Journal of Institutional Economics (2026), **22**, e15, 1–21
doi:10.1017/S1744137426100496

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

AI adoption in bureaucracies

Luca Pieroni[1] and Melcior Rossello Roig[2]

[1]Department of Political Science, Universita degli Studi di Perugia, Italy and [2]National Health Service England, United Kingdom
Corresponding author: Melcior Rossello Roig; Email: rosselloroigsion@gmail.com

(Received 17 May 2025; revised 5 March 2026; accepted 5 March 2026)

Abstract
This paper examines relationships between AI occupational exposure and workforce patterns in U.S. federal agencies from 2019–2024. Using administrative employment data, we document systematic associations between agencies' concentrations of AI-exposed occupations and employment dynamics. Agencies with higher AI exposure exhibit declining routine employment shares, expanding expert roles, and wage compression effects. We develop a theoretical framework incorporating institutional constraints distinguishing public organisations: employment protections, standardised compensation systems, and political oversight. The model features strategic interactions between budget-maximising directors and electoral-sensitive overseers, predicting workforce evolution under institutional constraints. Our identification exploits fixed occupational exposure scores, so observed changes in agency-level exposure reflect workforce composition shifts rather than measurement artefacts. Patterns suggest agencies with greater AI-susceptible occupations experience reallocation rather than displacement, providing insights for understanding technological change in institutionally constrained environments and informing governance frameworks balancing modernisation with democratic accountability.

Keywords: AI exposure; bureaucracy; institutional constraints; workforce composition

Introduction

Artificial intelligence is reshaping economies and labour markets, yet scholarly attention has focused primarily on the private sector. The public sector – where AI integrates into bureaucratic institutions responsible for governance, regulation, and service provision – remains comparatively under-explored. Understanding how AI interacts with these institutions' structures, incentives, and hierarchies is foundational for comprehending both technological change and democratic governance.

This paper examines AI occupational exposure (AIOE) and workforce patterns in U.S. federal agencies from 2019 to 2024. Using administrative employment data covering all federal agencies, we find that agencies with higher concentrations of AI-exposed occupations exhibit shifts from routine administrative positions towards expert professional roles, accompanied by wage compression. These patterns reflect the distinctive institutional environment of government organisations, characterised by civil service protections, standardised pay scales, and political oversight.

Our empirical approach leverages the AIOE measure from Felten et al. (2021), which assesses job susceptibility to AI by mapping AI capabilities onto workplace tasks. By combining occupation-level exposure scores with quarterly federal employment records, we construct agency-level measures capturing how workforce composition translates into technological exposure. Since occupational

© The Author(s), 2026. Published by Cambridge University Press on behalf of Millennium Economics Ltd. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

exposure scores remain fixed, observed changes in agency AI exposure reflect workforce composition changes rather than evolving measurement approaches.

Drawing on Garicano (2000), we conceptualise public bureaucracies as structured layers differentiated by complexity and cognitive demands. At lower tiers – where work is rule-bound and codifiable – AI might function as a substitute. At higher levels, where decision-making relies on contextual interpretation and expertise, AI might serve as a complement, augmenting rather than replacing expert capacity.

We develop a political economy model incorporating institutional constraints distinguishing public organisations from private markets. Unlike private firms that adjust employment and wages freely, government agencies navigate rigid employment protections, standardised compensation, and political oversight. We model strategic interactions between agency directors, maximising institutional scope, and political overseers, balancing service delivery against electoral costs.

While our model simplifies the bureaucratic agent as utility-maximising, we acknowledge Hodgson's (2012) critique that individuals operate under complex moral frameworks evolved from social interactions. Our model, therefore, recognises that bureaucratic agents may make workforce decisions based on moral deliberations about fairness, accountability, and legitimacy – incorporated through political cost functions and institutional constraint mechanisms.

Our framework generates four propositions aligning with observed patterns: substitution effects in routine employment; complementarity effects with expert employment expansion; wage compression from compositional shifts within rigid pay structures; and political equilibrium conditions affecting organisational responses. These demonstrate how institutional constraints shape workforce composition in ways serving democratic governance purposes beyond narrow efficiency.

The findings contribute to labour economics (how institutional constraints alter technological exposure-employment relationships), public administration (empirical documentation of workforce variation across federal agencies), and institutional economics (how organisational structures mediate technological pressures and workforce evolution).

The 'Literature review' section reviews literature on bureaucratic organisation, technological change, and political economy. The 'AI occupational exposure in federal government: empirical evidence' section presents empirical analysis. The 'A political economy model of AI adoption in bureaucracies' section develops the theoretical framework. The 'Discussion: theoretical insights and policy implications' section discusses insights, policy implications, and extensions. The 'Conclusion' section concludes.

Literature review

This paper contributes to three interconnected bodies of literature: the economics of technological change and labour markets, the political economy of bureaucratic decision-making, and institutional theories of organisational adaptation.

The foundational literature on skill-biased technological change provides theoretical groundwork for understanding AI's differential impact across occupational categories. Autor et al. (2003) show that technological change primarily affects routine cognitive and manual tasks whilst complementing non-routine analytical functions – an insight extended by Acemoglu and Restrepo (2020), who identify displacement effects for routine tasks alongside complementarities for complex problem-solving roles. Our findings of declining routine employment shares and expanding expert positions in AI-exposed agencies align with these predictions. Recent work by Brynjolfsson (2021) emphasises AI's particular capacity for pattern recognition, suggesting its organisational impact may differ from previous automation waves. Webb (2020) and Felten et al. (2021) develop occupation-level AI exposure measures upon which our analysis builds.

Public organisations face institutional constraints that private-sector models do not fully capture. The digital-government literature highlights accountability, legality, and equity concerns around algorithmic systems (Janssen and Kuk 2016); tensions between efficiency and democratic values

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

(Kernaghan, 2014; Wirtz and Müller, 2019); and the complex ways technological pressures affect bureaucratic work (Newman et al., 2022). Makridis (2021) shows that non-pecuniary motives weigh heavily in public-sector employment decisions, reinforcing our model's institutional frictions. These features suggest workforce composition changes in government may be mediated by political considerations rather than pure efficiency calculations (Downs, 1957; Ostrom, 1990).

Our multi-actor theoretical framework draws upon established theories of bureaucratic behaviour. Niskanen's (1975) budget-maximising bureaucrat model provides a starting point, though subsequent scholarship reveals more complex motivational structures. Moe (1984) demonstrates how political principals control bureaucratic agents through institutional design, whilst Epstein and O'Halloran (1999) show how political uncertainty affects bureaucratic discretion. The implementation literature, from Pressman and Wildavsky's (1973) classic study through Lipsky's (2010) street-level bureaucracy concept, illustrates how frontline discretion means workforce composition effects vary depending on how institutional pressures interface with existing practices. Winter (2012) emphasises organisational capacity and political support for successful change, whilst Allison (1971) bureaucratic politics model suggests workforce decisions emerge from bargaining between actors with different professional backgrounds and career incentives.

The public administration literature emphasises distinctive features of government employment. The public service motivation framework demonstrates that public employees are driven by commitment to public interest beyond material compensation (Perry and Wise, 1990; Rainey, 2009), suggesting workforce decisions may be influenced by professional norms absent from market-driven organisations. Persistent challenges in federal workforce management – recruitment difficulties in technical specialties and retention challenges – create pressures for adjustment, whilst civil service rules and political oversight constrain implementation approaches.

Institutional theory offers crucial insights into bureaucratic workforce evolution. DiMaggio and Powell's (1983) institutional isomorphism analysis demonstrates how organisations adopt similar practices due to coercive, mimetic, and normative pressures rather than efficiency alone. March and Olsen's (2010) logic of appropriateness suggests bureaucratic decision-making follows rules of proper behaviour rather than consequentialist calculation. Historical institutionalism provides tools for understanding gradual change: Pierson (1997) shows how initial choices create self-reinforcing dynamics constraining future options, whilst Mahoney et al. (2010) identify mechanisms of incremental adaptation. Carpenter (2001) demonstrates how agencies build capacity through reputation and expertise, and Fukuyama (2014) emphasises merit-based recruitment for effective governance – suggesting workforce patterns depend as much on organisational capabilities as external technological pressures.

The emerging algorithmic governance literature raises fundamental questions about democratic accountability that our model incorporates through political cost functions. Bovens and Zouridis (2002) identify tensions between efficiency and due process in automated government decision-making, whilst Citron (2008) argues that automated systems require procedural safeguards for legitimacy. Danaher (2016) examines how automated decision-making affects human agency and democratic participation. Zuboff's (2019) analysis of surveillance capitalism raises questions about power concentration applicable to government systems, and Eubanks' (2018) ethnographic study demonstrates how algorithmic tools can exacerbate inequalities whilst appearing neutral. The AI ethics literature increasingly emphasises procedural values: Floridi et al. (2018) stress explicability, fairness, and human oversight, whilst Jobin et al. (2019) reveal consensus on principles but variation in implementation.

Despite this substantial literature, significant gaps remain. Most empirical work on AI and employment focuses on private sector contexts where market mechanisms drive decisions. The public sector faces distinctive constraints – civil service protections, standardised pay scales, political oversight – that may fundamentally alter how occupational composition evolves. Existing government digitisation studies focus on service delivery rather than internal workforce patterns; bureaucratic politics literature has not incorporated AI exposure implications; institutional change literature examines policy adaptation rather than administrative workforce composition.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

This paper addresses these gaps through systematic empirical analysis of AIOE and workforce patterns in federal agencies. Our theoretical framework synthesises labour economics, public administration, and institutional theory to develop propositions about how institutional constraints shape employment composition, contributing to broader debates about technological change in institutional contexts.

AI occupational exposure in federal government: empirical evidence

Unlike private firms facing competitive pressures, government agencies must navigate rigid pay structures, civil service protections, political oversight, and institutional mandates, placing legitimacy alongside efficiency – suggesting workforce composition patterns may differ from competitive market observations.

Our approach draws on the AIOE measure from Felten et al. (2021),[1] which assesses job susceptibility to AI by mapping AI capabilities onto workplace tasks. This measure is constructed independently of employment outcomes, focuses on tasks rather than technology, and remains agnostic about whether AI replaces or augments workers. We combine these scores with quarterly employment records from the FedScope database maintained by the Office of Personnel Management,[2] spanning 23 quarters from early 2019 through late 2024.

Agency-level AI exposure is constructed by weighting occupational AIOE scores by employment share. Since underlying scores are fixed over time, observed changes must reflect workforce composition shifts rather than evolving measurement – central to our identification strategy. We supplement this with indicators distinguishing routine administrative work from expert professional roles, and wage structure measures proxying internal pay compression. The routine and expert employment shares derive from administrative classifications defined for personnel management using occupational series, grades, and appointment types – not AI-related criteria – while AI exposure measures derive from external task-capability mappings based on O*NET descriptions. These distinct data sources limit mechanical correlation between exposure and outcomes.

Importantly, our measure captures task susceptibility to automation or augmentation in principle – not whether AI is actively deployed. Recent evidence documents substantial heterogeneity in realised AI adoption conditional on task characteristics (Makridis, 2025). In public-sector settings, institutional constraints, procurement rules, and political oversight may further decouple potential exposure from actual utilisation. Our analysis is therefore intentionally framed in terms of exposure rather than use.

Our empirical strategy examines within-agency changes over time, controlling for fixed agency characteristics and using quarter fixed effects for government-wide developments. The analysis identifies associations between AI exposure changes and workforce composition patterns, though causal interpretation remains limited – occupational composition reflects historical mission requirements and institutional constraints rather than random assignment. Our estimates document systematic patterns consistent with differential adjustment across occupational groups, rather than causal treatment effects. Throughout, changes in agency-level AI exposure are treated as summary statistics for within-agency occupational reallocation rather than measures of AI deployment.

Compositional sources of changes in agency AI exposure

Before examining regression evidence, it is necessary to establish what drives variation in agency-level AI exposure. Our measure is constructed from fixed occupation-level scores developed by Felten et al. (2021)

[1]Our analysis relies on the AI Occupational Exposure (AIOE) measure developed by Felten et al. (2021), though alternative measures exist in the literature (Eloundou et al. 2024). As the index is constructed primarily from private-sector task descriptions, it may overstate automation potential in accountability-intensive public-sector roles. Section 3.5 and Online Appendix E present robustness checks using alternative measures.

[2]Federal employment counts are drawn from the publicly available FedScope Employment Cube maintained by the U.S. Office of Personnel Management (OPM). See the Data Availability statement for access details.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

that do not vary over time. Consequently, any observed change in agency AI exposure must arise entirely from workforce composition shifts – specifically, employment reallocation across occupations with different exposure levels – rather than from AI adoption, technological diffusion, or changes in task execution. The regression analysis therefore, relates employment pattern changes to exposure changes driven by occupational reweighting, documenting systematic workforce reallocation under institutional constraints. It does not, and cannot, identify causal effects of AI technology deployment.

The identification logic underlying this approach is made explicit by the construction of the exposure measure. We calculate agency AI exposure as a share-weighted average of occupational exposure scores:

$$A_{it} = \sum_o w_{iot}\, a_o,\ w_{iot} = \frac{\text{employment}_{iot}}{\sum_{o'} \text{employment}_{io't}},$$

where i indexes agencies, o occupations, and t quarters. The crucial feature is that a_o – the occupation-specific AI exposure score from Felten et al. (2021) – remains fixed over time in our data.[3] This means all changes in an agency's AI exposure must come from staff moving between occupations, not from changes in how we measure AI susceptibility.

We can formalise this by decomposing quarterly changes in exposure:

$$\Delta A_{it} = \underbrace{\sum_o a_o\, \Delta w_{iot}}_{\text{COMP}_{it}} + \underbrace{\sum_o w_{io,t-1}\, \Delta a_{ot}}_{\text{TECH}_{it}} + \underbrace{\sum_o \Delta a_{ot}\, \Delta w_{iot}}_{\text{CROSS}_{it}}$$

Since a_o is constant, $\Delta a_{ot} = 0$, which means $\text{TECH}_{it} = \text{CROSS}_{it} = 0$. The entire change in exposure comes from the composition term:

$$\text{COMP}_{it} = \sum_o a_o\, \Delta w_{iot}.$$

To visualise these patterns, we plot the cumulative composition changes:

$$C_{it} = \sum_{\tau \le t} \text{COMP}_{i\tau},$$

which tracks the net shift of employment weight toward or away from AI-exposed occupations within each agency over time.

Figure 1 shows cumulative changes for the twelve largest agencies, revealing considerable heterogeneity in workforce composition evolution. Several agencies – including the Air Force, Navy, and Veterans Affairs – show steady increases in AI-exposed employment, reflecting sustained shifts toward higher-exposure occupations. Others, like Agriculture and Interior, oscillate around zero with no clear trend, indicating temporary rather than systematic changes. Sharp discontinuities appear in some agencies – the Army's step change in 2020 and Defense's reversal around 2021 – likely reflecting discrete reorganisations or classification changes rather than gradual compositional shifts. Homeland Security, Health and Human Services, and Justice show sustained positive trends, whilst the Social Security Administration exhibits negative trends before partially recovering. These diverse patterns highlight how agencies with different missions, structures, and constraints respond differently to technological pressures.

Substitution effects: routine employment

We begin by examining how within-agency reallocation of employment across occupations with different AI exposure levels is associated with changes in routine employment shares. Because

[3]Occupation-level β exposure scores used in robustness checks were obtained from the official replication repository at https://github.com/openai/GPTs-are-GPTs/tree/main/data.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

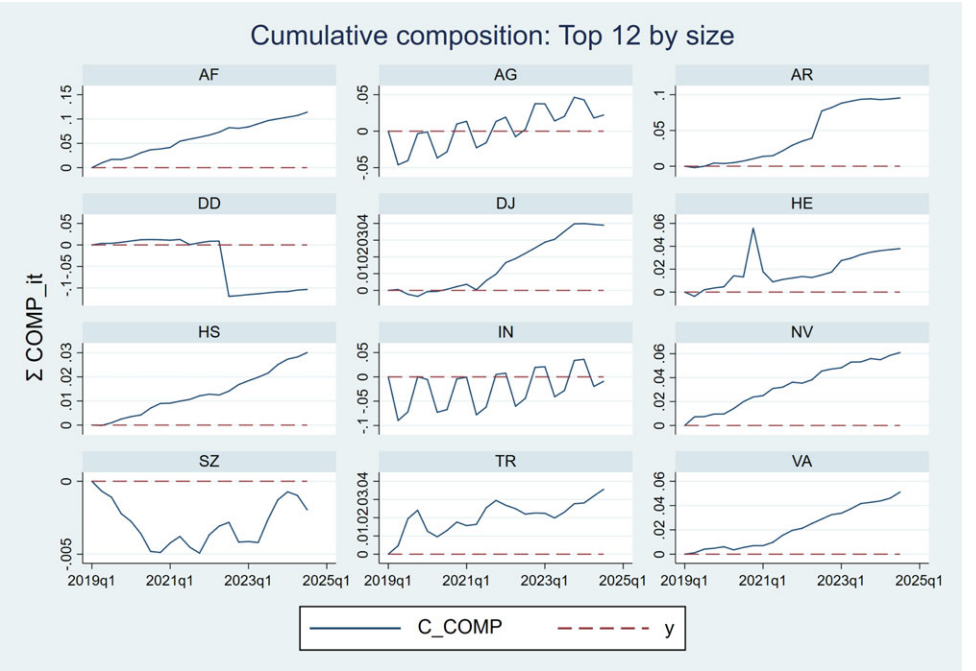


Figure 1. Cumulative composition: top 12 agencies by size. Each panel shows cumulative changes in AI exposure from occupational reweighting: $C_{it} = \sum_{\tau \le t} \sum_{o} a_o \, \Delta w_{iot}$, where a_o is the fixed occupational AIOE score and w_{iot} is the employment share. Positive values indicate net shifts toward more AI-exposed occupations since 2019Q1. Panel-specific scales emphasise within-agency dynamics. Agencies: Air Force (AF), Agriculture (AG), Army (AR), Defense (DD), Justice (DJ), Health and Human Services (HE), Homeland Security (HS), Interior (IN), Navy (NV), Social Security Administration (SZ), Treasury (TR) and Veterans Affairs (VA).

occupation-level exposure scores are fixed by construction, changes in agency-level AI exposure do not reflect technology adoption or diffusion. Instead, they summarise contemporaneous shifts in workforce composition across occupations. The empirical question addressed in this subsection is therefore not whether AI adoption causes changes in routine employment, but how compositional reweighting towards more AI-exposed occupations maps onto changes in the share of routine work within agencies.

Our approach estimates within-agency regressions that relate changes in the routine employment share to changes in agency-level AI exposure. This changes-on-changes specification is appropriate in this setting because both variables are generated by within-agency occupational reallocation, and the regression quantifies how that reallocation loads on routine versus non-routine categories under fixed institutional constraints. The baseline specification is as follows:

$$\Delta R_{it} = \beta_A \Delta A_{it} + \beta_w \Delta \ln\left(\text{salary}_{it}\right) + \beta_s \Delta \ln(\text{headcount}_{it}) + \lambda_t + \varepsilon_{it} \tag{1}$$

where R_{it} is the share of routine work – comprising administrative, clerical, and blue-collar positions – and A_{it} is the artificial intelligence exposure of agency i at quarter-year t. Quarter-year fixed effects λ_t control for period-specific shocks common across agencies, and standard errors are clustered at the agency level.

The results in Table 1 document a consistent negative association between AI exposure changes and routine employment shares. Across all specifications, higher within-agency changes in AI exposure are associated with declines in the routine share: the coefficient on ΔA_{it} is negative and precisely estimated at -0.301 (s.e. 0.030) in the baseline specification, attenuating to -0.198 (s.e. 0.049) as increasingly rich composition controls are added. The magnitude indicates that a one standard deviation increase in quarterly AI exposure (0.0718) is associated with a **1.42 percentage point** decline in routine

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Table 1. Substitution: Δ AIOE on Δ routine share

	(1) Baseline	(2) Δ + Age/education	(3) Δ + Tenure/appointment
ΔA_{it}	− 0.301***	− 0.251***	− 0.198***
	(0.030)	(0.028)	(0.049)
$\Delta\ln(\text{salary}_{it})$	0.131	0.067*-	0.055
	(0.097)	(0.038)	(0.052)
$\Delta\ln(\text{headcount}_{it})$	− 0.005	− 0.004	− 0.021*
	(0.019)	(0.016)	(0.012)
Observations	2703	2703	1670

Notes: The dependent variable is the within-agency quarterly change in the routine employment share, $\Delta R_{it} := R_{it} - R_{i,t-1}$. The key regressor is the change in AI exposure, ΔA_{it} ('AIOE'). All specifications are estimated on first differences with a full set of quarter-year fixed effects $\{\lambda_t\}$; agency time-invariant heterogeneity is removed by differencing. Core controls, included in every column, are $\Delta\ln(\text{salary}_{it})$ (change in average salary) and $\Delta\ln(\text{headcount}_{it})$ (change in reported headcount). Column (2) augments the baseline with vectors of changes in the age and education distributions (entered as changes in category shares; one category omitted per partition). Column (3) further adds vectors of changes in tenure and appointment/scheduling type shares. Compositional-control coefficients are included in the regressions but not reported. Standard errors, reported in parentheses, are clustered at the agency level. Observations are agency-quarter differences; agencies must appear in consecutive quarters to contribute an observation. Sample size declines in column (3) reflect missingness in tenure/appointment distributions. Changes in employment shares are measured in percentage points, while log-differenced variables represent proportional changes. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

employment shares, with this relationship remaining robust across specifications that control for compositional shifts in age, education, and employment arrangements. A 0.10 increase in ΔA_{it} (approximately 1.39 standard deviations) corresponds to a **1.98 percentage point** decline in the routine employment share.

Complementarity effects: expert employment

We next examine how the same within-agency compositional reallocation described in the '*Substitution effects: routine employment*' section.2 maps onto changes in expert employment shares. As before, changes in agency-level AI exposure summarise shifts in occupational employment weights rather than the adoption or diffusion of AI technologies. The empirical question here is whether reweighting towards more AI-exposed tasks is accompanied by expansion of expert roles, which comprise professional and technical positions within agencies.

Using a changes-on-changes specification that mirrors the routine employment analysis, we relate changes in the expert employment share to changes in agency-level AI exposure. The baseline specification is as follows:

$$\Delta\text{Expert}_{it} = \phi\Delta A_{it} + \gamma_w\Delta\ln(\text{salary}_{it}) + \gamma_s\Delta\ln(\text{headcount}_{it}) + \lambda_t + u_{it} \quad (2)$$

Table 2 documents a symmetric pattern for expert roles. Throughout all model specifications, increases in AI exposure exhibit positive associations with increases in the proportion of expert personnel: the coefficient estimate for ΔA_{it} demonstrates both positive magnitude and statistical precision across columns, ranging from 0.282 (s.e. 0.027) in the baseline to 0.208 (s.e. 0.050) with full controls. Using the same standard deviation for ΔA_{it} (0.0718), a one standard deviation increase in exposure is associated with a **1.50 percentage point** rise in the expert share; a 0.10 increase in ΔA_{it} corresponds to a **2.08 percentage point** rise. These patterns are consistent with what task-based models of technological change might predict: agencies with higher AIOE exhibit employment composition changes from administrative and clerical functions toward roles requiring specialised judgment and discretionary authority. In Online Appendix B,[4] we redefine expert definitions using tenure, wage, and

[4]The Online Appendix is available at https://github.com/MelciorRosselloRoig/AI-Adoption-in-Bureaucracies.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Table 2. Complementarity: Δ AIOE on Δ expert share

	(1) Baseline	(2) Δ + Age/Education	(3) Δ + Tenure/Appointment
ΔA_{it}	0.282***	0.245***	0.208***
	(0.027)	(0.025)	(0.050)
$\Delta \ln(\text{salary}_{it})$	− 0.059	− 0.057*	− 0.058
	(0.086)	(0.033)	(0.052)
$\Delta \ln(\text{headcount}_{it})$	0.006	− 0.000	0.019*
	(0.017)	(0.012)	(0.011)
Observations	2703	2703	1670

Notes: Dependent variable: ΔExpert_it (within-agency change in expert employment share). Key regressor: ΔAIOE_it. Estimation and controls as in Table 1. Column (2) adds age/education controls; column (3) adds tenure/appointment controls. Sample size declines in column (3) reflect missingness in tenure/appointment data. Significance: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

combined tenure-wage criteria. Tenure-based experts (professionals with ⩾10 years' service) exhibit stronger complementarity with AI exposure, whilst wage-based definitions (top quartile earners) yield negative coefficients, as high wages also capture managerial roles less amenable to AI augmentation. The combined definition similarly produces negative estimates, confirming that the baseline results reflect genuine technical complementarity rather than wage premia or occupational composition effects.

Wage compression associations

Beyond changes in employment composition, we examine how within-agency compositional adjustment towards more AI-exposed occupations is reflected in internal pay structures. Rather than treating AI exposure as a technology shock, this analysis uses changes in agency-level exposure as a summary statistic for occupational reweighting and asks how such reallocation correlates with shifts in the distribution of pay within agencies.

We measure wage compression using the ratio of median to mean salary within each agency and quarter:

$$\alpha_{it} = \frac{\text{median salary}_{it}}{\text{mean salary}_{it}} \tag{3}$$

where higher values indicate more compressed pay distributions. Our specification examines how this measure responds to AI exposure:

$$\Delta\alpha_{it} = \psi \Delta A_{it} + \theta_s \Delta \ln(\text{headcount}_{it}) + \lambda_t + e_{it} \tag{4}$$

This measure captures *compositional* wage compression arising from employment shifts across grades and occupations; within-grade wage dispersion is not observable in the administrative data. Table 3 shows that increases in AI exposure are associated with greater apparent compositional wage compression within agencies. Higher within-agency changes in AI exposure are associated with *increases* in wage compression: the coefficient on ΔA_{it} is positive and precisely estimated at 0.193 (s.e. 0.009) in the baseline, attenuating to 0.151 (s.e. 0.046) once full composition controls are included. With $\text{sd}(\Delta A_{it}) = 0.0718$, a one standard deviation increase in exposure raises the median-to-mean salary ratio by **0.0108**; a 0.10 increase in ΔA_{it} raises it by **0.0151**.

This pattern is particularly revealing given public sector institutional constraints. Unlike private firms, federal agencies cannot easily adjust individual wages in response to productivity changes,

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Table 3. Wage Compression: Δ AIOE on Δ wage compression

	(1) Baseline	(2) Δ + Age/education	(3) Δ + Tenure/appointment
ΔA_{it}	0.193***	0.193***	0.151***
	(0.009)	(0.016)	(0.046)
$\Delta \ln(\text{headcount}_{it})$	− 0.001	0.005	0.024
	(0.008)	(0.008)	(0.018)
Observations	2703	2703	1670

Notes: Dependent variable: Δα_it where α_it = median(salary_it)/mean(salary_it). Key regressor: ΔAIOE_it. Estimation and controls as in Table 1, except column (1) omits Δln(salary) to avoid denominator mechanics. Column (2) adds age/education; column (3) adds tenure/appointment. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

instead operating within standardised pay scales and civil service frameworks. The wage compression we observe may reflect compositional changes – employment patterns shifting toward occupations clustering at different points in the pay distribution. This aligns with earlier findings: agencies with higher AI exposure exhibit both declining routine employment (positions clustering near grade minimums) and expanding expert employment (higher pay grades). Rather than adjusting individual wages to reflect changing productivity as in private sector automation studies, government agencies operate within classification systems constraining such flexibility. Our wage compression results thus reflect institutional constraints of standardised pay scales combined with employment composition changes across occupations with different pay characteristics.

Robustness and alternative specifications

These patterns hold up well to a range of robustness checks, reported in full in Online Appendix E. Two-way clustering by agency and quarter leaves all three main coefficients significant (Table E1). A lead–lag specification finds near-zero and insignificant lead coefficients while contemporaneous effects retain the expected signs (Table E2), ruling out pre-existing trends. Substitution and complementarity results survive when AI exposure is orthogonalised relative to the routine-occupation mean (Tables E3 and E4) and when the Felten et al. (2021) baseline index is replaced by the Eloundou et al. (2024) exposure measure or a Principal component analysis (PCA)-based composite of both indices (Tables E5 and E6).

These findings contribute to understanding how employment composition relates to occupational characteristics in institutional settings. The systematic associations between AI exposure and employment composition suggest workforce evolution in federal agencies may be more structured than assumed, exhibiting patterns correlating with occupational characteristics even within rigid institutional frameworks. Wage compression effects highlight a key distinction from private sector contexts: whilst firms can adjust individual pay to reflect changing productivity, agencies must operate within classification systems limiting such flexibility – helping explain why employment composition changes appear to be the primary adjustment mechanism in government settings.

Occupational composition within agencies is not random; agencies differ systematically in mission, regulatory environment, and institutional constraints shaping workforce structure. Our results should therefore be interpreted as documenting patterns consistent with workforce reallocation rather than causal treatment effects of AI exposure. By exploiting within-agency changes over time and fixed occupational exposure scores, the analysis isolates compositional adjustment dynamics without claiming exogenous variation in AI adoption.

These patterns are consistent with task-based models of technological change yet operate within institutional environments quite different from competitive markets. Public agencies face political

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

oversight, budget constraints, and legitimacy concerns that may fundamentally alter workforce decisions in ways efficiency-focused models cannot capture. The wage compression suggests standard assumptions about flexible pricing may not apply in bureaucratic settings where pay scales and employment protections constrain adjustment mechanisms. This points toward theoretical frameworks explicitly accounting for political and institutional realities of government organisations, which we develop in the following sections. While AI exposure measures remain imperfect proxies for actual technological deployment, robustness checks suggest core associations are not dependent on specific measurement choices.

A political economy model of AI adoption in bureaucracies

The empirical patterns documented in the 'AI occupational exposure in federal government: empirical evidence' section reveal systematic associations between occupational AI exposure and employment composition changes across agencies. Understanding these patterns requires a deeper analysis of the institutional mechanisms governing AI adoption within government bureaucracies.

This section develops a theoretical framework for understanding how AI adoption proceeds within federal agencies, given their occupational compositions and distinctive institutional constraints. Unlike private firms in competitive markets, government agencies face unique political, legal, and administrative constraints that fundamentally alter how technological integration unfolds. We adopt a public-choice perspective in which agency heads seek to expand organisational scope and budgets whilst minimising political and reputational costs. Our model recognises that AI implementation emerges from strategic interactions between agency directors maximising institutional scope, political overseers balancing service delivery against electoral costs, and civil service institutions constraining adjustment mechanisms through employment protections and rigid pay scales.

Agencies begin with vastly different occupational compositions – some heavily weighted toward routine, AI-exposed tasks, others dominated by expert roles requiring human judgement. These initial conditions fundamentally shape how technological integration proceeds and what employment transitions to expect. We model AI adoption as ongoing institutional adaptation through strategic interactions between key actors (Paniagua and Rayamajhee, 2022). The Agency Director maximises institutional scope subject to performance requirements and political constraints, following the public choice tradition recognising bureaucrats as rational actors pursuing organisational objectives (Niskanen, 1968; Downs, 1965). The Political Overseer represents elected officials balancing technological modernisation against fiscal costs, electoral consequences, and interest group resistance. Civil Service Institutions operate through constraint functions reflecting employment protections, rigid pay scales, and procurement limitations, shaping feasible organisational responses.

This institutional environment fundamentally shapes AI adoption. Pay scale rigidity channels adjustment toward employment composition changes rather than individual wage adjustments, generating observed wage compression patterns. Employment protections prevent immediate workforce reductions characterising private sector automation, directing agencies toward complementarity strategies where AI augments rather than substitutes for protected workers. Political oversight subjects adoption decisions to legislative scrutiny, interest group pressure, and electoral timing, with no parallel in market-driven organisations.

The model provides frameworks for predicting which agencies face the greatest implementation challenges, expected employment transitions, and how institutional design shapes adoption outcomes. Rather than treating technological change as exogenous, we model AI adoption as emerging from complex bargaining between actors with divergent objectives, generating equilibrium outcomes that may deviate from narrow efficiency criteria to serve broader democratic governance purposes. The framework generates predictions about workforce composition evolution as AI capabilities mature, provides tools for understanding cross-agency heterogeneity and the role of institutional parameters, and examines welfare implications of alternative governance designs for managing technological change within democratic institutions.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Formal model: actors and objectives

We now formalise the strategic interactions that will govern AI adoption within government agencies as technological integration accelerates, focusing on the distinctive institutional features that differentiate public sector implementation from private market dynamics. Our framework centres on two key actors whose competing objectives and institutional constraints will shape how agencies with different occupational compositions navigate the adoption process.

The model features an *Agency Director* – typically a senior civil servant or department head who serves as the primary administrative decision-maker – and a *Political Overseer* – representing elected officials or ministers who provide strategic direction whilst remaining sensitive to electoral consequences. The Director chooses employment levels at two organisational layers (routine workers N_L and expert staff N_H) alongside AI capital stock A as technological capabilities become available. Agency output $Q(N_L, N_H, A)$ emerges from the evolving interaction of human labour and artificial intelligence, whilst costs $C(N_L, N_H, A)$ reflect both personnel expenses and technology investments.

The production function specifications directly reflect how different occupational compositions will interact with AI capabilities. Agencies with higher concentrations of routine, AI-exposed occupations face different technological opportunities and constraints than those dominated by expert roles requiring human judgement. This occupational starting point becomes a crucial state variable determining how adoption can proceed and what employment transitions will follow.

What distinguishes this public sector context from private enterprise is the presence of political costs – captured through functions $P(\cdot)$ and $R(\cdot)$ below – that will arise from legislative scrutiny, interest group opposition, and public concerns about algorithmic decision-making in government. These costs have no direct parallel in commercial settings but will prove central to determining how AI adoption unfolds across different agencies and policy domains.

The agency director

The Agency Director faces a complex optimisation problem that balances institutional survival, operational effectiveness, and political acceptability. Unlike private sector executives who can focus primarily on profit maximisation, public sector administrators must navigate competing objectives that frequently conflict:

$$U^D = B(N_L + N_H + \theta A) + \alpha\, Q(N_L, N_H, A) - C(N_L, N_H, A) - P(A) \tag{5}$$

The first component, $B(N_L + N_H + \theta A)$, reflects budget maximisation incentives well-established in public choice theory (Niskanen, 1968; Downs, 1965). Bureaucrats possess strong incentives to maintain or expand their institutional scope, as larger agencies command greater resources, provide enhanced career opportunities, and offer increased influence within government. The parameter $\theta \geq 0$ plays a crucial role, determining whether AI systems contribute to budget justification arguments. When $\theta > 0$, AI capital enhances the director's case for resource allocation, perhaps because visible technology investment signals institutional modernisation to political overseers and budget authorities. When $\theta = 0$, AI provides no such benefits, it is potentially because budget authorities focus exclusively on headcount metrics.

The function $B(\cdot)$ captures the political and budgetary value to the Agency Director of expanding organisational scope, rather than productive returns to scale. In the U.S. federal context, there are strong institutional reasons to expect this value to exhibit diminishing marginal returns. As agencies grow larger, incremental expansions typically attract disproportionately greater oversight, reporting requirements, and audit intensity, reducing the net marginal political benefit of further expansion. Moreover, appropriations and staffing ceilings are negotiated through incremental political processes in which marginal increases become harder to justify once the baseline scale is already large. The assumption $B'' \leq 0$ therefore, reflects declining marginal appropriability of budgetary rents as organisational size expands, rather than technological scale effects, and serves as a reduced-form representation of increasing oversight and political constraint at scale.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

The performance component $\alpha Q(N_L, N_H, A)$ captures the director's concern for operational effectiveness and service delivery quality, where $\alpha > 0$ weights performance relative to other objectives. This reflects genuine pressures to deliver effective public services arising from political oversight, professional norms within the civil service, media scrutiny of agency performance, and intrinsic motivation among public sector leaders.

Personnel and technology costs follow $C(N_L, N_H, A)$, but crucially, wages in the civil service are determined by standardised pay scales and collective bargaining agreements rather than marginal productivity. This creates fundamental rigidities absent from private labour markets, with implications we explore below.

Finally, $P(A)$ represents political and reputational costs unique to government contexts: legislative scrutiny over technology procurement decisions, media attention on algorithmic decision-making, potential public backlash over concerns about fairness and transparency, and compliance with government-wide AI governance frameworks. This function is assumed to be increasing and convex, reflecting that political opposition accelerates as automation scope and visibility expand.

The political overseer

Political Overseers – elected officials, political appointees, or senior ministers – face fundamentally different trade-offs from Agency Directors. Their primary concerns centre on electoral consequences, fiscal responsibility, and broader policy outcomes rather than institutional survival and bureaucratic autonomy:

$$U^P = \beta\, Q(N_L, N_H, A) - \gamma\, C(N_L, N_H, A) - R(A) - L(\Delta N) \tag{6}$$

The first term $\beta Q(N_L, N_H, A)$ captures electoral benefits from improved service delivery, where $\beta > 0$ reflects political rewards that accrue when politicians demonstrate tangible improvements in government performance. These may materialise through increased voter satisfaction, positive media coverage, or competitive advantage relative to political opponents.

The cost component $\gamma C(N_L, N_H, A)$ reflects taxpayer pressure for fiscal efficiency and legislative constraints on public expenditure, where $\gamma > 0$ captures sensitivity to fiscal costs that create direct political consequences for spending decisions.

The resistance function $R(A)$ represents organised opposition to automation from various interest groups: public sector unions concerned about employment security, professional associations worried about erosion of expert judgment, and civil liberties groups raising concerns about algorithmic fairness. We assume $R'(A), R''(A) \geqslant 0$, with the convex specification reflecting that opposition intensifies at an increasing rate as automation scope expands.

The final component $L(\Delta N)$ captures political costs associated with visible workforce changes, where $\Delta N = \left|N_L - N_L^0\right| + \left|N_H - N_H^0\right|$ measures absolute deviation from baseline employment levels. Government employment levels carry symbolic significance about political priorities, creating costs for both increases (suggesting fiscal irresponsibility) and decreases (raising concerns about service reduction).

Institutional constraints

Three key institutional features distinguish public sector technology adoption from private markets and prove central to explaining observed patterns:

Pay-scale rigidity. Unlike private enterprises, where wages can adjust to reflect changing marginal productivities, civil service compensation follows standardised grade classifications that create significant rigidities. Personnel costs are increasing and convex in headcount:

$$C(N_L, N_H, A) = \bar{w}_L N_L^{1+\eta_L} + \bar{w}_H N_H^{1+\eta_H} + w_A A, \quad \eta_L, \eta_H \geq 0,$$

where the convexity parameters $\eta_j \geqslant 0$ capture step increases within salary grades, promotion bottlenecks as agency size grows, and the tendency for larger organisations to employ more senior

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

personnel within each category. When $\eta_j = 0$, we obtain the linear case; when $\eta_j > 0$, grade progression effects emerge.

Employment protection. *Civil service rules, union agreements, and political commitments create employment floors that limit workforce flexibility:*

$$N_j \geq \underline{N}_j, \quad j \in \{L, H\} \tag{7}$$

When these constraints bind during AI adoption, agencies cannot substitute technology for labour directly but must instead pursue complementarity strategies where AI augments rather than replaces protected workers.

Procurement and adoption ceilings. Government procurement processes, security requirements, and institutional capacity constraints limit the speed and scope of AI deployment:

$$A \leq A_{\max}(t) \tag{8}$$

where $A_{\max}(t)$, and it grows over time as institutions develop necessary infrastructure, technical expertise, and governance frameworks to support larger-scale automation.

Production technology

The production function directly reflects our empirical findings, incorporating both substitution effects for routine tasks and complementarity effects for expert functions:

$$Q(N_L, N_H, A) = F(T_L(N_L, A),\ T_H(N_H, A)), \quad F_1, F_2 > 0,\ F_{11}, F_{22} \leq 0 \tag{9}$$

where output aggregates task completion at two organisational levels.

Routine tasks. Administrative processing, data entry, and rule-based decision-making exhibit perfect substitutability between human labour and AI systems:

$$T_L(N_L, A) = \min\left[\bar{T}_L,\ \phi_L N_L + \phi_A A\right], \quad \phi_L, \phi_A > 0 \tag{10}$$

This specification captures that AI can directly replace human effort in standardised functions up to the agency's capacity requirement $\bar{T}_L$, but some minimum human oversight remains necessary for handling exceptions and ensuring accountability.

Expert tasks. Policy analysis, legal interpretation, and complex case management exhibit complementarity rather than substitution:

$$T_H(N_H, A) = \phi_H N_H (1 + \sigma A), \quad \phi_H > 0,\ \sigma > 0 \tag{11}$$

The multiplicative form ensures that AI systems enhance rather than replace expert judgement, with $\sigma > 0$ capturing augmentation intensity. This reflects that whilst AI can rapidly process information and identify patterns, the interpretation and application require human insight that incorporates legal requirements, policy objectives, and contextual factors difficult to codify algorithmically.

Two-stage decision process

Decision-making occurs through a sequential process reflecting the hierarchical nature of government, whilst recognising substantial administrative discretion in implementation.

Stage 1: Political negotiation. The Political Overseer establishes the strategic framework within which the Agency Director operates, choosing constraints to maximise their utility function:

$$\max_{A^{\max}, \underline{N}_L, \underline{N}_H, \bar{B}} U^P \quad \text{subject to fiscal and political feasibility constraints} \tag{12}$$

The Overseer sets the maximum AI adoption level $A^{\max} \leq A_{\max}(t)$, potentially below the technical maximum due to political risk considerations; employment floor adjustments $\{\underline{N}_L, \underline{N}_H\}$ that may exceed baseline levels to provide workforce protection; and overall budget allocation $\bar{B}$ determining the resource envelope.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Stage 2: Agency response. Given political constraints from Stage 1, the Agency Director chooses optimal employment levels and AI adoption to maximise their utility:

$$\max_{N_L, N_H, A} U^D \text{ subject to budget, employment floors, and } A \leq A^{\max} \tag{13}$$

This sequential structure ensures that AI implementation decisions emerge from political-administrative bargaining rather than purely technocratic optimisation, whilst preserving meaningful administrative discretion in operational choices.

Analytical framework and equilibrium properties

To derive clear theoretical predictions while preserving the essential institutional features that distinguish public sector technology adoption, we impose several assumptions that ensure analytical tractability without compromising the model's core insights.

The production technology assumes separability between routine and expert task completion: $F(T_L, T_H) = f(T_L) + g(T_H)$ where both $f(\cdot)$ and $g(\cdot)$ exhibit positive but diminishing marginal returns ($f', g' > 0$ and $f'', g'' \leq 0$). This separability reflects the institutional reality that routine administrative functions and expert professional work operate through distinct processes within government agencies, with limited scope for substitution between task categories even when both contribute to overall agency output.

Budget benefits follow a concave specification $B'' \leq 0$, capturing diminishing political returns to agency size. While bureaucrats prefer larger budgets, the marginal political value of additional resources declines as agencies grow, reflecting legislative oversight, media scrutiny, and diminishing returns to empire-building. Personnel and technology costs $C(N_L, N_H, A)$ are increasing and convex in all arguments, consistent with civil service pay scales, where hiring additional staff within each category becomes increasingly expensive due to grade progression effects and promotional bottlenecks.

Political cost functions satisfy $P'(A), P''(A) \geqslant 0$ and $R'(A), R''(A) \geqslant 0$, reflecting the reality that political opposition to automation intensifies at an accelerating rate as deployment scope expands. Initial AI adoption may generate minimal political resistance, but large-scale automation triggers coordinated opposition from unions, professional associations, and advocacy groups concerned about employment displacement and algorithmic accountability.

These assumptions generate well-behaved optimisation problems that admit interior solutions where constraints do not bind, whilst preserving meaningful corner solutions when institutional limitations become active. The framework permits transparent comparative statics analysis whilst maintaining sufficient generality to capture the diverse institutional arrangements observed across different agencies and policy domains.

The equilibrium characterisation proceeds through backward induction. In Stage 1, the Political Overseer anticipates the Agency Director's optimal response to any set of constraints and chooses constraint parameters to maximise political utility. In Stage 2, the Director takes these constraints as given and optimises operational choices subject to the imposed limitations. This sequential structure ensures that AI adoption decisions emerge from strategic political-administrative bargaining rather than technocratic optimisation, generating predictions that align with the institutional realities of government technology procurement and deployment.

Results and testable implications

The framework generates four key propositions that correspond directly to the empirical patterns documented in the 'AI occupational exposure in federal government: empirical evidence' section sec: empirical. Formal proofs appear in Online Appendix A.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Proposition 1. (Substitution under flexible employment). *If routine employment floors are slack ($N_L > \underline{N}_L$) and the routine capacity constraint in 10 is not binding, optimal adjustment to AI adoption features $dN_L dA < 0$. See Online Appendix A for a full proof.*

This result captures the substitution effect observed in agencies with sufficient flexibility to adjust workforce composition in response to technological change. When institutional constraints do not prevent workforce adjustment, the perfect substitutability between AI systems and routine labour generates optimal reductions in human employment as automation capacity expands.

Proposition 2. (Complementarity and expert expansion). *If budget constraints permit and expert employment floors are slack ($N_H > \underline{N}_H$), AI adoption raises expert employment: $dN_H dA > 0$. See Online Appendix A for a full proof.*

This proposition reflects the complementary relationship between AI systems and expert human judgement in high-skill government functions. When agencies deploy AI for data analysis or pattern recognition, these tools amplify rather than replace expert capabilities. A policy analyst equipped with AI can process more background research, identify relevant precedents more quickly, and focus on interpretive work requiring human judgement about political feasibility, legal implications, and stakeholder concerns. As AI augmentation enhances expert productivity, agencies find it optimal to expand expert employment to exploit these enhanced analytical and decision-making capabilities. This complementarity specification reflects recent insights that while AI may reduce information processing costs, it does not resolve fundamental coordination challenges in complex organisations (Davidson, 2024).

The model's predictions become more complex when institutional constraints bind. When agencies have exhausted their capacity for routine task completion – meaning they have processed all the administrative work they can handle – they face a binding capacity constraint. At this point, any additional AI deployment must be perfectly offset by reductions in routine staff to maintain the same level of administrative output. This creates a direct trade-off: $\phi_L \, dN_L + \phi_A \, dA = 0$, where each unit of AI must replace an equivalent amount of human routine labour.

However, this substitution process stops when routine employment hits its protected floor – the minimum staffing level guaranteed by union agreements, civil service rules, or political commitments. Once agencies reach this employment floor, they cannot reduce routine staff further, regardless of how much AI capacity they add. At this point, the adjustment mechanism shifts entirely toward the complementarity margin. Further technology adoption must work through expert employment expansion and enhanced AI-expert collaboration rather than through routine labour substitution. This institutional reality helps explain why some agencies exhibit pure complementarity patterns even in domains where substitution might be technically feasible.

For wage compression analysis, let the agency-wide wage distribution be the mixture $G = s_L F_L + s_H F_H$ where $s_j = N_j/(N_L + N_H)$ represents employment shares and F_j denotes the wage distribution within skill category j. Define the compression index as $\alpha = \text{median}(G)/\mathbb{E}[G]$.

Proposition 3. (Wage compression under civil service constraints). *Suppose the expert wage distribution F_H first-order stochastically dominates the routine distribution F_L with weakly higher mean and median, the routine employment floor binds ($N_L = \underline{N}_L$), the expert floor remains slack, and the mixture median lies within the support of F_L. Along with the optimal response to AI adoption, the compression index increases: $d\alpha dA > 0$. See Online Appendix A for a full proof.*

This result explains the wage compression effects documented in our empirical analysis. When employment protection prevents routine workforce adjustment but allows expert employment expansion, AI adoption leads to workforce composition shifts toward higher-skilled positions, increasing the median-to-mean wage ratio.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Proposition 4. (Political equilibrium and adoption timing). *In Stage 1, Political Overseers permit positive AI adoption ($A^{\max} > 0$) if and only if marginal political benefits exceed marginal political costs at baseline conditions: $\beta\, Q_A - \gamma\, C_A > R'(0) + L'(0)$, subject to feasibility constraint $A^{\max} \leq A_{\max}(t)$ and anticipating Stage 2 administrative responses. (Implications: pre-election periods with higher L' delay adoption; fiscal stringency encourages adoption when $C_A < 0$, but discourages when $C_A > 0$; stronger interest group mobilisation raises R', dampening adoption.) See Online Appendix A for a full proof.*

This proposition characterises conditions under which Political Overseers provide necessary authorisation for AI adoption, linking adoption timing to electoral cycles, fiscal conditions, and interest group mobilisation.

The framework generates rich predictions about variation in AI adoption patterns across agencies and over time. Cross-agency responses vary with institutional parameters: pay rigidity coefficients (η_L, η_H) determine the strength of quantity versus price adjustments; budget salience of AI technology (θ) affects adoption incentives and resource reallocation; and augmentation intensity (σ) influences the magnitude of expert employment expansion.

Agencies operating under steeper pay schedules (higher η_j) should exhibit stronger employment quantity adjustments in response to AI adoption. Higher values of θ should accelerate both adoption and employment reallocation, whilst larger σ parameters amplify expert employment expansion. Additionally, adoption timing should co-vary with electoral cycles and budget calendar processes, and adjustment mechanisms differ systematically when employment protection floors bind asymmetrically across skill categories.

Taken together, Propositions 1–4 rationalise the substitution, complementarity, and wage compression patterns documented in our empirical analysis whilst delivering sharp, testable predictions about heterogeneity across agencies and variation in adoption timing. The theoretical framework demonstrates how institutional constraints and political considerations shape technology adoption in ways that may appear suboptimal from narrow efficiency perspectives but serve important broader purposes in democratic governance and public administration.

Comparative statics: institutional parameters and augmentation strength

This subsection summarises the directional comparative statics implied by the model with respect to key institutional parameters and the technological augmentation parameter σ. The objective is not to derive closed-form solutions, but to clarify how equilibrium outcomes respond qualitatively to changes in (i) employment floor tightness and pay-scale rigidity, and (ii) the strength of AI–expert complementarity. The signs reported below follow directly from the Kuhn–Tucker characterisation of the Agency Director's constrained optimisation problem and from the structure of the production technology.

Employment floor tightness. Consider first marginal increases in employment protection, captured by higher employment floors $\underline{N}_j$ for $j \in \{L, H\}$. When a given employment floor is slack at the optimum, the associated multiplier is zero, and marginal tightening has no first-order effect on optimal choices. When the constraint binds, tightening it shifts the feasible set mechanically.

In particular, if the routine employment floor binds ($\mathrm{N}_{\mathrm{L}}^{*} = \underline{N}_{\mathrm{L}}$), a marginal increase in $\underline{N}_{\mathrm{L}}$ raises routine employment one-for-one, $\partial \mathrm{N}_{\mathrm{L}}^{*} / \partial \underline{N}_{\mathrm{L}} > 0$. Because routine tasks are locally substitutable with AI capital, tighter routine employment protection reduces the scope for AI substitution and weakly lowers optimal AI adoption, $\partial \mathrm{A}^{*} / \partial \underline{N}_{\mathrm{L}} \leq 0$. Through the resulting budget and cost trade-offs, tighter routine floors also weakly reduce expert employment, $\partial \mathrm{N}_{\mathrm{H}}^{*} / \partial \underline{N}_{\mathrm{L}} \leq 0$ unless offset by sufficiently strong budgetary incentives.

Analogously, if the expert employment floor binds ($\mathrm{N}_{\mathrm{H}}^{*} = \underline{N}_{\mathrm{H}}$), tightening expert protection raises expert employment mechanically while limiting the extent to which AI can be deployed for augmentation purposes. In this case, the effect on routine employment is ambiguous and depends on whether the Director reallocates resources toward routine labour or contracts overall organisational scale.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Pay-scale rigidity. Pay-scale rigidity enters through the convexity parameters η_L and η_H, which govern how marginal personnel costs rise with employment. An increase in η_j steepens the marginal cost of expanding employment in category j.

Higher routine pay rigidity ($\eta_L \uparrow$) weakly reduces optimal routine employment, $\partial N_L^*/\partial \eta_L \leq 0$, and – when substitution is feasible – raises the relative attractiveness of AI capital as a substitute input, $\partial A^*/\partial \eta_L \geqslant 0$. The effect on expert employment is in general ambiguous, reflecting opposing forces: higher routine costs may reallocate resources toward expert–AI complementarity, but increased overall cost pressure may also contract total scale.

Higher expert pay rigidity ($\eta_H \uparrow$) weakly reduces expert employment, $\partial N_H^*/\partial \eta_H \leq 0$. Because AI capital primarily generates value through complementarity with expert labour in the production technology, steeper expert pay schedules also dampen AI adoption incentives, yielding $\partial A^*/\partial \eta_H \leq 0$ when the expert margin is operative.

Augmentation strength σ. Finally, consider changes in the augmentation parameter σ in the expert task technology $T_H(N_H, A) = \phi_H N_H(1 + \sigma A)$. This specification exhibits increasing differences in (N_H, A) and σ, implying that higher σ raises the marginal productivity of expert labour when AI is deployed and raises the marginal return to AI when expert employment is positive.

Provided the expert employment floor is slack and the AI adoption ceiling does not bind, stronger augmentation unambiguously increases expert employment, $\partial N_H^*/\partial \sigma > 0$, and weakly increases AI adoption, $\partial A^*/\partial \sigma \geqslant 0$. When routine employment is adjustable, the resulting increase in AI adoption further implies weak substitution away from routine labour, $\partial N_L^*/\partial \sigma \leq 0$. If employment floors or adoption ceilings bind, these effects are locally muted, with the corresponding derivatives equal to zero.

Taken together, these comparative statics highlight how institutional parameters govern the margins along which bureaucratic adjustment occurs. Employment protection and pay rigidity shift responses away from direct substitution and toward constrained reallocation and complementarity, while stronger augmentation amplifies expert expansion where institutional constraints permit. These mechanisms distinguish bureaucratic AI adoption from market-driven adjustment and motivate the empirical focus on compositional change.

Discussion: theoretical insights and policy implications

Our theoretical framework generates insights extending beyond conventional technology adoption models, particularly regarding the distinctive institutional environment within which public agencies operate.

The model predicts that agencies may expand expert employment even when AI could theoretically substitute for some expert tasks – appearing inefficient under standard economic reasoning but emerging naturally from institutional constraints. This result arises from the interaction between civil service employment protections and political budget dynamics. When routine employment floors bind due to union agreements or political commitments ($N_L = \underline{N_L}$ in equation 7), agencies cannot achieve optimal workforce reductions and must pursue alternative adjustment mechanisms. The complementarity specification in equation 11 ensures AI adoption increases expert workers' marginal productivity through the augmentation parameter $\sigma > 0$, making expert employment expansion attractive even when direct substitution might be technically feasible.

Consider tax administration agencies deploying AI for document processing and fraud detection. While such systems could potentially automate complex case analysis performed by tax attorneys and senior auditors, our model predicts agencies will instead use AI to augment these experts' capabilities – enabling more case processing, subtler pattern identification, and focus on higher-value interpretive work. This contradicts standard efficiency models suggesting substitution wherever technically feasible, but aligns with institutional realities where agencies must maintain political legitimacy and operate within civil service constraints, making wholesale workforce substitution infeasible. From the Agency Director's utility function in equation 5, the budget component $B(N_L + N_H + \theta A)$ creates incentives to maintain employment levels, while the political cost function $P(A)$ penalises visible workforce displacement.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

The model also predicts asymmetric responses to fiscal pressure, differing from private sector patterns. During budget constraints, agencies facing binding employment floors may accelerate AI adoption as a politically acceptable substitute for workforce expansion – 'defensive complementarity' emerging when political constraints force agencies toward second-best solutions combining protected human employment with augmentative technology.

The framework yields concrete insights for optimal institutional design. Employment protections improve AI adoption outcomes under several conditions. First, when AI capabilities remain incomplete or unreliable, employment floors prevent premature workforce reductions, compromising service quality. The constraint $N_j \geqslant \underline{N_j}$ in equation 7 forces agencies toward complementarity regimes where human workers provide essential oversight, error correction, and exception handling. The production specification in equations 10 and 11 reflects this: routine tasks require minimum human oversight through capacity constraint $\overline{T_L}$, while expert tasks exhibit multiplicative complementarity, preserving human judgment. This generates more robust service delivery than pure substitution, particularly in high-stakes domains like benefits administration or regulatory enforcement, where algorithmic errors carry significant social costs.

Second, employment protections can improve political sustainability by reducing interest group resistance. From equation 6, the Political Overseer's resistance function $R(A)$ decreases when workforce displacement concerns are mitigated through job security guarantees – a political economy effect potentially outweighing efficiency losses, particularly when adoption decisions require legislative approval or face union opposition.

However, employment protections hinder efficient adoption when preventing necessary workforce reallocation toward higher-value activities. Binding employment floors at routine skill levels can prevent agencies from exploiting AI's full productivity potential, forcing excessive investment in complementary technology whilst maintaining suboptimal employment in automatable tasks.

Our analysis suggests employment protections should be designed with task-based rather than occupation-based criteria – protecting workers performing essential human oversight whilst allowing reallocation away from purely routine activities. The model supports graduated protection schemes where employment floors decrease as AI capabilities mature:

$$\underline{N_{j,t}} = \underline{N_{j,0}} exp\left(-\phi_j \sum_{s=0}^{t-1} 1\left[\text{successful adoption}\right]\right) \tag{14}$$

where protection levels decline with accumulated successful implementations, with decline rates $\phi_j \geqslant 0$ potentially differing across skill categories. Understanding transition dynamics requires extending the analysis to multiple periods, addressing how learning and reputation effects shape adoption over time.

The static framework captures contemporaneous adjustment mechanisms but abstracts from how political-administrative bargaining evolves over time. Online Appendix D extends the model to incorporate dynamic learning and strategic adaptation across multiple periods.[5] This extension formalises how early adoption experiences shape subsequent beliefs, political tolerance, and investment decisions.

The dynamic framework generates path dependence in AI adoption: early implementation failures reduce long-run AI stock through persistent belief updating (Proposition 5, proved in Online Appendix D). Agency Directors update beliefs about Political Overseer responsiveness based on approval rates, whilst Overseers update beliefs about AI effectiveness based on observed performance. These learning mechanisms create adoption momentum once initial hurdles are overcome.

This dynamic extension provides a theoretical foundation for learning subsidies to overcome initial resistance and explains why institutional memory and reputation effects facilitate subsequent adoption.

[5]The dynamic extension includes formal proof of Proposition 5 on path dependence. While theoretically important, the static framework in the 'A political economy model of AI adoption in bureaucracies' section captures the core mechanisms underlying our empirical patterns.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

The gradual, history-dependent adjustment we document empirically aligns with these dynamic learning mechanisms.

Conclusion

This paper has examined AIOE and workforce patterns within federal government agencies through empirical analysis and theoretical modelling, documenting systematic associations distinguishing public sector dynamics from private market patterns. Agencies with higher concentrations of AI-exposed occupations exhibit employment composition shifts from routine administrative roles towards expert professional positions whilst experiencing wage compression – patterns emerging from the government's distinctive institutional environment.

The empirical analysis establishes three principal findings. Federal agencies with higher AIOE systematically exhibit declining routine employment shares, consistent with substitution effects operating through gradual workforce adjustment rather than immediate displacement. These agencies show expanding expert employment, suggesting complementarity between analytical capabilities and high-skilled work within government hierarchies. Wage compression effects reflect compositional shifts towards higher-skilled occupations within rigid civil service pay structures, providing novel evidence of how institutional wage constraints shape workforce evolution.

Our theoretical framework bridges labour economics, public administration, and institutional economics by exploring how political constraints alter the relationship between technological exposure and workforce composition in ways serving broader social purposes beyond efficiency maximisation. The multi-actor model extends task-based technological change models to incorporate budget-maximising bureaucrats, electoral considerations of political overseers, and civil service institutions constraining adjustment mechanisms – generating equilibrium outcomes deviating substantially from narrow efficiency criteria. By contrast, existing private-sector studies – such as Acemoglu and Restrepo (2020) showing industrial robots reducing employment and wages in competitive markets – highlight how technological change produces both substitution and task reallocation under different institutional conditions.

The systematic patterns suggest agencies with greater AIOE experience primarily reallocation rather than displacement within current institutional frameworks – likely reflecting employment protections preventing immediate workforce reductions whilst allowing compositional change through natural turnover. Such gradual adjustment may prove more politically sustainable than rapid market displacement, though it implies potential challenges for long-term talent retention as technological pressures increase the market value of analytical skills within rigid pay scales.

Our theoretical analysis indicates optimal workforce composition decisions may diverge from narrow cost minimisation due to political resistance, employment disruption, and legitimacy concerns, creating costs not fully internalised by individual organisations. This supports coordinated approaches to understanding workforce evolution whilst acknowledging implementation challenges within federal systems characterised by bureaucratic politics and agency autonomy.

Several methodological considerations qualify our findings. Our analysis relies on occupational exposure measures rather than direct observation of technological implementation, capturing potential vulnerability rather than actual deployment. The within-agency identification strategy cannot definitively establish that observed patterns reflect deliberate responses versus other organisational developments. Future research should test cross-agency heterogeneity predictions and explore theoretical mechanisms through detailed case studies. Cross-national comparison could identify how different administrative traditions affect relationships between occupational composition and workforce evolution. The rapidly evolving nature of AI technologies – potentially expanding beyond task-based applications towards general-purpose systems – suggests institutional constraints and political dynamics may themselves evolve, requiring theoretical refinement.

As AI advances, the institutional constraints we document – often characterised as inefficiencies – may prove essential for managing workforce transitions whilst preserving democratic accountability.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

The challenge lies not in eliminating such constraints but in designing institutions that balance technological adaptation with procedural values essential to legitimate public administration. Our findings provide analytical foundations for understanding how institutional arrangements shape workforce composition, though sustained research attention will be required to inform effective policy responses. Future research might explore novel governance mechanisms, including polycentric approaches to institutional adaptation (Makridis and Ammons, 2025).

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S1744137426100496

Data availability statement. The analysis in this article is based on publicly accessible datasets. Federal employment data were drawn from the FedScope Employment Cube, available via the U.S. Office of Personnel Management: https://www.opm.gov/data/datasets/. To construct the AI exposure index at the federal occupational level, we used the occupation–AI exposure mappings provided by Felten et al. (2021), 'Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and Its Potential Uses', accessible at: https://github.com/AIOE-Data/AIOE.https://github.com/AIOE-Data/AIOE Linking federal occupational categories to standard SOC codes required a crosswalk, which was sourced from the EEOC Federal Sector Occupation Cross-Classification Table, available at: https://www.eeoc.gov/federal-sector/management-directive/eeoc-federal-sector-occupation-cross-classification-table.

Conflict interests. The authors declare that they have no conflict of interest.

Funding. The author received no financial support for the research, authorship, or publication of this article.

References

Acemoglu D. and Restrepo P. (2020). Robots and jobs: evidence from us labor markets. *Journal of Political Economy* **128**(6), 2188–2244.

Allison G.T. (1971). *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown and Company.

Autor D.H., Levy F., Murnane R.J. (2003). The skill content of recent technological change: an empirical exploration. *The Quarterly Journal of Economics* **118**(4), 1279–1333.

Bovens M. and Zouridis S. (2002). From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review* **62**(2), 174–184.

Brynjolfsson E. (2021). The turing trap: the promise & peril of human-like artificial intelligence. *Daedalus* **150**(2): 272–287.

Carpenter D.P. (2001). *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862–1928*. Princeton, NJ: Princeton University Press.

Citron D.K. (2008). Technological due process. *Washington University Law Review* **85**, 1249.

Danaher J. (2016). The threat of algocracy: reality, resistance and accommodation. *Philosophy & Technology* **29**(3), 245–268.

Davidson S. (2024). The economic institutions of artificial intelligence. *Journal of Institutional Economics* **20**, e20.

DiMaggio P.J. and Powell W.W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* **48**(2), 147–160.

Downs A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy* **65**(2), 135–150.

Downs A. (1965). A theory of bureaucracy. *The American Economic Review* **55**(1/2), 439–446.

Eloundou T., Manning S., Mishkin P. and Rock D. (2024). GPTs are GPTs: labor market impact potential of LLMs. *Science* **384**(6702), 1306–1308.

Epstein D. and O'Halloran S. (1999). *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers*. Cambridge, UK: Cambridge University Press.

Eubanks V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY, USA: St. Martin's Press.

Felten E., Raj M. and Seamans R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: a novel dataset and its potential uses. *Strategic Management Journal* **42**(12); 2195–2217.

Floridi L., Cowls J., Beltrametti M., Chatila R., Chazerand P., Dignum V., Luetge C., Madelin R., Pagallo U. and Rossi F. (2018). AI4People – an ethical framework for a good ai society: opportunities, risks, principles, and recommendations. *Minds and Machines* **28**(4), 689–707.

Fukuyama F. (2014). *Political Order and Political Decay: From the Industrial Revolution to the Globalization of Democracy*. New York, NY, USA: Farrar, Straus and Giroux.

Garicano L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy* **108**(5), 874–904.

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press

Hodgson G.M. (2012). *From Pleasure Machines to Moral Communities: An Evolutionary Economics Without Homo Economicus*. Chicago, IL, USA: University of Chicago Press.
Janssen M. and Kuk G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly* **33**(3), 371–377.
Jobin A., Ienca M. and Vayena E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence* **1**(9), 389–399.
Kernaghan K. (2014). The rights and wrongs of robotics: ethics and robots in public organizations. *Canadian Public Administration* **57**(4), 485–506.
Lipsky M. (2010). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York, NY, USA: Russell Sage Foundation.
Mahoney J. and Thelen K. (2010). A theory of gradual institutional change, In Mahoney J and Thelen K (eds.) *Explaining Institutional Change: Ambiguity, Agency, and Power*. Cambridge: Cambridge University Press, pp. 1–37.
Makridis C. (2025). Organizational Transmission of AI: Managerial Influence on Generative AI Adoption. Available at SSRN 5401053.
Makridis C.A. (2021). (Why) is there a public/private pay gap? *Journal of Government and Economics* **1**, 100002.
Makridis C.A. and Ammons J.D. (2025). Governing the large language model commons: using digital assets to endow intellectual property rights. *Journal of Institutional Economics* **21**, e19.
March J.G. and Olsen J.P. (2010). *Rediscovering Institutions*. New York: Simon and Schuster.
Moe T.M. (1984). The new economics of organization. *American Journal of Political Science* **28**(4), 739–777.
Newman J., Mintrom M. and O'Neill D. (2022). Digital technologies, artificial intelligence, and bureaucratic transformation. *Futures* **136**, 102886.
Niskanen W.A. (1968). The peculiar economics of bureaucracy. *The American Economic Review* **58**(2), 293–305.
Niskanen W.A. (1975). Bureaucrats and politicians. *The Journal of Law and Economics* **18**(3), 617–643.
Ostrom E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.
Paniagua P. and Rayamajhee V. (2022). A polycentric approach for pandemic governance: nested externalities and co-production challenges. *Journal of Institutional Economics* **18**(4), 537–552.
Perry J.L. and Recascino Wise L. (1990). The motivational bases of public service. *Public Administration Review* **50**(3), 367–373.
Pierson P. (1997). Path Dependency, Increasing Returns, and the Study of Politics, Center for European Studies Working Paper Series, No. 7.7. Cambridge, MA: Center for European Studies, Harvard University.
Pressman J.L. and Wildavsky A.B. (1973). *Implementation: How Great Expectations in Washington Are Dashed in Oakland*. Berkeley, CA, USA: University of California Press.
Rainey H.G. (2009). *Understanding and Managing Public Organizations*. San Francisco, CA, USA: Jossey-Bass.
Webb M. (2020). The Impact of Artificial Intelligence on the Labor Market. Available at SSRN 3482150.
Winter S. (2012). Implementation, In Pierre J. and Peters BG. (eds), *The SAGE Handbook of Public Administration*. London: SAGE, pp. 255–263.
Wirtz B.W. and Müller W.M. (2019). An integrated artificial intelligence framework for public management. *Public Management Review* **21**(7), 1076–1100.
Zuboff S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Cite this article: Pieroni L and Rossello Roig M (2026). AI adoption in bureaucracies. *Journal of Institutional Economics* **22**, e15, 1–21. https://doi.org/10.1017/S1744137426100496

https://doi.org/10.1017/S1744137426100496 Published online by Cambridge University Press