimum to the minimization problem.

Under mild regularity conditions, the SMM estimator is asymptotically normal

$$\sqrt{N}(\widehat{\Theta} - \Theta) \xrightarrow{d} \mathcal{N}(0, V), \tag{IA.5}$$

where V is the covariance matrix as in Newey and McFadden (1994). Note that we use more grid points for state variables when calculating standard errors and conducting counterfactual experiments for increased precision.

IA.3 Identification

The Simulated Method of Moments estimators are identified when the empirical moments equal the simulated moments if and only if the structural parameters are at their true value. A sufficient condition for this is a one-to-one mapping between a subset of structural parameters and the selected moments. In other words, the moments must vary when the structural parameters vary. Note that all the moments are somewhat sensitive too all the parameters, because effort and exit decisions are intertwined with the wage dynamics. However, some relationships are strongly monotonic in the underlying parameters, and are thus useful for identifying the corresponding parameter. When structurally estimating the model, we include all the meaningful moments generated by our model to understand which features of the data it can and cannot explain. Concretely, we focus on exit rate; variance and serial correlation of wages; mean, variance and serial correlation of pay gaps (difference between the executive's pay and the maximum allowable pay); and covariance between wages and pay gaps. The public sector preference θ is identified by the exit rate. Intuitively, for any

level of salary we can find a value of θ such that the executive is indifferent between staying in the government or switching to the private sector. The volatility of productivity σ is identified by the variance of the detrended wages: when σ is higher and the executive is paid for performance, the variation in detrended wages increases. The productivity persistence ρ is identified by the serial correlation of detrended wages, since high persistence yields a higher serial correlation. The pay-for-performance parameter α_{P4P} is identified by the mean pay gap: the greater the exposure to pay-for-performance, the greater the incentive to exert effort, and thus the higher the chance of hitting the pay cap (low pay gap).[5] Finally, the cost of effort ξ is identified by the covariance between wages and pay gaps, and the serial correlation of wages and pay gaps. When the cost of effort is high, employees exert less effort, making the noise component of productivity more prominent. As a result, wages and pay gaps become less persistent, and the correlation between them weakens.

We compute the diagnostic measure of Andrews, Gentzkow, and Shapiro (2017) to investigate whether the model parameters are locally identified by the underlying moments. The benefit of the measure is that a reported high sensitivity means not only that the moment is sensitive to the underlying parameter, but also that the parameter is precisely estimated. The results are presented in Table IA.2. Each column corresponds to a structural parameter, and each row corresponds to a moment used in the estimation procedure. The sensitivities are scaled by the standard deviations of moments. The results confirm the intuition behind the identification of the structural parameters. For instance, higher turnover rate results in higher pay-for-performance α_{P4P} while larger persistence and volatility of productivity translate to higher to variance and serial correlation of residual log wages and pay gaps, respectively. It should be noted, however, that the elasticities are only local and, moreover, highly sensitive to the numerical properties of the gradient. Because of that it might appear that some moments are not informative about the underlying parameter while in reality they do provide substantial identifying information. Also note that the sign and magnitudes of the elasticities for α_{P4P} and σ are different, in line with the intuition outlined above.

[5] Importantly, α_{P4P} and σ have the opposite effect on effort, and thus the value of the outside option, and affect wage dynamics. Consequently, they influence the executive's optimal decisions differently. For example, they distinctively affect the mean of pay gaps, which helps separately identify the two parameters.

Table IA.1: **Structural Estimation: Parameters and Moments**

The table summarizes the values of parameters estimated outside of the model (Panel A) and the definitions of variables used to create the model-implied moments (Panel B). See Appendix IA.1.

Panel A: Parameters estimated outside of the model

Parameter	**Symbol**	**Full sample**	**After reform**
Risk aversion	γ	3	3
Risk-free interest rate	r	0.0275	0.0175
Maximum tenure	T	25	25
Years in retirement	N	15	15
Retirement replacement factor	λ	0.6	0.6
Wage trend intercept	l_0	0	0
Wage trend coefficient	l_t	0.0320	0.0266
Private sector premium mean	μ_ϕ	0.2197	0.2363
Private sector premium standard deviation	σ_ϕ	0.0591	0.0654
Incremental growth rate of private sector pay	p_t	0.0230	0.0186
Initial pay cap	$\bar{w}_0$	0.1001	0.1334
Pay cap coefficient	$\bar{w}_t$	0.0286	0.0161

Panel B: Definitions of variables

Moment	**Model**	**Data**
Log relative wage	$\alpha_{Base} Tenure_t + \alpha_{P4P} \max\{0,\ \log(z_t)\}$	$\log(\text{relative wage}_{it})$
Pay gap	$(\overline{w}_t - \widetilde{w}_t^g)/\widetilde{w}_t^g$	$(\text{pay cap}_{it} - Pay_{it})/Pay_{it}$
Turnover rate	$\sum$ quits / $\sum$ executives	$\sum_{it}$ quits$_{it}$ / $\sum_{it}$ executives$_{it}$

Table IA.2: **Structural Estimation: Diagnostics**

We present the sensitivities of the structural parameters to moments using the full-sample estimates and the diagnostic tool of Andrews, Gentzkow, and Shapiro (2017), which measures the local sensitivity of parameters to moments (scaled by the standard deviation of moments). α is the sensitivity to pay-for-performance; θ is the preference for public sector; ρ is the persistence of productivity; σ is the volatility of productivity; ξ is the cost of effort. See Appendix IA.2 and Appendix IA.3

	α_{P4P}	θ	ρ	σ	ξ
Variance of residual log wage	-0.0086	-0.0582	0.0533	0.0058	-0.0002
Serial correlation of residual log wage	-0.0045	-0.0311	0.0278	0.0024	-0.0001
Mean pay gap	0.0056	0.0473	-0.0519	-0.0027	0.0001
Variance of pay gap	-0.0200	-0.1391	0.1310	0.0151	-0.0004
Serial correlation of pay gap	0.0040	0.0232	-0.0196	-0.0033	0.0001
Mean turnover	0.0005	0.0039	-0.0031	-0.0004	0.0000
Covariance of wages and pay gaps	-0.0129	-0.0896	0.0831	0.0115	-0.0003