Lights, Camera, Transaction: Assessing management styles through studio vignettes*

Girum Abebe Marcel Fafchamps Michael Koelle Simon Quinn Tom Schwantje

February 14, 2025

Abstract

We introduce a new method for measuring managerial traits of young professionals: using management vignettes in a video studio. This method – analysed through the lens of a Bayesian hierarchical model – allows us to identify four distinct managerial archetypes (which we term 'authoritative', 'affiliative', 'coercive' and 'timid'). We find that past labour market exposure (including exposure induced through a previous field experiment) correlates strongly with the propensity to act as an authoritative manager. We then use the videos to run an incentivised experiment with firm managers, to elicit preferences over young professionals. Strikingly, we find that firms consistently prefer authoritative-type managers for entry-level managerial positions. Empirically, our results highlight an underexplored mechanism for labour market exclusion among young professionals. Methodologically, we demonstrate the value of controlled vignette scenarios for assessing managerial traits. Our findings underscore the importance of managerial training in shaping labor market outcomes, and offer new avenues for studying the development of managerial talent.

*We are grateful to Stephen Hansen, Andrea Prat, John van Reenen and Chris Woodruff for helpful comments. This project was funded by the Private Enterprise Development in Low-Income Countries programme (PEDL). We also want to thank Gezahegn Gbremedhin for outstanding research assistance. The project would not have been possible without the constant support of Rose Page and the Centre for the Study of African Economies (University of Oxford), nor without the support of EconInsights in Addis Ababa. IRB approval was obtained from the University of Oxford (ECONCIA21-22-19). We would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work: http://dx.doi.org/10.5281/zenodo.22558.

1 Introduction

The past two decades have witnessed a substantial expansion in economists' interest in the role of management. This has been driven by a new empirical focus on the measurement of management practices, spearheaded by the World Management Survey ('WMS') (Scur, Sadun, Van Reenen, Lemos, and Bloom, 2021). This approach has studied the role of management practices in a very wide variety of organisations, including large firms (Bloom and Van Reenen, 2007), microenterprises (McKenzie and Woodruff, 2017), hospitals (Bloom, Propper, Seiler, and Van Reenen, 2015; Chandra, Finkelstein, Sacarny, and Syverson, 2016), schools (Bloom, Lemos, Sadun, and Van Reenen, 2015), government bureaucracies (Rasul and Rogger, 2018), and universities (McCormack, Propper, and Smith, 2014).

To date, this approach has focused primarily on the measurement of management *practices* at the level of the *organisation* – in particular, with a focus upon practices in operations, monitoring, target setting, and personnel management. This approach has – quite deliberately – steered away from measuring the role of managerial *traits* at the level of the *individual*. Managerial traits at the individual level have been the focus of a much smaller literature in economics – which has studied decisions taken by managers, in their managerial capacity; that is, samples of respondents who (i) have already been selected by firms for their managerial abilities, and (ii) whose managerial challenges are determined by the particular managerial contexts in which they find themselves (see, in particular, Bertrand and Schoar (2003), Malmendier, Tate, and Yan (2011), Kaplan, Klebanov, and Sorensen (2012), Bandiera, Guiso, Prat, and Sadun (2015), Benmelech and Frydman (2015), and Bandiera, Prat, Hansen, and Sadun (2020)).

In this paper, we design and implement a novel style of controlled experiment, designed to measure managerial traits at the individual level. We work with a large sample of young professionals, chosen for their interest in business and entrepreneurship – many of whom are not (yet) working in managerial roles. We invite these young professionals to a video studio, where we explain that they will each watch a series of actors portraying different managerial scenarios. We ask the young professionals to record their responses – as if they are in an entry-level managerial position in a hypothetical firm. With the agreement of those young professionals, we then show these videos to human resources managers in established medium and large manufacturing firms – using an incentive-compatible mechanism to elicit the firms' assessments of the suitability of the young professionals both for wage jobs as entry-level managers and for running their own firms. By designing this controlled experimental setting, we are able to pose the same set of managerial challenges to every respondent in this sample – avoiding the endogenous assignment of managerial tasks to managers that is otherwise inherent in working with observational data on managerial decisions.

Central to our approach is the notion that managerial traits can be quite impressionistic; different managerial styles are often best understood as complex bundles of behaviours, rather than being straightforward to observe and record (Goleman, 2000). For this reason, we deliberately chose a diverse range of

managerial problems (rather than, for example, repeating a single class of scenarios), and we encouraged open-ended answers (rather than, for example, forcing our respondents to choose from a small menu of available actions). Similarly, we then encode respondents' answers in a high-dimensional space (specifically, we measure responses through the combination of managerial action, justification, tone, and source of authority). Additionally, we work with a proprietary video-analysis company to generate summary statistics on visual elements in the videos. To analyse this data, we build and estimate a bespoke Bayesian hierarchical model. This has the flavour of Latent Dirichlet Allocation ('LDA') models (Bandiera et al., 2020; Blei, Ng, and Jordan, 2003; Griffiths, 2004) and – like LDA models – allows us to characterise heterogeneity in managerial styles across a high-dimensional response space.

We have three key results. First, our method – that is, the combination of a studio exercise with a Bayesian hierarchical model – succeeds in identifying meaningful and substantial differences in managerial traits across young professionals. Specifically, we describe four 'pure types' of managerial style – which, following Goleman's (2000) classic work in managerial studies, we label as 'authoritative', 'affiliative', 'coercive' and 'timid'. These types differ quite radically in their conceptualisation of the role of a junior manager. For example, coercive managers seem to view their role in terms of implementing their personal objectives, and do so relying upon their personal authority. Affiliative managers seek shared ground – often yielding to the other side in their responses, and emphasising shared interests as they do so. Authoritative-type managers see themselves as implementing rules set by more senior managers: they are much more likely than other types to rely upon formal policy, and they emphasise the firm's interests in doing so.

Second, young professionals' managerial behaviour is strongly related to their past exposure to the labour market. In particular, we find that authoritative-type respondents are more likely than the other types to have been self-employed, to earn higher wages (conditional on wage employment), and to report higher reservation wages and reservation profits. Strikingly, they are also significantly more likely to have been in the treatment group in a previous field experiment that randomly increased some respondents' exposure to medium and large firms (see our related work, Abebe, Fafchamps, Koelle, and Quinn (2024)); this implies that the relationship between labour market exposure and managerial style is causal. Moreover, our findings suggest that this effect is entirely driven by individuals from who parents did not finish high school. This suggests that this treatment teaches these individuals something about how to behave as a manager that they might not have learnt otherwise.

Third, we find a clear preference across firms for authoritative-type managers over managers of the other three types. We designed our experiment with a view that some firms may (for example) have a clear preference for authoritative-type managers, whereas (for example) others may prefer more affiliative types (see, in particular, Bandiera et al. (2020)). It is striking that we do not find this kind of heterogeneity: while we do detect some heterogeneity in strength and ranking of different styles, firms seem largely united in their view that the authoritative style of management is best – both for entry-level managerial roles and for

self-employment.

Together, these results contribute to three bodies of literature. First, our findings add to our understanding of managerial traits. On the one hand, our result that managerial traits are related to past labour market exposure is broadly consistent with earlier results showing the relevance of past experiences in shaping managerial style (Benmelech and Frydman, 2015; Malmendier et al., 2011). In contrast, our result on the relative homogeneity of firm preferences provides a counterpoint to the seminal earlier work of Bandiera et al. (2020) – who find heterogeneous firm-manager match quality among different kinds of employed senior managers. This presents a novel insight on firm preferences over management: while different firms may have different preferences over the managerial styles of their senior leaders, firms seem reasonably homogeneous in their preferences for entry-level managers among a broad pool of young professionals.

Second, our results complement the recent literature studying labour market exclusion among young professionals. Our results show that individuals with greater labour market exposure are more likely to exhibit the authoritative style of management – and, conversely, that the authoritative style of management is in higher demand among prospective employers. Together, these two results imply that embodied managerial traits may act as a key mechanism by which some young professionals enjoy sustained labour market success while others face exclusion. This supports a key insight from our earlier work (Abebe et al., 2024) – and, indeed, our results here help to interpret the mechanisms identified in that earlier paper.

Third, methodologically, our results contribute to recent literature using innovative and open-ended data collection methods in economics. This includes – in particular – recent literature on the relevance of tone in economic communications (Gorodnichenko, Pham, and Talavera, 2023; Handlan and Sheng, 2023), and on the use of machine learning techniques to understand CEO performance (Borgschulte, Guenzel, Liu, and Malmendier, 2021). More generally, our approach complements recent advances in the use of open-ended survey questions in economics (Haaland, Roth, Stantcheva, and Wohlfart, 2024; Stantcheva, 2021). Separately, our paper illustrates and validates the use of studio vignettes as a credible method for eliciting managerial traits in field settings. This builds on a small literature in economics that has used a different kind of team-based 'management simulation' to test the impacts of managerial training (in particular, Macchiavello, Menzel, Rabbani, and Woodruff (2020), who ran a 'communications game' and a Lego-based 'production game' to measure managerial skills). More generally, we note that the use of video responses to pre-recorded questions is increasingly common as part of job recruitment exercises and university entrance processes, but has received relatively little empirical attention in economics.

2 The experiment

2.1 Vignettes in the studio

Our experiment is designed to measure management traits among young professionals across a series of realistic management scenarios. To do this, we used a studio setting, in which respondents participated in a role-play scenario. Specifically, we ran a series of separate vignettes; in each vignette, the respondent watched a video of a paid actor, who played the role of a counter-party in a managerial problem. For each vignette, we played the video (explaining that it showed an actor), and then asked the respondent to provide a short response – as if she or he were in a managerial role, responding to the character in the video. We explained to the respondents that their video recordings would be played to human resource managers in successful Ethiopian firms.

Specifically, we set each respondent five different scenarios:[1]

(i). *Line management of an employee:* The actor is an accounting clerk who has been absent for three days without warning. (S)he shows up for work on the fourth day, explaining that (s)he was unwell.

(ii). *Negotiating with a supplier:* The actor plays a supplier, who explains that (s)he cannot fulfil an order according to specifications, because of problems sourcing input materials. (S)he offers to supply a replacement of inferior quality instead. Respondents are told specifically that the firm they represent is known for producing and selling the highest quality products in the industry.

(iii). *Negotiating a pay rise:* The actor is a production worker, who comes to ask for a pay rise. The worker argues that (i) within her unit, (s)he has been the most productive worker during the last three months, and one of the few people who exceeded the personal productivity targets; and (ii) (s)he has been with the company for ten years, and even her mother used to be an employee of the company. Respondents are told that the firm they represent has no plan to increase any salaries this year.

(iv). *Negotiating an adjustment with the bank:* The actor is a bank manager, who calls to remind the firm about an unpaid loan installment (explaining that failure to pay increases interest payments and reduces the firm's credit rating). Respondents are told that the firm they represent will not have sufficient funds to pay the bank for another two weeks.

(v). *Negotiating with a client:* The actor is a client, who has not paid what you have invoiced, in spite of one reminder. The client comes to place a new order.

[1] The full scenario scripts are provided in Appendix A.

2.2 Incentivised elicitation of firm preferences

In the second part of our experiment, we played the studio recordings to human resources managers from successful Ethiopian firms.[2] For each vignette, the human resources manager watched three responses, and was asked to rank these separately based on (i) the manager's assessment of the respondents' suitability as entry-level managers at their firm, and (ii) the manager's assessment of the respondents' suitability to run their own small business.

Specifically, each human resources manager was assigned to assess video recordings from three different respondents, for each of the five vignette scenarios. Each triplet of young professionals was assessed by at least two different managers for each vignette. To elicit revealed preferences comparing different candidates, we implemented two complementary direct elicitation mechanisms: one to elicit the perceived suitability of candidates for a managerial position in the firm; the other to elicit candidates' perceived suitability to run their own business. We incentivised the assessment of suitability for managerial positions through the prospect of receiving respondents' contact details (with respondents' permission), and we incentivised the assessment of suitability to run a business through the prospect of the respondent being invited to a business plan competition. They were designed to be straightforward and to make truthful ranking an 'obviously dominant' strategy for respondents.[3] Separately, we asked each human resources manager directly to assess the individual respondents (again, both for their suitability as an entry level manager, and as an entrepreneur), using a series of stated-preference questions. We explain both the revealed-preference and stated-preference methods in detail in Appendix B.

2.3 Implementing our experiment

Table 1 provides descriptive statistics for the 982 young professionals in our sample. The respondents are primarily male (78%), highly educated (78%), and relatively young (with an average age of 31). Most of the young professionals are in wage employment (68%) or self-employment (17%). Of those in wage employment, 98% work in a professional position, and 18% in a managerial position. Conditional respectively on wage employment and self-employment, the median wage and profit are both 6000 ETB (although self-employment income varies substantially more across respondents). This is more than twice the average household expenditure of 2500 ETB in Addis Ababa.[4]

[2] We ran this assessment with the senior member of the firm who was responsible for hiring. We refer to this person here as 'the human resources manager' – though this was not always the actual title or job description.

[3] Both mechanisms closely resembled the 'OSP-RSD' ranking mechanism described by Li (2017).

[4] See the Ethiopian Socioeconomic Survey (Ethiopian Statistical Service and World Bank, 2023).

Table 1: Descriptive statistics of young professionals

Categorical Variables				
Gender (% Male)	78%			
University degree	78%			
Self-employed	17%			
Wage-employed	68%			
Conditional on Wage Employment				
In a professional position	98%			
In a managerial position	18%			
Continuous Variables				
	Median	Std. Dev.	5th Perc.	95th Perc.
Age (years)	31	2.8	27	36
Wage (ETB/Month)	6000	5979	2285	18000
Profit (ETB/Month)	6000	16099	0	42000
N	982			

Notes This table provides summary statistics of the respondents included in the sample. The wage and profit in Ethiopian Birr (ETB) are calculated conditional on being in wage- and self-employment, respectively. The percentages of individuals in professional and managerial positions are conditional on being in wage employment.

The sample of firms was drawn from a set of 713 established Ethiopian firms involved in a previous experiment (Abebe et al., 2024), supplemented by additional firms to account for attrition. These are medium-sized to large firms operating across a range of industries, in Addis Ababa and in the neighbouring towns of Bishoftu and Adama.

We conduct interviews with a sample of 576 firms. These firms have a median of 58 employees (mean 323) and employ a median of five managers (mean 12). The interviewed managers work primarily in human resources (40%) and administration (34%). They have been in their current position for, on average, eight years. 80% have a university degree and 72% have formal management training.

2.4 Experimental variation in labour market exposure

For our experiment, we invite all individuals who participated in a previous experiment conducted in 2016 and 2017 that placed young professionals in medium and large firms for a one-month placement

shadowing middle managers in their daily work (Abebe et al., 2024). During the placement, the treatment, they shadowed managers in their daily work, gaining direct exposure to managerial practices in a real-world business environment. This experiment allows us to study the causal effect of exposure to firm management on management styles.

As only around 60% of the original sample attends the studio, we study whether the sample that attended the studio is comparable to the original sample, and whether we can continue to think of treatment as randomly assigned, or whether for example treated individuals are more likely to attend. First, to compare the sample attending the studio compared to the original sample in the experiment, and the balance on treatment conditional on attending the studio, we run a number of multinomial logit regressions. In Table 2, we first examine whether some individuals were more or less likely to attend the studio. In column (1) we show that, compared to the baseline sample female participants are significantly less likely to attend, and that individuals with a BA degree are less likely to attend (although this is marginally significant). When we study attendance to the studio conditional on being in the four-year followup of the experiment during which we invited participants(column 2), we similarly find that female participants are less likely to participate, and that individuals with an above median reservation wage (conditional on employment) are less likely to attend. There are many potential explanations why female candidates are less likely to attend, for example, one might hypothesize they may be less willing to be recorded as part of the experiment. Importantly, we find no evidence for the hypothesis that treated individuals are more likely to attend the studio.

In column (3), we study whether any individual characteristics are predictive of treatment status conditional on attending the studio. We find suggestive evidence that female participants who are treated appear to be less likely to drop out of the experiment, and the sample appears to be well-balanced on all other characteristics. Importantly for our subsequent analysis, we find good balance on treatment status and whether either parent finished primary school.[5]

[5] In due course, we intend to further examine the robustness of our results to these issues.

Table 2: Logit Regression Results on Studio Attendance and Management Experience Treatment

	(1) Baseline sample	(2) 4-year follow-up sample	(3) Studio sample
	Attended Studio	Attended Studio	Treated
Treated	0.0208	0.0159	
	(0.0239)	(0.0241)	
Gender (1=Female)	-0.138***	-0.121***	0.0783*
	(0.0275)	(0.0283)	(0.0413)
Age at baseline	-0.00425	-0.000840	-0.00149
	(0.00435)	(0.00456)	(0.00589)
BA Degree	-0.109*	-0.0175	0.0142
	(0.0591)	(0.0545)	(0.0706)
Either parent finished primary school	-0.0135	-0.000208	-0.0128
	(0.0243)	(0.0252)	(0.0327)
Either parent owned a business	-0.0311	-0.0123	-0.00308
	(0.0247)	(0.0252)	(0.0333)
In wage employment		0.00745	0.0664
		(0.0318)	(0.0420)
Manages others while in wage emp.		0.0635	-0.0261
		(0.0420)	(0.0508)
In self-employment		-0.0127	0.0609
		(0.0395)	(0.0516)
Above median reservation wage		-0.0726**	0.00834
		(0.0294)	(0.0390)
Above median reservation profit		0.0470	0.0346
		(0.0299)	(0.0383)
Observations	1637	1429	993

Notes: This table reports the average marginal effects from a multivariate logit regression on attending the studio conditional of being part of the previous management experience experiment in column (1) and conditional on being part of the four-year follow-up in column (2). Finally, column (3) reports the average marginal effects from a multivariate logit regression on being treated conditional on attending the studio in column (2). As independent variable a set of variables related to socioeconomic background of studio participants, and labour market outcomes before attending the studio are included. Robust standard errors in parentheses. Significance levels: $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

3 Management styles in the studio

Our first experimental objective is to characterise the main management styles among young Ethiopian professionals. To do this, we build a bespoke Bayesian Hierarchical Model; this allows us to identify distinct clusters of managerial behaviour. To do this, we first employed two enumerators to watch all of the videos, and to encode each response along the following four dimensions:[6]

(i). *Action*: Did the respondent agree or disagree with the actor's request? (For example, in the third scenario, did the respondent agree to the employee's request for a pay rise?)

(ii). *Authority*: What source of authority did the respondent rely upon? Specifically, did the respondent rely upon (a) their formal authority as a manager, (b) their seniority or personal authority, (c) higher principles, such as an appeal to the appropriateness of the action, or (d) did the respondent not rely upon any source of authority.

(iii). *Justification*: How did the respondent justify her or his action? Specifically, did the respondent (a) provide no justification, (b) emphasise the interests of the firm they represent, (c) emphasise the interest of the other party, or (d) emphasise the shared interest of the firm and the other party?

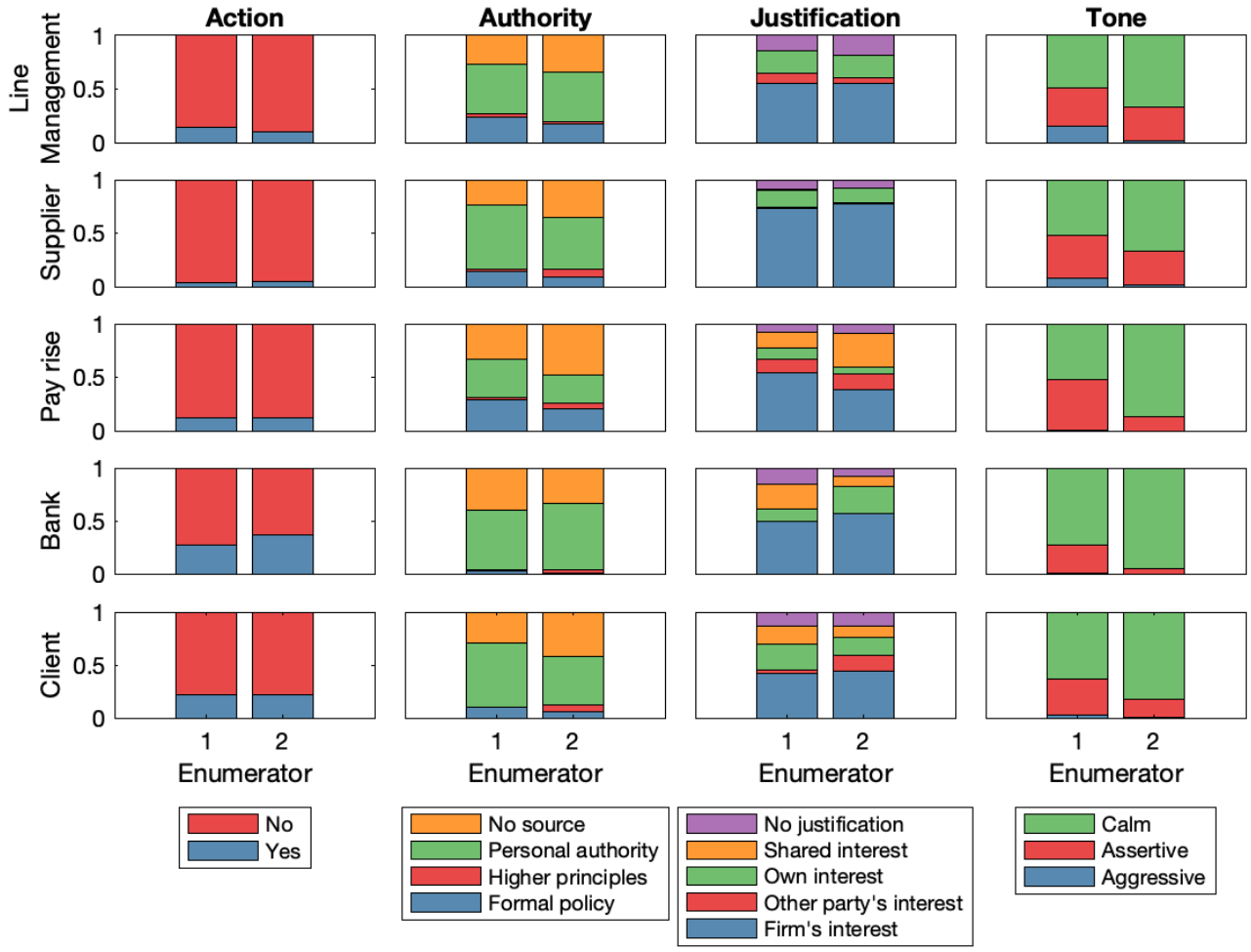
(iv). *Tone*: What tone did the respondent use? Specifically, was it (a) calm/assured, (b) assertive, or (c) aggressive?

3.1 Management traits among Ethiopian young professionals

We begin, in Figure 1, by describing the average behaviour across our five vignettes. Specifically, Figure 1 shows one row for each vignette; across that row, we characterise the average behaviour across each dimension of management (as discussed, these are action, authority, justification and tone). We disaggregate the data by the two enumerators. We draw two conclusions from the figure. First, we find substantial heterogeneity *across* different vignettes – and this is reflected across the bundle of different managerial dimensions. For example, when dealing with an employee seeking a pay rise (vignette 3), most respondents refuse the request; to do so, they rely heavily on formal authority and on personal authority, and their tone can be relatively assertive. In contrast, when dealing with their bank (vignette 4), respondents are much more likely to accede to the request, to place almost no reliance upon formal policy, and generally to use a less assertive tone. Similarly, when dealing with an employee seeking a pay rise (vignette 3), respondents frequently invoke their shared interest, but this rarely happens when dealing with a supplier failing to deliver a high-quality good (vignette 2).

[6] Appendix C provides a detailed description.

Figure 1: **The distribution of behaviours by vignette**



Notes This figure reports the distribution of the encoded behaviours by vignette and by enumerator. The bars show the probability that an enumerator encodes a specific behaviour for each dimension by vignette.

Second, there is substantial heterogeneity *within* different vignettes, across respondents. This is evident from the disaggregation across the separate enumerators. Despite assessing the videos independently, the enumerators describe the responses in broadly similar terms – implying that the observed differences in action, authority, justification and tone are driven by heterogeneous approaches by different young professionals, rather than by differences in enumerators' perceptions.

3.2 A Bayesian model of management styles

To extend this analysis, we now build a bespoke Bayesian Hierarchical Model. This model serves two related purposes. First, it allows us to cluster the different dimensions of managerial responses into dif-

ferent managerial styles. Second, the resulting classification will be a key input for understanding firm preferences in the following section.[7]

To do this, we specify a generative model in which each young professional – when responding to a given vignette – chooses a combination of action, authority, justification and tone. The model characterises 'pure types' that represent archetypal managerial styles; each young professional is then represented as a convex combination of those archetypes. A Bayesian Hierarchical framework is well suited for this task due to its ability to handle high-dimensional categorical data and to uncover latent themes. Very broadly, this is similar to several recent papers on heterogeneity in management styles (e.g. Bandiera et al., 2020).[8]

Specifically, we observe individuals $i \in \{1, \ldots, N\}$ performing on vignettes $v \in \{1, \ldots, 5\}$. For each individual assessment of a vignette, we have two enumerators, $e \in \{1, 2\}$. Each enumerator records a set of 'attributes' of the response (action, authority, justification and tone): $a \in \{1, \ldots, 4\}$. Each attribute a has $J(a)$ possible categorical responses, $y^a_{ive} \in \{1, \ldots, J(a)\}$.[9] For each vignette, each respondent draws their behaviour from one of K pure types, z_{iv}. The total probability of the model is as follows:

$$P(W, Z, \boldsymbol{\theta}, \boldsymbol{\phi}, \zeta, \eta) = \underbrace{\prod_{j=1}^{K} P(\boldsymbol{\phi_j}; \zeta)}_{\text{Dirichlet parameters}} \cdot \underbrace{\prod_{i=1}^{N} P(\boldsymbol{\theta_i}; \eta)}_{\text{Types}} \cdot \prod_{v=1}^{5} \sum_{k=1}^{K} \left(\underbrace{P(z_{iv} = k | \boldsymbol{\theta_i})}_{\text{Type assignment}} \cdot \underbrace{\prod_{e=1}^{2} \prod_{a=1}^{4} P(y^a_{ive} | \boldsymbol{\phi}_{ka}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ae})}_{\text{Studio behaviour}} \right), \quad (1)$$

where $P(\boldsymbol{\phi}_j; \zeta)$ and $P(\boldsymbol{\theta}_i; \eta)$ follow a Dirichlet distribution, $P(z_{iv} | \theta_i)$ follows a categorical distribution and $P(y^a_{ive} | \boldsymbol{\phi}_{ka}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ae})$ follows a Multinomial Logit distributions.

In this model, $\boldsymbol{\phi}_{ka}$ characterises the behaviour of the pure types and $\boldsymbol{\theta}_i$ characterises an individual as a convex combination of those archetypes. z_{iv} is an individual's type for a specific vignette. $\boldsymbol{\psi}_{av}$ is a vignette fixed effect that allows average behaviour to vary across vignettes, and $\boldsymbol{\chi}_{ae}$ is an enumerator fixed effect that allows average behaviour to vary across enumerators. We estimate this model using a Hamiltonian Monte Carlo algorithm in the Stan language; we discuss the parameterisation and estimation of this model in more detail in Appendix E.

[7] For this section, we only uses the data from the enumerator encoding of the vignettes to estimate our model.

[8] Our model has some analogy to the class of Latent Dirichlet Allocation ('LDA') models. In the language of LDA models, we could think of each respondent-vignette pair as comprising a 'document', each document comprising four 'words', and each word being drawn from one of four different dictionaries (namely, a dictionary for action, a dictionary for tone, a dictionary for source of authority and a dictionary for justification. Our inclusion of enumerator and vignette effects is a further variation on the typical LDA setup.

[9] Specifically, $J(1) = 2$ for the action; $J(2) = 4$ for source of authority, $J(3) = 5$ for justification and $J(4) = 3$ for tone.

3.3 Results

We estimate our model allowing for four 'pure' management styles.[10] To interpret these styles, we draw upon the classic work of Goleman (2000) – who described a set of archetypal leadership traits for large organisations. This has been a highly influential description of leadership for many corporate thinkers.[11] In his famous work, Goleman described six styles of *senior leadership*. We use this as a basis for describing four styles of *entry-level management* – the setting appropriate for our sample.[12]

Figure 2 shows the four estimated 'pure type' management styles and their behaviour in our five vignettes:

(i). Type 1: The *authoritative manager* implements the rules set by the firm. Specifically, they refuse the requests almost all of the time; in doing so, they are much more likely than the other types to rely upon formal policy – and much more likely to emphasise the firm's interests in doing so. They are more likely than the other three types to use an assertive tone; indeed, at times, they can even be aggressive.

(ii). Type 2: The *affiliative manager* seeks shared ground. Strikingly, the affiliative type is much more likely than the other three types to concede to their counterparty in the hypothetical discussion – notwithstanding that many of the vignette prompts suggested that such concessions were not in the interests of the hypothetical firm. The affiliative type tends to rely upon personal authority (or cites no specific authority), and is much more likely than the other three types to emphasise shared interests. Their tone is generally calm, but can be assertive.

(iii). Type 3: The *coercive manager* implements their personal objectives. They refuse the requests almost all of the time; in doing so, they are overwhelmingly likely to rely on personal authority. As justification, they tend to emphasise the firm's interest – but are much more likely than the other three types to justify their actions by reference to their own personal interest. Similar to the authoritative type, they tend to be assertive, or even aggressive.

(iv). Type 4: The *timid manager* does not provide a clear explanation for their decision. They refuse the requests almost all of the time – and, in doing so, are unlikely to rely on any source of authority, and

[10] In Appendix D, we consider alternative versions with two, three and five types and show that the general conclusions remain broadly similar.

[11] This seminal piece had been cited 5710 times at the time of writing according to Google Scholar and featured in the Harvard Business Review's classic series of groundbreaking ideas in 2017.

[12] Goleman's six styles were: (i) coercive, (ii) authoritative, (iii) affiliative, (iv) democratic, (v) pacesetting and (vi) coaching. Of these, we use the coercive, authoritative and affiliative styles; to this, we add a 'timid' style that is evident in our results. Goleman's other three styles – namely, democratic, pacesetting and coaching – are primarily about *leadership* rather than *management*. In particular, Goleman describes those styles with their emphasis on coordinating teams and developing junior staff; these roles are not directly applicable to our management vignettes, nor to our sample of young professionals. Goleman's 'authoritative' leader was defined in terms of mobilising teams towards a vision; in the context of entry-level management, we repurpose this term to refer to enthusiasm to implement shared policy.

often provide no specific justification for their decision. Of the four pure styles, timid managers are most likely to be calm in their tone.

In Figure 3, we describe the distribution of individuals over types. To do this, we draw a tetrahedron, in which each vertex represents probability one of belonging to a particular type. We learn three things from this figure about our model results. First, the model distributes individuals quite evenly across the tetrahedron, such that any given individual exhibits a combination of different 'pure type' behaviours. Second, very few individuals sit close to the affiliative vertex: this implies that affiliative behaviour, while empirically important, is something that individuals incorporate into their responses – rather than being a style that consistently describes any particular individual. Third, the stacked bar on the right of Figure 3 shows the observed vignette responses in terms of the average distribution of each pure type across individuals. We estimate that the most prevalent are the authoritative type (28.0%) and the coercive type (27.9%); this is followed by the timid type (25.2%) and then the affiliative type (19.0%).

3.4 Managerial types and vignette content

Prior to recording each video, we asked each respondent how they intended to answer, and why. Table 3 describes the average prevalence of different answers to these questions, across the four identified types.[13] Specifically, we report prevalence of (i) distrust as a stated motivation for the planned response, (ii) relationship maintenance as a stated motivation for the planned response, (iii) intention to explain the decision in terms of fixed company procedure, (iv) intention to set an example for future cases.

These reported intentions align with our earlier interpretation of the four types. The affiliative type, for example, is least likely to distrust their counterparty, and much more likely to emphasise the value of maintaining an ongoing relationship; in doing so, they are far less likely than the other types to emphasise the importance of formal procedure. In contrast, the authoritative type stands out for its relatively high distrust of the counterparty and – above all – for being much more likely than the other three types to list formal procedure and setting of an example as key motivators. The coercive and timid types are similar to each other in their stated intentions, and both place a higher emphasis on formal procedure than the affiliative type. In Appendix D.1, Table A.3 provides a breakdown of the average prevalence of these intentions by vignette.

[13] To generate these statistics, we use the point estimates of $\hat{\boldsymbol{\theta}}_i$ as an estimate of the probability of the individual to belong to each pure type. We then assign them to a pure type T_i using: $T_i = \arg\max \hat{\boldsymbol{\theta}}_i$.

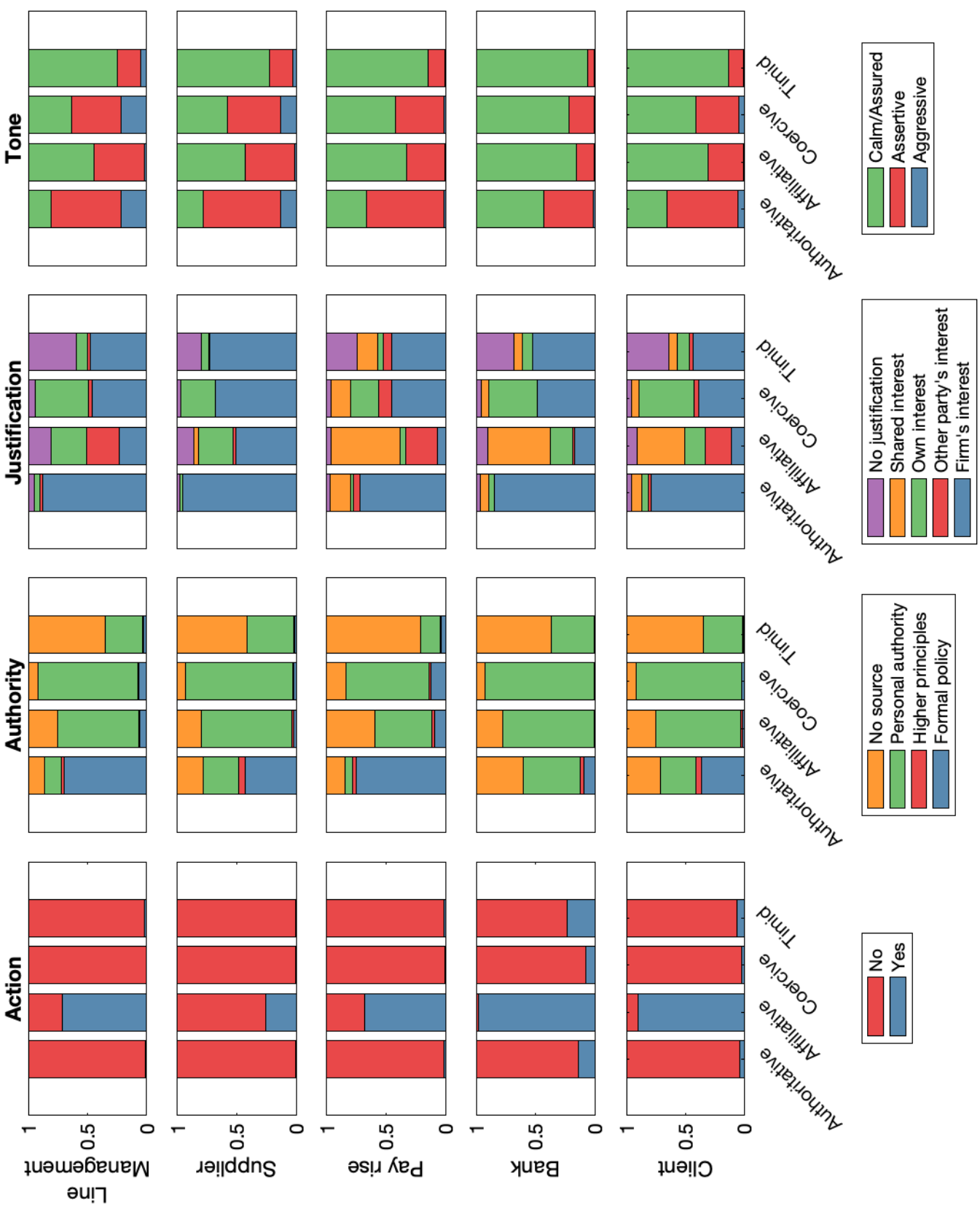


Figure 2: **'pure type' management styles amongst Ethiopian young professionals**

Figure 3: Distribution of types across individuals

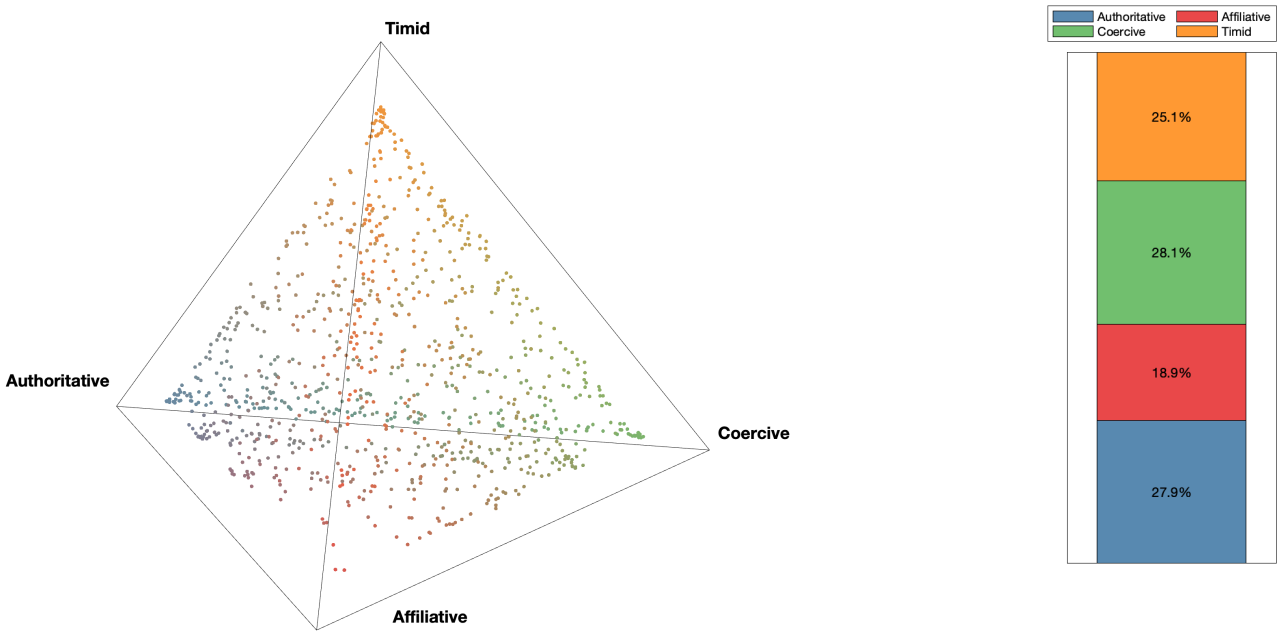


Table 3: **Reported intentions by type**

Type	Expresses distrust counterparty	Maintain relationship	Follow procedure	Set an example
Timid	.257	.218	.347	.241
Authoritative	.3	.259	.453	.332
Affiliative	.2	.332	.306	.25
Coercive	.274	.233	.361	.245
p-value	<0.001	<0.001	<0.001	<0.001

Notes This table describes the intentions of respondents', which are enumerated before the respondent actually responds to the vignette. This is based both an what the respondent intends to say, and their motivations for doing so. The four characteristics are (i) whether the respondent expresses they do not trust the other part, (ii) whether the respondent wishes to maintain a good relationship, (iii) whether the respondent mentions following procedure, and (iv) whether the respondent wishes to set an example for future interactions. These numbers are calculated by assigning each individual to the pure type for which they have the highest estimated $\hat{\boldsymbol{\theta}}_i$. Then, a conditional average is taken for each pure type. To test the significance of the differences between these conditional distributions, we conduct a joint test to assess whether the four-dimensional difference vector is credibly nonzero. Using an elliptical approximation, we compute the Mahalanobis distance of the posterior mean from zero, scaled by the posterior covariance, and compare it to a Chi-squared distribution with 3 degrees of freedom to calculate the p-value.

4 Characteristics of managerial types

4.1 Managerial types and labour market experience

Having estimated the distribution of individuals' across types, we now study the labour market outcomes and trajectories of these individuals conditional on their managerial traits. Table 4 describes the average characteristics of individuals, conditional on their type assignment. The authoritative type is most likely to be male and self-employed, and is more likely than the other types to have an above-median reservation wage and profit. The differences among the other three types are more subtle, but a few distinctions are worth noting: the affiliative type is the type most likely to be female, the timid type is particularly likely to have a below-median reservation wage, and the coercive type is likely to have an above-median reservation profit.

Next, we examine the relationship between labour market experience and manager types. Approximately 12 months after respondents' participation in the study, we asked them a series of questions regarding their labour market experiences over the past six years, including their employment status at six-month intervals. Table 5 provides an overview of the average labour market experiences across different manager types based on these responses. The table indicates that authoritative-type managers had spent more years in employment, particularly in permanent positions. Further, these managers exhibited more stable employment histories, transitioning between jobs approximately 0.2 fewer times on average compared to other types.

Table 4: **Characteristics and Types: Summary Statistics**

Type	Gender [1=male]	Wage employment indicator	Self-Employment indicator	Above Median Reservation Wage	Above Median Reservation Profit	Average Duration Response
Timid	.795	.666	.146	.266	.299	36.2
Authoritative	.846	.709	.201	.45	.427	55.3
Affiliative	.771	.678	.151	.336	.305	38.8
Coercive	.778	.679	.144	.332	.371	41.8
p-value	<0.001	.276	0.004	<0.001	<0.001	<0.001

Notes This table describes the average characteristics of individuals of each type. Specific, this includes their gender (1 indicating male, 0 female), a dummy for their wage- and self-employment status, and the probability they have an above-median reservation wage and profit based on data collected before respondents' attended the studio. We also include a dummy for whether or not the individual was treated in the management placement experiment, and the average duration of the responses of the candidate across the vignettes. Note that the median splits on reservation wage and profit do not yield a 50/50 split due to bunching in the underlying data at ETB10.000. These numbers are calculated by assigning each individual to the pure type for which they have the highest estimated $\hat{\boldsymbol{\theta}}_i$. Then, a conditional average is taken for each pure type. To test the significance of the differences between these conditional distributions, we conduct a joint test to assess whether the four-dimensional difference vector is credibly nonzero. Using an elliptical approximation, we compute the Mahalanobis distance of the posterior mean from zero, scaled by the posterior covariance, and compare it to a Chi-squared distribution with 3 degrees of freedom to calculate the p-value. For non-binary outcomes, we first create a binary split at the median to split the sample.

Table 5: **Labour market experience and Types: Summary Statistics**

Type	Employment	Permanent employment	Unemployed	Management Position	Number of transitions
	Years	*Years*	*Years*	*Share*	*Count*
Timid	5.226	3.541	.577	.109	.887
Authoritative	5.522	3.906	.329	.192	.659
Affiliative	5.306	3.671	.456	.124	.851
Coercive	5.231	3.626	.51	.139	.939
p-value	.003	<0.001	<0.001	<0.001	<0.001

Notes This table describes the labour market experience over the past 6 years of individuals of each type. Specific, this includes the number of years they have been in employment including both self- and wage-employment, the number of years they have been in permanent employment, the number of years they have been unemployed, whether they were in a management position before participating in the studio experiment, and finally the number of labour market transitions they have gone through. These numbers are calculated by assigning each individual to the pure type for which they have the highest estimated $\hat{\boldsymbol{\theta}}_i$. Then, a conditional average is taken for each pure type. To test the significance of the differences between these conditional distributions, we conduct a joint test to assess whether the four-dimensional difference vector is credibly nonzero. Using an elliptical approximation, we compute the Mahalanobis distance of the posterior mean from zero, scaled by the posterior covariance, and compare it to a Chi-squared distribution with 3 degrees of freedom to calculate the p-value. For non-binary outcomes, we first create a binary split at the median to split the sample.

4.2 The causal effects of managerial experience

The results in the previous section suggest that authoritative managers have better labour market outcomes, but do not allow us to study whether labour market experience influences the development of managerial traits, or whether managerial traits, instead, determine labour market outcomes. To answer this question, we turn to evidence from respondents' participation in a management experience experiment five years prior to attending the studio.

This experiment allows us to estimate the causal effect of management experience on different management styles among young professionals. For each management style – Timid, Authoritative, Affiliative, and Coercive – we compute the treatment effect as the difference between the estimated latent trait for treated versus untreated individuals. We calculate the average treatment effect as:

$$\Delta\theta_i^{(d)} = \theta_{i,\text{treated}}^{(d)} - \theta_{i,\text{control}}^{(d)},$$

where $\Delta\theta_i^{(d)}$ represents the treatment effect for style i for draw d. For each category, we report the mean and the 95% credible interval of the distribution of these draws. We separately conduct this exercise for three groups: the full sample, individuals whose parents did not complete primary school, and those with at least one parent who did.

We find, in Table 6, that the management experience treatment makes individuals on average more authoritative and less timid (column 2), and that this result is entirely driven by individuals whose parents' have not finished primary school (column 4). We find no evidence for any treatment effect for individuals

whose parents finished primary school (column 6).

Table 6: The causal effect of managerial experience on management style by parents education

	Full sample		Low parental education		High parental education	
	(1)	(2)	(3)	(4)	(5)	(6)
Timid (%)	27.1	-3.6	29.4	-5.3	24.9	-2.0
		[-5.8, -1.5]		[-8.3, -1.8]		[-5.3, 1.1]
Authoritative (%)	26.5	2.4	22.5	5.4	30.2	-0.3
		[0.2, 4.5]		[2.4, 8.7]		[-3.4, 2.6]
Affiliative (%)	18.4	1.2	19.1	1.0	17.7	1.4
		[-0.5, 2.9]		[-1.6, 3.6]		[-1.1, 4.0]
Coercive (%)	28.0	-0.1	29.0	-1.2	27.1	0.9
		[-2.3, 2.3]		[-5.0, 2.1]		[-2.5, 4.2]
N	479	500	229	239	250	261

Notes This table reports the treatment effect of the management experience experiment on the managerial traits of individuals. The treatment effect is calculated based on the distribution of the difference in the average value of θ for treated and untreated individuals. Columns (1), (3) and (5) report the average estimated value of $\boldsymbol{\theta}_i$ for individuals that were not treated in the management experience experiment for respectively all individuals, individuals whose parents did not finish primary school and for individuals for whom at least one parent did. Columns (2), (4) and (6) report the treatment effect of the management experience experiment on their managerial traits for these three groups respectively. In columns (2), (4) and (6) both the average treatment effect and the 95% credible interval, in square brackets, are reported.

We draw two conclusions from this result. First, management experience shapes an individual's management style. Since the treatment focused on shadowing a manager rather than directly managing, this suggests that the intervention teaches individuals what effective management looks like rather than specifically how to implement it. This interpretation is further supported by the heterogeneity in treatment effects between individuals whose parents did not complete primary school and those whose parents did. We suspect that individuals from the former group have had less exposure to firms and business practices in Ethiopia, resulting in a larger treatment effect on management styles. Strikingly, the treatment significantly reduces the gap in the probability of being an authoritative manager between these two subsamples, decreasing the difference in respondents' reliance on this management style from 7.7 percentage points for untreated individuals to 2.0 percentage points for treated individuals. These findings suggest that managerial experience plays an important role in shaping managerial traits, particularly for individuals with less prior exposure to formal firms.

4.3 Aspiring managers' perceptions

A recurring theme in organisational economics is that poorly managed firms may not even realise that they are poorly managed (see, for example, Bloom, Lemos, Sadun, Scur, and Reenen (2014); Gibbons and

Henderson (2012); Rivkin (2000)). Previous empirical measures on this issue (for example, in the World Management Survey) have tended to ask firms' subjective assessments of how well they are managed relative to others. Our design makes further progress on this issue by measuring aspiring managers' perceptions of others' performance.

To measure this, we developed one additional role-play scenario — in which workers were on strike demanding proper latrines, and a representative of the workers came to talk to the respondent as their manager. We recorded an actor portraying the worker representative. We then used different actors to record six separate possible managerial responses. We scripted these responses to represent six of the possible extreme points in the 'action/justification/authority/tone space' that we had described earlier; in a sense, these responses represented Weberian 'ideal types' capturing different bundles of management traits (Weber, 1904). For example, one actor might have (i) refused any change in company practice, (ii) justified her action purely through the company's interests, (iii) appealed to her authority, and (iv) spoken aggressively in doing so. Another actor might have (i) agreed a change in company practice, (ii) justified her action by appealing to the shared interests of company and workers in having a harmonious workplace, (iii) appealed to formal company policy, and (iv) spoken in a calm/assured manner in doing so. We provided the full script for each of these six scenarios in Appendix Section A.2.

We showed these responses to a senior human resources consultant at an Ethiopian firm (explaining that the participants were actors), and we asked him to agree to a ranking.[14] We showed each of the six recordings to each respondent. In each case, we asked the respondent to predict how the senior managers had ranked the same responses; we incentivised this response by promising a transfer of phone airtime in return for a correct answer. We also asked (unincentivised) that each respondent rank the six recordings in the order of how closely the recordings reflected their own management style.

Table 7 summarises each of the responses, ordered based on the ranking of the senior manager, with 'manager 1' being the best response according to the senior manager, and 'manager 6' the worst response. Around 60% of the respondents correctly predicted that the senior manager ranked manager one the highest, and respectively 30% and 45% of respondents thought the senior manager had ranked

[14] The senior manager was asked to create a single ranking of the twelve videos including those with both the male and female actors. He always ranked the male and female actor acting out the same script equivalently compared to actors ranking other scripts. The senior manager explained to us that, in deciding his rankings, he focused primarily on the message, next the managers' gestures and sympathy, and finally their tone.

responses 5 and 6 as the worst possible response.

Table 7: Summary of Managerial Responses to Hypothetical Workplace Conflict

	Response
Manager 1	**Apologetic, collaborative, empowering** - Apologizes for slow response. - Proposes forming a committee with workers' involvement to address issue. - Pledges resources and seeks cooperation to end strike.
Manager 2	**Acknowledging, proactive, realistic** - Accepts request as fair and promises improvements. - Details immediate and long-term actions needed. - Requests patience and asks for strike to end.
Manager 3	**Appreciative, conciliatory, blame-shifting, behavioral focus** - Thanks representative and recognizes issue. - Blames workers for misusing toilet breaks. - Suggests ending strike and working together to improve behavior before addressing facilities.
Manager 4	**Acknowledging, deflective, postponing, subtle pressure** - Recognizes importance of issue but cites current financial struggles. - Suggests postponing improvements and implies potential career consequences for pushing issue.
Manager 5	**Dismissive, threatening, defensive** - Claims company's facilities are sufficient. - Compares to other companies' and workers' home facilities. - Threatens workers to cease strike or face consequences.
Manager 6	**Legalistic, hostile, threatening, dismissive** - Refuses to discuss issue. - Cites labour law violations and threatens legal action. - Demands strike to end immediately.

Notes This table summarises the managerial responses to the hypothetical workplace conflict, ordered based on the ranking of the senior manager. The adjectives describing each of the vignettes were generated by ChatGPT.

We use Kendall's Tau as a measure of agreement for three comparisons: (i) the comparison between what respondents think senior managers will do and what senior managers actually do, (ii) the comparison between what respondents say they would do themselves and what senior managers actually do, and (iii) the comparison between what respondents say they would do themselves and what they think senior

managers would do. Table 8 summarizes these comparisons for the full sample of respondents, as well as for subgroups based on the pure type for which respondents have the highest estimated $\hat{\boldsymbol{\theta}}_i$. Kendall's Tau measures the strength of ordinal association between two rankings, based on the relative proportion of agreements and disagreements across pairs.

Table 8: Agreement in rankings between the respondents' and the HR consultant's rankings.

	(1) Prediction Senior Manager and Senior Manager	(2) Own Action and Senior Manager	(3) Internal Agreement
Overall	.511	.415	.793
Timid	.493	.395	.808
Authoritative	.566	.47	.797
Affiliative	.499	.416	.781
Coercive	.476	.375	.779

Notes Table 8 displays Kendall's Tau values comparing the respondents' rankings and the expert's rankings. The column "Prediction of Senior Manager versus Actual Senior Manager" compares the respondents' predictions of the senior manager's actions with the actual senior manager's actions. The column "Own Action versus the Senior Manager" compares the senior manager's actions with what the respondent would do themselves. The column "Internal Agreement" displays Kendall's Tau for the respondents' perception of their own actions versus what they expect a senior manager would do. Kendall's Tau is calculated as $\tau = \frac{\sum \text{agreements} - \sum \text{disagreements}}{\sum \text{pairwise comparisons}}$.

In comparison (1), where respondents predict what senior managers will do, we observe a moderate degree of agreement, with Kendall's Tau equal to 0.511 averaged over all respondents. This suggests that respondents can predict senior managers' decisions with reasonable accuracy. The authoritative type stands out with the highest alignment (0.566), indicating a stronger shared understanding with senior managers. In contrast, the timid type shows the weakest alignment (0.476), suggesting greater difficulty in predicting managers' decisions. The affiliative (0.499) and coercive (0.493) types are characterised by Kendall's Tau's these two extremes, implying they correct predict the senior manager's pairwise ranking just under 75% of the time.

In comparison (2), where we compare respondents' own decisions to those of the senior manager, agreement rates are consistently lower. The overall Kendall's Tau is 0.415, roughly 0.1 lower than in comparison (1). This indicates that while respondents have some sense of the managers' decisions, their own actions tend to diverge more. Once again, the authoritative type demonstrates the highest agreement (0.470), while the timid type struggles the most (0.375), showing the largest gap between their personal decisions and senior managers' decisions.

Finally, in comparison (3), internal consistency – how respondents' own decisions align with their expectations of senior managers – is high for all types. The overall Kendall's Tau is 0.793, highlighting strong internal agreement. This suggests that respondents are confident in the alignment between their

own decision-making processes and their expectations of senior managers. Across subgroups, internal agreement remains consistently high, with values ranging from 0.779 (timid) to 0.808 (coercive).

5 Firm preferences for management styles

The key next question is about how this is received: do these differences reflect (say) alternative approaches to management that may be valued differently by different kinds of firms (as in Bandiera et al. (2020)), or is one type clearly preferred by firms? In this section, we turn to the results from our incentivised elicitation of firm preferences over the vignette responses of our respondents.

To assess how heterogeneous firms are in their revealed preferences for aspiring manager, we first assess the probability that two managers agree with each other, both for the full sample and by individual vignette. Table 9 shows that the probability of agreement between two managers varies between 56% and 61% across the vignettes.

To quantify the level of agreement between firms, we implement a random-effects rank-ordered logit model without covariates; we specify the latent utility from firm f assessing candidate i as:

$$y^*_{fivs} = \mu_{ivs} + \varepsilon_{fivs},$$

where we specify $\mu_{ivs} \sim N(0, \sigma^2_\mu)$ and we assume that ε_{fivs} has a Type 1 Extreme Value distribution.[15] We assume that μ_{ivs} is independent across vignettes and types of employment (so, in effect, we estimate this model separately for each vignette and each type of employment).[16] This model reveals much about firm preferences and the notion of 'management as technology' because the variance σ^2_μ captures the strength of relative agreement among firms. In one limiting case, firm preferences are completely idiosyncratic: $\sigma^2_\mu = 0$. In another limiting case, preferences are common across all firms: $\sigma^2_\mu \to \infty$. An important intermediate case is where the proportion of variation due to common preferences equals the proportion due to idiosyncratic preferences: $\sigma^2_\mu = \pi^2/6$. In this way, estimation of σ^2_μ provides a specific figure — to our understanding, the first formal quantification in the literature — for heterogeneity in firm preferences over management traits.[17]

We find that, in line with the probability of agreement between managers, the share of the variation due to idiosyncratic preferences by far outweighs that share of the variation explained by common preferences. Focusing on the agreement on the ranking as a manager, we estimate that around 20% (for the line management vignette) to 35% (for the supplier vignette) of the total variation in the latent utility is

[15] We again estimate this using Hamiltonian Markov Chain Monte Carlo using the Stan language.
[16] We refer to self-employment and wage-employment as two distinct types of employment throughout.
[17] This approach is broadly analogous to the identification strategy of Bertrand and Schoar (2003), which involves comparing R^2 values when adding manager fixed effects; this is conceptually similar to the role played by μ_i.

explained by the common component μ.

Table 9: Dyadic agreement of managers across vignettes

	Vignette					
	Line Management	**Supplier**	**Pay rise**	**Bank**	**Client**	**Total**
Ranking as manager						
Probability of agreement	56.7	60.9	58.0	57.7	59.3	58.5
Variance μ	0.26	0.78	0.41	0.48	0.60	0.58
Ranking as entrepreneur						
Probability of agreement	57.9	57.9	56.3	60.2	60.5	58.5
Variance μ	0.43	0.53	0.25	0.70	0.74	0.52

Notes This table depicts the probability of agreement and the variance of μ. The first is the probability that two managers assessing the same pair of candidates agree on their ranking as an entrepreneur and as a manager, split by the two types of employment and across vignettes. The variance of μ provides a measure of the homogeneity of preferences across managers.

5.1 Agreement by dyadic characteristics

To decompose the heterogeneity in preferences for management traits across firms, we first analyse whether the dyadic agreement between firms can be predicted by observable characteristics. To do so, we implement median splits of the sample based on a number of observable characteristics. We then create three indicator variables: $\texttt{low}_{ij}$, equal to one if both firms are below median for that characteristic, $\texttt{different}_{ij}$ equal to one if one firm is below, and one firm above median, and $\texttt{high}_{ij}$ if both firms are above the median.

We then analyse the resulting data using the following regression model for pairs of firms $i \neq j$ with standard errors clustered at the pair level:

$$\texttt{agree}_{ij} = \alpha + \beta_1 \texttt{different}_{ij} + \beta_2 \texttt{high}_{ij} + \varepsilon_{ij} \tag{2}$$

Table 10 reports the results from this regression. These results further stress that agreement between the HR managers is relatively low, and show the relationship with observable characteristics is limited. We see effectively no differences in rates of agreement by management score as calculated based on MOPS-type questions, number of employees, share of employees in a management position at the firm and share of non-payroll employees (in respectively columns 1, 2, 4 and 5). We only observe a clear increase in two characteristics: the number of competitors and managers' self-reported trust. Firms in more competitive environments appear to be more likely to agree (column 1), where the effect seems to be additive although

only significant when both firms have a high number of competitors. The most clear relationships seems to be for high self-reported trust, an indicator for whether the manager reports to be generally trusting. We find that the rate of agreement drops sharply when either manager reports to be trusting (column 6). This suggests that non-trusting managers pick up on some dimension of the responses, and trusting managers ignore this dimension without coordinating on some other observable instead. This is a sharp difference, with either manager being trusting decreasing the probability of agreement from 62.8% to around 55% in a binary choice.

Table 10: Agreement between HR managers by firm and manager characteristics

Split by:	(1) Management Score	(2) Number of employees	(3) Number of competitors	(4) Share of employees in management	(5) Share of non-payroll employees	(6) Manager self-reported trust
One high, one low	0.00885	0.0252	0.0331	-0.00837	0.000646	-0.0712
	(0.0235)	(0.0245)	(0.0230)	(0.0261)	(0.0216)	(0.0223)
Both high	0.00226	0.0392	0.0687*	0.0218	-0.0166	-0.0831**
	(0.0271)	(0.0289)	(0.0326)	(0.0287)	(0.0317)	(0.0299)
Constant	0.580***	0.563***	0.558***	0.584***	0.587***	0.628 ***
	(0.0182)	(0.0202)	(0.0190)	(0.0217)	(0.0141)	(0.0163)
N	3948	3948	3849	3860	3948	3553

Notes This table reports the parameter estimates of regression 2. The column labels indicate the variables based on which the binary split of the sample for each regression is made. The dependent variable is the share of agreement between the two HR managers assessing the same pair of candidates. The variables on which the splits are made are the overall management score, the number of employees, the number of competitors, the share of employees in management positions, the share of non-payroll employees, and the self-reported trust of the manager. The number of observations in column (6) is lower due to some observations missing for this variable. The standard errors are in parentheses and the p-values are reported as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Heterogeneity in demand for managerial types

The LDA model identifies K types of aspiring managers/entrepreneurs. We now want to estimate the heterogeneity in demand of firms over these types, for which we use a Plackett-Luce model. This model is specified as follows. For firm f, the utility of choosing each candidate i in vignette v for employment type s is:

$$U_{fiv} = \boldsymbol{\alpha}_f \boldsymbol{x}_i + \boldsymbol{\beta}_f \boldsymbol{\theta}_i + \gamma_i + \varepsilon_{fiv}, \tag{3}$$

where θ_i is the vector of type probabilities from the Latent Dirichlet model,[18] $\boldsymbol{x}_i$ a vector of observable characteristics, and γ_i acts as an individual-level fixed effect. We omit $\boldsymbol{x}_i$ in our main specification, but add this later to include the candidates' gender, treatment status in the earlier management experience experiment, and the interaction of these two variables. We provide more detail on this model in Appendix E.

Table 11 provides the parameter estimates for the Plackett-Luce model.[19] Using the timid type as the baseline, we see that in particular the authoritative type is strongly preferred by managers, and that the coercive type is significantly, but less strongly preferred over the timid type. Hiring managers appear to have similarly strong preference for the timid and affiliative types on average. Turning to the version including covariates in the second and fourth columns, we first note that there is only a small change to the parameter estimates for β, indicating that the results are not driven by, for example, men being both more authoritative and on average preferred by managers. In terms of the other covariates, columns two and four show that untreated men appear to do worse on average, with the treatment in the management experience experiment equalising the performance of male and female candidates. There is no significant difference in these regressions in the ranking for a management position compared to the ranking for self-employment.

[18] We jointly estimate the Latent Dirichlet model and the Plackett-Luce model to explicitly incorporate uncertainty about the estimates for θ. See Appendix E for details.

[19] Appendix E shows that as expected the estimated types are the same in the joint model are nearly identical to those estimates in the Dirichlet model.

Table 11: Demand for management styles

	Manager		Self-employment	
	Baseline	With controls	Baseline	With controls
$\beta_{\text{Authoritative}}$	1.17	1.23	1.27	1.30
	[0.92, 1.43]	[0.97, 1.49]	[0.98, 1.56]	[1.01, 1.58]
$\beta_{\text{Affiliative}}$	-0.08	-0.06	0.05	0.05
	[-0.44, 0.27]	[-0.43, 0.33]	[-0.37, 0.44]	[-0.36, 0.45]
β_{Coercive}	0.31	0.34	0.39	0.38
	[0.03, 0.59]	[0.04, 0.63]	[0.07, 0.69]	[0.08, 0.70]
$\alpha_{\text{treatment}}$	.	-0.16	.	-0.09
	.	[-0.38, 0.08]	.	[-0.33, 0.15]
α_{gender}	.	-0.24	.	-0.25
	.	[-0.43, -0.06]	.	[-0.44, -0.05]
$\alpha_{\text{interaction}}$	.	0.26	.	0.25
	.	[0.01, 0.51]	.	[-0.01, 0.51]

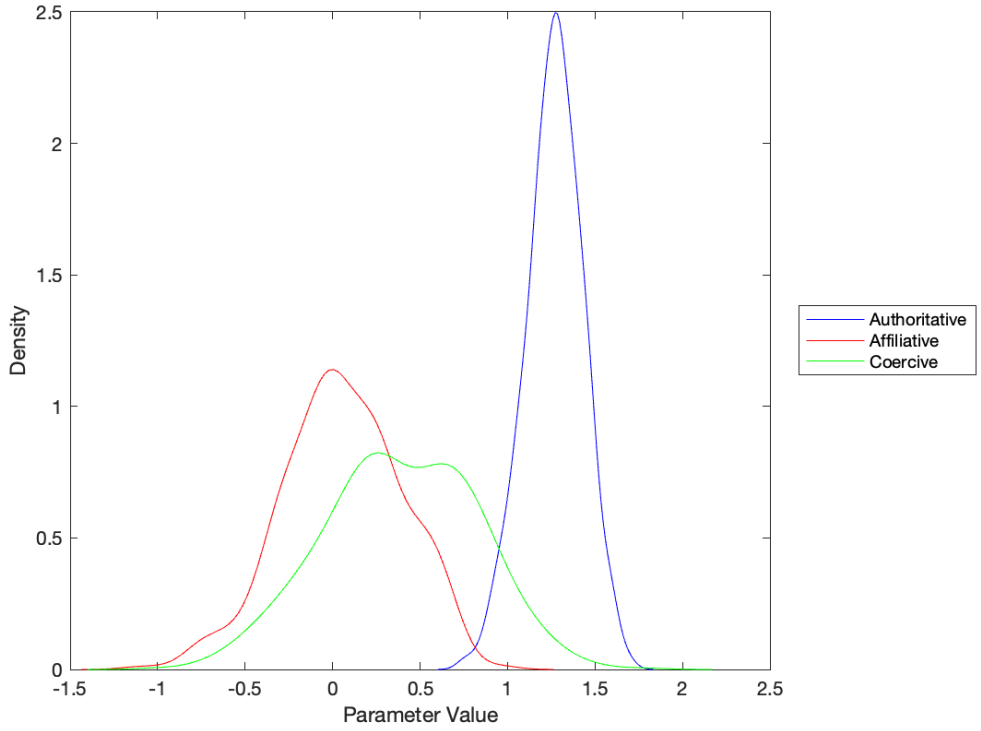
Notes This table reports the parameter estimates for the Plackett-Luce model. The reported estimates are the mean of the converged HMC chains, with the 5th and 95th percentile of these chains reported in square brackets as 90th percent credible interval.

To interpret these results, we consider what would happen if we shift the value of $\boldsymbol{\theta}$ to represent a 'pure' type manager for one member of each triplet. Specifically, for the simulation of the parameter $\beta_{\text{authoritative}}$, we adjust one individual in each triplet to have $\theta_{\text{authoritative},i} = 0.85$, setting the other elements of the vector $\boldsymbol{\theta}_i = [\theta_{\text{affiliative},i},\ \theta_{\text{coercive},i},\ \theta_{\text{timid},i}]$ to 0.05.[20] We then simulate the utilities of the HR managers using this updated $\boldsymbol{\theta}$ and calculate the new winning probability for this simulated individual. Initially, the winning probability of each individual in the pairwise comparisons is 50%. This probability increases to 62% for the authoritative type, decreases to 45% for the affiliative and timid types, and slightly decreases to just under 50% for the coercive type.

A benefit of the hierarchical model we estimate is that we estimate individual-level values for $\hat{\boldsymbol{\beta}}_{fs}$. We plot these firm-level estimates in Figure 4. This shows that in particular the strong preference for the authoritative type is uniform across firms, with virtually all firms preferring this type over the affiliative and timid type. This suggests that a specific management style is preferred by the majority of Ethiopian firms – authoritative management.

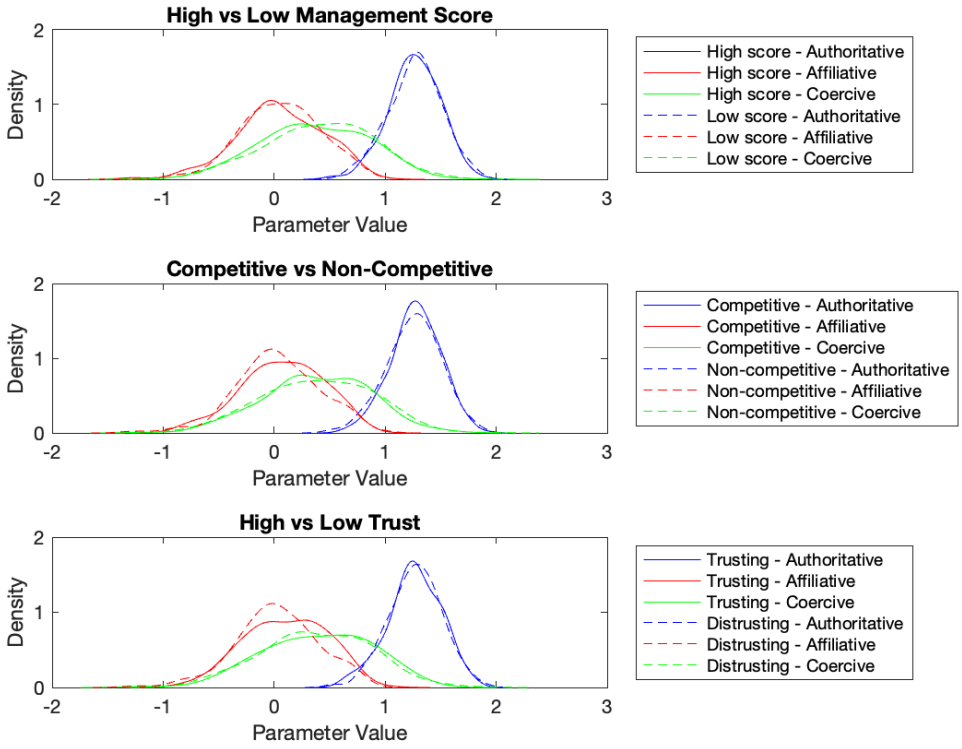
[20] This represents approximately the 95th percentile in the estimated distribution of $\hat{\boldsymbol{\theta}}$.

Figure 4: The distribution of the preferences for entry-level managers



Notes This figure shows the distribution of the estimates for $\hat{\boldsymbol{\beta}}_{fs}$ across all firms in terms of demand for entry-level managers. In this figure, the timid type is omitted and the estimates for $\hat{\beta_{\text{Authoritative}}}$, $\hat{\beta_{\text{Affiliative}}}$ and $\hat{\beta_{\text{Coervice}}}$ are plotted

Figure 5: The distribution of β_f conditional on covariates



Notes This figure shows the distribution of the estimates for $\hat{\boldsymbol{\beta}}_{fs}$ for a number of subsamples of the data. The top pane splits the sample by firms' management scores based on MOPS questions, the second pane splits the sample by the median number of competitors, and the third pane splits the sample by whether or not HR managers' identify themselves as being generally trusting

In Figure 5, we examine the distribution of the parameter estimates for β_f for subsets of the firms. We split the sample by management scores, the number of competitors a firm identifies, and whether or not the manager is trusting. As perhaps expected based on Table 10, there is no clear relationship between management scores and the distribution of β_f. Similarly, the link between competitiveness and preferences for traits is limited. The clearest, albeit still weak, link arises for the split by how trusting the manager is, with less trusting managers having a stronger preference not to choose the affiliative type.

Next, we decompose the variance of the parameter estimates of the model specified in Equation 3. We then calculate the variance of (a) the total non-random component of the utility function, (b) of just the variance of the $\boldsymbol{\beta}_f\boldsymbol{\theta}_i$ term and (c) the variance of γ_i.

For the rankings as a manager, we find that:

$$\widehat{\text{var}(\boldsymbol{\beta}_f\boldsymbol{\theta}_i + \gamma_i)} = 0.63$$
$$\widehat{\text{var}(\boldsymbol{\beta}_f\boldsymbol{\theta}_i)} = 0.34$$
$$\widehat{\text{var}(\gamma_i)} = 0.29$$

This shows that there is indeed a lot of idiosyncratic noise, but that over half of the non-idiosyncratic variation in preferences can be explained by the preferences over these types.

6 Discussion and conclusion

This study has developed a novel experimental approach to assessing managerial traits, using a controlled studio setting to observe and analyse how young Ethiopian professionals respond to realistic workplace challenges. By combining experimental data with Bayesian hierarchical modeling, we identified four distinct managerial types: authoritative, affiliative, coercive, and timid. These archetypes reflect meaningful variation in how young professionals approach decision-making and conflict resolution in managerial roles.

Our findings highlight several important insights. First, we show that managerial traits are not uniformly distributed but exhibit clear heterogeneity. The authoritative type, for instance, emerges as the most dominant managerial style, particularly valued by firms for entry-level managerial roles. This type tends to emphasize formal procedures, personal authority, and firm-centered interests, contrasting sharply with the affiliative type, which seeks shared ground and often concedes to counterparties even at the expense of the firm's interests. These managerial types capture not only differences in style but also broader distinctions in labor market experience and outcomes.

Notably, labor market exposure appears to play a role in shaping managerial traits. Young professionals with more substantial exposure to the labor market were more likely to exhibit authoritative traits. Our evidence based on random assignment to a previous management placement experiment suggests this is at least partially a causal link. This finding suggests that managerial styles are not purely innate, but can be shaped and developed through professional experience. These results align with previous studies showing the importance of early exposure to managerial environments in shaping long-term behavioural traits (Bertrand and Schoar, 2003; Malmendier et al., 2011).

Further, our analysis reveals that firms in Ethiopia, across a range of industries, tend to prefer authoritative-type managers for entry-level positions. This homogeneous preference for authoritative traits, as opposed to more affiliative or coercive traits, challenges the notion that firms have highly di-

verse preferences for managerial styles at the entry level. While some previous studies (Bandiera et al., 2020) have emphasized the heterogeneity in firm preferences for senior managers, our results suggest that firms may place greater value on uniformity in managerial traits at the lower levels of management. This homogeneity could reflect the need for clear, consistent, and rule-based decision-making among less-experienced managers, particularly in firms seeking to establish stable internal processes.

Methodologically, this study makes an important contribution to the literature by demonstrating the feasibility and utility of using studio-based vignettes to assess managerial traits. Unlike traditional observational studies, our controlled setting allows us to systematically compare managerial responses across a fixed set of scenarios, removing the endogeneity inherent in real-world managerial decision-making. This approach opens up new avenues for studying managerial behavior, particularly in contexts where direct observation of managerial actions may be impractical or biased by the organizational setting.

In addition to its methodological contributions, our study complements the growing literature on labor market exclusion faced by young professionals. The finding that authoritative-type managers are both in higher demand by firms and more likely to have greater labor market exposure suggests that managerial traits may serve as a key mechanism by which some young professionals achieve sustained labor market success, while others face exclusion. This reinforces earlier findings on the importance of management training and professional exposure in shaping labor market outcomes (Abebe et al., 2024).

Further steps in our analysis include studying the 'cognition versus implementation' gap amongst these managers – knowing what to do versus being able to do it (Gibbons and Roberts, 2013). We intend to do so based on data on what respondents intend to say versus what they actually say in their responses. Moreover, we have random variation in the gender of the actors that respondents and HR manager see which we intend to use to explore difference in both how managers act and in how these actions are perceived when responding to male versus female actors.

In conclusion, this paper offers important insights into the determinants of managerial behavior and firm preferences in developing country settings. It highlights the role of labor market experience in shaping managerial traits and provides evidence of the strong preference for authoritative managers at the entry level. Future research could expand on this work by exploring how these preferences evolve as managers move into higher-level roles and by investigating the role of cultural and institutional factors in shaping managerial behavior. Additionally, the scalability of our experimental method presents opportunities for further research across different contexts and labor markets, providing a valuable tool for understanding the development of managerial talent globally.

References

Abebe, G., M. Fafchamps, M. Koelle, and S. Quinn (2024). Matching, management and employment outcomes: A field experiment with firm internships. *Working paper*.

Bandiera, O., L. Guiso, A. Prat, and R. Sadun (2015). Matching firms, managers, and incentives. *Journal of Labor Economics 33*(3), 623–681.

Bandiera, O., A. Prat, S. Hansen, and R. Sadun (2020). Ceo behavior and firm performance. *Journal of Political Economy 128*(4), 1325–1369.

Benmelech, E. and C. Frydman (2015). Military ceos. *Journal of Financial Economics 117*(1), 43–59.

Bertrand, M. and A. Schoar (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics 118*(4), 1169–1208.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Bloom, N., R. Lemos, R. Sadun, D. Scur, and J. V. Reenen (2014). Jeea-fbbva lecture 2013: The new empirical economics of management. *Journal of the European Economic Association 12*(4), 835–876.

Bloom, N., R. Lemos, R. Sadun, and J. Van Reenen (2015). Does management matter in schools? *The Economic Journal 125*(584), 647–674.

Bloom, N., C. Propper, S. Seiler, and J. Van Reenen (2015). The impact of competition on management quality: evidence from public hospitals. *The Review of Economic Studies 82*(2), 457–489.

Bloom, N. and J. Van Reenen (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Borgschulte, M., M. Guenzel, C. Liu, and U. Malmendier (2021, March). Ceo stress, aging, and death. Working Paper 28550, National Bureau of Economic Research.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2016). Health care exceptionalism? performance and allocation in the us health care sector. *American Economic Review 106*(8), 2110–2144.

Ethiopian Statistical Service and World Bank (2023, September). Ethiopia socioeconomic panel survey (esps) report - wave 5, 2021/22. Published in September 2023.

Gibbons, R. and R. Henderson (2012). Relational contracts and organizational capabilities. *Organization science 23*(5), 1350–1364.

Gibbons, R. and J. Roberts (2013). *The handbook of organizational economics*. Princeton University Press.

Goleman, D. (2000). Leadership that gets results. *Harvard Business Review 78*(2), 78–89.

Gorodnichenko, Y., T. Pham, and O. Talavera (2023). The voice of monetary policy. *American Economic Review 113*(2), 548–584.

Griffiths, T. L. (2004). Finding scientific topics. *PNAS*.

Haaland, I., C. Roth, S. Stantcheva, and J. Wohlfart (2024). Measuring what is top of mind. *Working paper*.

Handlan, A. and H. Sheng (2023). Gender and tone in recorded economics presentations: Audio analysis with machine learning. *Working paper*.

Kaplan, S. N., M. M. Klebanov, and M. Sorensen (2012). Which CEO characteristics and abilities matter? *The Journal of Finance 67*(3), 973–1007.

Li, S. (2017). Obviously strategy-proof mechanisms. *The American Economic Review 107*(11), 3257–3287.

Macchiavello, R., A. Menzel, A. Rabbani, and C. Woodruff (2020, July). Challenges of change: An experiment promoting women to managerial roles in the bangladeshi garment sector. Working Paper 27606, National Bureau of Economic Research.

Malmendier, U., G. Tate, and J. Yan (2011). Overconfidence and early-life experiences: The effect of managerial traits on corporate financial policies. *The Journal of Finance 66*(5), 1687–1733.

McCormack, J., C. Propper, and S. Smith (2014). Herding cats? management and university performance. *The Economic Journal 124*(578), F534–F564.

McKenzie, D. and C. Woodruff (2017). Business practices in small firms in developing countries. *Management Science 63*(9), 2967–2981.

Rasul, I. and D. Rogger (2018). Management of bureaucrats and public service delivery: Evidence from the nigerian civil service. *The Economic Journal 128*(608), 413–446.

Rivkin, J. W. (2000). Imitation of complex strategies. *Management science 46*(6), 824–844.

Scur, D., R. Sadun, J. Van Reenen, R. Lemos, and N. Bloom (2021). The world management survey at 18: lessons and the way forward. *Oxford Review of Economic Policy 37*(2), 231–258.

Stantcheva, S. (2021). Understanding tax policy: How do people reason? *The Quarterly Journal of Economics 136*(4), 2309–2369.

Weber, M. (1904). Die" objektivität" sozialwissenschaftlicher und sozialpolitischer erkenntnis. *Archiv für sozialwissenschaft und sozialpolitik 19*(1), 22–87.

Lights, Camera, Transaction!
Assessing management styles through studio vignettes

Appendix for online publication

February 14, 2025

Contents

A Vignette scripts **3**
A.1 Respondent management vignettes . . . 3
A.2 Assessment of management styles: Hypothetical workplace conflict . . . 5

B Experiment **8**
B.1 Structure of the assessments . . . 8
B.2 Matching firms to candidates . . . 9
B.3 A revealed preference exercise with human resource managers . . . 9
B.4 A stated preference exercise with human resource managers . . . 10

C Details: Encoding of Responses **10**

D Estimated type characteristics for different numbers of types **11**
D.1 Further details on reported intentions . . . 17

E Bayesian modelling in STAN **17**
E.1 Dirichlet model . . . 17
E.2 Placket Luce model . . . 20
E.3 Estimation . . . 21
E.3.1 The Dirichlet model . . . 22
E.3.2 The Placket-Luce model and joint-estimation . . . 22
E.3.3 Stability of type parameters with joint estimation . . . 23

A Vignette scripts

A.1 Respondent management vignettes

Vignette 1: An absent employee

NARRATOR: *In this video, an accounting clerk who has been absent for three days without advance notice to her immediate boss is shown. (S)he shows up for work on the fourth day, explaining that (s)he was unwell. Do watch the video carefully.*

THE CLERK: I realize that I did not come to work for the last three days. I could not also give you any advance notice as this was something beyond my control. I was not feeling well. I had a bad flu that made me see a doctor. The doctor recommended that I should take it easy and perhaps take a few days of rest at home. I thought that, with my conditions, coming to the office and working on our accounts would be difficult [actor coughs slightly]. Also, I did not want others to catch my flu as well, so I stayed away.

Boss, you remember Almaz, my colleague, who had a bad flu three weeks ago. Instead of staying at home, she chose to come to work while being very sick. The next day, almost half of the staff in our department had a terrible headache and were coughing. They caught whatever she had. I certainly did not want that to happen [actor shakes head in the negative].

Vignette 2: Negotiating with a supplier

NARRATOR: *Here we have a supplier, who explains to you that (s)he cannot fulfil an order according to specifications, because of problems sourcing input materials. (S)he offers to supply a replacement of inferior quality instead. Note that quality is the most important feature of the firm's product and that the firm is known in the market for producing and selling the highest quality products in the industry. Do watch the video carefully.*

THE SUPPLIER: As you know, the main roads that we use to transport our shipments from Djibouti have been blockaded in the past several weeks. It has been three weeks since we got our last shipment of the materials that we promised to supply to you. During the course of the past three weeks, our inventory of these materials has been depleted. I am thus sorry to inform you that we cannot supply you with the materials as specified in the original supply agreement.

However, the good news is that we have a pretty good stock of alternative materials that we could supply you at a cheaper rate. We cannot, however, guarantee that these materials currently in stock meet the quality and specification requirements that were stipulated in the original supply agreement. Nevertheless, I think what we have for you is a good deal that is offered with the cheapest possible price. Of course, once our sourcing problem is resolved, we will continue to supply you with materials that comply with the specifications that you set.

Vignette 3: An employee seeks a pay rise

NARRATOR: *In this video, we will show you a production worker, coming to you and asking for a pay rise. Your company does not have a plan to increase any of your workers' salary this year. Do watch the video carefully.*

THE PRODUCTION WORKER: Boss, you know how dedicated I am to my work. In the past three months, our department has done very well and we have met the quarterly targets set by the management. I'm very happy about that. As you know, I have been the most productive worker in the department. I am one of the few people in the department that exceeded the personal productivity targets that you set for us at the beginning of the year. Of course, I also did not have even a single day of absence recorded in the completed quarter. If you don't believe me, you can check the attendance repository and the productivity (or KPI) sheet. As you know, I also worked for the company for the past ten years, not so many people can say that! I am not sure whether you know my mother, Feleku Kuma; she is widely known around here, she also worked here for a very long time.

I think these speak volumes to my dedication and loyalty to this company. But honestly, I don't think my current pay reflects the effort that I exert and the loyalty that I and my family have for the company. In fact, I spoke to people who work with our competitors at the same type of occupation as I am; surprisingly, their pay is much larger than mine. I am sure you agree with me that I deserve a pay rise. Can you let me know as to when and how much I should expect to get as a pay rise?

Vignette 4: Negotiating an adjustment with the bank

NARRATOR: *This video shows a bank manager, who calls to remind you about an unpaid loan instalment. But you will not have sufficient funds to pay the bank for another two weeks. Take a watch.*

THE BANK MANAGER: Thanks a lot for taking my call. As you know, you are one of our clients who missed the latest installment on your outstanding loan from our bank. Our bank treats its customers with utmost respect and encourages timely payment to avoid breaches to loan terms and conditions. But also, we have quite a lot of pressure from the board of directors to ensure that scheduled loan payments are made on time. Failure to pay loan installments not only increases the interest payment you owe us; it also reduces you credit rating with us. Next time you ask for loan, you may be denied access altogether – or may be eligible to only a fraction of the amount that you have applied for. I thus cannot stress enough the importance of settling your outstanding installment as soon as you can. That said, I am wondering whether you will be able to pay the unpaid loan instalment very soon. I will follow-up this call with a written reminder to alert you to the periodic payment that we expect to be paid from your company.

Vignette 5: Negotiating with a client

NARRATOR: *The next video shows a client of your business, who has not paid what you have invoiced, in spite of one reminder. The client comes to place a new order.*

THE CLIENT: I really like the service I got from your company last time. I don't think I had the chance to thank you for that – so thank you. Oh, I did not forget about the payment and I confirm that I have received the invoice you issued. But you know how it is these days, business is not great, and our product is not exactly flying off the shelves. We are really short on cash at the moment. I will thus have to ask you to bear with me a little to pay for the previous order. With the risk of stretching your generosity, I'm here today to place another order. In a spirit of our long-standing relationships, I appreciate if you could deliver my new order soon without waiting for payment for the previous order. I promise we will meet our payment obligations for the previous and new orders once we sort out our cash flow problem.

A.2 Assessment of management styles: Hypothetical workplace conflict

Setting up the scenario

NARRATOR: You are faced with a situation where workers go on strike demanding proper latrines. A representative of the workers comes to talk to you. Take a look at the video carefully.

THE WORKER REPRESENTATIVE: In my capacity as a representative of workers in this factory, I would like to speak to you about the ongoing strike in the factory. I am sure that by now you realize what workers are striking about. But with the slow response from the line supervisors and the deafening silence from you, I am not certain that the management in the factory understood the severity of the problem. That is why I'm here.

Look, the strike is about the poor toilet facilities in the factory. As you recall, during the worker/management meeting, many workers complained about the lack of sufficient toilet facilities; there are very few toilets in the compound and workers have to walk to the other end of the compound to access them. Even when they reach the toilets, they have to wait their turn, sometime for several minutes, to use them. In recent months, the problem has gotten worse. The few toilets that are available either do not have water or the flushing system does not work and tissue papers are not available. To make things worse, sometimes the toilets are locked and workers are not told why they are locked and who is keeping the keys. Further, even when workers try to use the toilets, they face restrictions on the number of toilet breaks they are entitled to take.

The attitude of your supervisors towards toilet use is also appalling. Our female workers suffer a great deal particularly when they are experiencing their period. Studies indicate that failing to use toilets on the course of a day lead to several health problems including damage to the bladder and several types of infections that have serious implications on the workers'

wellbeing. We as workers are very much aware of the implications of poor toilet facilities. Please do not underestimate our intelligence, we know about all this and our rights. Through our discussion with fellow workers, we came to the conclusion that our voice needs to be heard. We are particularly troubled by the dominant view of the management that workers shirk using toilet breaks as an excuse. Such views are not only unfounded, we reject the thought process behind it as it deprives the dignity of workers as human beings. Unless our demand for improved toilet facilities is not met, our strike will continue.

Six possible responses to the scenario

We then use different actors to record six separate possible managerial responses. The responses are designed to represent different styles of engagement. In this summary, we suggest several descriptors for each managerial response; these descriptors were not provided to the respondents.

NARRATOR: After listening to the complaint by the worker representative, six different managers respond in different ways. Please do watch the videos.[1]

MANAGER 1 [*apologetic; collaborative; empowering*]: We understand workers' complaints about the poor state of toilet facilities in our factory. We admit that we have been slow to act and hence invited the strikes onto ourselves. On behalf of the company, I would like to apologize for our failure to nip this problem in its bud before it reached to this level. We will take immediate actions to rectify the situation. For a start, we will constitute a committee that will include workers' and management representatives. The committee will work on identifying the key problems in the existing toilet facilities and would suggest solutions for improving access and quality of toilet facilities. In parallel, the management will meet and earmark the necessary resources required to improve the facilities. We actually want the whole process to be led by workers and their representatives. We believe that there is no substitute for better working conditions to get the best out of our workers. With this pledge, we would like you to convince the workers to go back to work.

MANAGER 2 [*acknowledging; proactive; realistic*]: Honestly, this is a fair request. We should have seen this coming. As we promised during the worker/management meeting, the management takes all workers' requests very seriously. We are now making preparations to invest in improving the toilet facilities. Some of the improvements related with maintaining the flushing system of the existing facilities can be readily made. But increasing the number of toilets and digging septic tanks could take longer time. As you know, obtaining construction permit in this part of the city takes quite a long time. So even if we want, we may not be able to overhaul the whole facility at once. We thus ask for patience on the workers end. With this in mind and with the promise to make feasible changes immediately, we request you to stop the workers'

[1] For clarity, these are ordered based on the senior managers' ranking which was used to incentivise the respondents. Respondents' instead saw these in a random order.

strike at once.

MANAGER 3 [*appreciative; conciliatory; blame-shifting; behavioural focus*]: Thanks for bringing this to our attention before things go out of hand. We really appreciate your effort to be a voice of workers. The way your represent the concerns of your workers has been exemplary – and, on behalf of the company, I would like to express my gratitude. Of course, even if you are the representative of workers, we always felt that you are one of us, the management team. As both a worker representative and as someone who works closely with us, you understand how it has been difficult to manage workers, particularly their use of toilet facilities. We agree that facilities should improve, and that can be done.

But this issue is not that important. What worries us the most is workers have been abusing their toilet breaks by sitting idle and chattering with fellow workers during the breaks. Some of them spend more than 50 minutes inside the toilet, locked inside with their phone, checking Facebook or playing games. We told them repeatedly that this is not an acceptable behavior and it ought to change. When they refused, we started locking some of the toilet rooms. I think you should help us to get their act together. We can work together to change their behavior use their rightful toilet breaks only for intended purposes. Explain to them that the strike should come to an end and they should stop abusing their breaks. We can then discuss about ways to improve the quality of the facilities.

MANAGER 4 [*acknowledging; deflective; postponing; subtle pressure*]: We understand that toilet facilities are important and workers have the right to complain about problems in accessing these facilities in our compound. But everyone knows that we are struggling at the moment: business is not good and we are making losses. Asking the management to invest in improving the toilet facilities at the moment is not a good idea. This simply is not the right time to do that.

As a representative of workers, you should be aware that management does not like to be ambushed or cornered into making decisions in difficult times. I am sure you also do not want to take the blame for creating trouble in the company while conditions are already hard on us. I know you are also concerned about your career prospects. This is really a tough time for everyone, so let us weather the storm together and we can resume the discussion about improving the toilet facilities afterwards.

MANAGER 5 [*dismissive; threatening; defensive*]: As you said, we have discussed this during the worker-management meeting. During the meeting, we were very clear. We can only meet demands that are convincing and consistent with the company's policy. We told the workers that we know what kind of toilet facilities other employers provide for their workers.

Did you have a chance to visit other company's toilet facilities? If you or your colleagues had done that, you would have realized that ours is not of sub-standard quality compared to them. If anything at all, we provide flushing toilets and fresh water, which many don't. We also know the kind of toilet facilities workers have at their place of residence, mostly pit latrines

that are unsafe and often dirty. We certainly provide better facilities than those at home. So if you don't like our toilets here, you and your comrades can go and work elsewhere. And believe me, this is what is going to happen if you don't cease striking in the next 24 hours.

MANAGER 6 [*legalistic; hostile; threatening; dismissive*]: Who do you think you are, barging into my office and putting unwarranted demands on this company? I am not going to discuss about toilets or about any related issues with you. No, not at all. Forget toilets, let us talk about the strikes. Do you know what the labor law says about strikes? I am sure you don't! Had you known, you wouldn't be striking! Let me enlighten you. You see this [actor shows a copy of a labour proclamation]: this is Labour Proclamation No. 377/2003.

According to Article 158 (1) of the proclamation, before any strike happens, workers must give advance notice to the company explaining in detail the reasons for the strike. As far as I know, no such notices have been submitted to my office. Further, the law says that, for the strikes to be carried out, it should be supported by the majority of workers with a formal meeting. To my knowledge, no such formal meeting has happened. These two violations by you and your comrades indicate that the current strike is clearly in the contravention of the labor law. If the strikes does not stop before the end of the day, you will hear from our lawyers. We will sue you personally for initiating and implementing the strikes in defiance of the country's law.

B Experiment

B.1 Structure of the assessments

Groups of two firms assess the same set of 15 candidates:

- Each group of 15 interns is assessed by 10 firms.
- Each pair of firms assesses 15 candidates together, 3 for each vignette.
- They rank the three candidates as entrepreneurs and managers.
- We have 66 of these groups.

	Firm Group (groups of 2)				
Candidate Group (3 candidates)	1	2	3	4	5
1	V1	V2	V3	V4	V5
2	V5	V1	V2	V3	V4
3	V4	V5	V1	V2	V3
4	V3	V4	V5	V1	V2
5	V2	V3	V4	V5	V1

B.2 Matching firms to candidates

We match firms to candidates using a random sequential dictator algorithm. The matching procedure is structured as follows:

(i). We divide firms into groups of 10 and, separately, candidates into groups of 15.

(ii). Each group of candidates is assigned a random preference ordering over all groups of firms.

(iii). Groups of candidates are matched sequentially based on their preference rankings:

- Each group of candidates is assigned to their most preferred group of firms that has not yet been matched.

(iv). We impose an additional constraint: firms cannot assess candidates who have previously interned at their firm, as part of the experiment conducted in (Abebe, Fafchamps, Koelle, and Quinn, 2024).

Within each group-level match, triplets of candidates are matched to pairs of firms for the first round of assessments using the same algorithm. After this, triplets of candidates rotate across pairs of firms using the mechanism described in Section B.1.

B.3 A revealed preference exercise with human resource managers

We designed these mechanisms to achieve three objectives: (i) the mechanisms are intuitively simple for respondents to understand, and truthful reporting should be an 'obviously dominant strategy' (within the formal definition of Li (2017)); (ii) as far as possible, the two mechanisms are the same (so that any differences in reporting are attributable to differences in preferences, rather than to the mechanisms themselves), and (iii) we do not mislead the respondents in any way (therefore, any promises that we make in order to incentivise the decision are feasible for us to implement). Our mechanisms are as follows:

(i). Suitability as an employed manager: We tell the respondent:

> *I'm about to ask the computer to randomly choose a number: 1 or 2. This is the number of candidates whose details I will be asked to pass to you. I need you to commit, in advance, to which candidates you would like to see in each circumstance. You have seen three candidates. Suppose that the computer tells me that you may receive the contact details of two. In that situation, of the three candidates, whose details would you then not want to see? Now, suppose that the computer tells me that you may receive the contact details of only one. Of the two candidates remaining, whose details would you then not want to see?"*

This process continues until just one candidate remains.[2]

[2] In a follow-up survey with the candidates, twelve report having been contacted by firms, none of these were eventually hired.

(ii). Suitability as an entrepreneur: We explain to the respondent that we plan to run a series of business plan competitions – in which a set of candidates is judged by experienced business managers, with a winner receiving US$1000 to start (or to support) his or her own business. To elicit a ranking over candidates' entrepreneurial abilities, we nest the previous mechanism in a simple 'random dictator' mechanism. We tell the respondent:

> *You are one of two business people who will review this set of three aspiring managers. We will randomly choose just one of the two of you to decide which of the three should be candidates at the business plan competition. Suppose that person is you. The computer will, again, randomly choose a number: 1 or 2. This time, this number will tell us how many candidates will be invited to the business plan competition. You have seen five candidates. Suppose that you are chosen to send candidates to the business plan competition, and that the computer decides that you should send two candidates. In that situation, of the three candidates, who should not be sent to the competition?*

As above, this process continues until just one candidate remains.

These mechanisms are each similar to the 'OSP-RSD' ranking mechanism described in **?**; they are simple to understand, and it is an 'obviously dominant' strategy for the respondent to rank truthfully.

B.4 A stated preference exercise with human resource managers

After each recording, we ask the manager the following questions (each of which uses a five-point Likert scale):

(i). Suitability as an employed manager: *Think about the lowest-level managerial position in your firm. Imagine that a firm – either your firm or another firm like yours – is interested in hiring for this position. Based on the recording that you have just heard, how likely are you to recommend hiring this person for a managerial position like that?*

(ii). Suitability as an entrepreneur: *Suppose that the person you have just seen is acting in the capacity as a founder and managing director of his/her own firm. In your opinion, how likely is it that this person would be a successful entrepreneur?*

These answers record managers' stated preferences about each candidate separately.

C Details: Encoding of Responses

* This is almost never used.

Dimension	Enumerator encoding	Encoding for analysis
Action	1 Agrees	1 Agrees
	2 Disagrees	2 Disagrees
Justification	Varies by vignette; see following table	1 Firm's Interest
		2 Other person's interest
		3 Respondent's own interest
		4 Shared Interest
		5 No justification
Authority	1 Formal Authority	1 Formal Authority
	2 Higher Principles	2 Higher Principles
	3 Personal Authority	3 Personal Authority
	4 Personal Relationship*	4 No Source
	5 No Source of Authority	
Tone		1 Aggressive
		2 Assertive
		3 Calm/Assured
		4 Timid

Table A.1: Encoding of Responses

* We manually determined "Other" implies no justification.

D Estimated type characteristics for different numbers of types

Figures A.1, A.2, A.3 and A.4 show the characteristics of the estimated types for respectively the two, three, four and five-type model.

Encoding Enumerator	Encoding for analysis
Line Management of an Employee	
The action is to protect/help the respondent's company	1. Firm's Interest
The action is in the interests of the other employees.]	
The action is to protect/help the employee	2. Other party's interest
The action is in the respondent's own interest	3. Respondent's own interest
The respondent did not justify his or her action	5. No justification
Negotiating with a Supplier	
The action is to protect/help the respondent's company	1. Firm's Interest
The action is to protect/help the supplier	2. Other party's interest
The action is in the respondent's own interest	3. Respondent's own interest
Reference to the relationship between the respondent's firm and the supplier	4. Shared interest
The respondent did not justify his or her action	5. No justification
Other*	
Negotiating a Pay Rise	
The action is to protect/help the respondent's company	1. Firm's Interest
The action is to protect/help the employee	2. Other party's interest
The action is in the respondent's own interest	3. Respondent's own interest
Reference to the relation with the employee and his/her family	4. Shared interest
Reference to the employee's good performance	
Reference to the effect it might have on other employees	1. Firm's Interest
The respondent did not justify his or her action	5. No justification
Other*	5. No justification
Negotiate an Adjustment with the Banks	
The action is to protect/help the respondent's company	1. Firm's Interest
The action is to protect/help the bank	2. Other party's interest
The action is in the respondent's own interest	3. Respondent's own interest
Reference to the relationship between the firm and the bank	4. Shared interest
Reference to the inability of the company to repay the loan	1. Firm's Interest
The respondent did not justify his or her action	5. No Justification
Other*	5. No justification
Negotiate with a Client	
The action is to protect/help the respondent's company	1. Firm's Interest
The action is to protect/help the client	2. Other party's interest
The action is in the respondent's own interest	3. Respondent's own interest
Reference to the relationship between the respondent's firm and the client	4. Shared interest
The respondent did not justify his or her action	5. No justification
Other*	5. No justification

Table A.2: Encoding Justification for Various Scenarios

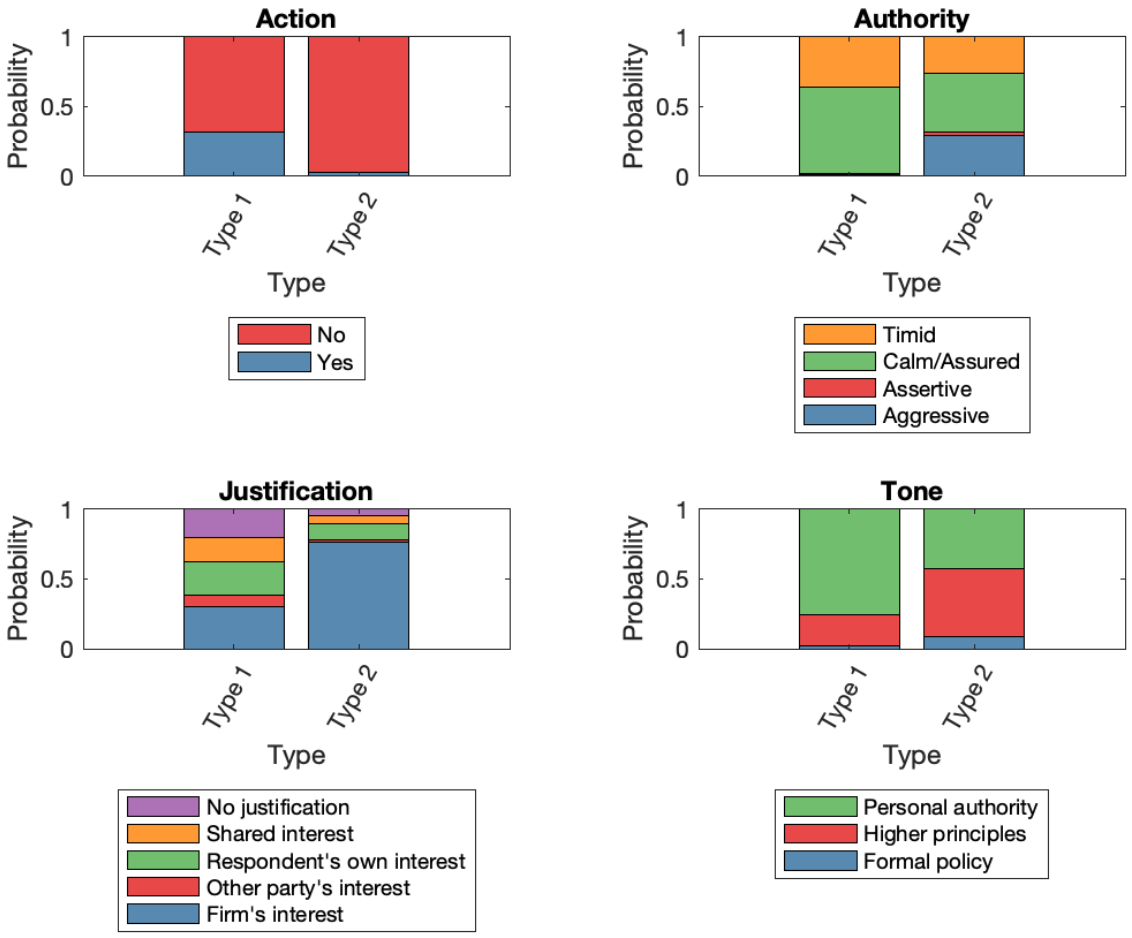


Figure A.1: Type estimates in the two-type model

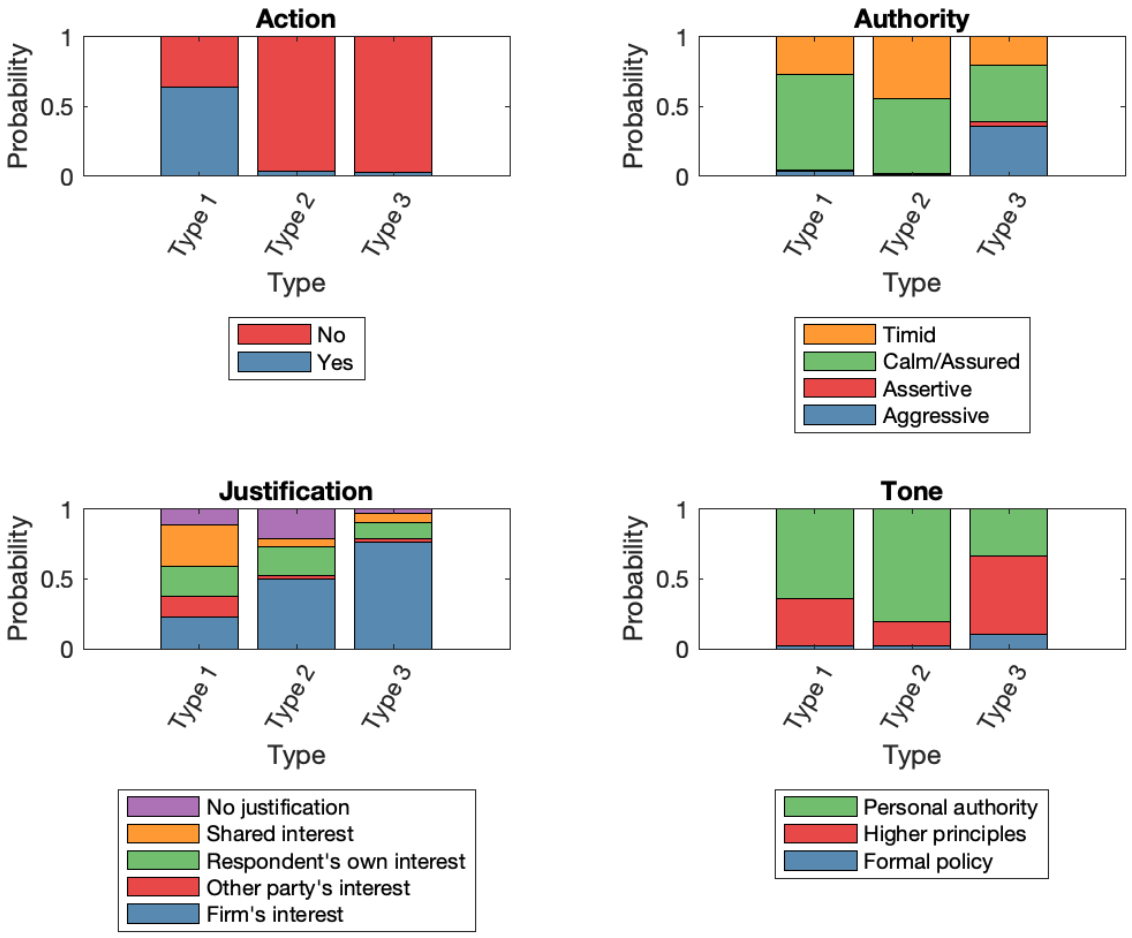


Figure A.2: Type estimates in the three-type model

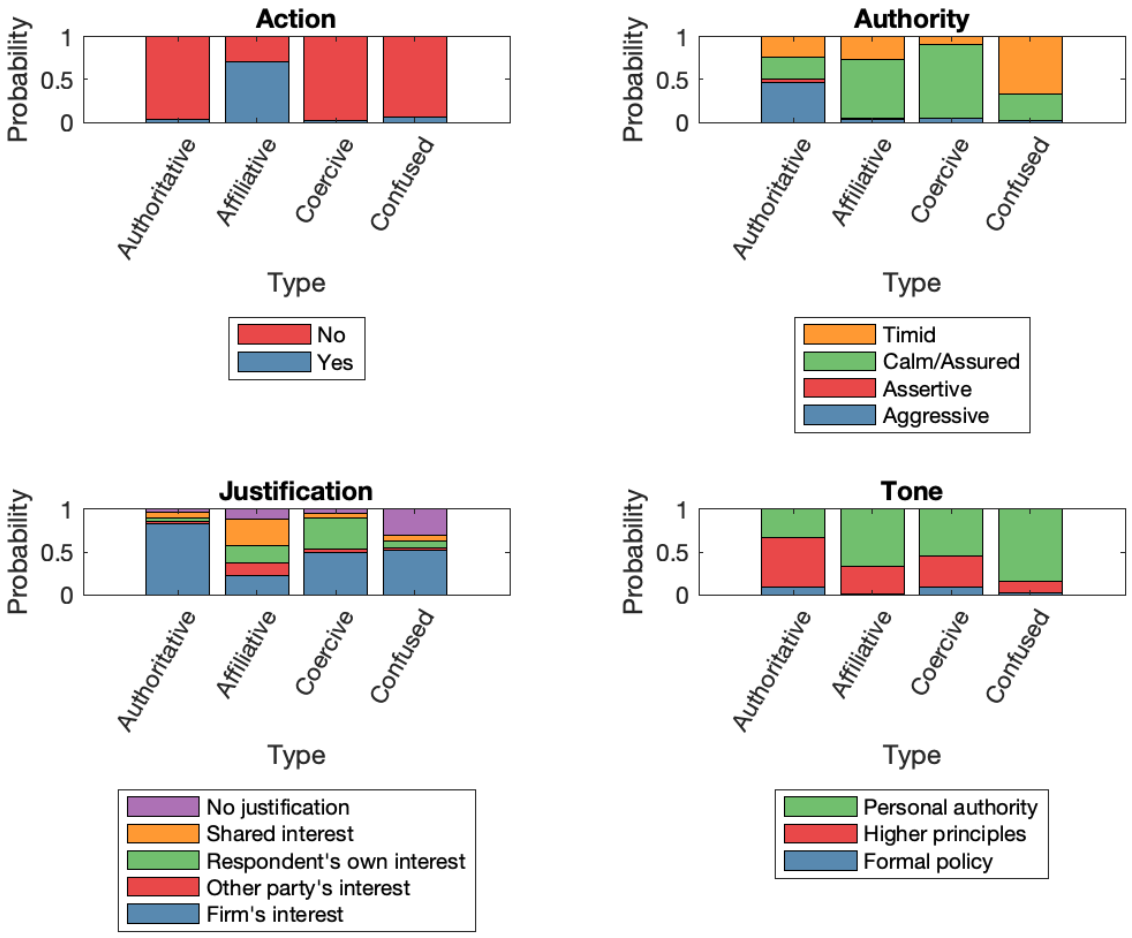


Figure A.3: Type estimates in the four-type model

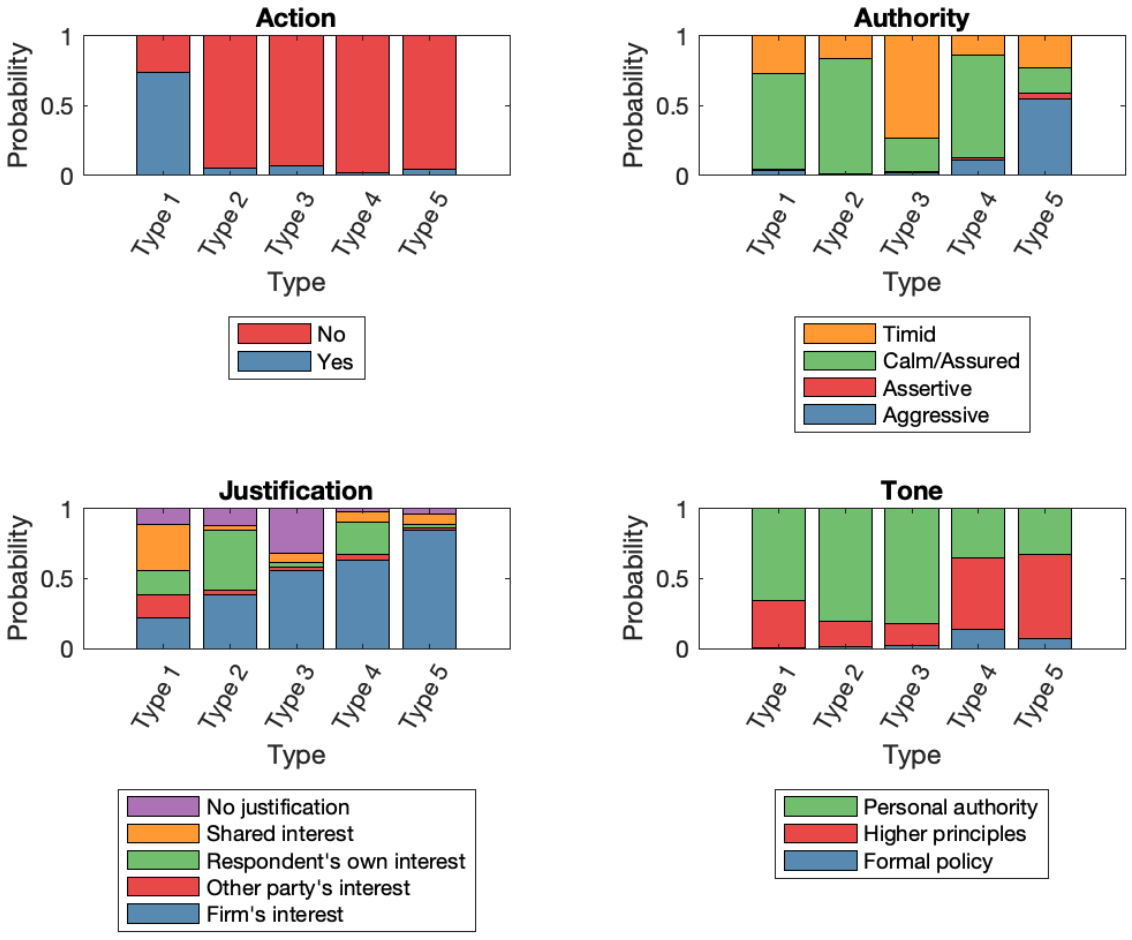


Figure A.4: Type estimates in the five-type model

D.1 Further details on reported intentions

E Bayesian modelling in STAN

E.1 Dirichlet model

We observe individuals $i \in \{1, \ldots, N\}$ performing on vignettes $v \in \{1, \ldots, 5\}$. For each individual assessment of a vignette, we have a set of two enumerators, $e \in \{1, 2\}$. Each enumerator records a set of 'attributes' of the response (action, authority, justification and tone): $a \in \{1, \ldots, 4\}$. Each attribute a has $J(a)$ possible categorical responses.[3] $y^a_{ive} \in \{1, \ldots, J(a)\}$ is the response recorded by enumerator e for vignette v for individual i for attribute a.
For attribute a, consider a 'pure type' $k \in \{1, \ldots, K\}$, having a vector $\boldsymbol{\phi}_{ka}$ of dimension $J(a)$, such that $\boldsymbol{\phi}_{ka} \in (0, 1)^{J(a)}$. Denote by $\tilde{\boldsymbol{\phi}}_{ka}$ the inverse Multinomial Logit transformation of $\boldsymbol{\phi}_{ka}$.[4] For simplicity and tractability, we assume that, conditional on $\boldsymbol{\phi}_{ka}$, the attributes are realised independently as:

$$y^a_{ive} \,|\, \boldsymbol{\phi}_{ka}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ev} \sim_{iid} \texttt{Multinomial Logit}\left(\tilde{\boldsymbol{\phi}}_{ka} + \boldsymbol{\psi}_{av} + \boldsymbol{\chi}_{ae}\right), \tag{1}$$

where $\boldsymbol{\psi}_{av}$ and $\boldsymbol{\chi}_{ae}$ here represent $J(a)$-element vectors.[5][6]
This is essentially a Latent Dirichlet Allocation model, with a few adjustments for our setting:

(i). $\boldsymbol{\psi}_{av}$ and $\boldsymbol{\chi}_{ae}$ provide a measurement scheme; this is similar in spirit to Item Response Theory. (That is, this captures the notion that different enumerators will have different underlying tendencies to assess different attributes on different vignettes.) Note that, in the special case that $\boldsymbol{\psi}_{av} = \mathbf{0}$ and $\boldsymbol{\chi}_{ae} = \mathbf{0}$, this collapses to a more standard Latent Dirichlet Allocation.

(ii). By having independent 'attributes', we reduce the state space.

(iii). Each attribute has a different "dictionary"

(iv). For each vector $\boldsymbol{\psi}_{ka}$ and $\boldsymbol{\chi}_{ea}$ the first element is normalised to zero.

We can then posit the following generative model:

(i). Draw $\boldsymbol{\phi}_{ka}$ independently for $k \in \{1, \ldots, K\}$ and for $a \in \{1, \ldots, A\}$ from $\texttt{Dirichlet}(\zeta)$;

[3] Specifically, $J(1) = 2$ for the action; $J(2) = 4$ for source of authority, $J(3) = 5$ for justification and $J(4) = 3$ for tone.

[4] That is, for the sth element of $\boldsymbol{\phi}_{ka}$ – which we denote $\boldsymbol{\phi}^{(s)}_{ka}$ – we have:

$$\boldsymbol{\phi}^{(s)}_{ka} = \frac{\exp\left(\tilde{\boldsymbol{\phi}}^{(s)}_{ka}\right)}{1 + \sum_{m=1}^{J(a)} \exp\left(\tilde{\boldsymbol{\phi}}^{(t)}_{ka}\right)} \iff \tilde{\boldsymbol{\phi}}^{(s)}_{ka} = \ln\left(\frac{\boldsymbol{\phi}^{(s)}_{ka}}{1 - \sum_{m=1}^{J(a)} \boldsymbol{\phi}^{(t)}_{ka}}\right).$$

[5] Where the first element of these vectors are normalised to zero, and the whole vector is normalised to zero for the first respectively vignette and enumerator.

[6] $\boldsymbol{\psi}_{av}$ and $\boldsymbol{\chi}_{ae}$ provide a measurement scheme; this is similar in spirit to Item Response Theory. That is, this captures the notion that different enumerators will have different underlying tendencies to assess different attributes, and similarly that for different vignettes respondents have different underlying tendencies to display traits. Note that, in the special case that $\boldsymbol{\psi}_{av} = \mathbf{0}$ and $\boldsymbol{\chi}_{ae} = \mathbf{0}$, this collapses to a more standard Latent Dirichlet Allocation.

Table A.3: **Reported intentions of types by vignette**

Vignette	Type	Expresses distrust counterparty	Maintain relationship	Follow procedure	Set an example
Line Management	Authoritive	.58	0	.51	.38
	Affiliative	.41	0	.39	.32
	Coercive	.49	0	.36	.28
	Timid	.42	0	.44	.28
Supplier	Authoritative	.03	0	.30	.53
	Affiliative	.03	.02	.21	.44
	Coercive	.03	0	.21	.46
	Timid	.04	0	.16	.46
Pay rise	Authoritative	0	.12	.61	.34
	Affiliative	0	.21	.28	.24
	Coercive	0	.08	.50	.18
	Timid	0	.06	.47	.20
Bank	Authoritative	0	.71	0	0
	Affiliative	0	.75	0	0
	Coercive	0	.65	0	0
	Timid	0	.59	0	0
Client	Authoritative	.61	.20	.45	.07
	Affiliative	.23	.50	.18	.02
	Coercive	.60	.21	.35	.02
	Timid	.55	.19	.29	.02

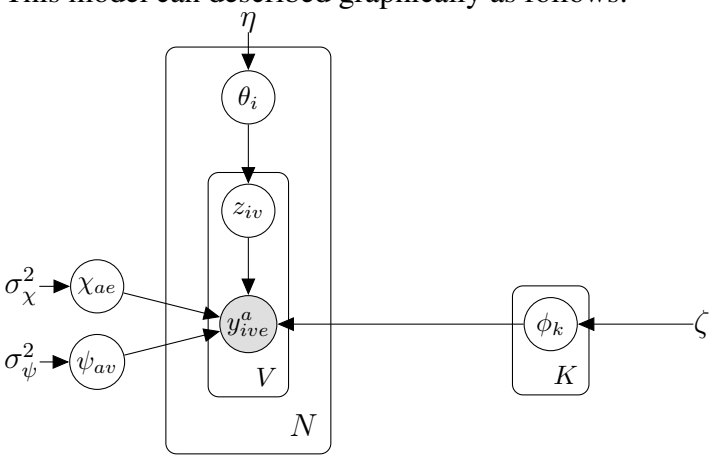
Notes This table describes the intentions of respondents' for each vignette, which are enumerated before the respondent actually responds to the vignette. The four characteristics are (i) whether the respondent expresses they do not trust the other part, (ii) whether the respondent wishes to maintain a good relationship, (iii) whether the respondent mentions following procedure, and (iv) whether the respondent wishes to set an example for future interactions. These numbers are calculated by assigning each individual to the pure type for which they have the highest estimated $\hat{\boldsymbol{\theta}}_i$. Then, a conditional average is taken for each pure type.

(ii). Draw $\boldsymbol{\theta}_i$ independently for $i \in \{1, \ldots, N\}$ from $\texttt{Dirichlet}(\eta)$, where $\boldsymbol{\theta}_i$ is a K-dimensional vector;

(iii). Draw $\boldsymbol{\psi}_{av}$ and $\boldsymbol{\chi}_{ev}$ from some suitable prior (respectively $\mathcal{N}(0, \sigma_\psi^2)$ and $\mathcal{N}(0, \sigma_\chi^2)$);

(iv). For each vignette, draw a type z_{iv} from $\boldsymbol{\theta}_i$ and draw an attribute from $y_{aive} \mid \boldsymbol{\phi}_{z_{iv},a}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ae}$ (using the Multinomial Logit formula above).

This model can described graphically as follows:



The total probability of the model is as follows, denoting by W^q the qth attribute within a vignette and by β^q the set of parameters for that attribute.[7]

$$P(W, Z, \boldsymbol{\theta}, \boldsymbol{\phi}, \eta, \zeta) = \prod_{j=1}^{K} P(\boldsymbol{\phi_j}; \zeta) \prod_{i=1}^{N} P(\boldsymbol{\theta_i}; \eta)$$
$$\prod_{v=1}^{V} \sum_{k=1}^{K} \left(P(z_{iv} = k | \boldsymbol{\theta_i}) \prod_{e=1}^{2} \prod_{a=1}^{A} P(y^a_{ive} | \boldsymbol{\phi}_{ka}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ae}) \right)$$

Where $P(\boldsymbol{\phi_j}; \zeta)$ and $P(\boldsymbol{\theta_i}; \eta)$ follow a latent Dirichlet distribution, $P(z_{iv} | \theta_i)$ follows a categorical distribution and $P(y^q_{ive} | \boldsymbol{\phi}_{z_{iv}}, \boldsymbol{\psi}_{av}, \boldsymbol{\chi}_{ae})$ follow multinomial logit distributions.
To fully specify the model, we need to impose an additional set of priors, specifically:

$$\alpha = 0.3$$
$$\beta = 1$$
$$\boldsymbol{\chi}_{ae} \sim \mathcal{N}(0, 3)$$
$$\boldsymbol{\psi}_{av} \sim \mathcal{N}(0, 3)$$

In practice, the model is highly insensitive to the choice of the prior β, as the data is highly informative for the parameter $\boldsymbol{\phi}_k$. However, the model is more sensitive to the choice of the parameter α since the parameter θ_i is estimated based only on data from a single candidate. The primary effect of the choice of α is the distance to the edge of the simplex of resulting parameter estimates.

[7] For clarify of exposition we omit the normal prior for ψ and χ from this equation. These enter the log-likelihood linearly.

E.2 Placket Luce model

To analyse the ranking data from the HR managers we develop the following Placket-Luce model. The model employs a Bayesian hierarchical structure where preferences for types and attributes can vary across firms. We include an individual fixed effect for each individual to improve the precision of the resulting estimates.

$$U_{fiv} = \boldsymbol{\beta}_f \boldsymbol{\theta}_i + \boldsymbol{\alpha}_f x_i + \gamma_i + \epsilon_{fiv} \tag{2}$$

Where we have the following independent variables:

$$\begin{aligned} \boldsymbol{\theta}_i \implies & \quad \text{an estimated parameter from the first-stage model; } \tilde{\boldsymbol{\theta}}_i \\ x_i \implies & \quad \text{“Known”} \end{aligned}$$

Specify the error term as follows:

$$\epsilon_{fiv} \sim EV1$$

Then we employ a Bayesian hierarchical structure for each set of parameters. For the individual mean γ_i, where we specify the error structure for the variance on a bounded uniform distribution, we specify:[8]

$$\begin{aligned} \gamma_i &\sim \mathbf{N}(0, \sigma_\gamma) \\ \sigma &\sim \mathbf{U}(0, 10) \end{aligned}$$

For identification, we set $\gamma_i = 0$ for the first member of every triplet.
Then for the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ parameters we use the following structure:

$$\begin{aligned} \boldsymbol{\alpha}_f &\sim \mathcal{N}(\boldsymbol{\alpha}, \sigma_\alpha^2) \\ \boldsymbol{\alpha} &\sim \mathcal{N}(0, 10) \\ \sigma_\alpha^2 &\sim \mathbf{U}(0, 10) \end{aligned}$$

and:

$$\begin{aligned} \boldsymbol{\beta}_f &\sim \mathcal{N}(\boldsymbol{\beta}, \sigma_\beta^2) \\ \boldsymbol{\beta} &\sim \mathcal{N}(0, 10) \\ \sigma_\beta^2 &\sim \mathbf{U}(0, 10) \end{aligned}$$

Noting this will be a BH model we get the following structure, denoting by θ the parameters, by y the data, and by λ the hyperparameters.
The resulting total probability, denoting $U(\boldsymbol{\beta}_f, \boldsymbol{\alpha}_f, \gamma_i, \boldsymbol{\theta_i}, \boldsymbol{x}_i) \equiv U_{fiv}$.

[8] In setting the prior for the variance as a uniform distribution we follow Gelman (2006).

$$P(y|\boldsymbol{\beta}_f, \boldsymbol{\alpha}_f, \gamma_i, \boldsymbol{\theta_i}, \boldsymbol{x}_i) =$$
$$\prod_{f=1}^{F}\prod_{v=1}^{5} \frac{\exp\left(U^1_{fiv}\right)}{\sum_{r=1}^{3}\exp\left(U^r_{fiv}\right)} \cdot \frac{\exp\left(U^2_{fiv}\right)}{\sum_{r=2}^{3}\exp\left(U^r_{fiv}\right)}.$$

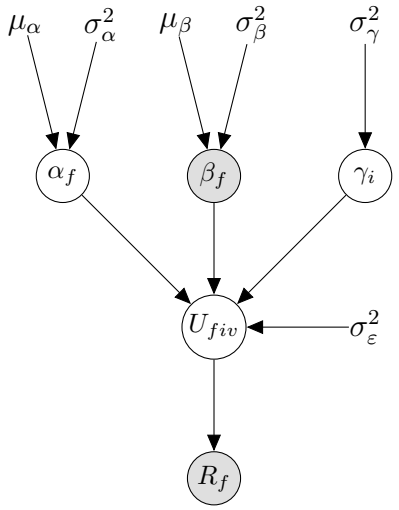
Then we specify the distribution of α_f, γ_i and β_f as:

$$P(\boldsymbol{\beta}_f, \boldsymbol{\alpha}_f, \gamma_i|\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\gamma, \sigma_\alpha, \sigma_\beta) =$$
$$\prod_{i=1}^{N}(\phi(\boldsymbol{\gamma}_i|0, \Sigma_\gamma) \prod_{f=1}^{F} \phi(\boldsymbol{\alpha}_f|\boldsymbol{\alpha}, \sigma_\alpha) \cdot \phi(\boldsymbol{\beta}_f|\boldsymbol{\beta}, \sigma_\beta)$$

And finally we define the hyperprior as:

$$P(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\gamma, \sigma_\alpha, \sigma_\beta) =$$
$$\phi(\boldsymbol{\alpha}|0, 10)\phi(\boldsymbol{\beta}|0, 10)$$
$$U(\sigma_\alpha, 0, 10)U(\sigma_\beta, 0, 10)U(\sigma_\gamma, 0, 10)$$

The full probability of drawing a set of parameters is the product of these three elements following Bayes' rule.



E.3 Estimation

We use Hamiltonian Monte Carlo (HMC) implemented in the STAN language to estimate our model. (e.g., Neal, Gelman, Betancourt as references). HMC is an advanced algorithm for Markov Chain Monte Carlo (MCMC) simulations. Its key feature is the use of the gradient of the posterior distribution in determining

Table A.4: Convergence statistics for the Dirichlet model with four types

Parameter	Average PSRF	Highest PSRF	Average effective sample size	Lowest effective sample size
Theta	0.999	1.009	2000	247
Phi Action	1.010	1.011	738	661
Phi Tone	0.999	1.001	1980	1359
Phi Justification	1.001	1.004	1462	965
Phi Authority	1.002	1.006	1116	791
Chi Action	0.999	0.999	2000	2000
Chi Tone	0.998	0.998	2000	2000
Chi Justification	0.999	1.000	2000	1748
Chi Authority	1.001	1.001	2000	1863
Psi Action	1.008	1.013	831	676
Psi Tone	0.999	1.002	2000	2000
Psi Justification	0.999	1.003	1777	1043
Psi Authority	1.001	1.006	1311	1228

Table A.4 shows the average and highest PSRF, and the average and lowest effective sample size for each subset of the parameters in the Dirichlet model. Crucially, the highest PSRF is low, showing good convergence of the model.

step sizes. This allows HMC to consider how the probability of the parameters changes, taking more informed steps during the simulation. This method is particularly useful because it often converges to high-probability regions faster and more efficiently than traditional MCMC methods. This is especially true in complex models with many parameters, like the one we implement. This efficiency is due to its ability to avoid getting stuck in less probable regions, a common issue with simpler MCMC methods.

E.3.1 The Dirichlet model

We first estimate the Dirichlet model, and use these results for Section 3. A few things are worth noting: Firstly, the commonly used Gibbs Sampler (Steyvers, Smyth, Rosen-Zvi, and Griffiths, 2004) is not applicable due to the additional parameters in our model. However, the small number of types and attributes makes our estimation procedure feasible. Secondly, a K-component mixture distribution is invariant to permutations of component labels, resulting in K! modes in the posterior distribution of the mixture parameters. This is particularly problematic when using HMCMC, as the leapfrog estimator may attempt to jump between different configurations of the component labels. In practice, the distinct modes of our posterior distribution prevent label switching in our draws. Consequently, we can re-label the types for consistent labeling across chains before performing standard convergence tests for MCMC simulations.

E.3.2 The Placket-Luce model and joint-estimation

We estimate the Plackett-Luce model using Hamiltonian Monte Carlo (HMC) in STAN, implementing a non-centered reparameterization of the hierarchical parameters. This technique results in faster con-

vergence and increases the precision of the model estimates. In such a nonlinear model, incorporating uncertainty in the first-stage parameters—the parameters of the Latent Dirichlet Allocation (LDA) model—presents a challenge. Properly propagating this uncertainty to the second-stage model provides a more realistic representation of the model's confidence in its predictions and estimates. To address this, we adopt an integrated approach, jointly estimating the LDA and Plackett-Luce models. This joint estimation approach leverages the curvature of the log-likelihood to guide the level of uncertainty about the type parameters, implicitly weighting the observations in the second-stage model by the uncertainty about θ_i.[9]

E.3.3 Stability of type parameters with joint estimation

Figures A.5 and A.6 show that the estimated type probability and distribution over type is consistent in the individually estimated Dirichlet model and the jointly estimated model.

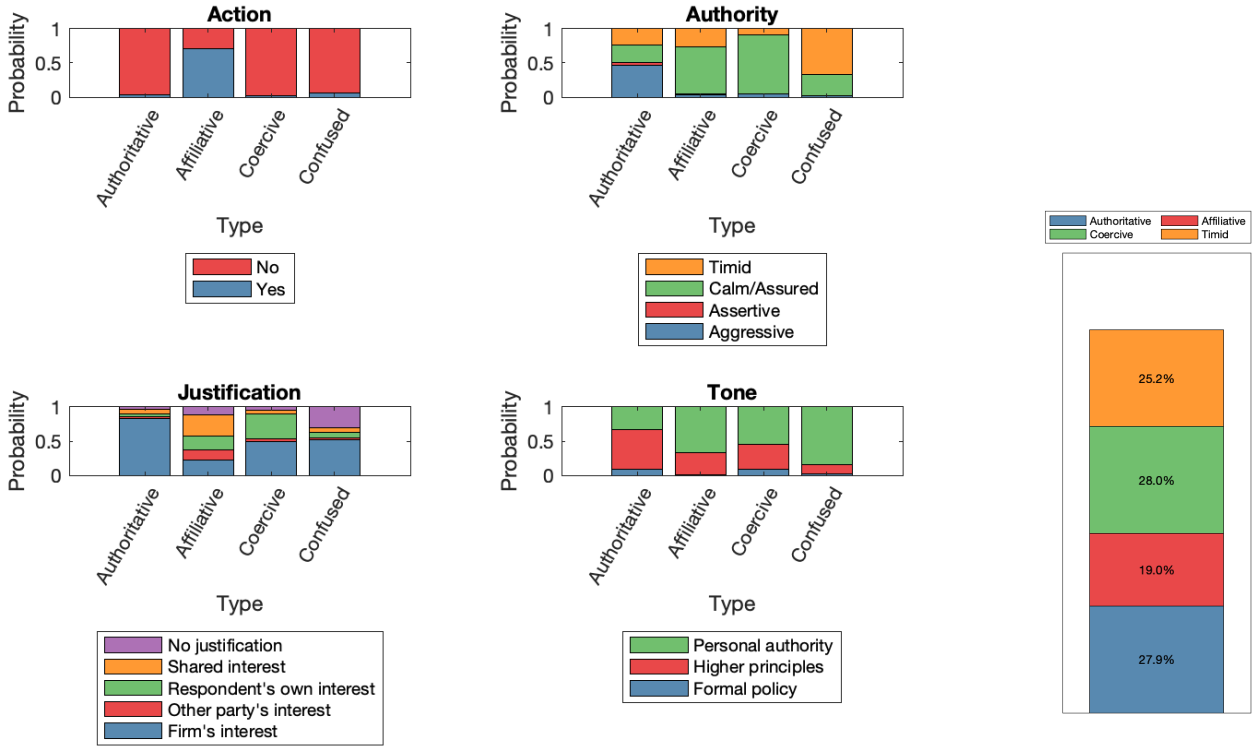


Figure A.5: The type parameters and distribution over type in the Dirichlet model

[9] Following Battaglia et al. (2024)

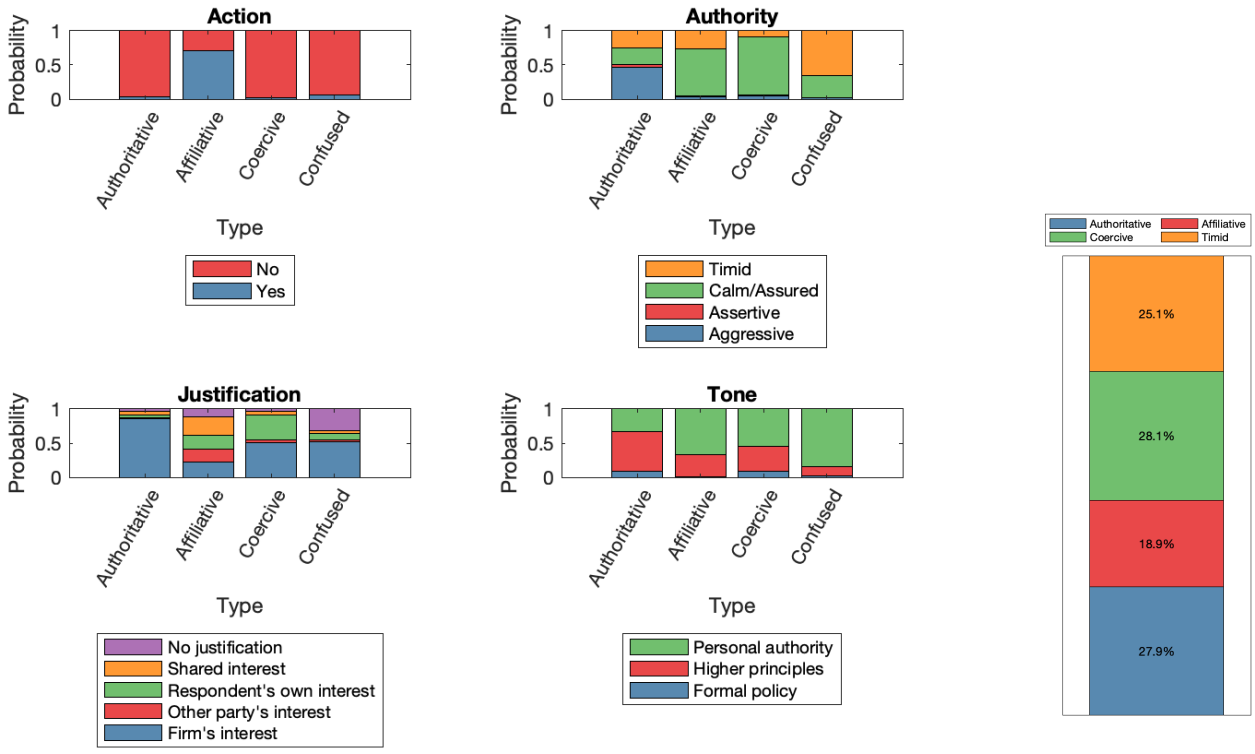


Figure A.6: The type parameters and distribution over type in the joint model

References

Abebe, G., M. Fafchamps, M. Koelle, and S. Quinn (2024). Matching, management and employment outcomes: A field experiment with firm internships. *Working paper*.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis 1*(3), 515 – 534.

Li, S. (2017). Obviously strategy-proof mechanisms. *The American Economic Review 107*(11), 3257–3287.

Steyvers, M., P. Smyth, M. Rosen-Zvi, and T. Griffiths (2004). Probabilistic author-topic models for information discovery. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 306–315.