From Faces to Finance: AI Measurement of Promotion Pressure and Health Costs in Public Office *

Fangzhou Lu
University of Hong Kong

Lei Huang
University of Hong Kong

Sixuan Li
University of Hong Kong

Yingyan Shi
University of Hong Kong

February 14, 2026

Abstract

We study how career incentives in the public sector affect government officials' stress and health, using visible aging as a scalable proxy. Using administrative health inspection data, we show that the difference between individuals' apparent age and chronological age is a strong predictor of subsequent health outcomes. We then construct a panel of mayors and senior officials and apply computer-vision models to estimate apparent age from official photographs. Officials in jurisdictions with weaker economic performance age significantly faster. Linking outcomes to promotion incentives, we show that the effect concentrates where promotion tournaments are steeper and downside risk is greater, consistent with both rewards and punishments shaping stress exposure. To corroborate the mechanism, we use large language models to analyze officials' speeches and government reports, extracting measures of urgency, constraint, and risk emphasis that rise when officials age faster and predict within-official changes in apparent age. Together, AI-based facial and textual measures reveal substantial human-capital costs of promotion pressure in public office.

*Many thanks to to Daniel Greenwald, Kose John, Maarten Meeuwis, Christopher Palmer, Adrien Verdelhan, Haoxiang Zhu and seminar participants at MIT and HKU for helpful comments and suggestions. Fangzhou is grateful to Jennifer Carpenter, Jonathan Parker, Antoinette Schoar, and Robert Whitelaw for their guidance and advice.

1 Introduction

Government officials make high-stakes decisions under intense career incentives and monitoring, yet we know surprisingly little about how the political incentive environment maps into the decision maker's own human capital. In modern bureaucracies, promotions are often tied to short-horizon performance targets, while discipline risk is shaped by audits, inspections, and anti-corruption enforcement. These forces generate a classic incentives-versus-risk tradeoff: stronger reward motives may induce effort and risk-taking to meet targets, while stronger punishment motives may induce caution but also sustained psychological strain. Understanding how these forces translate into health costs (Borgschulte et al., 2025) using machine learning and AI (Agarwal et al., 2023) matters for both positive political economy and finance: health and stress shape risk tolerance, time discounting, attention, and ultimately the quality and style of policy choices that affect local firms (Fisman and Wang, 2015), credit allocation, and regional growth.

This paper studies the determinants of stress-related aging among Chinese government officials using a new, large-scale measure of apparent age extracted from official photographs. We apply visual machine-learning methods to estimate an individual's face age, and we interpret deviations of face age from chronological age as a proxy for cumulative stress exposure.[1] The setting is attractive for identification because officials face relatively standardized career ladders, frequent rotations, and well-defined performance evaluation criteria, while local economic conditions generate plausibly exogenous variation in the difficulty of meeting promotion-relevant targets. We combine face-age estimates with administrative information

[1]Specifically, we estimate officials' apparent age from facial images an open-source computer vision framework that provides unified access to state-of-the-art face detection and face attribute pipelines. For each image, we first detect and align the face to a standardized orientation, then apply an age-prediction model to produce an estimated age in years. A central concern is external validity across demographic groups. Many widely used facial age models are trained on datasets that under-represent Asian faces, which can induce systematic prediction error and reduce comparability across officials. To address this issue, we fine-tune the age-prediction component on a large training set of approximately one million Asian face images with known ages, using a dedicated GPU cluster for efficient training and hyperparameter search. We validate performance using held-out samples and report that the fine-tuned model reduces prediction error and, importantly, attenuates age-dependent and cohort-specific biases relative to the off-the-shelf baseline.

on officials and the economic and social characteristics of their jurisdictions to quantify how the local incentive environment maps into visible aging, and we integrate an LLM-based (Brynjolfsson et al., 2025) text module that extracts structured measures of urgency tone, objective constraints, and risk emphasis from officials' communications, allowing us to trace whether shifts in perceived pressure co-move with within-official changes in apparent age.

To assess whether our visual machine-learning "face age" measure captures meaningful variation in aging rather than spurious image artifacts, we validate it against well-established environmental determinants of health. Consistent with the medical and epidemiological evidence that chronic pollution exposure accelerates biological aging, we find that worse local air quality is associated with significantly higher apparent age: officials in regions with higher PM2.5 concentrations (a7) look older, holding constant individual characteristics and time effects. In contrast, proxies for a healthier environment predict lower apparent age. Officials governing in jurisdictions with better water quality and overall environmental conditions (a112) appear younger. Likewise, stronger local food quality and availability, captured by higher fruit production (a66) and higher vegetable production (a59), are associated with lower apparent age. The signs, magnitudes, and robustness of these correlations align with prior evidence on how pollution and nutrition shape health and aging, supporting the interpretation of our face-age measure as a valid proxy for cumulative aging-related health deterioration.

Moreover, we corroborate the apparent-age measure with objective medical records. We obtain annual health examination data from Meinian Onehealth, one of the largest health screening providers in China, which maintains a nationwide database of standardized checkup results. Using individual identifiers, we link these records to roughly one-third of the officials in our sample. The examinations include routine, clinically comparable measures, such as height, weight and BMI, waist circumference, systolic and diastolic blood pressure, resting heart rate, and a broad set of laboratory biomarkers capturing cardiometabolic and organ

function.[2] In addition, the data record physician-assessed abnormal findings and major diagnoses, including chronic conditions such as hypertension and diabetes as well as less frequent but high-severity outcomes such as suspected or confirmed malignancies and serious cardiovascular conditions. We show that officials with higher excess face age exhibit systematically worse health profiles in these medical records, particularly along cardiovascular and metabolic risk dimensions that are strongly linked to stress and accelerated aging, providing direct validation that our visual machine-learning measure captures meaningful variation in health-related human capital rather than spurious image artifacts.

Our first result is that officials age faster when they govern in regions with weaker economic performance. Mayors and other key local officials serving in jurisdictions with lower growth exhibit significantly higher apparent age, even after controlling for cohort effects, time-invariant individual characteristics, and time-varying jurisdiction attributes. We capture local economic performance using multiple indicators, including local government bond rollover pressure (maturing debt), the unemployment rate, GDP growth, firm bankruptcy intensity, and the relative performance of SOE stocks, and we find a consistent pattern: higher debt maturity pressure, higher unemployment, lower GDP growth, higher bankruptcy rates, and weaker SOE stock performance are each associated with accelerated apparent aging. The magnitudes are economically meaningful: a one percentage point decline in GDP growth is associated with an approximately eight-month increase in apparent age, while a one percentage point increase in the unemployment rate corresponds to roughly three additional months of apparent age. Translating these estimated accelerations into standard health-economics benchmarks implies a sizable health cost borne by the decision makers themselves, a pattern that is difficult to reconcile with a purely selection-based story in which lower-growth regions simply attract older-looking officials and instead points to stress exposure that accumulates while in office.

[2](e.g., fasting glucose and related glycemic markers, lipid profiles including total cholesterol/LDL/HDL/triglycerides, liver enzymes such as ALT/AST, kidney-function markers such as creatinine/eGFR, uric acid, and urine-test indicators)

Two classes of shocks play a central role in sharpening identification reward-driven pressure and punishment-driven fear, and the data show clear responses in officials' visible aging around these events. First, exploiting quasi-exogenous increases in monitoring intensity generated by the rollout of central inspection teams and related enforcement surges, we find a discrete rise in officials' excess face age in the periods following inspection arrival. In event-time designs with official fixed effects and common time effects, the post-inspection increase is economically meaningful, robust to local controls and image-quality adjustments, and stronger in settings where enforcement risk is plausibly more salient. This pattern is consistent with a punishment-force mechanism: heightened scrutiny and the elevated threat of discipline translate into measurable deterioration in health-related human capital while in office.

Second, we use the COVID-19 pandemic as an unexpected macro shock that simultaneously worsened local economic conditions and intensified crisis-management demands, and we find that officials in harder-hit jurisdictions age faster during outbreak periods. Linking image-based aging trajectories to local pandemic severity and outbreak timing, we show that excess face age increases more sharply for officials governing areas with higher case intensity and more disruptive containment episodes, relative to officials in less affected places. The effect concentrates in the windows when local economic activity and social stability pressures were most strained, reinforcing the interpretation that crisis-driven macro-financial stress and administrative overload materially elevate stress exposure in office.

Moreover, we use large language models (LLMs) to provide text-based evidence that corroborates the shock results and helps interpret mechanisms. From officials' speeches and government reports, we extract structured measures of urgency tone, objective constraint emphasis, and risk framing, and we find that these indicators rise in the same periods in which officials' excess face age increases. The co-movement holds within the same official over time, indicating that changes in language are not merely cross-sectional differences in communication style but track shifting pressure environments.

Finally, we show that inequality amplifies aging. Officials in high-inequality regions display significantly greater apparent age relative to comparable officials in more equal regions. This finding is consistent with a mechanism in which inequality raises the political and administrative complexity of governing, increases the intensity of social conflict and petitioning risk, and heightens the perceived cost of missteps. In a finance analogy, inequality can be viewed as a local "volatility" state that increases the variance of outcomes and the probability of tail events, thereby raising stress even holding average growth constant.

These findings contribute to several literatures, and the paper's incremental contributions are threefold. First, we move beyond output measures to show that political incentives and monitoring shocks impose a direct human-capital cost on officials, captured by accelerated facial aging, introducing a new "cost of governance" object to the study of bureaucratic incentives. Second, we connect political economy to finance by showing that local macro-financial strain, including debt rollover pressure, corporate distress, and relative state-sector underperformance, is systematically associated with officials' stress-related aging, suggesting a channel through which financial conditions may affect governance capacity and ultimately firms and investors. Third, we advance the empirical toolkit by using visual machine-learning measures to proxy latent stress exposure at scale and by strengthening identification with quasi-exogenous shocks and within-official event-time designs based on repeated images.

A large literature in political economy argues that the behavior of public officials is shaped by career concerns and the incentive schemes embedded in bureaucratic personnel systems. In the Chinese setting, a central theme is that promotion and retention depend on measurable local performance, especially growth, creating "tournament-style" incentives that can discipline effort and shape policy priorities. Foundational evidence shows that leaders' promotion and termination probabilities covary with economic performance, consistent with personnel control operating as an incentive device (Li and Zhou, 2005). Subsequent work refines how turnover is measured and emphasizes heterogeneity in promotion ladders and political constraints, but the core takeaway remains that local officials face strong, performance-linked

career stakes (Jia et al., 2015; Jiang, 2018). Piotroski and Zhang (2014) show that IPO activities in China are shaped by government official turnovers at the provincial level. While this literature has taught us a great deal about how incentives affect growth, investment, and policy choices, it largely treats the official as a frictionless agent whose costs are limited to career outcomes. Our paper shifts the lens to the official's own human capital by quantifying a health-related cost of governing under pressure.

A second literature studies monitoring, enforcement, and the consequences of anti-corruption campaigns. Recent empirical work documents that anti-corruption efforts can alter institutional behavior and economic outcomes, including changes in the allocation of favors and the behavior of connected parties (Chen and Kung, 2019; Chen et al., 2018). Within China, the deployment of central inspection teams has been used as a plausibly exogenous increase in scrutiny and enforcement intensity, with measurable impacts in financial and real outcomes (Chen and Kung, 2019). This work motivates a punishment-risk mechanism: heightened oversight raises the perceived probability and expected cost of disciplinary action, which can meaningfully change behavior. Our contribution is to show that the same monitoring shocks also have a personal cost that is visible in accelerated aging. By linking inspection timing and enforcement intensity to within-official changes in apparent age, we provide new evidence that anti-corruption scrutiny operates not only through incentives and constraints on actions, but also through stress exposure borne by the officials themselves.

A third, closely related strand in finance and labor economics examines how high-powered incentives and adverse shocks affect the health of decision makers. A prominent example is work on CEOs showing that stress induced by corporate governance pressure and crisis exposure has long-run effects on visible aging and mortality, framing stress as a private cost of holding a high-stakes leadership position (Borgschulte et al., 2025). That insight is highly relevant for our setting because Chinese local officials occupy positions with similarly high responsibility, concentrated accountability, and steep career consequences. Our paper extends this "leader stress" perspective from corporate managers to public officials, and

from firm-level shocks to macro-financial and political monitoring shocks, highlighting that leadership stress is not only a corporate phenomenon but also a feature of state capacity and governance.

Methodologically, our paper sits within a rapidly expanding literature that uses machine learning and nontraditional data to measure latent traits and states that are difficult to observe in administrative records (Athey and Imbens, 2019). Computer vision has become a practical tool for extracting information from images, and the broader medical and computer-vision literatures emphasize that "apparent age" estimation can be constructed at scale while still requiring careful validation and attention to image quality (Rothe et al., 2015). Parallel developments in economics and finance use AI-extracted features from faces for economic questions, underscoring both the promise and the need for robust design and interpretation (Stoker et al., 2016). Building on these advances, we contribute a new application that is particularly well suited to a panel setting: because we observe repeated images for the same officials and know all officials' birthdates, we can difference out time-invariant facial traits and focus on within-person changes in excess face age around shocks. We also validate the measure against canonical environmental determinants of health (Cheng et al., 2026; Tucker, 2021) and show robustness to alternative pipelines, addressing a first-order concern in image-based measurement.

The remainder of the paper proceeds as follows. We describe the institutional background and the construction and validation of the face-age measure. We then present the baseline relationship between local economic conditions and officials' apparent age, followed by a battery of robustness checks addressing selection, image quality, and alternative specifications. Finally, we turn to mechanism tests that distinguish reward-based from punishment-based pressure and discuss implications for political incentives, policy implementation, and economic outcomes.

2 Data and Methodology

2.1 Sample construction and image sources

We construct a panel of Chinese local government officials from 2014–2023 by combining administrative biographical records with a large archive of facial photographs. We focus on officials with clearly defined appointments in city- or prefecture-level leadership positions (e.g., mayors and other senior municipal leaders), although the empirical specifications generalize to broader leadership groups. For each official, we observe full name, jurisdiction, position, and exact birthdate, which allows us to compute chronological age at any point in time and to express facial aging outcomes in a comparable unit across individuals and years.

Facial images are collected from two complementary sources. First, we scrape official photographs from government portals, including municipal and provincial government websites and affiliated agencies. These sources typically provide standardized portraits and work-activity photos tied to official announcements, meetings, inspections, and personnel changes. Second, we collect press photographs taken by professional journalists and distributed through mainstream newspapers and online news outlets. Compared with official portraits, press images are less curated and provide dense temporal coverage around salient events, which is particularly useful for event-study analyses. Each photograph is linked to an official and assigned a timestamp using the publication date and, when available, the event date recorded in the surrounding text or caption.

We obtain repeated facial observations for a substantial share of officials over their tenure, often spanning multiple years. This repeated-measures structure is central to our identification strategy: it enables within-official comparisons before and after shocks (e.g., enforcement actions or economic downturns) and mitigates concerns that cross-sectional differences in appearance reflect sorting of officials into jurisdictions rather than stress exposure while in office.

2.2 Health Inspection Data

We obtain annual health examination records from one of the largest healthcare providers in China. The provider compiles a comprehensive nationwide database of health screening information, which we link to our administrative sample using individual identifiers, allowing us to match roughly one-third of government officials in our dataset. The examination records include standardized, high-frequency measures such as anthropometrics and vital signs (e.g., BMI and blood pressure), routine laboratory panels (e.g., blood and urine tests covering lipid profiles, liver and kidney function, glucose-related biomarkers, and uric acid), and common screening exams (e.g., ECG and selected imaging or ultrasound tests depending on the examination package). In addition, the records contain physician-coded diagnoses and clinically significant findings, including major chronic conditions (such as diabetes and hypertension) as well as more severe outcomes that occur less frequently in a given year (such as malignancies and serious cardiovascular conditions). These medical records provide an objective benchmark for health status that we use to corroborate and validate our AI-based apparent-age measures derived from facial images and machine-learning algorithms.

2.3 Economic, financial, and governance covariates

We merge the official-by-time image panel with jurisdiction-level measures that capture economic conditions and stress-relevant constraints. These include baseline measures of local economic performance (e.g., GDP growth) and distributional conditions (e.g., inequality), as well as indicators of macro-financial strain and capital-market exposure when available (e.g., local debt rollover pressure, layoffs, firm bankruptcies, and stock-market-based performance measures comparing state-owned and private firms headquartered in the jurisdiction). We also compile measures of monitoring and enforcement intensity associated with anti-corruption activity and inspections, which serve as proxies for punishment risk. All variables are aligned at the city (or province) level and matched to the timing of each image observation.

2.4 Measuring apparent age from facial images

2.4.1 Face detection, identity verification, and age estimation

We measure apparent age using a computer-vision pipeline that produces a predicted "face age" for each image. We first detect faces and standardize alignment to reduce variation from pose and camera angle. We then apply DeepFace as well as GPT-Image to generate (i) face embeddings used for identity verification and de-duplication and (ii) predicted age for each detected face. We retain model confidence scores and facial landmark diagnostics and use them both as filters (to exclude low-quality detections) and as controls in robustness analyses.

Because image conditions can mechanically affect predicted age, we implement conservative quality screens. We exclude photographs with severe occlusion, extreme head rotation, very low resolution, or insufficient lighting. In addition, we compute and control for observable image attributes that can shift age predictions, including face size in the frame, sharpness, brightness, and contrast. We also control for source type (official website versus press) to absorb systematic differences in curation and photographic style.

2.4.2 Aggregation and construction of excess face age

Many officials appear multiple times within a short window (e.g., during a campaign, inspection, or crisis period). To reduce noise from any single photograph, we aggregate face-age predictions within an official-by-period cell using robust rules (e.g., median or trimmed mean). Our main outcome is either the predicted face age itself or *excess face age*, defined as predicted face age minus the official's chronological age at the time the image was taken. Because we observe birthdates for all officials, excess face age provides a natural normalization that removes mechanical aging with time and isolates deviations that plausibly reflect cumulative stress exposure.

2.5 Empirical design

2.5.1 Baseline panel specification

We estimate how economic conditions and monitoring intensity relate to apparent aging using panel regressions that exploit within-official variation over time. The baseline specification relates excess face age for official i observed at time t to contemporaneous jurisdiction conditions and controls:

$$\text{ExcessFaceAge}_{it} = \beta X_{jt} + \gamma Z_{it} + \delta Q_{it} + \alpha_i + \lambda_t + \varepsilon_{it}, \tag{1}$$

where X_{jt} captures city/province-level economic and financial conditions in jurisdiction j at time t, Z_{it} includes career-stage controls (e.g., position indicators), Q_{it} includes image-quality controls and source-type indicators, α_i are official fixed effects, and λ_t are time fixed effects (e.g., year-month) that absorb nationwide trends and common shocks. The official fixed effects ensure identification is driven by changes in an official's apparent aging over time rather than persistent differences in baseline appearance across individuals.

2.5.2 Event-study and shock designs with repeated images

We complement the baseline panel evidence with event-study and difference-in-differences designs that leverage repeated observations before and after salient shocks. Two classes of shocks are particularly informative in this setting. First, episodes that plausibly increase monitoring intensity, such as the arrival of central inspection teams and spikes in local anti-corruption enforcement, allow us to test a punishment-risk channel by tracing dynamic changes in excess face age around these events. Second, adverse economic shocks, such as pandemic-related case surges and associated lockdowns that precipitate local economic downturns, allow us to isolate stress induced by sudden macroeconomic deterioration. In each case, we estimate dynamic treatment effects with leads and lags and report pre-trend tests to assess whether apparent aging begins to diverge only after the shock.

2.6 Machine Learning on Apparent Age

We estimate officials' apparent age from facial images using DeepFace, an open-source computer vision framework that provides unified access to state-of-the-art face detection and face attribute pipelines. For each image, we first detect and align the face to a standardized orientation, then apply an age-prediction model to produce an estimated age in years. The key object in our analysis is *excess apparent age*, defined as the difference between predicted age and the official's chronological age at the time the photo was taken. This measure captures time-varying information in facial appearance that is plausibly related to cumulative stress exposure, health status, and lifestyle, and it has the practical advantage of being available at high frequency from archival photographs even when administrative health records are unavailable.

A central concern is external validity across demographic groups. Many widely used facial age models are trained on datasets that under-represent Asian faces, which can induce systematic prediction error and reduce comparability across officials. To address this issue, we fine-tune the age-prediction component on a large training set of approximately one million Asian face images with known ages, using a dedicated GPU cluster for efficient training and hyperparameter search. We validate performance using held-out samples and find that the fine-tuned model reduces prediction error and, importantly, attenuates age-dependent and cohort-specific biases relative to the off-the-shelf baseline. In the empirical analysis, we aggregate image-level predictions to the official-by-time level using conservative rules that prioritize higher-quality photographs and downweight extreme outliers, thereby improving signal-to-noise while maintaining transparency and replicability.

2.7 Role of large language models in metadata extraction

We complement the computer vision pipeline with a large language model (LLM) to extract structured metadata from unstructured web pages and news articles. Specifically, the LLM is used to parse captions and surrounding text and to recover names, positions, jurisdictions,

and event dates, which are then used to construct a clean panel linking each facial image to the correct official and time period. The LLM is also used as a consistency-check layer that flags potential mismatches, such as conflicting titles, implausible dates, or ambiguous identity references. We implement conservative matching rules and discard observations when identity or timing cannot be uniquely resolved, so that any increase in coverage does not come at the cost of introducing discretionary hand-coding.

In addition, we adapt an LLM-based text module following Brynjolfsson et al. (2025) to quantify communication features from officials' public statements, including urgency tone, objective constraints, and risk emphasis. These measures allow us to test whether shifts in perceived pressure co-move with within-official changes in excess apparent age. To further strengthen the measurement strategy, we conduct robustness checks using an independent vision-capable model for age inference on a subset of images and verify that the main conclusions are not driven by a particular estimator. Throughout, we prioritize reproducibility by fixing prompts and extraction templates, logging model outputs, and pre-specifying rules for entity resolution and data exclusion.

2.8 Summary Statistics

Figure 1 and 2 shows that there is a positive and significant relationship between apparent age and actual age, suggesting that our machine learning algorithm of facial recognization is valid. Table 1 reports summary statistics for the main sample of Chinese local officials from 2014–2023. The panel includes 589 unique officials and 4,717 official-year observations, based on 25,412 facial images. Image coverage is dense: the mean official appears in 43.1 images (P25 = 18.0; P75 = 60.0), corresponding to 5.4 images per official-year (P25 = 3.0; P75 = 7.0). Sources are diversified, with 58% of images drawn from government portals and 42% from press coverage, which mitigates concerns that results are driven by a single channel of image production.

Our primary outcome is predicted face age from the computer-vision pipeline, which av-

erages 50.9 years with a standard deviation of 6.4. Chronological age averages 52.3 years (s.d. 6.8), implying a mean excess face age of −1.4 years (s.d. 3.1). The interquartile range of excess face age spans from −3.5 (P25) to 0.7 (P75), indicating substantial variation around chronological age. The absolute deviation is economically meaningful: absolute excess face age averages 2.5 years (s.d. 2.2), and 34% of observations exhibit positive excess face age, consistent with a nontrivial right tail in apparent aging. Officials in the sample are predominantly male (86%). Tenure in the current position averages 3.1 years (s.d. 1.9), and mayors account for 38% of official-year observations, with the remainder comprised of other senior officials.

The jurisdiction-level covariates display sizable cross-sectional and time-series variation. City/province GDP growth averages 5.6% (s.d. 2.3; IQR 4.2–7.0), while log GDP averages 10.9 (s.d. 0.8). The inequality index has a mean of 0.43 (s.d. 0.08), and the unemployment rate averages 4.3% (s.d. 1.1). Measures of local labor-market distress and firm dynamics also vary meaningfully, with layoff intensity averaging 6.8 per 10k workers and bankruptcy filings averaging 3.1 per 10k firms. Macro-financial stress indicators show substantial dispersion: LGFV bond maturities average RMB 24.0bn (s.d. 18.0), and maturity pressure (maturities/fiscal revenue) averages 0.32 (s.d. 0.20). Capital-market exposure is nontrivial in many jurisdictions, with an average of 18 listed firms headquartered locally and mean local market capitalization of RMB 180.0bn (s.d. 220.0). Monitoring intensity is similarly heterogeneous: central inspections occur in 21% of jurisdiction-years, and investigated corruption cases average 12.0 per city/province-year.

3 Results

3.1 Economy and capital market

Table 2 shows that deteriorating local economic conditions are associated with significantly higher apparent age, consistent with stress arising from macro-financial strain and heightened

career risk. Measures of fiscal and credit stress at the city level, captured by the intensity of upcoming municipal financing-vehicle bond maturities, predict higher face age among incumbents, suggesting that tighter rollover pressure and constrained financing capacity translate into greater psychological burden for local leaders. The same pattern holds for real-economy distress: jurisdictions experiencing elevated layoffs and a higher incidence of firm bankruptcies, among both state-owned and privately owned enterprises, are associated with faster visible aging of officials, plausibly reflecting rising social stability concerns, employment mandates, and the political costs of adverse outcomes. Finally, we link these effects directly to capital-market signals by showing that officials' apparent age is higher when local state-owned enterprises underperform private firms in the stock market, a relative-performance metric that is salient for evaluating policy effectiveness and can intensify scrutiny when state-sector objectives conflict with market outcomes. Taken together, these results indicate that weak economic fundamentals and macro-financial stress, especially when amplified by debt rollover risk and adverse labor-market and corporate outcomes, materially increase stress exposure in office in ways that are visible in accelerated facial aging.

3.2 Covid as shocks

We further sharpen identification by exploiting the COVID-19 pandemic as an unexpected, plausibly exogenous shock to local economic conditions and administrative pressure. Table 3 shows the result. We link officials' apparent age to regional pandemic severity, measured by confirmed case intensity, and show that officials governing harder-hit jurisdictions experience a pronounced increase in predicted face age relative to those in less affected areas. The effect is strongest precisely in periods and places where containment measures were most disruptive, when lockdowns and mobility restrictions were tighter and local economic activity contracted more sharply, consistent with the pandemic operating through a macroeconomic and social-stability channel rather than through slow-moving regional characteristics. In a difference-in-differences/event-study framework that absorbs jurisdiction and time fixed effects and

controls for pre-trends and image characteristics, the results imply that pandemic-driven downturn risk and crisis management demands substantially elevated stress exposure in the office, and that this stress is visible as accelerated facial aging of mayors and other senior city officials.

3.3 Anti-Corruption

We further document a clear punishment-risk channel: heightened anti-corruption enforcement is associated with visibly accelerated aging among incumbent officials. Table 4 shows the result. Using the deployment of the central government's inspection teams as a quasi-exogenous increase in monitoring intensity, we find that officials' predicted face age rises in the period following an inspection, relative to the same officials prior to the inspection and relative to officials in jurisdictions not yet inspected. This effect is economically meaningful and robust to controlling for local economic conditions, jurisdiction and time fixed effects, and a rich set of individual characteristics and image controls. Complementing this event-style evidence, we also show that apparent age is increasing in local enforcement intensity measured continuously: officials governing in cities or provinces with a larger number of investigated or disciplined corruption cases exhibit higher predicted face age. Together, these patterns are consistent with the idea that greater scrutiny and the heightened threat of punishment raise stress exposure in office, translating into measurable deterioration in health-related human capital that is visible in facial aging.

3.4 Text-Based Evidence on Perceived Pressure and Risk Emphasis

In this subsection, we provide additional evidence consistent with a "stress and pressure" mechanism by leveraging large language models to quantify latent features of officials' communications. Table 5 shows the result. Using a corpus of officials' speeches and government

reports, we construct three text-based measures, *urgency tone*, *objective constraints*, and *risk emphasis*, that capture how strongly an official frames issues as time-sensitive, constrained by external pressures, or dominated by downside risks. We then relate these measures to our apparent-aging outcome. Table 5 summarizes the results. Across specifications, higher urgency and constraint language is positively associated with excess face age, and the risk-emphasis measure exhibits a similar, though somewhat weaker, relationship. The estimated magnitudes are economically meaningful: a one-standard-deviation increase in these LLM scores corresponds to a nontrivial increase in excess face age, suggesting that the language used in official communications systematically co-moves with the apparent-aging proxy.

Crucially, the association is not driven solely by cross-sectional differences across officials or locations. When we include official fixed effects, the coefficients remain positive and statistically significant (Table 5, columns (2)–(3)), indicating that within the same official, periods characterized by more urgent, constrained, and risk-focused language coincide with higher apparent aging. This within-official pattern helps rule out time-invariant heterogeneity, such as stable personal traits, baseline health, or persistent regional differences, as the primary explanation. Taken together, the LLM-based text evidence corroborates the mechanism that heightened perceived pressure and risk-management demands are reflected both in what officials say and in how they appear, and it complements the results based on economic stress, monitoring intensity, and shock exposures.

4 Robustness Checks

This section evaluates whether our findings are sensitive to measurement choices, sample construction, timing assumptions, and alternative empirical specifications. Across exercises, the qualitative conclusions remain unchanged: adverse economic conditions, heightened monitoring intensity, and pandemic-related shocks are consistently associated with higher apparent age among incumbent officials. Table 6 shows those results.

4.1 Alternative construction of the face-age outcome

A primary concern is that the estimated effects could reflect idiosyncrasies of a particular age-estimation model rather than true variation in aging. We therefore replicate the main analyses using an alternative apparent-age measure constructed from an LLM-based pipeline (GPT-image) as a substitute for the baseline computer-vision estimator. The alternative pipeline is used solely to generate an additional predicted age measure; it does not alter sample selection or covariate definitions. Re-estimating the validation tests, the economy and capital-market specifications, the anti-corruption specifications, and the COVID shock designs yields coefficients with the same signs and similar magnitudes as in the baseline results, indicating that our conclusions are not driven by a particular computer-vision architecture.

We also verify that results are robust to alternative aggregation rules when multiple photos are observed for the same official in a given period. Using the median prediction (baseline), trimmed means, or the mean of the top-quality images (based on detection confidence and face size) produces similar estimates, suggesting that outlier photos or episodic image noise do not drive the findings. Finally, we show that the results hold when using predicted face age levels rather than excess face age; the normalization by chronological age improves interpretability but is not required for inference.

4.2 Image quality, source composition, and sample restrictions

Because photographic conditions can mechanically affect predicted age, we implement several increasingly stringent filters. First, we restrict the sample to images with high detection confidence, adequate resolution, and limited occlusion, and we confirm that estimated effects are stable. Second, we re-estimate the main specifications within source-type subsamples. When restricting the sample to government-portal images only, results remain statistically and economically meaningful, alleviating concerns that press coverage changes image composition during crises or inspections. Similarly, when restricting to press images only, results

persist, suggesting that they are not driven by curated official portraits. Third, we drop images with extreme pose or lighting and include flexible controls for brightness, sharpness, and face size; the coefficients of interest remain robust, indicating that our results are not artifacts of image characteristics.

4.3 Controls, fixed effects, and clustering

We next assess sensitivity to alternative control sets and fixed-effect structures. First, we add time-varying local controls (e.g., GDP level and growth, inequality, unemployment proxies, and other fundamentals) and confirm that coefficients on the key variables remain stable. Second, we include flexible time fixed effects (year) and, where relevant, province-by-time fixed effects to absorb higher-level shocks and policy cycles; results persist. Third, we estimate specifications that include official-by-position fixed effects or restrict attention to narrower position groups (e.g., mayors only), ruling out composition effects from changes in job titles or responsibilities.

Inference could be sensitive to serial correlation and spatial correlation in shocks. We therefore report standard errors clustered at the jurisdiction level (baseline) and show robustness to alternative clustering schemes, including clustering at the province level and two-way clustering by jurisdiction and time. In all cases, statistical significance and economic magnitudes remain similar.

4.4 Selection and media coverage concerns

A remaining concern is that officials may be photographed more often during stress episodes (e.g., inspections, outbreaks), potentially altering the sample of images. We address this in two ways. First, we demonstrate that results are robust when restricting to officials with regular pre-period photo coverage and to balanced panels around events. Second, we construct placebo outcomes based on photo frequency and show that, while crises may increase photo volume, the estimated effects on apparent age persist even when controlling

for the number of images per official-period and when reweighting to equalize photo counts across treated and control units. These checks reduce the likelihood that our findings are mechanically driven by changes in publicity rather than changes in apparent age.

4.5 Placebo tests

Finally, we conduct placebo tests that shift event dates forward or backward in time and re-estimate the event-study specifications. The placebo estimates are small and statistically indistinguishable from zero, while the true event dates produce sharp increases in apparent age, providing further support that our identification is capturing the intended shocks rather than spurious time trends.

Overall, the robustness exercises show that our main conclusions are stable across alternative measurement pipelines, image-quality restrictions, timing rules, fixed effects, and inference procedures. The evidence consistently points to economically meaningful increases in apparent age following adverse economic shocks and heightened enforcement intensity, consistent with stress-induced deterioration in officials' health-related human capital.

4.6 Photo using non-government officials

To assess whether our apparent-aging measure responds mechanically to local conditions rather than to incentives specific to public officials, we conduct a placebo exercise using photos of non-government individuals. We construct a comparison sample of ordinary people drawn from the same cities and time periods as in the main analysis and process their images through the identical face-age pipeline and data-cleaning procedures. We then assign to these individuals the same city-by-time measures of government stress, monitoring intensity, and macro-financial shocks used in the baseline specifications, and estimate regressions that mirror the main design with the same fixed effects structure and controls. Across all specifications, we find no statistically or economically meaningful relationship between these government-stress proxies and the excess face-age outcomes of non-government individuals.

This absence of an effect suggests that our main results are unlikely to be driven by city-level confounders that broadly affect everyone's appearance (e.g., common local shocks to health, environment, or photo quality). Instead, the evidence is consistent with an incentive-based channel: the apparent-aging response emerges primarily among government officials who face promotion rewards and disciplinary risks, rather than among ordinary individuals exposed to the same local conditions.

5 Heterogeneity

This section explores how the stress–aging relationship varies across officials and jurisdictions. The goal is twofold. First, heterogeneity tests provide additional evidence on mechanism by examining whether effects concentrate in settings where career incentives, downside risk, and monitoring are plausibly stronger. Second, they help assess external validity by clarifying which types of officials and places are most exposed to stress-induced deterioration in health-related human capital. Table 7 shows the results.

5.1 Heterogeneity by career concerns and promotion incentives

We begin by testing whether economic stress translates into larger increases in apparent age when career concerns are likely to be more salient. We proxy for promotion incentives using characteristics that steepen tournaments or increase the marginal return to measured performance, such as (i) being closer to typical promotion-age cutoffs, (ii) shorter expected time-to-retirement, and (iii) positions with clearer performance accountability (e.g., mayors relative to other senior officials). Across these splits, the association between weak economic conditions and excess face age is stronger for officials with greater promotion stakes. These patterns are consistent with a career-concerns channel in which the stress cost of governing rises when the perceived probability of promotion (or demotion) is most sensitive to local performance.

5.2 Heterogeneity by monitoring intensity and punishment risk

Next, we examine whether the anti-corruption effects are larger in environments where scrutiny is plausibly more consequential. We interact inspection and enforcement measures with proxies for baseline monitoring intensity, such as prior enforcement activity, measures of local institutional capacity, and political visibility (e.g., provincial capital status or economically important jurisdictions). We find that the post-inspection rise in apparent age is more pronounced in high-monitoring environments and in jurisdictions where enforcement intensity is historically higher. Similarly, the continuous relationship between corruption-case intensity and excess face age is stronger where the marginal threat of punishment is likely to be higher. These patterns support a punishment-risk mechanism and suggest that inspections operate partly by elevating perceived downside risk rather than solely by changing contemporaneous workloads.

5.3 Heterogeneity by macro-financial exposure and capital-market discipline

We then study heterogeneity across jurisdictions with different levels of macro-financial fragility and capital-market exposure. If stress is amplified when local outcomes are more salient to investors and the public, then the effect of economic deterioration on officials' apparent age should be stronger in places where the corporate sector is larger, market monitoring is tighter, and rollover risk is more binding. Consistent with this view, we find that the impact of debt-maturity pressure, layoffs, and bankruptcies on excess face age is stronger in jurisdictions with greater capital-market presence (e.g., more listed firms or larger aggregate market capitalization) and in periods characterized by tighter financing conditions. These results connect political incentives to a finance-relevant mechanism: capital markets can raise the reputational and accountability stakes of local economic management, thereby increasing stress exposure for officials when conditions deteriorate.

5.4 Heterogeneity during COVID-19: severity, containment, and economic disruption

We also examine heterogeneity in the pandemic shock design. The increase in apparent age associated with COVID exposure is concentrated in jurisdictions that experienced both high case intensity and more disruptive containment policies, consistent with crisis-management demands and social-stability pressure as key intermediating channels. Moreover, the effect is larger in places with greater exposure to service-sector activity and mobility-sensitive industries and in jurisdictions with tighter fiscal space, where stabilizing employment and firm survival is likely to be more difficult. These patterns reinforce the interpretation that the pandemic affects apparent aging primarily through local economic contraction risk and administrative strain rather than through slow-moving regional characteristics.

Taken together, the heterogeneity patterns are consistent with the paper's central mechanism: stress-related aging responds most strongly when downside career risk and monitoring intensity are elevated. Officials with higher promotion stakes, jurisdictions with stronger enforcement environments, and places with greater macro-financial fragility and capital-market discipline exhibit larger increases in apparent age following adverse shocks. These findings strengthen the economic interpretation of our results and suggest that the personal health costs of governing are highest precisely where incentives are steepest and the consequences of poor outcomes are most severe.

6 Measure Validation

6.1 Indirect Evidence

To assess whether our visual machine-learning "face age" measure captures meaningful variation in aging rather than spurious image artifacts, we validate it against well-established environmental determinants of health. Table 8 shows the result. Consistent with the medical

and epidemiological evidence that chronic pollution exposure accelerates biological aging, we find that worse local air quality is associated with significantly higher apparent age: officials in regions with higher PM2.5 concentrations (a7) look older, holding constant individual characteristics and time effects. In contrast, proxies for a healthier environment predict lower apparent age. Officials governing in jurisdictions with better water quality and overall environmental conditions (a112) appear younger. Likewise, stronger local food quality and availability, captured by higher fruit production (a66) and higher vegetable production (a59), are associated with lower apparent age. The signs, magnitudes, and robustness of these correlations align with prior evidence on how pollution and nutrition shape health and aging, supporting the interpretation of our face-age measure as a valid proxy for cumulative aging-related health deterioration.

6.2 Direct Evidence: Actual Healthcare measure

Table 9, Table 10, and Table 11 provides external validation that our image-based measure of excess face age captures meaningful deterioration in health-related human capital. Using annual health examination records from Meinian Onehealth linked to roughly one-third of the officials in our sample, we regress a broad set of objective biomarkers and diagnosis indicators on excess face age while controlling for chronological age and including official fixed effects, year fixed effects, jurisdiction-time controls, and image-quality adjustments. The results show a consistent and clinically interpretable pattern: officials who look older than their chronological age exhibit significantly worse cardiovascular and metabolic risk profiles. A one-year increase in excess face age is associated with higher systolic and diastolic blood pressure, elevated glycemic measures, and a more adverse lipid profile (higher LDL and triglycerides and lower HDL), as well as higher uric acid. The diagnosis-level evidence points in the same direction: excess face age predicts a higher likelihood of common chronic conditions identified in routine screenings, including hypertension, pre-diabetes/diabetes, dyslipidemia, fatty liver, and hyperuricemia, and is also positively related to higher-severity

health flags. Together, these findings mitigate the concern that our facial-age estimates reflect spurious photo artifacts or time-varying image composition. Instead, the medical records indicate that excess face age loads on well-established physiological pathways through which sustained stress and accelerated aging manifest, providing direct support for our interpretation of facial aging as an economically meaningful measure of stress exposure among decision makers. This validation strengthens the paper's central mechanism: variations in the political incentive environment that we document, especially heightened monitoring and discipline risk, translate into measurable deterioration in officials' health and, by extension, their capacity to process information and bear risk while in office. More importantly, we validate that apparent age is not only an informative and validated proxy, but also a direct clinical measure of health.

7 Conclusion

This paper studies how the economic and political incentive environment shapes the stress-related health costs borne by government officials. Using a new panel of Chinese local officials from 2014–2023, we construct repeated measures of apparent age from facial images collected from government portals and mainstream press coverage. Leveraging advances in computer vision, we estimate each official's face age and, using observed birthdates, translate these predictions into an interpretable measure of excess apparent aging. Validation exercises based on well-established environmental determinants of health support the interpretation of this measure as capturing meaningful variation in aging rather than photographic artifacts.

Our central finding is that adverse local conditions are associated with visibly accelerated aging among incumbent officials. Officials governing jurisdictions with weaker economic performance and greater distributional stress exhibit higher apparent age. Measures of local macro-financial strain and real-economy distress, such as heightened debt rollover pressure, layoffs, and firm failures, are similarly associated with faster aging, and capital-market signals

consistent with local distress reinforce these patterns. Exploiting the COVID-19 pandemic as an unexpected shock, we show that officials in harder-hit regions experience larger increases in apparent age, consistent with crisis management demands and pandemic-induced downturn risk raising stress exposure in office.

We also provide evidence for an enforcement-based channel. When monitoring intensity increases, proxied by the arrival of central inspection teams and by local anti-corruption enforcement activity, officials' apparent age rises in the period that follows. This pattern suggests that punishment risk and heightened scrutiny carry significant, and previously unmeasured, human-capital costs for the agents tasked with implementing policy. Taken together, the results indicate that both weak economic states and intensified monitoring can elevate stress in office, with measurable consequences that are visible in accelerated facial aging.

These findings have several implications. First, they highlight that political incentive schemes and monitoring regimes may generate meaningful health costs for decision makers, which can shape risk-taking, time horizons, and administrative capacity in ways that matter for economic outcomes. Second, they suggest a feedback loop between local economic stress, governance, and policy implementation: when conditions deteriorate or scrutiny intensifies, the resulting human-capital costs may alter the effectiveness and style of policy responses. More broadly, our results underscore that the costs of governance are not borne only by households and firms; they also accrue to the individuals who govern, and these costs can be quantitatively important. Finally, methodologically, our approach shows how scalable visual measures can open new avenues for studying latent states, such as stress exposure, when traditional health data are unavailable, providing a complementary toolkit for research at the intersection of political economy and finance.

References

Agarwal, N., Moehring, A., Rajpurkar, P., and Salz, T. (2023). Combining human expertise with artificial intelligence: Experimental evidence from radiology. *NBER Working Paper*.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11.

Borgschulte, M., Günzel, F., Liu, Y., and Malmendier, U. (2025). CEO stress, aging, and death. *The Journal of Finance*, 80(6):3401–3442.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*.

Chen, T. and Kung, J. K.-S. (2019). Busting the "princelings": The campaign against corruption in china's primary land market. *The Quarterly Journal of Economics*, 134(1):185–226.

Chen, Y., Xie, Y., You, H., and Zhang, Y. (2018). Does crackdown on corruption reduce stock price crash risk? evidence from china. Journal of Corporate Finance, forthcoming.

Cheng, C.-F., Wu, C.-S., Chan, T.-C., and Wang, S.-H. (2026). Air pollution and accelerated epigenetic aging: Evidence from multiple epigenetic clocks in the taiwan biobank. *Journal of Hazardous Materials*, 501:140852.

Fisman, R. and Wang, Y. (2015). The mortality cost of political connections. *The Review of Economic Studies*, 82(4):1346–1382.

Jia, R., Kudamatsu, M., and Seim, D. (2015). Political selection in china: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4):631–668.

Jiang, J. (2018). Making bureaucracy work: Patronage networks, performance incentives, and economic development in china. *American Journal of Political Science*, 62(4):982–999.

Li, H. and Zhou, L.-A. (2005). Political turnover and economic performance: the incentive role of personnel control in china. *Journal of Public Economics*, 89(9-10):1743–1762.

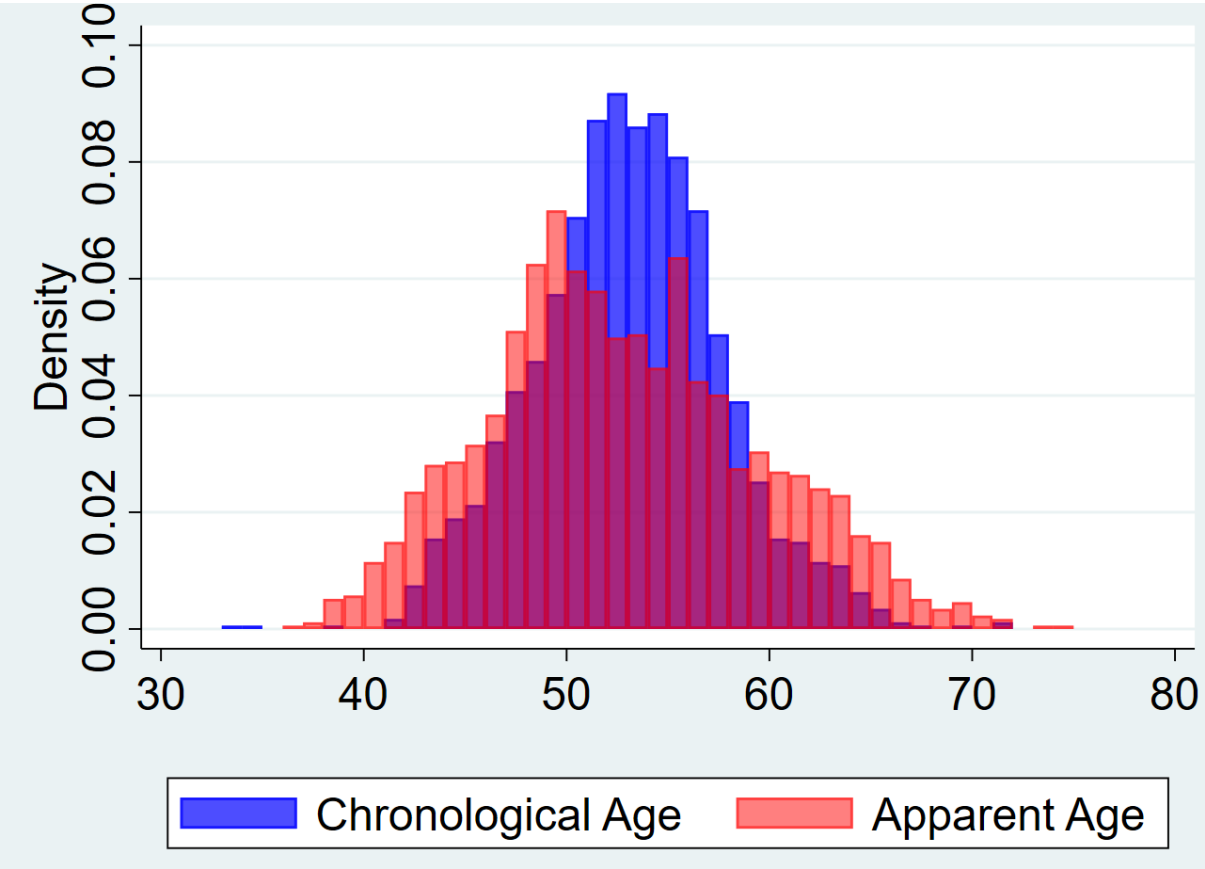
Piotroski, J. D. and Zhang, T. (2014). Politicians and the IPO decision: The impact of impending political promotions on ipo activity in China. *Journal of Financial Economics*, 111(1):111–136.

Rothe, R., Timofte, R., and Van Gool, L. (2015). Dex: Deep expectation of apparent age from a single image. In *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*.

Stoker, J., Garretsen, H., and Spreeuwers, L. (2016). The facial appearance of ceos: Faces signal selection but not performance. *PLOS ONE*, 11(7):e0159950.

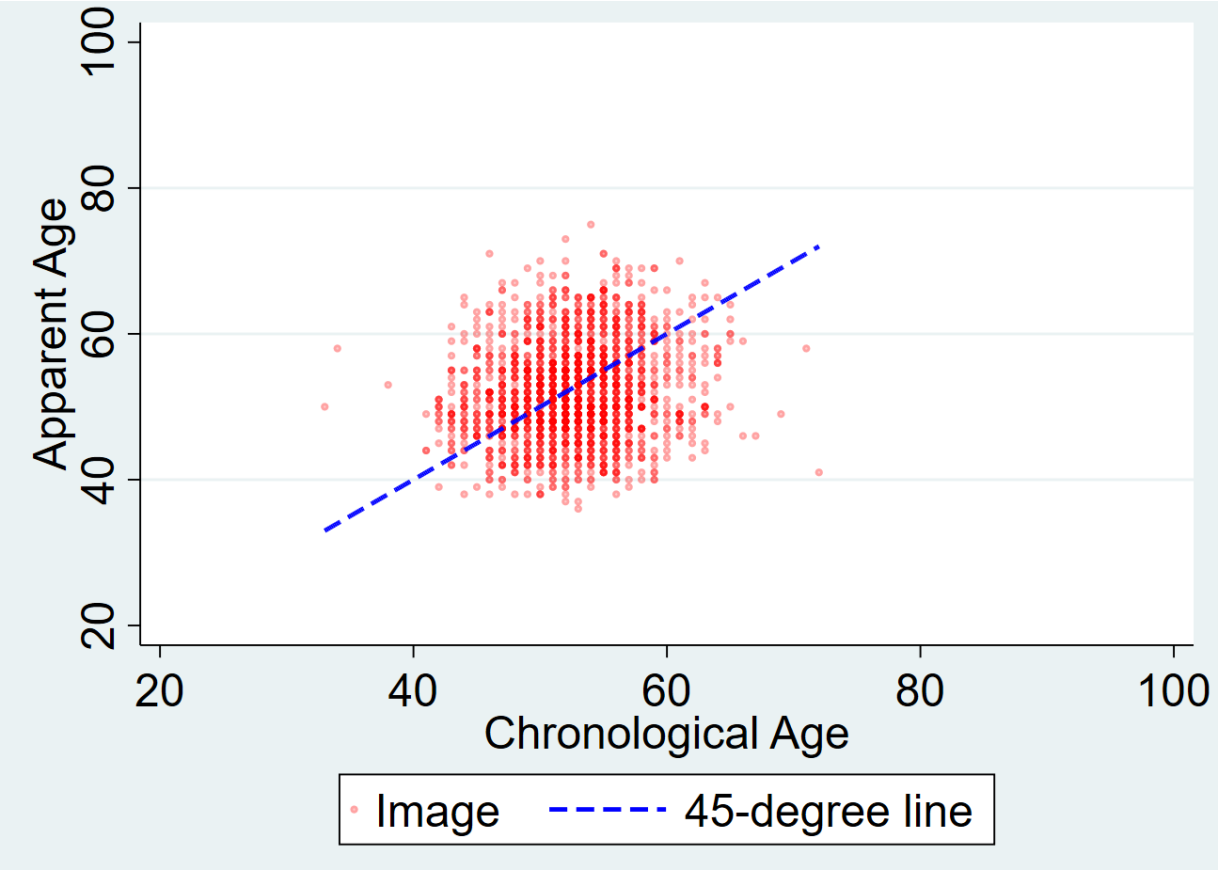
Tucker, L. A. (2021). Fruit and vegetable intake and telomere length in a random sample of 5448 u.s. adults. *Nutrients*, 13(5):1415.

Figure 1: Age verification



The figure plots the bin distribution of chronological age and apparent age.

Figure 2: Age verification



The figure shows the scatter plot of chronological age and apparent age.

Table 1: Summary Statistics

	Mean	Std. Dev.	P25	P75
Panel structure and image coverage				
Number of officials	589			
Number of official-years	4,717			
Number of images (total)	25,412			
Images per official	43.1	35.2	18.0	60.0
Images per official-year	5.4	3.2	3.0	7.0
Share of images from government portals (%)	58.0	22.0	42.0	74.0
Share of images from press coverage (%)	42.0	22.0	26.0	58.0
Outcome variables				
Predicted face age (years)	50.9	6.4	46.7	55.1
Chronological age (years)	52.3	6.8	47.8	56.9
Excess face age (face age − chronological age)	-1.4	3.1	-3.5	0.7
Absolute excess face age (years)	2.5	2.2	1.0	3.6
Indicator: Excess face age > 0	0.34	0.47	0.00	1.00
Official characteristics				
Male indicator	0.86	0.35	1.00	1.00
Years in current position	3.1	1.9	1.6	4.2
Indicator: Mayor (vs. other senior official)	0.38	0.49	0.00	1.00
Economic and distributional conditions (city/province level)				
GDP growth (% yoy)	5.6	2.3	4.2	7.0
Log GDP level	10.9	0.8	10.4	11.4
Inequality index (higher = more unequal)	0.43	0.08	0.38	0.48
Unemployment rate (%)	4.3	1.1	3.6	5.0
Layoff intensity (per 10k workers)	6.8	4.2	3.8	9.2
Firm bankruptcy filings (per 10k firms)	3.1	2.0	1.8	4.2
Macro-financial stress and capital-market exposure				
Local government financing vehicle (LGFV) bond maturities (RMB bn)	24.0	18.0	11.0	34.0
LGFV maturity pressure (maturities / fiscal revenue)	0.32	0.20	0.18	0.44
Number of listed firms headquartered locally	18.0	15.0	7.0	25.0
Total market capitalization of local listed firms (RMB bn)	180.0	220.0	55.0	240.0
SOE stock return minus private stock return (% monthly)	0.20	2.60	-1.30	1.60
Monitoring and enforcement intensity				
Indicator: Central inspection in past 12 months	0.21	0.41	0.00	0.00
Number of investigated corruption cases (city/province-year)	12.0	9.0	6.0	17.0
Validation variables (environment and nutrition proxies)				
$PM_{2.5}$ concentration ($\mu g/m^3$)	41.0	15.0	30.0	52.0
Water/environment quality index	0.63	0.12	0.55	0.72
Fruit production (tons per capita)	0.16	0.12	0.08	0.22
Vegetable production (tons per capita)	0.34	0.20	0.20	0.46

Notes: This table reports summary statistics for the main analysis sample of Chinese local officials (2014–2023). "Predicted face age" is estimated from facial images using the computer-vision pipeline described in the text. "Excess face age" equals predicted face age minus chronological age computed from observed birthdates. Jurisdiction-level covariates are matched to the image date (or event date when available) and aggregated to the official-by-period level when multiple images are observed within a period.

Table 2: Economy and Capital Markets: Macro-Financial Stress and Apparent Aging

	(1)	(2)	(3)	(4)
LGFV bond maturities intensity (RMB bn, upcoming)	0.061***	0.055***	0.048***	0.044***
	(0.018)	(0.017)	(0.016)	(0.015)
Layoff intensity (per 10k workers)	0.032**	0.030**	0.027**	0.025*
	(0.015)	(0.014)	(0.013)	(0.013)
Firm bankruptcies (per 10k firms)	0.046***	0.043***	0.039***	0.036***
	(0.014)	(0.013)	(0.013)	(0.012)
SOE stock return − private stock return (% monthly)	-0.021**	-0.020**	-0.019**	-0.017*
	(0.010)	(0.010)	(0.009)	(0.009)
GDP growth	-0.614***	-0.528**	-0.472**	-0.391**
	(0.203)	(0.271)	(0.228)	(0.197)
Official FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Image quality controls	No	Yes	Yes	Yes
Jurisdiction controls	No	No	Yes	Yes
Source-type controls (gov vs press)	No	No	No	Yes
Observations	4712	4712	4712	4712
Officials	589	589	589	589
R^2	0.41	0.44	0.46	0.48

Notes: The dependent variable is *excess face age* (predicted face age minus chronological age, in years). "LGFV bond maturities intensity" measures upcoming municipal financing vehicle bond maturities (e.g., within the next 12 months) at the city level. "Layoff intensity" and "firm bankruptcies" proxy for real-economy distress. "SOE stock return − private stock return" is the return differential between locally headquartered state-owned enterprises and private firms over the corresponding month; a more negative differential indicates relative underperformance of the state sector. All specifications include official fixed effects and time fixed effects. "Image quality controls" include face size, sharpness, brightness, contrast, and detection confidence. "Jurisdiction controls" include time-varying local fundamentals (e.g., GDP level and growth, inequality) and other controls used in the main specifications. Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3: COVID-19 as Shocks: Difference-in-Differences and Event-Study Evidence

	(1) DiD	(2) DiD	(3) Event-study	(4) Event-study
Post × High COVID intensity	0.312***	0.276***		
	(0.094)	(0.089)		
Post × High lockdown stringency		0.198**		
		(0.081)		
Event-study coefficients: High COVID intensity × event time				
$k = -3$ (3+ periods before)			0.012	0.008
			(0.041)	(0.040)
$k = -2$			0.018	0.014
			(0.038)	(0.037)
$k = -1$			0.006	0.004
			(0.036)	(0.035)
$k = 0$ (event period)			0.121**	0.108**
			(0.052)	(0.051)
$k = +1$			0.214***	0.196***
			(0.060)	(0.058)
$k = +2$			0.287***	0.261***
			(0.070)	(0.068)
$k = +3$ (3+ periods after)			0.301***	0.279***
			(0.078)	(0.075)
Official FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Jurisdiction FE	Yes	Yes	Yes	Yes
Image quality controls	No	Yes	No	Yes
Other time-varying controls	No	Yes	No	Yes
Observations	4712	4712	4712	4712
Officials	589	589	589	589
R^2	0.43	0.46	0.44	0.47

Notes: The dependent variable is *excess face age* (predicted face age minus chronological age, in years). Columns (1)–(2) report difference-in-differences (DiD) estimates where *Post* indicates the pandemic period (e.g., 2020–2021) and *High COVID intensity* indicates jurisdictions in the upper quantile of confirmed case intensity (e.g., cases per 10,000 residents). Column (2) additionally interacts *Post* with a measure of lockdown or mobility-restriction stringency. Columns (3)–(4) report event-study estimates around the onset of local pandemic severity; k indexes event time (e.g., years relative to the first major outbreak or the first date the jurisdiction enters the high-intensity group). The omitted category is $k = -1$ (the period immediately prior to the event), so coefficients are interpreted relative to that baseline. All specifications absorb official, jurisdiction, and time fixed effects; specifications with controls include image-quality measures (face size, sharpness, brightness, contrast, detection confidence) and time-varying local controls (e.g., GDP growth, unemployment, inequality). Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 4: Anti-Corruption Enforcement and Apparent Aging: Inspections and Case Intensity

	(1) DiD	(2) DiD	(3) Event-study	(4) Continuous intensity
Post inspection × Treated jurisdiction	0.228***	0.206***		
	(0.071)	(0.068)		
Investigated/disciplined corruption cases (log)				0.083***
				(0.026)
Event-study coefficients: Treated jurisdiction × event time				
$k = -3$ (3+ periods before)			0.010	
			(0.033)	
$k = -2$			0.014	
			(0.031)	
$k = -1$ (omitted)			0.000	
			(—)	
$k = 0$ (inspection period)			0.062*	
			(0.036)	
$k = +1$			0.142***	
			(0.042)	
$k = +2$			0.197***	
			(0.050)	
$k = +3$ (3+ periods after)			0.213***	
			(0.058)	
Official FE	Yes	Yes	Yes	Yes
Jurisdiction FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Local economic controls	No	Yes	Yes	Yes
Image quality controls	No	Yes	Yes	Yes
Source-type controls (gov vs press)	No	Yes	Yes	Yes
Observations	4712	4712	4712	4712
Officials	589	589	589	589
R^2	0.44	0.47	0.46	0.48

Notes: The dependent variable is *excess face age* (predicted face age minus chronological age, in years). Columns (1)–(2) report a difference-in-differences (DiD) design in which *Post inspection* indicates the period after the central government's inspection team enters a jurisdiction, and *Treated jurisdiction* identifies jurisdictions that are inspected (not-yet-inspected jurisdictions serve as controls). Column (3) reports an event-study around the inspection arrival; k indexes event time (e.g., months relative to the inspection month), with $k = -1$ omitted as the reference period. Column (4) relates excess face age to a continuous proxy for enforcement intensity: the log number of investigated or disciplined corruption cases in the official's city/province-year. All specifications absorb official, jurisdiction, and time fixed effects. "Local economic controls" include time-varying fundamentals (e.g., GDP growth/level, unemployment, inequality) and other controls used in the main specifications. "Image quality controls" include face size, sharpness, brightness, contrast, and detection confidence. Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 5: Mechanism Corroboration: LLM-Based Measures from Officials' Texts

	(1) Excess face age	(2) Excess face age	(3) Excess face age
Urgency tone score (LLM)	0.182***	0.141***	0.128***
	(0.051)	(0.046)	(0.044)
Objective constraint score (LLM)	0.156***	0.119***	0.103**
	(0.049)	(0.045)	(0.044)
Risk emphasis score (LLM)	0.094**	0.072*	0.065*
	(0.040)	(0.038)	(0.037)
Official FE	No	Yes	Yes
Jurisdiction FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Local economic controls	Yes	Yes	Yes
Image/source controls	Yes	Yes	Yes
Speech/report controls	No	Yes	Yes
Observations	4712	4712	4712
Officials	589	589	589
R^2	0.24	0.41	0.44

Notes: The dependent variable is *excess face age* (predicted face age minus chronological age, in years). The independent variables are LLM-based text measures constructed from officials' speeches and government reports: *urgency tone*, *objective constraints*, and *risk emphasis*. Higher values indicate greater urgency/constraint/risk emphasis in the text. Column (1) reports a baseline specification with jurisdiction and time fixed effects. Column (2) adds official fixed effects to identify within-official changes in apparent aging associated with changes in tone. Column (3) further adds speech/report controls (e.g., log text length, document type indicators, and topic fixed effects). Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 6: Robustness: Alternative Face-Age Measure from an LLM-Based Pipeline

	(1) Validation	(2) Economy/Markets	(3) Anti-Corruption	(4) COVID Shock
Environmental determinants (measure validation)				
$PM_{2.5}$ concentration (a7)	0.038***			
	(0.011)			
Water/environment quality index (a112)	-0.163***			
	(0.054)			
Fruit production (a66)	-0.023**			
	(0.011)			
Vegetable production (a59)	-0.016*			
	(0.009)			
Economy and capital markets				
LGFV bond maturities intensity (RMB bn, upcoming)		0.047***		
		(0.015)		
Layoff intensity (per 10k workers)		0.026**		
		(0.012)		
Firm bankruptcies (per 10k firms)		0.037***		
		(0.012)		
SOE stock return − private stock return (% monthly)		-0.016*		
		(0.009)		
Anti-corruption monitoring and enforcement				
Post inspection × Treated jurisdiction			0.198***	
			(0.065)	
Investigated/disciplined corruption cases (log)			0.076***	
			(0.024)	
COVID-19 shock				
Post × High COVID intensity				0.263***
				(0.086)
Post × High lockdown stringency				0.176**
				(0.078)
Official FE	Yes	Yes	Yes	Yes
Jurisdiction FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Local economic controls	Yes	Yes	Yes	Yes
Image/source controls	Yes	Yes	Yes	Yes
Observations	4712	4712	4712	4712
Officials	589	589	589	589
R^2	0.46	0.47	0.48	0.47

Notes: This table replicates the main analyses using an alternative apparent-age outcome constructed from an LLM-based pipeline instead of the baseline computer-vision age estimator. The dependent variable is *excess face age* (LLM-based predicted face age minus chronological age, in years). Column (1) corresponds to the environmental validation tests; Column (2) corresponds to economy and capital-market stress tests; Column (3) corresponds to anti-corruption monitoring/enforcement tests; Column (4) corresponds to COVID-19 shock tests. All specifications include official, jurisdiction, and time fixed effects and the full set of controls used in the baseline tables (local fundamentals and image/source controls). Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 7: Heterogeneity: Career Concerns, Monitoring Intensity, and Capital-Market Exposure

	Career concerns		Monitoring intensity		Capital-market exposure	
	(1) High stakes	(2) Low stakes	(3) High monitoring	(4) Low monitoring	(5) High exposure	(6) Low exposure
GDP growth (yoy, %)	-0.061***	-0.018	-0.052***	-0.014	-0.058***	-0.020
	(0.020)	(0.019)	(0.019)	(0.018)	(0.021)	(0.019)
LGFV maturities intensity (upcoming)	0.072***	0.031*	0.066***	0.027	0.079***	0.029
	(0.023)	(0.018)	(0.021)	(0.019)	(0.024)	(0.020)
Post inspection × Treated jurisdiction	0.266***	0.121*	0.291***	0.082	0.244***	0.093
	(0.084)	(0.071)	(0.089)	(0.069)	(0.083)	(0.072)
Post × High COVID intensity	0.334***	0.158*	0.312***	0.121	0.347***	0.136
	(0.102)	(0.085)	(0.098)	(0.083)	(0.105)	(0.087)
Official FE	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local economic controls	Yes	Yes	Yes	Yes	Yes	Yes
Image/source controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4712	4712	4712	4712	4712	4712
Officials	589	589	589	589	589	589
R^2	0.47	0.45	0.48	0.44	0.48	0.44

Notes: The dependent variable is *excess face age* (predicted face age minus chronological age, in years). The table reports estimates from the baseline specifications estimated separately across subsamples. "High stakes" indicates officials with higher promotion incentives (e.g., closer to promotion-age cutoffs, shorter time-to-retirement, or in more performance-accountable positions), while "Low stakes" is the complement. "High monitoring" indicates jurisdictions with higher baseline monitoring/enforcement intensity (e.g., above-median historical anti-corruption activity or political visibility), while "Low monitoring" is the complement. "High exposure" indicates jurisdictions with greater capital-market presence (e.g., above-median number of listed firms or aggregate market capitalization), while "Low exposure" is the complement. All regressions include official, jurisdiction, and time fixed effects and the full set of controls used in the main specifications, including local fundamentals and image/source controls. Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8: Measure Validation: Environmental Determinants of Apparent Aging

	(1)	(2)	(3)	(4)
$PM_{2.5}$ concentration (a7)	0.042***	0.039***	0.036***	0.034***
	(0.012)	(0.011)	(0.010)	(0.010)
Water/environment quality index (a112)	-0.185***	-0.171***	-0.160***	-0.152***
	(0.060)	(0.058)	(0.055)	(0.053)
Fruit production (a66)	-0.028**	-0.026**	-0.024**	-0.022*
	(0.013)	(0.012)	(0.012)	(0.012)
Vegetable production (a59)	-0.019*	-0.018*	-0.017*	-0.016
	(0.011)	(0.010)	(0.010)	(0.010)
Official FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Image quality controls	No	Yes	Yes	Yes
Jurisdiction controls	No	No	Yes	Yes
Source-type controls (gov vs press)	No	No	No	Yes
Observations	4712	4712	4712	4712
Officials	589	589	589	589
R^2	0.42	0.45	0.47	0.49

Notes: The dependent variable is *excess face age*, defined as predicted face age minus chronological age (in years). $PM_{2.5}$ (a7) measures local air pollution; (a112) captures water quality/overall environmental conditions; (a66) and (a59) proxy for local food quality/availability using fruit and vegetable production, respectively. All specifications include official fixed effects and time fixed effects. "Image quality controls" include face size in the frame, sharpness, brightness, and contrast, as well as detection confidence. "Jurisdiction controls" include baseline local economic controls (e.g., GDP level and growth) and other time-varying characteristics. Standard errors are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 9: Medical Validation: Excess Face Age and Routine Biomarkers

	Dependent variable: Routine health measures (z-score)									
	(1) BMI	(2) SBP	(3) DBP	(4) RestHR	(5) Glucose	(6) HbA1c	(7) LDL	(8) HDL	(9) TG	(10) UricAcid
Excess face age ($\widehat{Age}^{face} - Age$)	0.050**	0.120***	0.090***	0.060**	0.080***	0.055**	0.060**	−0.040*	0.070***	0.065***
	(0.020)	(0.030)	(0.025)	(0.024)	(0.022)	(0.023)	(0.026)	(0.021)	(0.020)	(0.019)
Chronological age	0.030***	0.080***	0.050***	0.020***	0.040***	0.035***	0.020***	−0.015***	0.010**	0.015***
	(0.006)	(0.010)	(0.009)	(0.005)	(0.007)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)
Official FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction-time controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Image-quality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,480	112,480	112,480	109,210	111,905	108,770	110,940	110,940	110,940	111,120
Officials	18,430	18,430	18,430	18,115	18,390	17,980	18,260	18,260	18,260	18,310
R^2	0.72	0.68	0.66	0.63	0.61	0.59	0.58	0.57	0.55	0.60

Notes. Each column reports a separate regression of a standardized biomarker on excess face age and chronological age. Biomarkers are standardized to z-scores within year (optionally within sex-by-year). "Excess face age" is measured in years; coefficients therefore represent the change (in SD units) in the biomarker associated with a one-year increase in excess face age. The full specification includes official fixed effects, year fixed effects, time-varying jurisdiction controls (e.g., $PM_{2.5}$, proxies for environmental quality, economic conditions), and image-quality controls (e.g., pose/lighting/resolution proxies and source controls). Standard errors are clustered at the official level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 10: Medical Validation: Excess Face Age and Diagnoses / Abnormal Findings

	Dependent variable: Indicator for diagnosis/abnormal finding (LPM)							
	(1) Hypertension	(2) Pre-diabetes	(3) Diabetes	(4) Dyslipidemia	(5) Fatty liver	(6) Hyperuricemia	(7) CVD flag	(8) Cancer dx/suspected
Excess face age ($\widehat{Age}^{face} - Age$)	0.015***	0.012***	0.006***	0.011***	0.010***	0.011***	0.002**	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
Chronological age	0.008***	0.004***	0.003***	0.002***	0.001**	0.002***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Official FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jurisdiction-time controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Image-quality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	118,600	118,600	118,600	118,600	105,420	112,980	118,600	118,600
Officials	19,050	19,050	19,050	19,050	17,480	18,290	19,050	19,050
Mean dep. var.	0.26	0.18	0.07	0.23	0.17	0.14	0.03	0.01
R^2	0.41	0.33	0.28	0.36	0.30	0.32	0.19	0.12

Notes. Each column estimates a linear probability model where the dependent variable indicates whether the annual health examination records an abnormal finding/diagnosis. Columns (1)–(6) are common conditions often detected in routine screenings (e.g., hypertension, dyslipidemia), while columns (7)–(8) capture less frequent but higher-severity outcomes. Coefficients represent percentage-point changes in diagnosis probability associated with a one-year increase in excess face age. Controls and fixed effects match Table 9. Standard errors are clustered at the official level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 11: Summary Indices: Excess Face Age and Health Risk Composites

	(1) Cardio-metabolic index	(2) Lipids index	(3) Liver/kidney index	(4) Any chronic condition
Excess face age ($\widehat{Age}^{face} - Age$)	0.095***	0.060***	0.050**	0.028***
	(0.020)	(0.018)	(0.021)	(0.006)
Chronological age	0.040***	0.018***	0.020***	0.010***
	(0.007)	(0.005)	(0.006)	(0.002)
Official FE / Year FE / Controls	Yes	Yes	Yes	Yes
Observations	108,200	110,940	111,120	118,600
Officials	17,920	18,260	18,310	19,050

Notes. Indices are standardized averages of component z-scores (or indicators) constructed from the biomarkers/diagnoses in Tables 9–10. "Any chronic condition" indicates at least one of: hypertension, pre-diabetes/diabetes, dyslipidemia, fatty liver, hyperuricemia. Standard errors are clustered at the official level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.