CSAE Working Paper WPS/2026-09

Individual Demand for Building State Effectiveness

Harry J. Dienes ⓡ Daniel Rogger ⓡ Yetsedaw Emagne ⓡ Giulio Schinaia*

October 2, 2026

Abstract

Investments in the human capital of public sector workers can improve service delivery and state effectiveness, yet little is known about what motivates public servants to undertake such investments. We elicit Ethiopian public servants' willingness-to-pay (WTP) for professional development using an incentive-compatible mechanism, and experimentally vary the anticipated benefits of investing along two dimensions: the perceived managerial environment in which new skills would be deployed, and whether returns to training are framed as accruing to the worker or to others. Demand is positive but low, with median WTP roughly one-tenth the cost of provision. Increasing the salience of an enabling managerial environment raises WTP by an amount approximately equal to the baseline mean one year later. By contrast, emphasising benefits to citizens rather than to the worker's own career leaves demand largely unchanged. These findings highlight that civil servants' demand for professional development behaves like an investment decision governed by expected self-regarding returns, rising when civil servants expect their organisation to let them use new skills.

Keywords: human capital investment, public sector workers, narrative interventions

JEL Classification: J24, M53, O12

Acknowledgments: We gratefully acknowledge financial support from the International Growth Centre, the World Bank's Development Impact Fund, Bureaucracy Lab and Governance Global Practice. We would like to thank Adebabay Abay, Temidayo Taiwo Falade, Prarthna Iyer, Devvrat Raghav, and Getnet Tesfaw for excellent research assistance, and the staff at the Ethiopian Management Institute and the Civil Service University for guidance and support. We also thank Regan Alsup and Synergy Habesha, especially Bewketu Achenef and Moges Tafesse, for the development of the videos used in this experiment. We are grateful to James Berry, Stefan Dercon, Benedetta Lerva, Vinayak Nagaraj, Victor Orozco-Olvera, Sherin Varkey, Torsten Figueiredo Walter, Martin J. Williams, and seminar participants at Advances with Field Experiments 2024, KDI School-WB DIME Development Impact 2024, the World Bank, the Centre for the Study of African Economies Conference 2025, and the Society for Institutional and Organizational Economics 2025, who provided very useful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organisations, or those of the Executive Directors of the World Bank or the governments they represent. Ethics has been approved by the University of Oxford with identifier ECONCIA22-23-21. PAP registration is at the American Economic Association's registry under AEARCTR-0012407 [Dienes et al., 2024]. We randomised author order using the American Economic Association's tool, confirmation code: IGuTQtTyqIu9. *Author affiliations and emails:* Dienes: Cornell University (`hjd45@cornell.edu`); Rogger: The World Bank (`drogger@worldbank.org`); Emagne: University of Messina and The World Bank (`ybekele1@worldbank.org`); Schinaia: University of Chicago (`gschinaia@uchicago.edu`).

Centre for the Study of African Economies
Department of Economics · University of Oxford · Manor Road Building · Oxford OX1 3UQ
T: +44 (0)1865 271084 · F: +44 (0)1865 281447 · E: csae.enquiries@economics.ox.ac.uk · W: www.csae.ox.ac.uk

1 Introduction

A longstanding question in economics concerns who should bear the cost of investments in human capital [Becker, 1964]. Much of the literature has focused on settings where the benefits of human capital investments accrue to either the worker or firm [see, e.g., Acemoglu and Pischke, 1999; Leuven, 2005; Maffioli et al., 2023; Cefala et al., 2024; Abebe et al., 2025]. However, employee investment decisions in organisational contexts where production has wider social benefits – such as in public bureaucracies – have received limited attention. It remains unclear whether public sector workers internalise these broader benefits, what drives their investment decisions, or how much they are willing to pay for such investments. Yet understanding these choices is key to strengthening state capability and informing our broader understanding of human capital investments.

Investments in the human capital of public workers have important implications for state capacity and public service delivery. Competent public servants translate policy into action and manage the personnel and resources that determine government performance [Besley et al., 2022]. In practice, governments invest heavily in capacity building – over half of public employees worldwide report recent participation in training programmes [Schuster et al., 2023], and the World Bank alone spends roughly USD 720 million annually on such initiatives [Belman Inbal and Chin, 2008]. Such investments in professional development can improve public sector outcomes [Jakobsen et al., 2019; Fornasari et al., 2025; Mehmood et al., 2024], complementing recruitment and incentive reforms [Dahis et al., 2025; Ashraf et al., 2020]. Yet unlike in the private sector, where returns to skill accrue largely to the individual, public sector professional development yields social returns through enhanced effectiveness and service quality.

This paper provides the first experimental evidence on how willing public workers are to invest in their own capacity, examining the role of individual motivations and organisational context in shaping these decisions. We conduct a field experiment with 500 mid-career public servants in Ethiopia, using an incentive-compatible Becker–DeGroot–Marschak (BDM) mechanism to elicit willingness-to-pay (WTP) for professional development. We interpret WTP as a revealed measure of underlying demand for investing in one's own human capital – capturing not only monetary investment but also a wider orientation towards professional development, encompassing time, effort, and attention costs.

To understand what drives WTP, we run two randomised experiments that vary features central to human capital investment in public bureaucracies. The first varies the perceived managerial environment in which new skills would be deployed, a channel emphasised by models of dele-

gation. These models predict that workers have stronger incentives to acquire skills when they expect their capabilities to be used productively [Prendergast, 1993; Aghion and Tirole, 1997; Dessein, 2002; Gailmard and Patty, 2007].[1] Before the initial elicitation, participants were randomly assigned to watch video-based narratives portraying a senior manager who either grants autonomy to their team, including over the application of their skills, or emphasises close monitoring. The second varies whether the returns to training are framed as self- or other-regarding in their benefits [Bénabou and Tirole, 2003; Ashraf et al., 2014]. In the follow-up survey a year later, the same participants were randomly assigned to hear one of two recorded messages, emphasising either career advancement or benefits to citizens. The two randomisations are independent.

Demand for professional development is widespread but modest. Almost all participants (94%) are willing to pay a positive amount for at least one professional development activity offered, in a context where most *monetary* costs of in-service training are not borne by participants. However, the median WTP is 100 Ethiopian Birr, or one-tenth of the 1,000 ETB cost of a full day of executive training. Respondents in the top decile of bids, who offered more than 450 ETB, are older, better paid, and have roughly two more years of service than the rest of the sample. Only 6% would pay nothing, and they are not observably distinctive.

We find that the perceived organisational environment in which skills will be used substantially increases demand for training. Participants shown a video portraying an enabling management style bid 87 ETB ($p=0.029$) more for training almost one year later than the control group, who were shown a placebo video about the training institution with no management content. This effect is close to the mean bid in the initial elicitation, although even treated participants' bids remain an order of magnitude below the cost of provision.[2]

Demand appears to reflect self-regarding rather than other-regarding returns to professional development. Three pieces of evidence support this. First, participants wrote a short justification of their own bids at the end of the elicitation, and 80% of these cited benefits to themselves (e.g., career progression) rather than benefits to their team, organisation, or the public. Second, in the one-year follow-up each participant bid on four otherwise identical training programmes that differed only in the framing of who would benefit, and WTP was highest for the training framed around their own career prospects. Third, the audio-based message experiment, randomly

[1] Existing evidence is consistent with this theoretical prediction: managerial autonomy improves firm performance in state-owned enterprises [Kala, 2024], whilst training programmes generate larger productivity gains in firms with more supportive management practices [Diaz et al., 2025].

[2] Participants shown a monitoring management style also bid more than the control group, but the effect is about one-third as large and not statistically significant. The difference between the enabling and monitoring treatments is suggestive ($p=0.11$).

highlighting private career benefits rather than benefits to citizens, increases WTP by 7 ETB ($p=0.064$). This is a small difference relative to both median WTP and the cost of provision.

Thus, demand for professional development in this public bureaucracy behaves like an investment decision governed by expected self-regarding returns. Officials pay for skills they expect to be able to use, and highlighting the social value of their work does little to change what they will pay.

Our paper makes three contributions to the literature. First, we advance the personnel economics of the state [Finan et al., 2017; Rasul and Rogger, 2018] by providing the first incentive-compatible measure of public servants' demand for professional development. While prior work focuses on the *impact* of training [Mehmood et al., 2024; Fornasari et al., 2025; Jakobsen et al., 2019; Seidle et al., 2016], we examine individuals' *willingness to invest* their own time and resources in capacity building. Because in-service training is normally employer-funded in this setting, we interpret willingness to pay as a measure of officials' engagement with capacity building rather than as a spending decision. It therefore provides a common metric for comparing how motivational and organisational features shape investment in state capability.

Second, we provide new evidence on the drivers of human capital investment. While rule-based mandates can increase attendance at required professional development activities, improvements in state capacity depend on public servants adopting and applying new practices [Grindle and Hilderbrand, 1995; Mikkelsen et al., 2022]. A large literature on private sector workers shows that workers invest in training when labour markets allow them to capture its productivity gains [Maffioli et al., 2023; Abebe et al., 2025; Cefala et al., 2024; Acemoglu and Pischke, 1998, 1999]. We instead study public sector workers, who often hold other-regarding motivations [Ashraf et al., 2020], and find little evidence that these motives influence demand for capacity building.

Third, we contribute to the organisational economics literature by showing that perceptions of the managerial environment within one's own organisation substantially shape investment in professional development. While theory predicts that hierarchy and managerial autonomy influence learning and effort [Prendergast, 1993; Aghion and Tirole, 1997; Garicano, 2000; Dessein, 2002], there is little causal evidence on how these factors affect individual investment decisions. We provide experimental evidence showing that willingness to invest in one's own capacity is malleable, and depends on the organisational environment workers expect to return to. Our results are consistent with recent non-experimental findings [Diaz et al., 2025; Bloom et al., 2012] and complement field experiments on incentives, training, and management practices in firms and public organisations [e.g. Bandiera et al., 2011, 2021; Ashraf et al., 2025; Fornasari et al., 2025]. Our results also illustrate that low-cost behavioural interventions

targeted at these features of the public service can change investment decisions in state capacity [cf. Banerjee et al., 2019*b*; Khan, 2025; Bernard et al., 2026].

This paper continues as follows. Section 2 introduces the study context. Section 3 presents the experimental design and data collection. Section 4 shows our estimated demand curves for professional development. Section 5 explores mechanisms driving demand using our experimental interventions. Section 6 concludes.

2 Experimental Setting

Ethiopia's large public sector makes it a valuable setting to better understand human capital investment decisions in public administration. It is Africa's second most populous country, with 130 million people and a public service that employs half of paid workers [Baig et al., 2021]. The public sector wage premium is 12%, placing Ethiopia mid-range globally and suggesting typical patterns of selection into government jobs. While government effectiveness is relatively weak, Ethiopia is not an outlier, ranking 102$^{\text{nd}}$ out of 141 countries in the 10-year average Government Effectiveness index [Kaufmann and Kraay, 2023].

Capacity building is core to Ethiopia's public sector and often takes the form of in-service training. In a 2024 survey of Ethiopian public servants, separate from the experimental sample in our study, 43% stated they had received training in the past year, reflecting large investments of time and financial resources.[3] The training is primarily on public service laws and regulations (73%), administrative processes (32%), ethics (29%), and ministry-specific subjects (27%). This focus reflects the rationale of aligning staff on procedures to improve overall effectiveness, and this in-service training is viewed positively: 74% of officials rate it as highly relevant and 92% report that it improves their productivity. Importantly, the monetary costs of in-service training are typically borne by the employer, with workers often receiving per diem payments to cover attendance costs rather than contributing financially themselves.

In many bureaucratic settings, public servants may have low extrinsic incentives to invest in skills.[4] In the same 2024 survey, only half of respondents believed they would be promoted for strong performance, and formal qualifications were rated as a far more important predictor of advancement than skill upgrading. Echoing public services around the world, capacity building

[3]Details on the surveys cited in the paper are in Appendix Section E.

[4]Incentives to invest in capacity building may differ in distinctive and important ways between the public and private sectors. In the private sector, skill acquisition is commonly rewarded through performance-based promotions and wage gains. While similar mechanisms could operate in the public sector, recent work shows that public servants often face promotion systems driven more by tenure or political connections than by demonstrated ability [Deserranno et al., 2025; Bertrand et al., 2020; Colonnelli et al., 2020].

in Ethiopia's public service focuses on service-specific skill upgrading and has limited linkages to extrinsic rewards.

If bureaucratic management practices support the generation of value from human capital investments, the incentives to invest in that capital will likely be higher. A separate 2016 survey of Ethiopian public servants that asked about management practices points to limited staff autonomy and a relatively more prevalent top-down management style. It finds 86% of directors reporting that decisions are driven or made exclusively by senior management, and 94% indicating no systematic mechanism for staff to contribute ideas or identify improvements. These patterns suggest that the expectation of being able to apply new skills may be a meaningful constraint on the value workers place on capacity building, and one that varies with how individual managers operate.

3 Estimating Demand for Capacity Building

This section describes our experimental approach to measuring demand for training among Ethiopian civil servants using a Becker-DeGroot-Marschak mechanism. The study comprises two rounds of data collection from the same sample: an in-person elicitation at the Civil Service University (CSU) in November 2023, and a follow-up phone survey approximately one year later. Each round paired a willingness-to-pay elicitation with a narrative intervention designed to shift perceptions of the benefits of investing in training.

Narrative interventions aim to shape beliefs and preferences by encouraging individuals to identify with characters and situations that feel personally or professionally relevant [Banerjee et al., 2019*a*; Berg and Zia, 2017]. The first, delivered in person, implicitly varied the managerial environment in which new skills would be used; the second, delivered by phone, explicitly varied whether the returns to training were framed as self- or other-regarding. The two treatments were independently randomised, and we achieved full compliance in delivering both since each was brief and integrated into data collection. Appendix Tables A.4 and A.5 (Panel B) confirm balance in respondent characteristics, including income, across treatment groups. Figure 1 illustrates the full study timeline and design, and Appendix Section B provides further detail on both interventions.

3.1 Round One: In-Person Elicitation at the Civil Service University

We measure individual demand for professional development activities through a 'framed' field experiment [Harrison and List, 2004]. In November 2023 we elicited WTP for three professional development opportunities from 500 mid-career public servants, each of whom first watched one of three short films varying the managerial environment.

Sample. Our sample comprises Ethiopian public servants who are broadly representative of the national public sector. Participants were enrolled at CSU in Addis Ababa, with roughly half employed in a *woreda*-level government office.[5] They are predominantly male, in their mid-thirties, have nearly a decade of public service experience, and were pursuing postgraduate degrees at CSU. Such degrees, such as 'accounting', are distinct from in-service training and are typically not related to the service-oriented topics of in-service training. Observable characteristics closely match a nationally representative 2016 survey of Ethiopian civil servants (Appendix Table A.3), suggesting that the sample is plausibly representative of the public administration, excluding street-level staff such as teachers and doctors.

Professional development activities. Participants could purchase three activities as part of the experiment: a one-day executive training, half-day one-to-one professional coaching, and a one-day work shadowing observing a senior official in their role. The description of these activities emphasised practical managerial skills that aimed to improve service functioning, rather than the regulatory and procedural content that dominates standard in-service training. Drawing on the World Bank's Bureaucracy Lab experience, we designed these opportunities to develop managerial competency through sharing international experience and direct observation of senior colleagues. All three activities offered completion certificates.

The management videos. Before bidding, participants individually watched one of three short fictional movies (under 15 minutes) that made salient a key feature of the organisational environment: how individuals are managed. Management practices shape the returns to human capital investments by determining how effectively workers' skills are deployed, and the practices depicted have been empirically linked to bureaucratic behaviour [e.g., Rasul and Rogger, 2018; Rasul et al., 2021; Bandiera et al., 2021].[6] The movies are set in the Ethiopian public service and feature the same cast and plot, and were developed with an experienced Ethiopian production company and international media advisor to combine realism with our experimental components. Two treatment arms highlighted either an *enabling* or a *monitoring* managerial environment through specific scenes involving a middle manager, their senior manager, and their team, designed to trigger reflection on the participant's own professional setting and on how supported or constrained they feel by their superiors. Enabling managers support autonomy through trust and informal feedback; monitoring managers emphasise oversight and formal accountability. The control video introduced the history and services of

[5] *Woredas* are administrative units comparable to US counties.

[6] The two styles depicted – enabling and monitoring – were further identified through factor analysis of survey data from Ethiopia's public sector based on questions derived from the World Management Survey for bureaucracies.

the training institute, designed as a placebo with no reference to management practices.

Eliciting willingness to pay. We adapt a Becker-DeGroot-Marschak (BDM) mechanism to our setting, building on applications of the method in similar contexts [e.g., Berry et al., 2020; Cole et al., 2020; Burchardi et al., 2021; Maffioli et al., 2023]. Participants first stated their maximum WTP for each activity, in randomised order; if a participant reported a WTP of zero, we asked for the minimum they would require to participate. Each participant was then assigned one activity (with probabilities: shadowing, 1/9; training, 1/3; coaching, 5/9) and a price, drawn from a mixed uniform distribution between -500 and 3,000 Birr (with 80% of prices below 1,500 Birr). Assignments and prices were pre-randomised but concealed from both participants and enumerators until bidding was complete. If the drawn price for the selected activity fell below the participant's bid, they purchased the activity at that price; otherwise they did not. Finally, participants wrote a short (150-character) reflection explaining their bidding rationale, which we coded as reflecting either self-regarding or other-regarding motivation.

Implementation and fidelity. We took several steps to ensure participants understood the mechanism, and that responses reflected true valuations. First, we designed both the activities and their presentation based on feedback from focus groups to ensure appeal, and presented them as delivered by the country's leading management institute in partnership with international organisations. To mitigate potential opportunity costs of missing work, participants were informed the activities would take place at a central location on a weekend two weeks later, so that attendance would not require time away from work.[7] Second, we pre-verified that participants could pay, excluding thirteen of the 513 sampled individuals who lacked a payment method and yielding a sample of 500. Third, participants completed a practice BDM round involving a benchmark item, a pen; 86% bid at or below market price, indicating strong comprehension of the mechanism. Finally, to maintain incentive compatibility, we emphasised that payment would occur on the same day, and nearly all participants who were required to pay did so (with only four exceptions). Appendix A provides further implementation details and fidelity checks, including the arrangements for administering the videos during dedicated class time; our pre-analysis plan contains the protocol and the description of the professional development opportunities [Dienes et al., 2024].

3.2 Round Two: Follow-Up Phone Survey

Re-contact and elicitation. We conducted a follow-up phone survey approximately one year after the initial elicitation, allowing participants to have experienced a full annual cycle of public service. This timing enables us to examine how perceptions of the managerial

[7]Work shadowing took place over two half-days in the working week following the training and coaching.

environment shape investment decisions after participants have returned to their workplace and experienced that environment directly. We successfully re-contacted 88% of the original sample (438 out of 500), with no evidence of differential attrition by age, gender, government tier, location, tenure, organisational experience, or income (Appendix Tables A.4 and A.5, Panel A). Participants were reminded of the original BDM process, told their initial bids for the training activity, and informed about a potential new round of training. They were told that their bids would determine their eligibility and payment if the training proceeded, and that they would pay no more than the amount they stated. While this was not a direct BDM elicitation, the process was consistent with respondents' earlier experience and referred explicitly to future payments, making bids behaviourally incentivised at the point of elicitation, though this additional round was ultimately not implemented.[8] Recent evidence suggests that incentivised and hypothetical WTP distributions are broadly similar [Dizon-Ross and Jayachandran, 2022].

Training framings. We elicited WTP for four versions of training, each framed to highlight a different objective. The four framings, presented in random order, emphasised (i) skills to support the respondent's own career advancement in the public or private sector; as well as three other-regarding framings: (ii) skills specific to improve the productivity within the respondent's organisation; (iii) skills to enhance coordination and performance across the public sector as a whole; and (iv) skills to increase the respondent's impact on citizens. All four were presented to every respondent, so comparisons across framings are made within individual.

The audio message. Before restating their bids, respondents were randomly assigned to hear one of two recorded messages from the Civil Service Commission, which oversees public sector workers. One emphasised how professional development advances personal growth and career progression across government and non-government roles; the other, read by the same narrator, emphasised its social returns and potential to improve outcomes for the public. Both were framed as official communications from the Commission and introduced as such by the enumerator. Appendix Table A.1 reproduces the text.

4 The Demand for Capacity Building

This section presents our baseline estimates of public sector workers' demand for capacity building. We first document how demand varies across training, coaching, and shadowing opportunities (Section 4.1) and explore individual characteristics and stated motivations that

[8]The phone-survey responses are not influenced by prior attendance at a professional development activity offered in the experiment, consistent with accurate valuation: using the random price as an instrument, we find that attendance has only a small, not statistically significant effect on WTP in the phone survey. See Appendix Table A.13.

correlate with willingness-to-pay (Section 4.2).

4.1 Low but Mostly Positive Demand for Capacity Building

Figure 2 shows the inverse demand curve resulting from the BDM exercise. The figure indicates the share of subjects that have a WTP less than p, for $p \in [-300,700]$, across coaching, shadowing, and training. Median bids were modest: 100 Birr for training, 90 for shadowing, and 80 for coaching – equivalent to 1.40 USD, 1.60 USD and 1.80 USD respectively, or the value of 3–4 pens. The median bid is equivalent to about 1% of median monthly income and only 2% of participants were willing to pay more than 10% of their monthly income for any activity.[9] This pattern is consistent with findings from other settings, where socially beneficial goods face low demand at even small positive prices [Cohen and Dupas, 2010; Mobarak et al., 2012]. Demand for training was slightly higher than for coaching or shadowing, though all curves show similar patterns.[10] With a full day of training at the Ethiopian Management Institute costing approximately 1,000 ETB (17 USD), our estimated demand for capacity building is low relative to the cost of provision.

Despite low overall demand, the high share of positive bids suggests that participants placed positive value on the professional development activities offered. Almost all participants (94%) bid a positive amount for at least one activity, while the remainder asked to be compensated to attend. This finding is noteworthy since the costs of most in-service professional development activities in Ethiopia's public sector are not borne by the participants. Because the activities were scheduled on a weekend to avoid taking participants away from their duties, attendance came at the cost of personal time rather than work time. Respondents in the top decile of the distribution of WTP reported more than 450 ETB – enough to cover about half the training cost – and show relatively inelastic demand. These respondents are more experienced (10.7 vs. 8.9 years, $p=0.031$), earn higher monthly salaries (roughly 9,000 vs. 7,700 ETB, $p=0.040$), and are older (34 vs. 32.2 years, $p=0.029$) than the rest of the sample. At the other end, only 6% were unwilling to pay anything. This group does not differ significantly in characteristics such as age, gender, wage, or government tier.[11]

Low willingness-to-pay is unlikely to reflect perceptions of low quality of the activities. At least qualitatively, participants regarded the activities as high quality: attendance among purchasers

[9] The nominal exchange rate was 60 ETB = 1 USD. At the time of the study, the market value of a pen was 29 Birr.

[10] Differences between training and the other activities are statistically significant (Kolmogorov-Smirnov test $p=0.00$), but economically small – approximately the value of a pen. Correlations in individual bids across activities are positive and high (0.65–0.7).

[11] Overall, income is the main observable characteristic positively associated with demand for professional development (Appendix Figure A.1).

was 80% (with no conditionality to attend for those that had "paid" a negative price), with all attendees expressing high interest for the certificates of completion, and providing positive feedback throughout.

4.2 Qualitative Motivations and Demand

We leverage participants' qualitative reflections to describe how self-regarding or other-regarding motivations correlate with bid values. Each participant wrote an open-ended reflection on what drove their bids for continuous professional development at the end of the first elicitation. Research assistants coded the reflections into two categories, highlighting either self-regarding benefits (e.g., career concerns) or other-regarding benefits. Because reflections were written after the video intervention, we treat them as descriptive rather than as a pre-treatment measure of motivation; we find no significant effect of the video treatments on the coded categories.

Eighty percent reflected on their bids in self-regarding terms, focused on career and status. One respondent stated that "training is for promotion and other performance... to develop my career position" (respondent 1161); another that training "improves my status, so I need it" (1430). Some looked to the private sector, with one emphasising that "training is important for working in any company in the future" (1484).[12]

A minority reflected on benefits to their team, the civil service, or society. Respondents with other-regarding reflections had maximum bids 28 ETB higher across all activities than those with self-regarding reflections, though the difference was not significant ($p=0.291$).[13] Some emphasised organisational benefits beyond their own skills – that training "is not only needed for me, but for my team... with cooperative work" (1196) – while a few cited broader societal benefits, such as "contributing to the development plan of my country" (1256).

5 Experimental Evidence on Drivers of Demand

We study drivers of demand through two experiments that use narratives to vary the perceived benefits of training. The first implicitly shifted perceived benefits by varying the organisational environment through videos portraying different managerial styles. The video treatments were implemented during the initial elicitation at CSU (see Figure 1). The second experiment explicitly framed training as either advancing careers or benefiting society, with treatment implemented during the phone-based follow-up. Our analysis follows pre-analysis plans reg-

[12]Quotes have been edited for readability; originals are in the replication files.

[13]The gap is not different when looking at training specifically (34 ETB; $p=0.238$) and when including controls, including income (30 ETB for all and 34 ETB for training). 11% of the sample did not report a reflection or could not be coded.

istered before data collection ended; Appendix C describes other pre-registered analyses and any changes from the original plan.

5.1 Highlighting an Enabling Organisational Environment Substantially Increases Demand

We examine whether features of the organisational environment influence demand by varying exposure to different managerial styles. As described in Section 3.1, respondents were randomly assigned to watch one of three video treatments prior to submitting their first bid. The videos presented a narrative of a public sector manager adopting either an enabling management style; a monitoring, directive approach; or a placebo without reference to management style. Because managers shape organisational incentives through their routine practices, making this aspect of the environment salient may influence public sector workers' willingness to invest in their own skills, as they weigh how those skills could be applied and recognised within the organisation.

To estimate the impact of the narrative intervention over time, we estimate variations of the following regression:

$$Y_{t,i} = \mathbf{T_i}\boldsymbol{\beta} + \mathbf{S}\boldsymbol{\delta_1} + \varepsilon_i \quad \{t = CSU, Foll\} \tag{1}$$

where $Y_{t,i}$ is the WTP for a training with a specific emphasis elicited at round t. $Y_{CSU,i}$ is the WTP elicited in the initial in-person bid, and $Y_{Foll,i}$ denotes the bid elicited during the follow-up phone survey, before the audio experiment. $\mathbf{T_i}$ is an indicator vector for the nature of the video treatment (video with enabling manager, video with monitoring manager, and the placebo movie). $\mathbf{S}$ is a vector of fixed effects for training framing, which is not included for the CSU-elicited WTP.

The timing of the two rounds distinguishes immediate from delayed responses. The videos highlighted a feature of the organisational environment – the extent to which managers support workers in applying their skills – rather than providing new information about the returns to training. Such perceptions are likely to shape investment decisions only once respondents return to work and experience their management environment directly, so any effect would be expected to emerge with a delay.

While the video narratives had no immediate effect on bids in the initial elicitation, we find a substantial impact of the enabling manager narrative nearly one year later. Table 1 presents results across elicitation rounds: CSU baseline (Column 1) and the follow-up survey (Columns 2 and 3). The coefficients on "Manager Monitors Movie" and "Manager Enables Movie" capture the effect

of each video relative to the control. We observe no treatment effect immediately after the intervention, potentially because the features highlighted in the video may have become salient once respondents returned to work (Table 1, Column 1).[14] By the follow-up, exposure to the enabling manager narrative substantially raises WTP by approximately 87 ETB ($p=0.029$), suggesting that highlighting an enabling management environment increases the perceived value of capacity-building (by close to the baseline mean WTP; and by 9 percent of the cost of a day's training). In contrast, the monitoring management narrative has a smaller and not statistically significant effect, compared to the placebo video. The difference between the treatment videos has a p-value of 0.11. Column 3 tests heterogeneity by respondent motivation, interacting treatments with a "social reflection" indicator. While statistical power is limited, we find that other-regarding respondents respond less to these narratives, possibly due to their already higher baseline WTP.[15]

To interpret this effect, we present additional analyses of the potential channels for the video-treatment to affect willingness-to-pay. The lack of an immediate effect in the initial elicitation suggests that respondents continued to reflect on the video afterwards. They may draw on existing work experience, or embed capacity building in a new narrative of their management environment on returning to work.[16] Consistent with this idea, we find suggestive evidence that work experience amplifies the effect of the videos (Appendix Figure A.5). The video treatments may also operate through increasing the perceived promotion returns to capacity building, though we do not find evidence for this channel. Appendix Table A.11 shows no effect on the expected impact of training on perceived promotion chances. Our interpretation is that the enabling video raised WTP by reinforcing the expectation that workers would be able to apply new skills with more autonomy.

5.2 Demand Reflects Self-Regarding Rather Than Other-Regarding Returns

The preceding result suggests that officials treat professional development as an investment whose value depends on whether they expect their organisation to let them use new skills. But officials may value that investment for self-regarding reasons, such as expected career benefits,

[14]The different survey mode across rounds could affect the level of follow-up bids, but it cannot explain the difference across arms.

[15]As robustness checks, we include controls such as income – which, though balanced across groups, significantly predicts bids (Appendix Table A.10). Following Dizon-Ross and Jayachandran [2022], we also control for a benchmark good (pen). Full disaggregated results in Appendix Table A.8 are also consistent.

[16]Impacts of narrative interventions that target social dynamics, such as the managerial environment at work, may activate later as social reinforcers trigger reflection. This is consistent with other studies that find that video-based behavioural interventions can have effects after a period of time, or that grow [e.g., Peterman, 2025; Bernard et al., 2026], in contrast to interventions that target knowledge [e.g., Orozco-Olvera et al., 2019].

or because they internalise the social returns to their work. We attempt to disentangle these two drivers, and find evidence pointing towards self-regarding motives.

During the follow-up, participants heard one of two recorded messages from the Civil Service Commission, emphasising either career advancement or benefits to citizens, before restating their bids. We compare post-message bids across these two arms, controlling for each respondent's pre-message bid.[17] Since the pre-message bid already captures any effect of the video treatment, we use it as a control.[18] The career-focused message increases WTP by 7 ETB ($p=0.064$), or 35% of the control-arm mean, relative to the social-impact message. This effect is small relative to the median bid of 100 ETB and the 1,000 ETB cost of provision. The effect does not vary with the framing of the training, as shown in Table 2, Column 2.

The experimental evidence is consistent with two other findings. First, when asked to bid for training that emphasised different goals, officials bid more for training when the stated benefits accrue to themselves. They bid 23–27 ETB more for the version emphasising their own career prospects than for any of the three versions emphasising benefits to others (Appendix Figure A.3). These comparisons are within respondents because each official bid on all four variants. Second, officials' written justifications point in the same direction. As noted in Section 4.2, eighty percent refer to benefits to themselves rather than to their team, organisation, or the public. Taken together, these patterns provide little evidence that demand for capacity building is driven by internalised other-regarding returns.

6 Conclusion

Investments in public sector human capital are central to strengthening state effectiveness, as institutional performance rests on the skills of individual workers. But governments can mandate training without securing the engagement it requires: public servants must still absorb and apply what they learn, and where their incentives to bear those costs are limited, capacity building may do little for governmental effectiveness. This paper tests the channels shaping those incentives in the Ethiopian Civil Service.

Our findings shed some light on who might bear the cost of investing in human capital for state effectiveness. We find modest baseline willingness-to-pay for the professional development opportunities. In our setting, individual demand for capacity building in the public

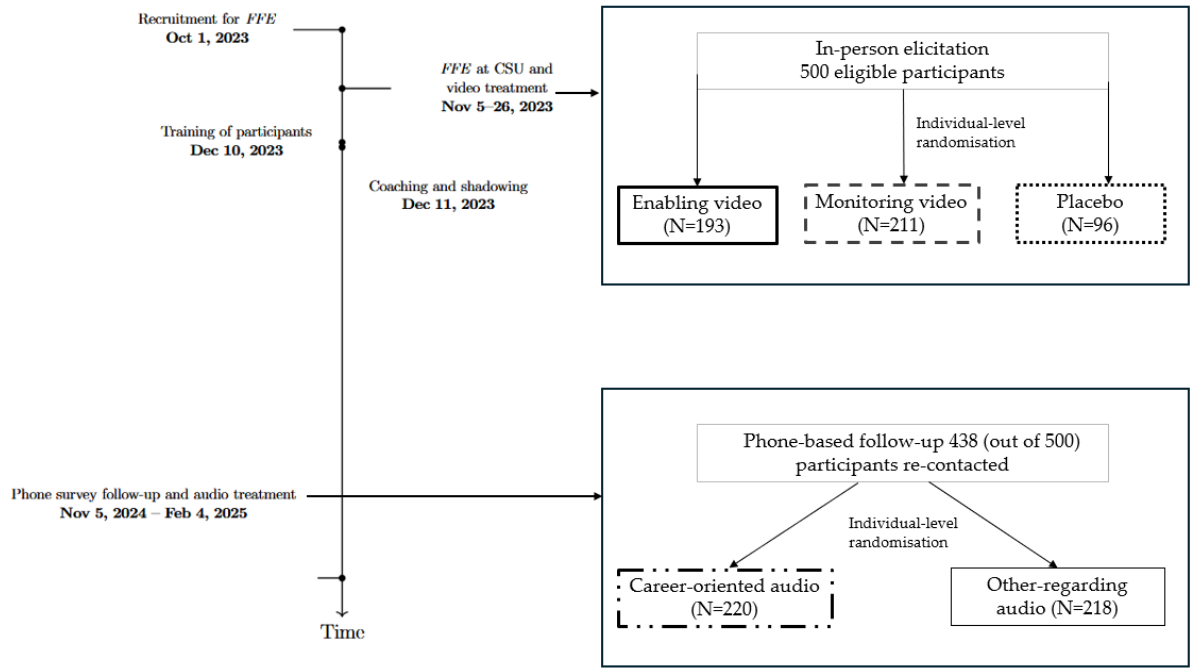
[17] Our pre-analysis plan specified difference-in-differences when bids were highly autocorrelated and ANCOVA otherwise. Appendix Table A.6 reports both specifications and the results are similar.

[18] The audio treatment is orthogonal to the video treatment, with $p=0.27$. We also do not find evidence of any interaction effects between the audio and video treatment arms (Appendix Table A.12).

sector appears to be driven primarily by expected self-regarding rather than other-regarding returns. Although officials may internalise some broader social benefits, our results suggest that willingness-to-pay for capacity building contains a meaningful self-regarding component in the perceived usefulness of new skills. Governments could increase demand for professional development by creating work environments in which newly acquired skills are expected to be used. At the same time, even the most motivated officials are unwilling to invest enough to cover the full cost of professional development, implying a central role for government financing regardless of how attractive such investments may be to individuals.

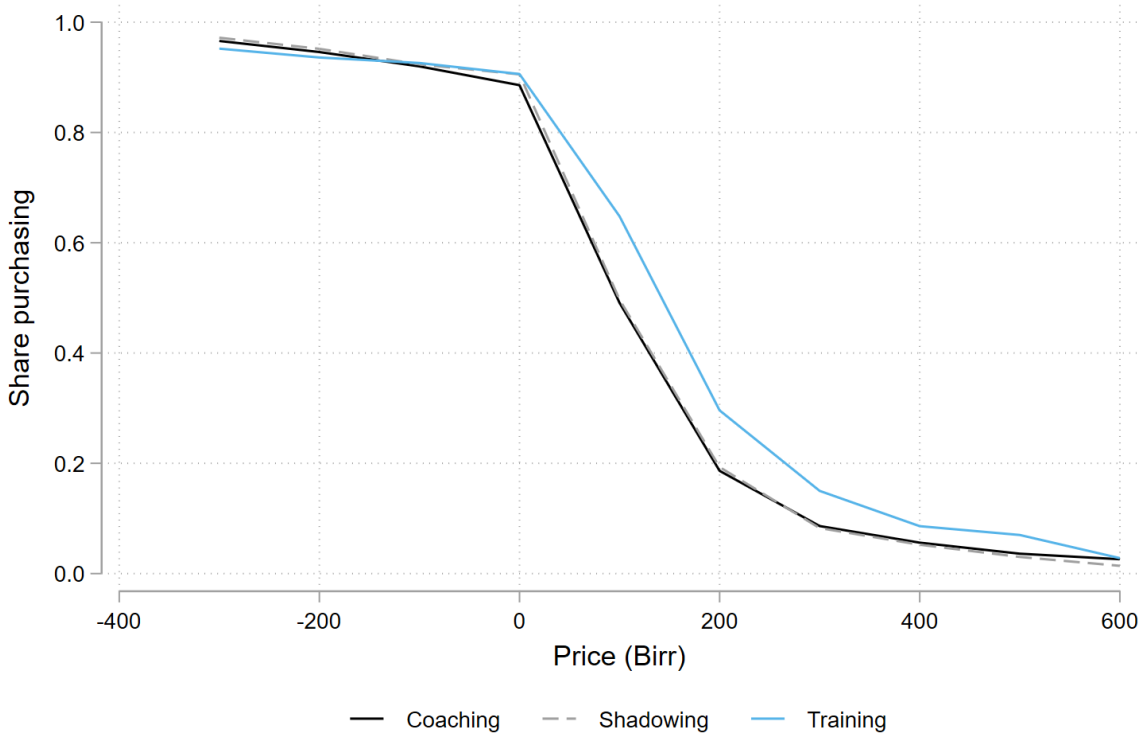
We leave several avenues for future research into the microfoundations of state effectiveness. Future work could examine how professional development that emphasises different goals – career, social, or institutional – or different applications affect downstream outcomes in learning and performance. Another promising direction involves studying how to shift managerial behaviour in a way that enhances the incentives of workers to invest in their capabilities. Viewing management as more or less supportive of workers using their skills may change organisational norms, motivation, and ultimately service delivery. Motivating stronger investments in public service capabilities may be a function of existing capabilities in ways that are not yet well understood.

Figures



Note: Timeline of the framed field experiment (FFE) at Civil Service University (CSU). The figure shows the sequence of data collection and treatments. Round one of the experiment took place in person at CSU in November 2023 (upper box) and included the video intervention; round two of the experiment involved the follow-up WTP elicitation that was conducted by phone approximately one year later (lower box) and included the audio intervention. Before the final elicitation at CSU, participants were randomly exposed to the different video treatments. WTP was elicited using a Becker–DeGroot–Marschak (BDM) mechanism: participants stated their maximum WTP, a binding price was drawn, and activities were delivered to purchasers. In the phone-survey follow-up, participants were reminded of the original BDM procedure and informed that their stated bids would determine eligibility and payment if a new round of training were to proceed. Participants were also randomised into audio-treatment conditions varying the messaging (Career-oriented vs. Other-regarding) and the number of participants from the respondent's organisation. The boxes on the right describe the randomisation design of the video- and audio-based treatments.

Figure 1: Study Timeline and Design



Notes: Estimates of willingness-to-pay (WTP) in Ethiopian Birr (ETB) and share of the sample purchasing at different elicited prices. Data collected during the in-person experiment at Civil Service University (CSU) and estimated using the sample that also completed the follow-up phone survey ($N = 438$). Bids winsorised at the 99th percentile.

Figure 2: Inverse Demand Curve for Coaching, Shadowing and Training

Tables

Table 1: Treatment Effects of the Management Videos on Willingness to Pay

	(1) CSU	(2) Follow-Up	(3) Follow-Up
Treatment: Manager Monitors Movie	27.16	33.27	26.52
	(33.09)	(38.01)	(43.00)
Treatment: Manager Enables Movie	-5.40	87.18**	101.80**
	(33.98)	(39.72)	(45.58)
Social Reflection	31.38	20.07	27.46
	(31.44)	(40.53)	(62.63)
Manager Monitors × Social Reflection			43.96
			(92.26)
Manager Enables × Social Reflection			-70.59
			(91.86)
Clusters		438	438
Observations	438	1752	1752
Predicted Control Mean	103.44	-31.64	-32.44
Monitor vs. Enable p-value	0.21	0.11	0.04

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training. "Treatment: Manager Monitors Movie" indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of a monitoring and directive senior manager. "Treatment: Manager Enables Movie" indicates they were shown a movie that showed an enabling senior manager. "Social Reflection" is a binary that takes the value 1 when the respondent provided a socially-oriented reflection. Column (1) reports the WTP measure collected in person at Civil Service University (CSU). Column (2) and (3) use the WTP measure from the follow-up phone survey about one year later, before the audio treatments were delivered (Follow-Up). Models (2) and (3) control for training type. Where social reflection is missing, the value is imputed with the median and an indicator of missingness included. Predicted control mean is the predicted WTP of the control group, given the model controls, which is similar to the unconditional mean. Standard errors in parentheses are clustered across bids at the individual-level for models (2) and (3). For model (1) standard errors are robust. Bids winsorised at 99th percentile. Monitor vs. Enable p-value displays the test statistics comparing the coefficient for Monitor Treatment against Enable treatment. Model (3) shows this p-value for respondents with individual (i.e., not "social") reflections, comparing coefficients in the first two rows.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 2: Relative Treatment Effects of Career Message

	(1)	(2)	(3)
Treatment: "Training is good for your career"	6.994*	8.264**	9.600**
	(3.771)	(4.089)	(4.310)
Other-Regarding Training	2.163	3.013	2.172
	(2.260)	(2.608)	(2.264)
Treatment x Other-Regarding Training		-1.692	
		(4.513)	
Social Reflection			4.807
			(6.235)
Treatment x Social Reflection			-13.84
			(9.535)
Clusters	438	438	438
Observations	1752	1752	1752
Predicted Control Mean	19.695	19.695	19.588

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training. Regressions control for the pre-treatment WTP elicited during the in-person elicitation in 2023 and the phone-survey elicitation conducted before the treatment. "Treatment" is an indicator for whether participants received a message emphasizing career advancement benefits before stating their final WTP bids. "Other-Regarding Training" is an indicator for whether the training content emphasised benefits to the team, organisation, or society at large. "Social Reflection" is an indicator for whether the participant, in an open-ended response explaining their WTP, focused on benefits to the team, organisation, or society (rather than self-regarding benefits). Interaction terms are combinations of these indicators. Where social reflection is missing, the value is imputed with the median and an indicator of missingness included. Predicted reference mean is the predicted WTP of the group treated with the other-regarding audio message, given the model controls – similar to the unconditional mean. Video treatment is independently randomised; pre-phone treatment WTP elicited in phone survey is post-video treatment. Standard errors in parentheses are clustered across bids at the individual-level. Bids winsorised at 99th percentile.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process.

During the preparation of this work the authors used ChatGPT and Claude to condense and edit text, to check internal consistency across sections, and to identify redundancy. These tools were not used for data collection, analysis, or interpretation of results. The authors reviewed and edited all content and take full responsibility for the content of the published article.

References

Abebe, Girum, Stefano Caria, Pascaline Dupas, and Marcel Fafchamps. 2025. "Competition and Management Upgrading: Experimental Evidence from Ethiopia." *National Bureau of Economic Research Working Paper 33886*, 1–72.

Acemoglu, Daron, and Jorn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113(1): 79–119.

Acemoglu, Daron, and Jorn-Steffen Pischke. 1999. "Beyond Becker: Training in Imperfect Labour Markets." *The Economic Journal*, 109(453): 112–142.

Aghion, Philippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy*, 105(1): 1–29.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics*, 120: 1–17.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 2020. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review*, 110(5): 1355–1394.

Ashraf, Nava, Oriana Bandiera, Virginia Minni, and Luigi Zingales. 2025. "Meaning at Work." *National Bureau of Economic Research Working Paper 33843*, 1–66.

Baig, Faisal Ali, Xu Han, Zahid Hasnain, and Daniel Rogger. 2021. "Introducing the Worldwide Bureaucracy Indicators: A New Global Dataset on Public Sector Employment and Compensation." *Public Administration Review*, 81(3): 564–571.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2011. "Field Experiments with Firms." *Journal of Economic Perspectives*, 25(3): 63–82.

Bandiera, Oriana, Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat. 2021. "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats." *The Quarterly Journal of Economics*, 136(4): 2195–2242.

Banerjee, Abhijit, Eliana La Ferrara, and Victor H. Orozco-Olvera. 2019*a*. "The Entertaining Way to Behavioral Change: Fighting HIV with MTV."

Banerjee, Abhijit, Eliana La Ferrara, and Victor H. Orozco-Olvera. 2019*b*. "Entertainment, Education, and Attitudes Toward Domestic Violence." *AEA Papers and Proceedings*, 109: 133–137.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition.* Chicago, IL:University of Chicago Press.

Belman Inbal, Aliza M., and Helen Chin. 2008. "Using Training to Build Capacity for Development: An Evaluation of the World Bank's Project-based and WBI Training." World Bank, Washington, DC.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *The Review of Economic Studies*, 70(3): 489–520.

Berg, Gunhild, and Bilal Zia. 2017. "Harnessing Emotional Connections to Improve Financial Decisions: Evaluating the Impact of Financial Education in Mainstream Media." *Journal of the European Economic Association*, 15(5): 1025–1055.

Bernard, Tanguy, Stefan Dercon, Kate Orkin, Giulio Schinaia, and Alemayehu Seyoum Taffesse. 2026. "The Future in Mind: Aspirations and Long-Term Outcomes in Rural Ethiopia." *The Quarterly Journal of Economics*, 141(2): 1383–1447.

Berry, James, Greg Fischer, and Raymond Guiteras. 2020. "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana." *Journal of Political Economy*, 128(4): 1436–1473.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu. 2020. "The Glittering Prizes: Career Incentives and Bureaucrat Performance." *The Review of Economic Studies*, 87(2): 626–655.

Besley, Timothy, Robin Burgess, Adnan Khan, and Guo Xu. 2022. "Bureaucracy and Development." *Annual Review of Economics*, 14: 397–424.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "Americans Do IT Better: US Multinationals and the Productivity Miracle." *American Economic Review*, 102(1): 167–201.

Burchardi, Konrad B., Jonathan de Quidt, Selim Gulesci, Benedetta Lerva, and Stefano Tripodi. 2021. "Testing Willingness to Pay Elicitation Mechanisms in the Field: Evidence from Uganda." *Journal of Development Economics*, 152: 102701.

Cefala, Luisa, Pedro Naso, Michel Ndayikeza, and Nicholas Swanson. 2024. "Under-Training by Employers in Informal Labor Markets: Evidence from Burundi." *Mimeo*, 1–91.

Cohen, Jessica, and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *The Quarterly Journal of Economics*, 125(1): 1–45.

Cole, Shawn, A. Nilesh Fernando, Daniel Stein, and Jeremy Tobacman. 2020. "Field Comparisons of Incentive-Compatible Preference Elicitation Techniques." *Journal of Economic Behavior & Organization*, 172: 33–56.

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso. 2020. "Patronage and Selection in Public Sector Organizations." *American Economic Review*, 110(10): 3071–3099.

Dahis, Ricardo, Laura Schiavon, and Thiago Scot. 2025. "Selecting Top Bureaucrats: Admission Exams and Performance in Brazil." *The Review of Economics and Statistics*, 107(2): 408–425.

Deserranno, Erika, Philipp Kastrau, and Gianmarco León-Ciliotta. 2025. "Promotions and Productivity: The Role of Meritocracy and Pay Progression in the Public Sector." *American Economic Review: Insights*, 7(1): 71–89.

Dessein, Wouter. 2002. "Authority and Communication in Organizations." *The Review of Economic Studies*, 69(241): 811–838.

Diaz, Brayan S., Andrea Neyra Nazarrett, Julian Ramirez, Raffaella Sadun, and Jorge A. Tamayo. 2025. "Training within Firms." *National Bureau of Economic Research Working Paper 33670*, 1–93.

Dienes, Harry, Yetsedaw Emagne, Dan Rogger, and Giulio Schinaia. 2024. "Managerial Narratives and the Demand for Management Training in Public Administration." *AEA RCT Registry: https://doi.org/10.1257/rct.12407-2.0.*

Dizon-Ross, Rebecca, and Seema Jayachandran. 2022. "Improving Willingness-to-Pay Elicitation by Including a Benchmark Good." *AEA Papers and Proceedings*, 112: 551–555.

Finan, Fred, Ben A. Olken, and Rohini Pande. 2017. "Chapter 6 - The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments*. Vol. 2 of *2*, , ed. Abhijit Vinayak Banerjee and Esther Duflo, 467–514. Amsterdam, Netherlands:North-Holland.

Fornasari, Margherita, Imran Rasul, Daniel Rogger, and Martin J. Williams. 2025. "Ideas Generation in Hierarchical Bureaucracies: Evidence from a Field Experiment and Qualitative Data." *CEPR Discussion Paper No. 20388*, 1–81.

Gailmard, Sean, and John W. Patty. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science*, 51(4): 873–889.

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy*, 108(5): 874–904.

Grindle, Merilee S., and Mary E. Hilderbrand. 1995. "Building Sustainable Capacity in the Public Sector: What Can Be Done?" *Public Administration and Development*, 15(5): 441–463.

Harrison, Glenn W, and John A List. 2004. "Field Experiments." *Journal of Economic Literature*, 42(4): 1009–1055.

Jakobsen, Morten, Christian Bøtcher Jacobsen, and Søren Serritzlew. 2019. "Managing the Behavior of Public Frontline Employees through Change-Oriented Training: Evidence from a Randomized Field Experiment." *Journal of Public Administration Research and Theory*, 29(4): 556–571.

Kala, Namrata. 2024. "The Impacts of Managerial Autonomy on Firm Outcomes." *Econometrica*, 92(6): 1777–1800.

Kaufmann, Daniel, and Aart Kraay. 2023. "Worldwide Governance Indicators, 2023 Update." *http://www.govindicators.org*.

Khan, Muhammad Yasir. 2025. "Mission Motivation and Public Sector Performance: Experimental Evidence from Pakistan." *American Economic Review*, 115(7): 2343–2375.

Lerva, Benedetta. 2023. "The Monetary Value of Externalities: Experimental Evidence from Ugandan Farmers." *World Bank Policy Research Working Papers No. 10521*, 1–60.

Leuven, Edwin. 2005. "The Economics of Private Sector Training: A Survey of the Literature." *Journal of Economic Surveys*, 19(1): 91–111.

Maffioli, Alessandro, David McKenzie, and Diego Ubfal. 2023. "Estimating the Demand for Business Training: Evidence from Jamaica." *Economic Development and Cultural Change*, 72(1): 123–158.

Mehmood, Sultan, Shaheen Naseer, and Daniel L. Chen. 2024. "Training Policymakers in Econometrics." *Mimeo*, 1–88.

Mikkelsen, Kim Sass, Christian Schuster, Jan-Hinrik Meyer-Sahling, and Magdalena Rojas Wettig. 2022. "Bureaucratic Professionalization Is a Contagious Process Inside Government: Evidence from a Priming Experiment with 3,000 Chilean Civil Servants." *Public Administration Review*, 82(2): 290–302.

Mobarak, Ahmed Mushfiq, Puneet Dwivedi, Robert Bailis, Lynn Hildemann, and Grant Miller. 2012. "Low Demand for Nontraditional Cookstove Technologies." *Proceedings of the National Academy of Sciences*, 109(27): 10815–10820.

Orozco-Olvera, Victor, Fuyuan Shen, and Lucie Cluver. 2019. "The Effectiveness of Using Entertainment Education Narratives to Promote Safer Sexual Behaviors of Youth: A Meta-Analysis, 1985-2017." *PLOS ONE*, 14(2): e0209969.

Peterman, Amber. 2025. "Edutainment to Prevent Violence Against Women and Children." *The World Bank Research Observer*, lkaf002.

Prendergast, Canice. 1993. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *The Quarterly Journal of Economics*, 108(2): 523–534.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *The Economic Journal*, 128(608): 413–446.

Rasul, Imran, Daniel Rogger, and Martin J. Williams. 2021. "Management, Organizational Performance, and Task Clarity: Evidence from Ghana's Civil Service." *Journal of Public Administration Research and Theory*, 31(2): 259–277.

Schuster, Christian, Kim Sass Mikkelsen, Daniel Rogger, Francis Fukuyama, Zahid Hasnain, Dinsha Mistree, Jan Meyer-Sahling, Katherine Bersch, and Kerenssa Kay. 2023. "The Global Survey of Public Servants: Evidence from 1,300,000 Public Servants in 1,300 Government Institutions in 23 Countries." *Public Administration Review*, 83(4): 982–993.

Seidle, Brett, Sergio Fernandez, and James L. Perry. 2016. "Do Leadership Training and Development Make a Difference in the Public Sector? A Panel Study." *Public Administration Review*, 76(4): 603–613.

Online Appendix Materials

In this Online Appendix, we provide details of several aspects of the study. Appendices A and B describe the design and implementation of the two rounds of data collection. Appendix C describes the analyses pre-specified in our pre-analysis plan. Appendix D briefly discusses additional analyses on the drivers of demand that complement the results section in the paper. Appendix E describes the secondary data sources used in the paper. Appendix F provides additional tables and figures.

A Round One: Field Experiment at Civil Service University

In this section we provide further details about the in-person elicitation and enrolment of participants in the experiment, which began at Civil Service University in November 2023.

A.1 Sampling and Implementation

The Ethiopian Management Institute prepared a formal request to the Office of the University President at Civil Service University, which authorised the study and directed it to the relevant college and departments. With their agreement, we used simple random assignment to select the participating class sections and scheduled sessions during dedicated class hours, with professors allocating their teaching time to the exercise.

At each session the research team explained the protocol and obtained informed consent. Participants then watched one of the three videos on individual tablets, with assignment randomised at the individual level and most using headphones for audio clarity. Participants completed the questionnaire on the same tablets immediately afterwards, and then took part in the BDM protocol described in Section 3.1. Enumerators verified each participant's practice-round responses before submission to confirm that the mechanism had been understood, and payments were settled the same day by mobile money transfer or in cash.

A.2 Willingness-To-Pay Elicitation Procedure

We used a Becker-DeGroot-Marschak (BDM) mechanism to elicit participants' willingness-to-pay (WTP) for three professional development opportunities: training, coaching, and shadowing. Participants stated their maximum WTP for each opportunity, in a randomised order. One of the three was then randomly selected, and a price was randomly drawn for that particular opportunity. If a participant's bid met or exceeded the random price, they purchased the opportunity at the drawn price; otherwise, no transaction occurred. Participants were reminded that bidding truthfully was in their best interest, as the price would not be influenced by their bid. To reinforce understanding and promote accurate valuation, respondents completed a practice round

using a pen, answered comprehension checks, and were required to confirm their ability to pay the amount they bid, either via the most common mobile money platform (*TeleBirr*) or in cash on the same day. The full script that enumerators used in the elicitation procedure is reported in our pre-analysis plan (`https://www.socialscienceregistry.org/trials/12407`).

A.2.1 Non-payment

Non-payment can distort revealed preferences, particularly if participants anticipate that payment will not be enforced. Messaging the need to pay may itself reduce expectations that opting out is acceptable. While existing studies vary in rates of non-payment, our experience compares favourably. Maffioli et al. [2023], for example, report that over half of selected participants refused to pay the full amount. In contrast, studies that request payment immediately after bidding tend to see much lower default rates [Berry et al., 2020; Burchardi et al., 2021]. A review by Maffioli et al. [2023] finds that most such studies report single-digit non-payment rates.

To mitigate non-payment and ensure bids reflect true demand, we followed best practices in implementing the willingness-to-pay mechanism. Participants were required to pay upfront if they drew a randomised price below their stated bid. To reinforce this, they received two follow-up prompts after submitting their bids: first, whether they would still pay a slightly higher price; and second, a reminder that payment would be required if the random draw fell below their bid. They then reviewed and confirmed all three bids side by side. Our implementation method was effective – only 4 out of 111 selected buyers ultimately declined to pay.

A.2.2 Coding of Reflections of Rationale for Bids

As part of the initial elicitation of bids, we asked each participant to write a short reflection on what drove their pricing decisions. 89% of respondents provided a reflection that could be analysed. Two research assistants independently coded each reflection by identifying its primary focus as individual, team, organisation, or society. They then reconciled any differences. For analysis, we grouped the latter three into a single category reflecting other-regarding motives.

A.3 Details on the Management Narrative Intervention

Before eliciting the original set of bids for capacity building activities, we had public officials watch a fictional movie about a team in the Ethiopian public service trying to solve a problem. We developed two sets of treatment videos (as well as a control) that emphasised different features of management practice prevalent in the public service. This section begins by outlining how we identified the major approaches to management in Ethiopia's public administration, and then how we developed the movies related to those themes.

A.3.1 Factor Decomposition of Management Practices in Ethiopia

We capitalised on pre-existing measures of management practice in Ethiopia's Government based on survey data from 2016, described in Section E. We undertook a factor analysis of these measures and found that they dominantly clustered onto two factor loadings in the sense that they were the only two with an eigenvalue greater than 1. These two factors jointly explain 58% of the variation before rotation. A summary of results is presented as Table A.2. We interpret these dimensions of management as follows: the extent to which managers pay close attention to measuring performance and utilising accountability mechanisms to ensure performance (which we label 'monitoring'); and the extent to which managers focus on providing enabling and supportive management practices to support their team to undertake tasks independently, based on trust and informal managerial training and feedback (which we label 'enabling').

A.3.2 Development of Management Movies with Randomised Components

We developed movies that had both common components (broad storyline, setting, actors) and varied the emphasis on an enabling or monitoring approach to management. In order to develop the content of the movies, we developed a script based on core elements of the Ethiopian public administration such as the Ethiopian Civil Service Code, the new competency framework being drafted by the Civil Service University, existing training material used by the Ethiopian Management Institute, and so on. We conducted interviews and focus group discussions with middle-managers in the civil service to collect anecdotes and professional stories to determine an engaging and relatable narrative of the script. Next, we developed the video intervention with an experienced local production company (Synergy Habesha) in partnership with the Ethiopian Management Institute, the Civil Service Commission and with support by a media expert at an international firm that develops videos for behaviour change (ImpactEd) to stimulate engagement in the training sessions. Finally, we piloted the different movie-based interventions and refined them based on feedback from viewers before starting the full-scale rollout.

B Round Two: Follow-up Phone Survey

In this section we provide further details about the phone-based elicitation and additional experiments, which took place between November 2024 and February 2025.

B.1 Survey logistics

We undertook follow-up phone-based surveys with respondents between November 5, 2024 and February 4, 2025. To minimise attrition rate, enumerators offered each participant the option to complete the 15-minute survey immediately or reschedule for a more convenient time; if a participant could not be reached, enumerators attempted up to three callbacks. Participants

received 50 ETB (0.50 USD) in airtime credit as compensation. The order in which respondents were contacted was randomised conditional on the above procedure. Assignment to treatment in the survey experiments was based on a pre-assigned randomisation at the individual-level.

B.2 Phone-based WTP Elicitation

For the follow-up phone survey, we mimicked the BDM mechanism employed in the first part of our study in November 2023, while stating that we would not collect money on that day. We stated that if we were to go ahead with the proposed training, the amounts they stated would be the amounts they would have to pay. This procedure allows us to measure WTP in a way that is consistent with our first experiment in November 2023, while allowing us to investigate many more variations in the characteristics of training in the finite interview time. We find no impact on bids in the follow-up survey from the random price drawn in the initial experiment, or whether respondents attended an activity offered (Appendix Table A.13).

B.3 Training Framings

The follow-up survey presented four versions of training, each framed to highlight a different objective, in randomised order: (i) skills to support the respondent's own career advancement in the public or private sector; and three other-regarding framings: (ii) skills to improve productivity within the respondent's organisation; (iii) skills to enhance coordination and performance across the public sector as a whole; and (iv) skills to increase the respondent's impact on citizens. These framings span motivational targets from personal benefit to broader social value, enabling us to identify how different types of perceived benefits shape demand.

B.4 Survey Experiments

After assessing demand for training with different framings, we undertook additional randomised survey experiments:

B.4.1 Audio-based Intervention

We primed the participant towards a private career intention for training, or public contribution intention, by having them listen to a recording from the Civil Service Commission that describes why training is important. The text of these recordings is reported below. Which of the two recordings was played to the participant is randomised. The audio recording was introduced by the enumerator as from the Civil Service Commission, with whom we have collaborated. These were read by a female ex-senior civil servant to contrast with the male enumerator, differentiating the message from the interview. Note that all participants received one of the two messages. The full audio script is reported in Appendix Table A.1.

B.4.2 Varying Hypothetical Saturation

We measured the extent to which individuals change their WTP conditional on the fact that other members of their unit are being trained. Specifically, we randomised the number of individuals that we state will be trained with the participant. The question line is as follows:

> Thank you. We are interested in whether the fact that from your department, [*RANDOMISED: 1 other; 2 others; 3 others; the rest of your department*] going to a training affects your willingness to participate.
>
> Imagine that from your unit, [*RANDOMISED: 1 other; 2 others; 3 others; the rest of your department*] registered to attend the training that emphasises [`RANDOMLY CHOOSE ONE: skills to strengthen your career prospects across any kind of organisation; skills that are identified by your team as crucial for your organisation's productivity; skills that make the public service as a whole work more effectively together; skills that make your direct impact on citizens larger`].
>
> Without that consideration, you had bid [X] Birr. How much are you willing to pay now that [*RANDOMISED: 1 other; 2 others; 3 others; the rest of your department*] are attending?

Civil servants' demand may be internalising some of the externalities that attending training can have on their productivity, on the productivity of the organisation, or on the government's capacity to serve the public as a whole. Given the prevalence of teamwork in public service, the decisions of a respondent's colleagues to invest in capacity building, and the externalities this generates, may influence individual demands. For example, we may expect free-riding to occur if individuals weigh their private costs against team performance gains, potentially reducing WTP when more colleagues attend. Conversely, complementarities in team skills may increase WTP if training returns improve with team-wide participation (Lerva [2023] explores the potential for such externalities in WTP in the context of agricultural pest-control technology and finds evidence of large spillovers). To test these ideas, we estimate the following equation:

$$Y_{post,i} = \alpha_1 Y_{pre,i} + \mathbf{C}\boldsymbol{\gamma} + \mathbf{S}\boldsymbol{\delta_1} + \varepsilon_i \tag{2}$$

where $Y_{pre,i}$ corresponds to the WTP before the audio-treatment, $\mathbf{C}$ is a vector of indicators for group sizes ("1 other", "2 others", "3 others", or "department-wide"), with "1 other" as the omitted category, and the other elements are described in the specification in Section 5.1.

B.5 Perceived Returns

As an additional secondary outcome in the follow-up phone-survey, we asked participants about their anticipated promotion prospects with the following question:

> We're interested in how training impacts your chances of promotion. Imagine you went on a training that emphasises `RANDOMLY CHOOSE ONE: skills to strengthen your career prospects across any kind of organisation; skills that are identified by your team as crucial for your organisation's productivity; skills that make the public service as a whole work more effectively together; skills that make your direct impact on citizens larger.` Do you think that increases the percentage chance of receiving a promotion? By how much? *ANSWER OPTIONS: 0% - 100%*

C Analyses from Pre-Analysis Plan

Our core design and primary outcomes were pre-specified in two pre-analysis plans (AEARCTR-0012407): an initial plan for the initial elicitation (November 14, 2023), and an addendum for the follow-up survey data (December 7, 2024). Both plans were lodged before data collection concluded. The pre-analysis plans specified several robustness checks and alternative specifications, all of which we report in the appendices and none of which alter our main findings. Specifically we planned to also report:

- A difference-in-differences (DID) estimator as a robustness check, conditional on high autocorrelation. Our main specification uses ANCOVA (Table 2); DID results are reported in Appendix Table A.6 and are very similar.
- Post-double LASSO estimation, reported in Appendix Table A.6 for the audio-based intervention and Appendix Table A.7 for the video-based intervention.
- Effects of the video-based intervention on willingness-to-pay for coaching, shadowing, and training, as well as an index of bids (Appendix Table A.9 Panel A).
- Treatment effects on the probability of a positive bid (Appendix Table A.9, Panels B and C).
- Spillover effects, estimated by varying the hypothetical saturation, i.e., the number of colleagues attending training (Appendix Section D.2).

Our main analysis departs from the pre-analysis plan in one way. Due to a coding error, participants were randomised across five videos (four treatment, one placebo) in equal proportions,

rather than a single treatment and placebo in equal proportions. As a result, we pool the four treatment videos into two groups in our main analysis to balance statistical power with interpretability. We report fully disaggregated effects for all five videos in Appendix Table A.8 Panel A, and binary treatment effects in Panel B; both are consistent with our main specification.

D Additional Analyses on the Drivers of Demand

D.1 Variation in WTP from Training Framings

In this subsection, we present the results on how training framings alter willingness-to-pay. Respondents value career-framed training most. As shown in the top half of Appendix Figure A.3, respondents are willing to pay 23–27 ETB more for trainings emphasising career benefits – nearly double the average – than for any other emphasis. There are no significant differences among the three public-service-oriented framings.

The bottom of Figure A.3 further splits these means by respondents' stated motivations from the qualitative reflections, coded as either "individual" or "social". Three patterns emerge. First, individuals with other-regarding motivations are willing to pay more for all trainings – including those emphasising career benefits. Second, across all groups, WTP declines as training emphasises skills that benefit a wider group of people; beyond the career-framed training, respondents value more trainings that build skills with more proximate social benefits (e.g., organisational) rather than more distant ones (e.g., society).[19] While those reporting other-regarding motivations are willing to bid a higher price for all training types, they also bid the highest amount for career-framed trainings. Third, even the highest WTP – among individuals coded as socially-oriented for the training emphasising career-specific skills – falls short of covering training costs.

Together, these results suggest that public officials are dominantly driven in their bidding behaviour by the private benefit of public sector capacity building. More socially-minded officials are happy to sacrifice more for such training, but even they reduce their bids as the object of training moves away from private benefit and closer to abstract citizen welfare.

D.2 No Evidence of Saturation Effects

Appendix Figure A.4 illustrates the results from the saturation variation experiment discussed in Appendix Section B.4.2. We find limited evidence of differential demand by group size. We find

[19]We note that these patterns are consistent with different interpretations of "other-regarding preferences". The first, that they have a lower valuation of benefits that are farther removed from them. The second that the benefits of training on social groups less proximate to the self are more ambiguous. Both imply that public servants are less likely to invest in public service capacity building for the benefit of less proximate social groups, but for distinct reasons.

suggestive evidence consistent with complementarities, rather than free-riding, as WTP is higher when the entire department is expected to attend (34 ETB) versus 3 other colleagues (15 ETB), significant at the 10% level. This overall result, of limited spillovers or the influence of saturation, implies our core results are not heavily influenced by the perception of colleagues' actions.

D.3 No Impact of Interventions on Perceived Promotion

We present treatment effects of the experimental interventions on the perceived impact of training on promotion in Appendix Table A.11, where we do not find that the experiments affected the perceived probability of promotion.

D.4 Interaction with Work Experience

As expected from a lifecycle perspective, WTP declines by years of experience in the control group. To assess how career stage interacts with the video-based treatment, we interact the treatment arms with years of experience and display the linear interaction model (Figure A.5). In general, the impact of the videos increases with years of experience.

E Additional Data

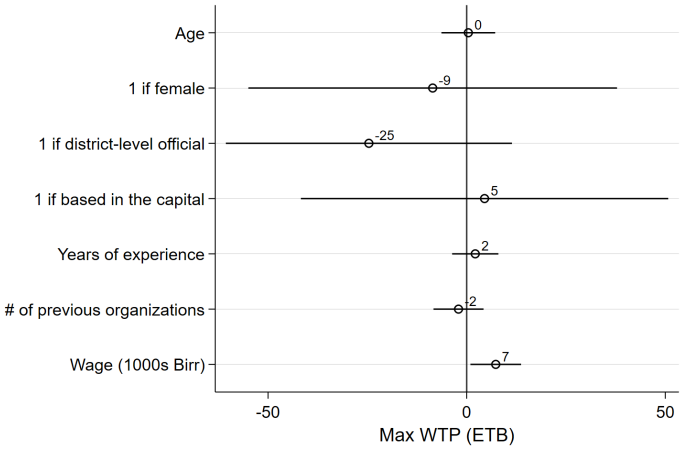
Besides the data collected during the in-person elicitation at CSU and the follow-up phone-survey, additional data is drawn from two sources: i) the Ethiopian Survey of Public Servants 2024, and ii) the Ethiopian Survey of Public Servants 2016.

Ethiopia Survey of Public Servants 2024 This data was used for information on current Ethiopian civil servants' perspectives on and experience of professional development in Section 2. A survey of $N = 3{,}321$ face to face interviews was implemented in February to April 2024 across the three tiers of Ethiopian government. Civil servants in non-managerial positions (employee-level track), civil servants in managerial positions (director-level track), and organisation heads (political appointee-level track) were included. Frontline service delivery staff such as teachers or nurses were not included. The survey especially sampled civil servants in the sectors of Finance, Revenue, Labour and Skill, Health, Education, Agriculture, and the Civil Service itself.

Ethiopia Survey of Public Servants 2016 This data was used to form the baseline representative group to compare our willingness-to-pay sample to in Section 3. A survey of $N = 2{,}164$ face-to-face interviews was implemented in 2016. The survey used random sampling techniques at all levels of sample selection to gain a representative sample of officials across the government sectors studied. Interviews were carried out in June to September 2016 across the three tiers of Ethiopian government; at the federal level, in all 11 regions existing at that time, and in a selection of 66 woredas. Civil servants in non-managerial positions (employee-level

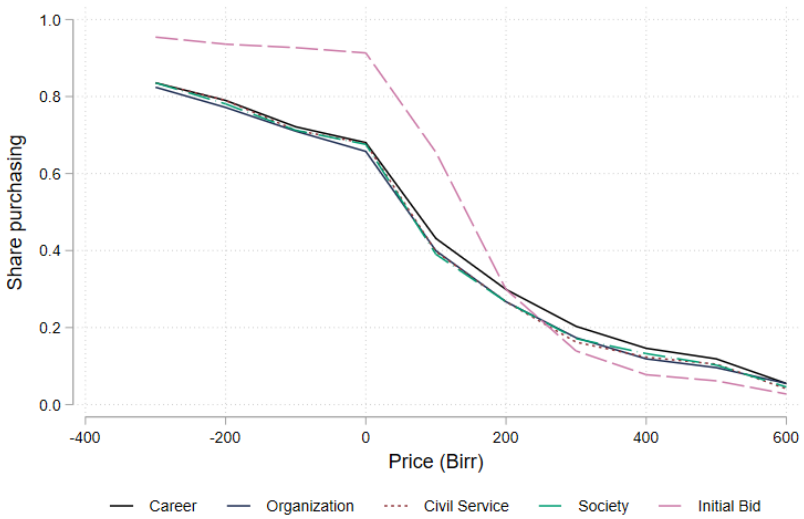
track), civil servants in managerial positions (director-level track), and organisation heads (political appointee-level track) were included. Frontline service delivery staff such as teachers or nurses were not included.

F Additional Figures and Tables



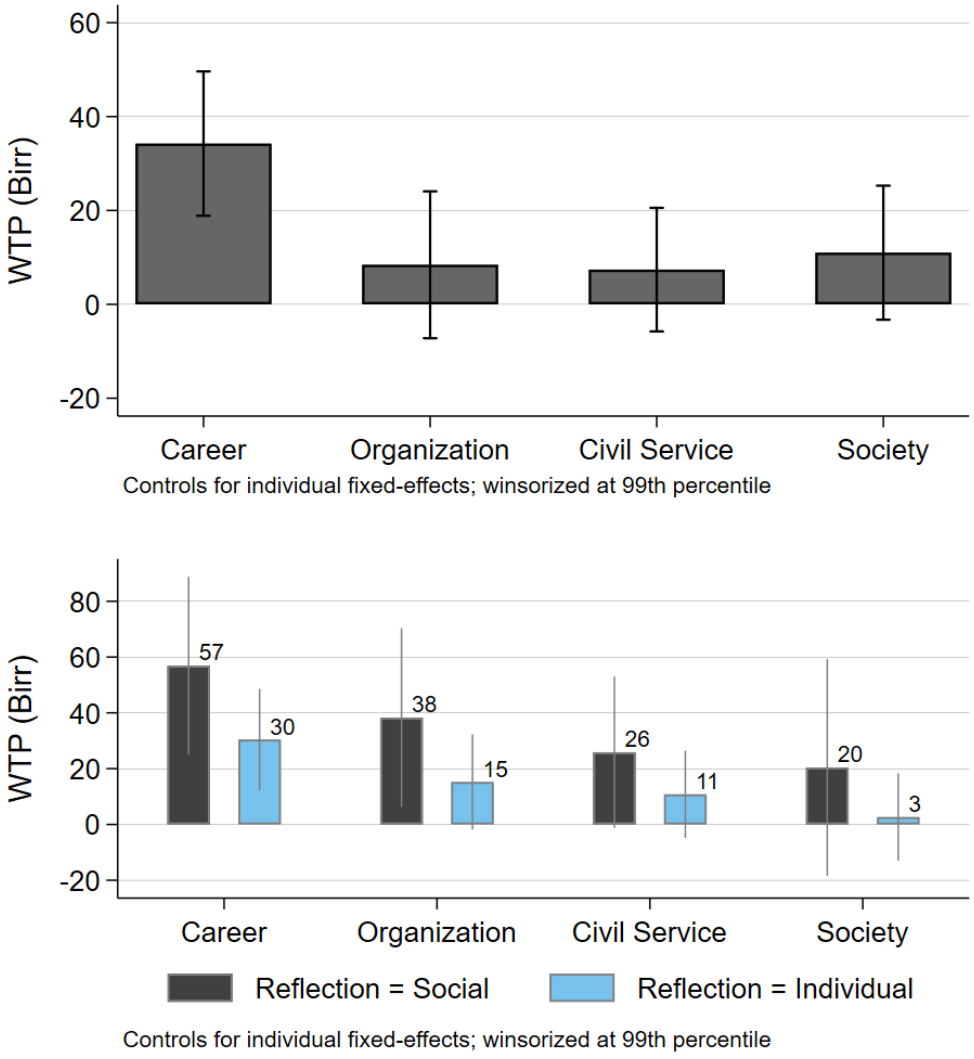
Note: Coefficients from a linear regression of various covariates on the maximum willingness-to-pay (WTP) among the three incentivised bids elicited at Civil Service University as part of the in-person field experiment. Bars display 95% confidence intervals. $N = 437$, results are the same when imputing for missing control variables.

Figure A.1: Relationship between individual characteristics and willingness-to-pay



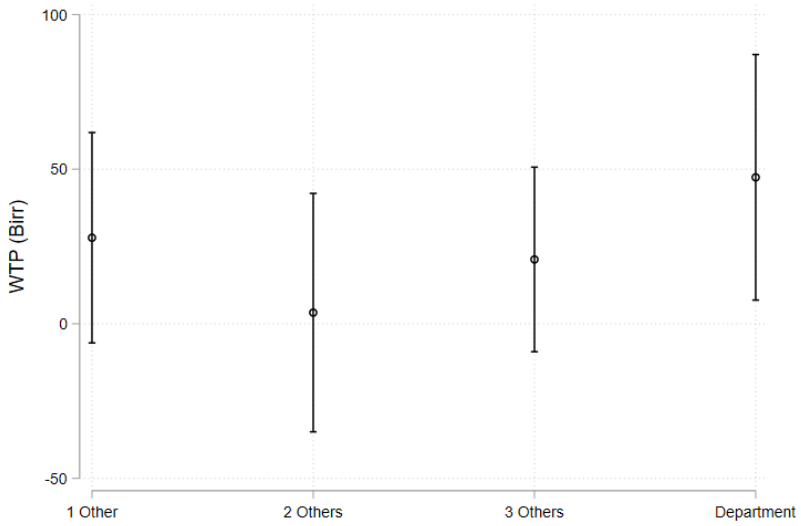
Note: Willingness-to-pay for training, including the initial bid in November 2023 and follow-up bids by framings in November 2024-February 2025. Demand curves estimated with $N = 438$. Bids winsorised at the 99th percentile.

Figure A.2: Inverse Demand: Training by Framings



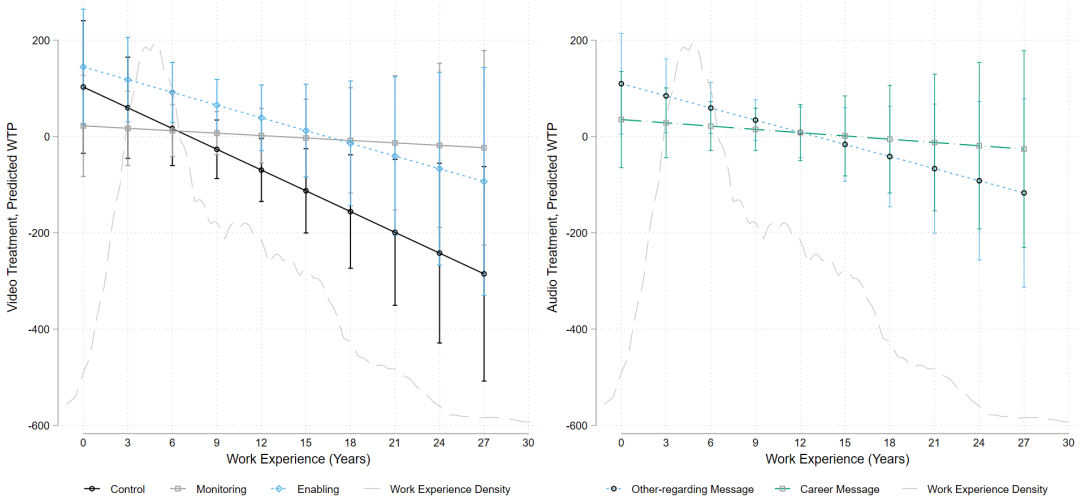
Notes: Predicted means of Willingness-to-Pay for training by framings and motivation, after controlling for individual fixed-effects, 95% confidence intervals are reported and constructed from robust standard errors. Elicited demands are winsorised at the 99th percentile. Motivations coded from reflections on rationale for initial bids. The top panel is estimated using 438 respondents, while the bottom panel is estimated using the 385 respondents for which we have measures of other-regarding motivations.

Figure A.3: Average WTP for Training by Framings and Motivation



Note: Predicted effects on willingness-to-pay for training elicited during the follow-up phone survey of the number of colleagues stated in question as also attending the course. Variation is between-subjects, phone survey sample, $N = 438$. Predicted values from a linear regression on WTP (winsorised at the 99th percentile) on the randomised number of colleagues controlling for previous bids, training framing. Reported 95% confidence interval are robust.

Figure A.4: Relationship Between the Hypothetical Number of Colleagues Attending and WTP



Notes: Predicted post-audio treatment willingness-to-pay (WTP) by video treatment (left) and audio treatment (right), interacted with years of work experience. Line displays density of work experience. Controls not shown: Constant, trial item (pen) bid, training framing, age, gender, whether office is in Addis Ababa, number of organisations, wage bracket indicators, and governmental tier indicators. 95% confidence intervals use robust standard errors. Bids winsorised at the 99th percentile.

Figure A.5: Predicted WTP by video and audio treatment, interacted with years of experience

Table A.1: Audio Treatments

Treatment	Description
Private—Career	"The Civil Service Commission wants to support the careers of civil servants, and so are implementing a competency framework to make clear the steps needed for career advancement. Training is a central part of our reform process to help all civil servants reach these standards. We want civil servants to perform and progress in their own personal careers."
Public—Other-regarding	"I've seen first-hand how training can make a difference in the work civil servants do. I recently visited a team in the service, and saw how by implementing the new practices they learnt at a training course, they made their processes much more efficient, and were able to serve more members of the public. Citizens I spoke to were very happy about their experience of the team, which is the feedback I love to hear. Improving skills and ways of working through training can make a difference for Ethiopia."

Table A.2: Management Style Factor Loadings from Principal Component Analysis

	Monitoring	Enabling
Clarity of roles and responsibilities	.6710113	.1453058
Directorate targets guide work schedule	.8190121	.0232894
Communication of targets and measures	.8975451	-.1163584
Tracking of service delivery	.8842074	-.0814437
Follow-up on meeting plans	.5914227	.2126669
Tolerance of underperformance	.6232459	.0978324
Staff discretion in assignments	.5907584	.2780989
Staff contribution to policy	.2149784	.6343335
Right staff for right job	.3385354	.2833133
Flexibility in work practices	.0735017	.7499799
Responsiveness to stakeholder needs	-.10624	.8419977
Attraction and retention of talent	.0238238	.7108054

Note: Table reports the factor loadings from principal component analysis (PCA) conducted on survey measures of management practices in Ethiopian Government directorates (2016). Each number in the table indicates the correlation or the relative contribution of a given management practice statement with each of the two principal factors identified: "Monitoring" and "Enabling.", which jointly explain 58% of the variance in management practices across directorates prior to rotation.

Table A.3: Descriptive Statistics

	Benchmark (2016) (N=2,164)	Sample (2023) (N=500)
Female	0.178 (0.383)	0.180 (0.385)
Age	35.178 (8.553)	32.360 (5.467)
Wage (Birr per month)		
0-2,000	0 (.%)	10 (2.2%)
2,001-5000	0 (.%)	83 (17.9%)
5,001-10,000	0 (.%)	308 (66.5%)
10,001-20,000	0 (.%)	58 (12.5%)
20,001-50,000	0 (.%)	4 (0.9%)
Work Experience (years)	13.063 (8.549)	9.057 (5.768)
Education		
Grade 8 completion	2 (0.1%)	0 (0.0%)
Grade 10 completion	4 (0.2%)	0 (0.0%)
Grade 12 completion	2 (0.1%)	5 (1.0%)
Diploma / TVET / Post-high school	376 (17.7%)	4 (0.8%)
Undergraduate degree	1,533 (72.2%)	445 (89.0%)
Masters degree	201 (9.5%)	46 (9.2%)
PhD	4 (0.2%)	0 (0.0%)
Program Engaged		
Bachelor Degree	0 (.%)	34 (6.9%)
Masters	0 (.%)	456 (91.9%)
Ph.D.	0 (.%)	6 (1.2%)
Administrative Tier		
Federal	285 (13.2%)	68 (13.7%)
Regional	443 (20.5%)	132 (26.6%)
Zone	0 (0.0%)	61 (12.3%)
Woreda	1,436 (66.4%)	232 (46.8%)
Other	0 (.%)	3 (0.6%)
Office in Addis Ababa	0 (.)	0.246 (0.431)

Note: The 2016 Benchmark is based on the Ethiopian Civil Service Survey (CSS), a representative sample of the Ethiopian Civil Service described in Appendix Section E. The 2023 sample are respondents in field experiment, who attended Civil Service University (CSU), and took part in the first incentivised willingness-to-pay elicitation. Standard deviation or percentage of category (%), for categorical variables, shown in parentheses.

Table A.4: Audio-based Treatment Experimental Balance Table (within Follow-up Sample)

	Audio-based Treatment			
	(1)	(2)	(3)	(4)
	Career	Other-regarding	Total	p-value
CSU Training Bid	99.045 (255.045)	118.610 (472.188)	108.783 (378.695)	0.589
Age	32.201 (5.481)	32.588 (5.533)	32.393 (5.504)	0.464
Female	0.156 (0.364)	0.202 (0.402)	0.179 (0.384)	0.225
Admin tier				
Federal	34 (15.6%)	28 (13.0%)	62 (14.3%)	0.156
Regional	54 (24.8%)	59 (27.3%)	113 (26.0%)	
Zone	33 (15.1%)	18 (8.3%)	51 (11.8%)	
Woreda	95 (43.6%)	110 (50.9%)	205 (47.2%)	
Other	2 (0.9%)	1 (0.5%)	3 (0.7%)	
Location Addis Ababa	0.258 (0.439)	0.258 (0.439)	0.258 (0.438)	1.000
Work Experience (years)	8.806 (5.795)	9.340 (5.796)	9.072 (5.795)	0.340
Number Orgs	2.083 (1.500)	2.284 (2.046)	2.182 (1.793)	0.244
Wage (Birr)				
0-2,000	5 (2.4%)	4 (2.0%)	9 (2.2%)	0.827
2,001-5000	37 (17.8%)	33 (16.5%)	70 (17.2%)	
5,001-10,000	143 (68.8%)	135 (67.5%)	278 (68.1%)	
10,001-20,000	21 (10.1%)	27 (13.5%)	48 (11.8%)	
20,001-50,000	2 (1.0%)	1 (0.5%)	3 (0.7%)	
N	220 (50.2%)	218 (49.8%)	438 (100.0%)	

Note: The table displays means for baseline variables among the sample that responded to the follow-up phone survey, by audio experiment assignment. The fourth column shows p-values for equal means in the first two columns, evaluating balanced attrition. Standard deviation, or percentage of category (%) for categorical variables, shown in parentheses. Number of observations reported in the bottom row.

Table A.5: Experimental Balance Table (Video, Realised Sample and CSU Sample)

	(1) Control	(2) Monitoring	(3) Enabling	(4) Total	(5) Test
Panel A: Realised Sample					
Age	32.000 (5.150)	32.948 (5.759)	31.938 (5.341)	32.393 (5.504)	0.177
Female	0.198 (0.401)	0.175 (0.381)	0.174 (0.381)	0.179 (0.384)	0.890
Admin tier					
Federal	10 (12.2%)	32 (16.7%)	20 (12.5%)	62 (14.3%)	0.891
Regional	20 (24.4%)	50 (26.0%)	43 (26.9%)	113 (26.0%)	
Zone	11 (13.4%)	20 (10.4%)	20 (12.5%)	51 (11.8%)	
Woreda	41 (50.0%)	89 (46.4%)	75 (46.9%)	205 (47.2%)	
Other	0 (0.0%)	1 (0.5%)	2 (1.2%)	3 (0.7%)	
Location Addis Ababa	0.205 (0.406)	0.298 (0.459)	0.237 (0.427)	0.258 (0.438)	0.202
Work Experience (years)	8.778 (5.050)	9.458 (6.254)	8.755 (5.577)	9.072 (5.795)	0.464
Number Orgs	2.220 (1.750)	2.171 (1.660)	2.177 (1.973)	2.182 (1.793)	0.978
Wage (Birr)					
0-2,000	0 (0.0%)	3 (1.7%)	6 (3.9%)	9 (2.2%)	0.223
2,001-5000	15 (19.5%)	23 (12.8%)	32 (21.1%)	70 (17.2%)	
5,001-10,000	52 (67.5%)	127 (70.9%)	99 (65.1%)	278 (68.1%)	
10,001-20,000	9 (11.7%)	24 (13.4%)	15 (9.9%)	48 (11.8%)	
20,001-50,000	1 (1.3%)	2 (1.1%)	0 (0.0%)	3 (0.7%)	
N	83 (18.9%)	193 (44.1%)	162 (37.0%)	438 (100.0%)	
Panel B: CSU Sample					
Age	31.895 (5.003)	32.943 (5.850)	31.959 (5.219)	32.360 (5.467)	0.129
Female	0.183 (0.389)	0.181 (0.386)	0.178 (0.384)	0.180 (0.385)	0.996
Admin tier					
Federal	12 (12.6%)	33 (15.7%)	23 (12.0%)	68 (13.7%)	0.923
Regional	25 (26.3%)	55 (26.2%)	52 (27.2%)	132 (26.6%)	
Zone	14 (14.7%)	23 (11.0%)	24 (12.6%)	61 (12.3%)	
Woreda	44 (46.3%)	98 (46.7%)	90 (47.1%)	232 (46.8%)	
Other	0 (0.0%)	1 (0.5%)	2 (1.0%)	3 (0.6%)	
Location Addis Ababa	0.198 (0.401)	0.278 (0.449)	0.236 (0.425)	0.246 (0.431)	0.298
Work Experience (years)	8.766 (4.896)	9.571 (6.224)	8.632 (5.625)	9.057 (5.768)	0.230
Number Orgs	2.211 (1.688)	2.175 (1.645)	2.296 (2.948)	2.228 (2.237)	0.862
Wage (Birr)					
0-2,000	0 (0.0%)	3 (1.5%)	7 (3.9%)	10 (2.2%)	0.208
2,001-5000	18 (20.2%)	26 (13.3%)	39 (21.9%)	83 (17.9%)	
5,001-10,000	60 (67.4%)	138 (70.4%)	110 (61.8%)	308 (66.5%)	
10,001-20,000	10 (11.2%)	27 (13.8%)	21 (11.8%)	58 (12.5%)	
20,001-50,000	1 (1.1%)	2 (1.0%)	1 (0.6%)	4 (0.9%)	
N	96 (19.2%)	211 (42.2%)	193 (38.6%)	500 (100.0%)	

Note: The table displays means for baseline variables for the sample that responded to the follow-up phone survey (Panel A), and for the full sample that was originally sampled for the WTP elicitation at Civil Service University (Panel B), by assignment to the video experiment. The fourth column shows p-values for equal means in the first two columns, evaluating balanced attrition. Standard deviation, or percentage of category (%) for categorical variables, shown in parentheses. Number of observations reported in the bottom rows of each panel.

Table A.6: Relative Treatment Effects of Career Message

	(1) OLS	(2) OLS	(3) LASSO	(4) DiD
Treatment: "Training is good for your career"	6.994*	7.036*	6.816*	
	(3.773)	(3.786)	(3.772)	
Other-regarding Training: Organisation	4.799	4.797	4.821	-25.82*
	(3.367)	(3.368)	(3.367)	(13.30)
Other-regarding Training: Civil Service	-0.959	-0.962	-0.936	-26.85**
	(2.332)	(2.333)	(2.341)	(12.00)
Other-regarding Training: Society	2.648	2.645	2.668	-23.24*
	(3.345)	(3.345)	(3.351)	(12.21)
Post				3.088
				(2.135)
Post × Treatment				6.675*
				(3.811)
Post × Organisation				4.632
				(3.308)
Post × Civil Service				-1.132
				(2.255)
Post × Society				2.498
				(3.351)
Clusters	438	438	438	438
Observations	1752	1752	1752	3504

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training elicited during the phone-survey after audio-treatment was administered. Models (1)-(4) control for the phone-survey elicitation conducted before the treatment. Model (2) controls for the benchmark item (pen) bid. Models (3) uses the LASSO via plugin and double selection to choose control variables, using imputed values where missing, and indicators for missingness are included in the group of control variables for selection. Model (4) is a two-period Difference-in-Differences specification. Period 1 is the phone survey elicitation before the treatment, Period 2 is after treatment. "Treatment: Training is good for your career" is an indicator for whether participants received a message emphasising career advancement benefits before stating their final WTP bids. The other-regarding trainings "Organisation", "Civil Service" and "Society" are indicators for the training content framing.

For LASSO specifications, the phone-survey pre-treatment WTP is always included as a control. Control variables for selection include: pre-treatment WTP elicited during the in-person elicitation in 2023, age, gender, whether office is in Addis Ababa, years of work, number of organisations, wage bracket, governmental tier, and (in model 3) missingness indicators. Standard errors are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.7: Treatment Effects of Manager Style, LASSO

Dependent variable: Willingness-to-pay for training	(1) CSU	(2) Pre	(3) Post
Treatment: Manager Monitors Movie	-0.109	31.28	31.75
	(0.308)	(37.85)	(38.90)
Treatment: Manager Enables Movie	-0.126	84.82**	89.40**
	(0.294)	(39.28)	(41.24)
Clusters	438	438	438
Observations	1752	1752	1752
Monitor vs. Enable p-value	0.94	0.12	0.10

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training. The dependent variable in column (1) is the WTP measure collected in-person at Civil Service University (CSU). "Treatment: Manager Monitors Movie" indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of a monitoring and directive senior manager. "Treatment: Manager Enables Movie" indicates they were shown a movie that showed an enabling senior manager. Column (1) reports the WTP measure collected in person at Civil Service University (CSU). Column (2) uses the WTP measure from the follow-up phone survey about one year later, before respondents received additional treatments (Pre). Column (3) uses the WTP measure from the same phone survey after respondents received additional treatments (Post). All models use the LASSO via plugin and double selection to chose control variables. An indicator for training type is always included as a control in models (2)-(4).
Control variables for selection in LASSO specifications include: pre-treatment WTP elicited during the in-person elicitation in 2023, trial item (pen) bid, age, gender, whether office is in Addis Ababa, years of work, number of organisations, wage bracket, governmental tier, and missingness indicators. Variables with coefficients displayed in the table were included as variables of interest. Standard errors in parentheses are robust. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.8: Disaggregated and Pooled Treatment Effects of Manager Videos

Dependent variable: Willingness-to-pay for training				
	(1) CSU	(2) CSU (Phone Sample)	(3) Pre	(4) Post
Panel A: Disaggregated Effects				
Video 1	15.51	0.481	17.14	16.86
	(38.45)	(37.68)	(43.53)	(44.54)
Video 2	-26.90	-54.07	77.96*	80.65*
	(41.23)	(43.44)	(44.82)	(46.97)
Video 3	22.19	34.42	91.16*	96.92*
	(35.13)	(34.56)	(48.32)	(50.76)
Video 4	39.82	49.69	45.69	47.70
	(38.61)	(36.94)	(44.48)	(45.47)
Panel B: Pooled Effects				
Shown Treatment Video	13.48	10.08	55.78	58.20
	(31.86)	(30.81)	(34.53)	(35.90)
Clusters			438	438
Observations	500	438	1752	1752

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training. In Panel A, Video 2 and 3 indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of an enabling senior manager. In addition, Video 2 had a monitoring main character, Video 3 had a enabling main character. Video 1 and Video 4 indicates they were shown a movie that showed a monitoring and directive senior manager. Video 1 had an enabling main character, Video 4 had a monitoring main character. In Panel B, all treatment video assignments are pooled. The placebo video is the reference category. Column (1) reports the WTP measure collected in person at Civil Service University (CSU). Column (2) uses the WTP measure from the follow-up phone survey about one year later, before respondents received additional treatments (Pre). Column (3) uses the WTP measure from the same phone survey after respondents received additional treatments (Post). Constant and training framing (for Pre and Post models) not shown. Standard errors in parentheses are robust (models 1-2) or clustered at individual level (models 3-4). * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.9: Treatment Effects of Manager Style by Professional Development Opportunity

	(1) Coaching	(2) Shadowing	(3) Training	(4) Index
Panel A: Pooled Effects on WTP				
Shown Treatment Video	5.660	28.03	13.48	13.20
	(29.64)	(25.32)	(31.86)	(28.32)
Panel B: Disaggregated Effects				
Video 1	0.006	0.035	0.041	0.038
	(0.052)	(0.051)	(0.045)	(0.040)
Video 2	-0.002	-0.002	-0.012	0.073**
	(0.054)	(0.054)	(0.051)	(0.036)
Video 3	0.024	0.034	-0.007	0.053
	(0.052)	(0.051)	(0.050)	(0.038)
Video 4	0.052	0.071	0.050	0.057
	(0.049)	(0.048)	(0.045)	(0.038)
Panel C: Pooled Effects on Making a Positive Bid				
Shown Treatment Video	0.0206	0.0355	0.0191	0.0547*
	(0.0420)	(0.0417)	(0.0385)	(0.0331)
Observations	500	500	500	500

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for different professional development activities offered during the in-person elicitation. In Panel A and C, all treatment video assignments are pooled. In Panel A, the outcome is the raw WTP in ETB. Index is a simple mean of WTP. In Panel B and C, estimates come from a linear probability models on whether the respondent made a positive bid for the professional development opportunities. Video 2 and 3 indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of an enabling senior manager. In addition, Video 2 had a monitoring main character, Video 3 had a enabling main character. Video 1 and Video 4 indicates they were shown a movie that showed a monitoring and directive senior manager. Video 1 had an enabling main character, Video 4 had a monitoring main character. The placebo video is the reference category. The Index is an indicator for whether at least opportunity bid is positive. Constant not shown.
Standard errors in parentheses are robust.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.10: Treatment Effects of Manager Style (with controls including pen bid)

Dependent variable: Willingness-to-pay for training	(1) CSU	(2) CSU	(3) CSU	(4) Pre	(5) Post	(6) Post
Treatment: Manager Monitors Movie	27.61	28.80	26.79	39.04	35.34	28.35
	(34.24)	(33.94)	(31.73)	(36.67)	(38.90)	(43.98)
Treatment: Manager Enables Movie	-1.97	3.49	4.55	85.11**	90.67**	101.18**
	(34.10)	(34.34)	(32.56)	(38.10)	(42.36)	(48.78)
Social Reflection		42.36	41.93	26.00	33.81	36.02
		(28.07)	(30.77)	(40.14)	(42.59)	(64.73)
Manager Monitors × Social Reflection						42.16
						(96.47)
Manager Enables × Social Reflection						-54.75
						(98.47)
Clusters				438	438	438
Observations	500	500	438	1752	1752	1752
Predicted Control Mean	102.61	100.00	99.92	-33.41	-25.90	-26.04
Monitor vs. Enable p-value	0.23	0.31	0.40	0.17	0.13	0.06
Sample	CSU	CSU	Phone	Phone	Phone	Phone

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training. "Treatment: Manager Monitors Movie" indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of a monitoring and directive senior manager. "Treatment: Manager Enables Movie" indicates they were shown a movie that showed an enabling senior manager. "Social Reflection" is a binary that takes the value 1 when the respondent provided a socially-oriented reflection. Columns (1) to (3) report the WTP measure collected in person at Civil Service University (CSU). Column (4) uses the WTP measure from the follow-up phone survey about one year later, before respondents received additional treatments (Pre). Columns (5) and (6) use the WTP measure from the same phone survey after respondents received additional treatments (Post). Models (1) and (2) use the full sample to analyse data from CSU elicitation, model (3) uses the restricted phone sample to analyse the CSU elicitation. Models (4), (5) and (6) control for training type. Models (5) and (6) control for "Career Treatment", an indicator for whether participants received a message emphasising career advancement benefits before stating their final WTP bids. Where social reflection is missing, the value is imputed with the median and an indicator of missingness included. Controls not shown: Constant, trial item (pen) bid, training framing, age, gender, whether office is in Addis Ababa, years of work, number of organisations, wage bracket indicators, and governmental tier indicators. Standard errors in parentheses are clustered at individual level. For model (1) standard errors are robust. Bids winsorised at 99th percentile. Monitor vs. Enable p-value displays the test statistics comparing the coefficient for Monitor Treatment against Enable treatment. Model (6) shows this p-value for respondents with individual (i.e., not "social") reflections, comparing coefficients in the first two rows. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.11: Treatment Effects of Manager Style on Perceived Returns

Dependent variable: Perceived promotion return from training	(1) No Controls	(2) Controls	(3) Promoted
Panel A: Audio Treatment			
Treatment: Career Message	0.62	0.11	2.86
	(2.80)	(2.93)	(3.90)
Not Promoted			-1.89
			(5.87)
Promoted			3.34
			(4.88)
Career Treatment x Not Promoted			-7.04
			(8.58)
Career Treatment x Promoted			-6.49
			(6.71)
Panel B: Video Treatment			
Treatment: Manager Monitors Movie	-4.61	-3.87	-3.24
	(3.78)	(3.90)	(5.05)
Treatment: Manager Enables Movie	0.44	1.29	1.36
	(3.87)	(4.03)	(5.25)
Not Promoted			0.81
			(10.80)
Promoted			-2.46
			(7.49)
Manager Monitors x Not Promoted			-7.80
			(12.31)
Manager Monitors x Promoted			2.55
			(9.14)
Manager Enables x Not Promoted			-5.61
			(12.80)
Manager Enables x Promoted			2.97
			(9.10)
Observations	438	438	438

Notes: Treatment effects on a measure of perceived probability of being promoted if (hypothetically) the respondent attended one of the training opportunities offered during the follow-up phone survey. Panel A: "Treatment: Career Message" is an indicator for whether participants received a message emphasising career advancement benefits before stating their final WTP bids. Panel B: "Treatment: Manager Monitors Movie" indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of a monitoring and directive senior manager. "Treatment: Manager Enables Movie" indicates they were shown a movie that showed an enabling senior manager. All models control for randomly selected training type used for this outcome. Controls not shown for models (2) and (3): Constant, social reflection, age, gender, whether office is in Addis Ababa, years of work, number of organisations, wage bracket indicators, and governmental tier indicators, and whether they have been promoted since baseline. Standard errors in parentheses are robust.

Table A.12: Interaction Effects of Manager Style and Audio Message

	(1) Main	(2) Interaction	(3) Controls
Treatment: Manager Monitors Movie	32.51	20.53	28.10
	(39.26)	(49.20)	(50.23)
Treatment: Manager Enables Movie	89.44**	119.86**	124.68**
	(40.88)	(53.61)	(55.86)
Treatment: Career Message		-3.91	-1.08
		(63.46)	(63.16)
Manager Monitors x Career Message		24.51	14.46
		(78.09)	(78.21)
Manager Enables x Career Message		-56.90	-64.56
		(83.07)	(85.99)
Observations	1752	1752	1752

Notes: Treatment effects on a measure of willingness-to-pay (WTP) for training, including interactions between video and audio treatments. "Treatment: Manager Monitors Movie" indicates the respondent was shown a movie during the CSU framed field experiment that provided a narrative of a monitoring and directive senior manager. "Treatment: Manager Enables Movie" indicates they were shown a movie that showed an enabling senior manager. "Treatment: Career Message" is an indicator for whether participants received a message emphasising career advancement benefits before stating their final WTP bids. The audio message highlighting benefits to citizens is the omitted reference. Columns use the WTP measure from the follow-up phone survey about one year later. Model (1) displays the main effects of the video experiment. Model (2) adds the audio treatment and the interaction with the videos. Model (3) adds control variables. Controls not shown: training framing. Model (3) additional controls not shown: Constant, trial item (pen) bid, age, gender, whether office is in Addis Ababa, years of work, number of organisations, wage bracket indicators, and governmental tier indicators. Where missing, controls imputed, with indicator variable for missingness. Standard errors in parentheses are clustered at individual level. Bids winsorised at 99th percentile. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.13: Effect of Course Attendance on Follow-Up Bid

	(1) 2SLS
Attended Course	37.17
	(62.50)
Prior Bid	0.17**
	(0.07)
First-stage F-statistic	157.56
Observations	438

Notes: 2SLS estimates of the impact of attending a course after the first round on second round bids, averaging over training framing. Instrument for attendance is the randomised price in the first round.

Standard errors in parentheses are robust. Bids winsorised at 99th percentile. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.