

The Embedded State Appendix A: Model and Robustness

A Performance of the Embedded State: Details

This section proceeds in three steps. First, we demonstrate that our main results closely match the structural model estimates reported in Appendix Section A.1.

Next, we evaluate the robustness of our findings to alternative model specifications:

- Appendix Section A.2 examines robustness with respect to within- and cross-team spillovers.
- Appendix Section A.3 examines robustness with respect to the estimation of returns to team size.

Finally, we analyze the identification of the regression results within the structural framework, as discussed in Appendix Sections A.4 and O.C.

A.1 Model

Inputs

Denote the task-specific input of position bp to performance of team bt as α_{bpt} . To address our research question, we decompose individual inputs without loss of generality as follows:

$$\alpha_{bpt} = \alpha + \gamma \cdot u_{bp} + \varepsilon_{bpt}, \quad (8)$$

where α is the expected input of a paid position, $\alpha + \gamma$ is the expected input of an unpaid position, and ε_{bpt} is a position-team-specific component. Our primary parameter of interest is γ , the expected difference in team production inputs between unpaid and paid positions.

Team Production

Following canonical models of team production (Abowd et al., 1999; Bonhomme, 2025), we assume that team output is additively separable in individual position inputs:

$$Y_{bt} = \lambda_{N_{bt}} \sum_{bp \in P_{bt}} \alpha_{bpt} + \varepsilon_{bt}, \quad (9)$$

where Y_{bt} denotes the output of team bt , P_{bt} are the positions on team bt , ε_{bt} captures shocks to team output unrelated to position inputs, $N_{bt} = |P_{bt}|$ is the size of team bt , and $\lambda_{N_{bt}}$ is a team-size-dependent scaling factor (normalized so that $\lambda_1 = 1$), which governs returns to scale in team production.

Plugging individual-level inputs from equation (8) into the team production equation (9), we get

$$\begin{aligned} Y_{bt} &= \lambda_{N_{bt}} \sum_{bp \in P_{bt}} (\alpha + \gamma \cdot u_{bp} + \varepsilon_{bpt}) + \varepsilon_{bt} \\ &= \alpha \lambda_{N_{bt}} N_{bt} + \gamma \lambda_{N_{bt}} U_{bt} + \underbrace{\lambda_{N_{bt}} \sum_{bp \in P_{bt}} \varepsilon_{bpt}}_{\varepsilon_{bt}} + \varepsilon_{bt} \end{aligned} \quad (10)$$

where U_{bt} is the number of unpaid positions on team bt .

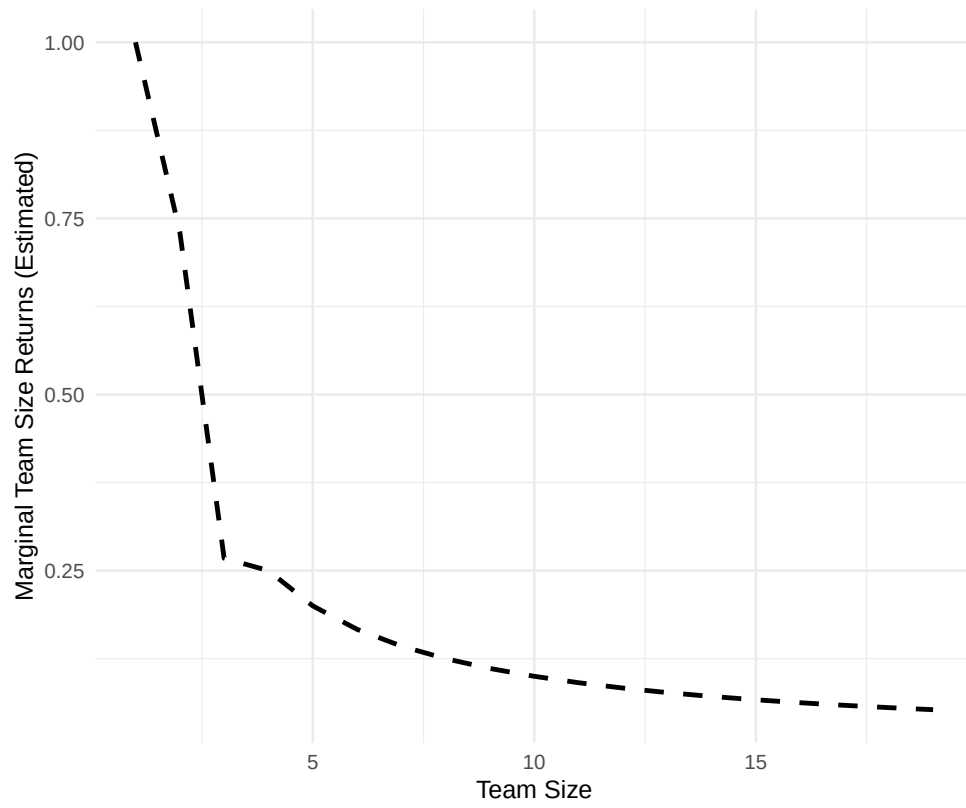
Estimating $\lambda_{N_{bt}}$

We first estimate the parameter that governs returns to team size $\lambda_{N_{bt}}$. Drawing on the algorithms proposed in Bonhomme (2025), we use our data on position-team assignment and team performance to estimate returns to team size.

This approach, which is rooted in a canonical ‘movers’ design (Abowd et al., 1999), relies on observing individuals across teams. Given the limitations of our historical data, we can estimate the returns to team sizes up to three positions: $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) = (1, 0.73, 0.27)$. This covers 72% of all teams. We then extrapolate the estimated sequence by assuming that the median effective team size remains constant as team size grows beyond our estimated range.

Appendix Figure A.1 plots the estimated sequence of $\hat{\lambda}_{N_{bt}}$ as a function of team size. In Appendix Section A.3, we provide robustness to alternative paths of $\lambda_{N_{bt}}$.

Figure A.1: Estimated Values of $\lambda_{N_{bt}}$ Across Team Sizes



Note: The figure displays the estimated and extrapolated values of $\lambda_{N_{bt}}$, as a function of team size N_{bt} . Team sizes range from 1 to 45.

Position-Level Decomposition

Then, we estimate the team production function in a setup symmetric to our baseline estimating equation (1). Define as w_{bpt} an indicator equal to 1 if position bp is assigned to team bt and 0 otherwise. Rewrite equation (9) to isolate the contribution of a single position bp :

$$Y_{bt} = \alpha \cdot w_{bpt} \lambda_{N_{bt}} N_{bt} + \gamma \cdot w_{bpt} u_{bp} \lambda_{N_{bt}} U_{bt} + \alpha \cdot (1 - w_{bpt}) \lambda_{N_{bt}} N_{bt} + \gamma \cdot (1 - w_{bpt} u_{bp}) \lambda_{N_{bt}} U_{bt} + \epsilon_{bt}, \quad (11)$$

which holds for every position bp and can thus be estimated in our borough-task panel. Appendix Table A.1 shows results of estimating the empirical analog of equation (11), using the estimated $\hat{\lambda}$ and a full set of fixed effects. Results are very similar to the baseline results in Table 4.

Table A.1: Unpaid and Performance (Model Estimation)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.18*** (0.04)	0.21*** (0.03)	0.21*** (0.03)
Observations	39,970	39,970	39,970
R^2	0.35	0.41	0.41
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.61	0.61	0.61
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Model Controls	✓	✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), modified according to equation (11). Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

A.2 Team Production: Spillovers

Clustering Because team performance may be correlated within boroughs, we assess the robustness of our results to alternative clustering choices. Appendix Table A.2 reports results from estimating equation (1) with standard errors clustered at various higher aggregation levels. Across all specifications, our main findings remain statistically robust.

Relative Performance To limit the scope for cross-team spillovers biasing our results, we construct alternative outcome variables:

1. Team performance relative to the leave-out mean of other teams in the borough,
2. Team performance relative to the leave-out max of other teams in the borough, and
3. A binary indicator of whether team performance exceeds the leave-out mean of other teams in the borough.

Appendix Tables A.3, A.4, and A.5 show results from estimating equation (1) using the outcome variables defined in (1)–(3) above. Our findings qualitatively hold up across these increasingly restrictive outcome variables.

Subsets To limit the scope for within-team spillovers, we restrict our sample to borough-tasks with just one position. Appendix Table A.6 shows results. To assess robustness to including very large teams, we can drop our restriction of truncating team size at ten positions: Appendix Table A.7 shows results when instead using the full team size distribution.

Aggregation We next assess the robustness of our findings to aggregating to the team level, to abstract from within-team spillovers. Equation (10) motivates a team-level empirical specification in which we estimate γ by regressing team performance on the number of unpaid positions, controlling for overall team size. Specifically, we estimate:

$$Y_{bt} = \beta_u \hat{\lambda}_{N_{bt}} U_{bt} + \beta_w \hat{\lambda}_{N_{bt}} N_{bt} + X_b \mathbb{1}(\tau = t)' \theta + \delta_b + \delta_t + \delta_{ct} + e_{bt}, \quad (12)$$

where all variables are as defined above, and δ_{ct} are county-by-task fixed effects. Standard errors are clustered at the borough level.

Appendix Table A.8 presents the results. The coefficient on U_{bt} is positive and statistically significant, indicating that a greater number of unpaid positions is associated with higher team performance, even after accounting for total team size and fixed effects.

To account for potential spillovers *between* teams, we aggregate our data to the borough level. We thus estimate

$$Y_b = \beta_u U_b + \beta_w N_b + X_b' \theta + \delta_c + e_b, \quad (13)$$

with $U_b = \sum_{t=1}^T \hat{\lambda}_{N_{bt}} U_{bt}$, $N_b = \sum_{t=1}^T \hat{\lambda}_{N_{bt}} N_{bt}$, and $Y_b = \frac{1}{T} \sum_{t=1}^T Y_{bt}$ the mean performance of the borough across tasks and all other variables as borough-level aggregates of the variables in equation (12). Appendix Table A.9 shows results, which again qualitatively affirm the association between unpaid status and performance.

Table A.2: Unpaid and Performance (Standard Errors)

	Performance						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unpaid Position on Team	0.12*** (0.02)	0.12*** (0.02)	0.12*** (0.02)	0.12*** (0.02)	0.12*** (0.02)	0.12*** (0.02)	0.12*** (0.02)
Standard-Errors	Borough $\times$ Task	Task Scope	Borough	County	50km	100km	200km
Observations	39,970	39,970	39,970	39,970	39,970	39,970	39,970
R^2	0.39	0.39	0.39	0.39	0.39	0.39	0.39
Outcome Mean	0.61	0.61	0.61	0.61	0.61	0.61	0.61
Borough fixed effects	✓	✓	✓	✓	✓	✓	✓
Task fixed effects	✓	✓	✓	✓	✓	✓	✓
Position Title fixed effects	✓	✓	✓	✓	✓	✓	✓

Note: The table presents results of estimating equation (1), using varying standard errors as indicated in the table. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.3: Unpaid and Relative Performance (Leave-Out Mean)

	Relative Performance (Leave-Out Mean)		
	(1)	(2)	(3)
Unpaid Position on Team	0.11*** (0.02)	0.13*** (0.02)	0.14*** (0.02)
Observations	39,970	39,970	39,970
R^2	0.18	0.25	0.25
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.00	0.00	0.00
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1). Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, subtracting the leave-out mean of borough-task performance in the respective borough.

Table A.4: Unpaid and Relative Performance (Leave-Out Max)

	Relative Performance (Leave-Out Max)		
	(1)	(2)	(3)
Unpaid Position on Team	0.09*** (0.02)	0.11*** (0.02)	0.12*** (0.02)
Observations	39,970	39,970	39,970
R^2	0.30	0.36	0.36
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	-0.39	-0.39	-0.39
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1). Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, subtracting the leave-out max of borough-task performance in the respective borough.

Table A.5: Unpaid and Outperformance (Leave-Out Mean)

	Outperform Leave-Out Mean		
	(1)	(2)	(3)
Unpaid Position on Team	0.09*** (0.02)	0.11*** (0.02)	0.12*** (0.02)
Observations	39,970	39,970	39,970
R^2	0.35	0.40	0.40
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.57	0.57	0.57
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1). Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is a binary indicator of whether borough-task performance exceeds the leave-out mean of borough-task performance in the respective borough.

Table A.6: Unpaid and Performance (Single-Position Teams)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.12*** (0.03)	0.08*** (0.02)	0.08*** (0.02)
Observations	14,684	14,684	14,684
R^2	0.51	0.63	0.63
Boroughs	248	248	248
Positions	5157	5157	5157
Tasks	7	7	7
Outcome Mean	0.50	0.50	0.50
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), restricting the sample to single-position teams. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.7: Unpaid and Performance (Unrestricted Team Size)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.07** (0.04)	0.11*** (0.03)	0.12*** (0.03)
Observations	62,874	62,874	62,874
R^2	0.34	0.39	0.39
Boroughs	258	258	258
Positions	1912	1912	1912
Tasks	7	7	7
Outcome Mean	0.62	0.62	0.62
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), not truncating the size of exceptionally large teams. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.8: Unpaid and Performance (Team Level)

	Performance		
	(1)	(2)	(3)
Unpaid Positions	0.10*** (0.03)	0.12*** (0.03)	0.10*** (0.04)
Observations	1,332	1,332	1,332
R^2	0.38	0.44	0.63
Boroughs	258	258	258
Tasks	7	7	7
Outcome Mean	0.61	0.61	0.61
Cluster	Borough	Borough	Borough
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
County $\times$ Task fixed effects			✓

Note: The table presents results of estimating equation (12). Observations are at the level of a borough-task. The number of borough-tasks and boroughs in each sample is indicated in the table. Standard errors are clustered at the borough level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.9: Unpaid and Performance (Borough Level)

	Performance (Borough)		
	(1)	(2)	(3)
Unpaid Positions	0.12** (0.05)	0.13** (0.06)	0.10* (0.06)
Observations	258	258	258
R^2	0.39	0.53	0.59
Boroughs	258	258	258
Outcome Mean	3.57	3.57	3.57
Cluster	County	County	County
Controls			✓
County fixed effects		✓	✓

Note: The table presents results of estimating equation (13). Observations are at the level of a borough. The number of boroughs in each sample is indicated in the table. Standard errors are clustered at the county level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is the mean performance across all tasks in a borough, as introduced in Section 4.

A.3 Team Production: Returns to Team Size

Estimated Returns to Team Size Due to data limitations, we extrapolate estimated returns to team size beyond teams of three members. While this captures three quarters of all teams, we show results when restricting our sample to just those teams in Appendix Table A.10.

Plausible Returns to Team Size To assess the robustness of our results across all team sizes and all plausible returns to team size, we evaluate a family of functions for $\lambda_{N_{bt}}$ of the form:

$$\lambda(N_{bt}) = \zeta + (1 - \zeta) \frac{1}{N_{bt}}, \quad (14)$$

where $\zeta \in [0, 1]$. This specification interpolates between constant returns to scale ($\zeta = 1$) and a setting where the marginal contribution of each team member decreases inversely with team size ($\zeta = 0$).

Appendix Figure A.2 plots the resulting values of $\lambda_{N_{bt}}$ against observed team sizes, which range from 1 to 45. We consider $\zeta \in [0, 1]$ in step sizes of $\frac{1}{1000}$.

We then re-estimate regression equation (1), scaling both w_{bpt} and $w_{bpt}u_{bp}$ by the corresponding $\lambda_{N_{bt}}$. Across the full range of $\lambda_{N_{bt}}$ values considered, the estimated coefficient $\hat{\beta}_u$ remains positive and significant at the one percent level, indicating a robust positive relationship between unpaid status and team output.

Estimating Paid-Unpaid Differences in Inputs Finally, we can apply the methodology from Bonhomme (2025) to assess whether our general findings also hold up in a sample of estimated *individual* team inputs. Note that relying on the ‘connected set’ reduces our sample to 433 positions.¹

We proceed in two steps:

1. We estimate individual position productivity inputs $\hat{\alpha}_{bp}$.
2. We then regress the estimated inputs $\hat{\alpha}_{bp}$ on the unpaid status of each position u_{bp} .

We thus test for statistically significant differences in individual inputs between unpaid and paid positions. Appendix Table A.11 reports the results of Step 2, both with and without the inclusion of borough fixed effects. Among unpaid positions, estimated inputs are significantly higher than among paid positions.

¹As outlined in equation (8), our parameter of interest relies on the mean of the distribution of individual inputs, and does not require characterizing the full distribution of individual inputs. Hence, our main analyses use the entire sample.

Table A.10: Unpaid and Performance (Estimated Returns)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.10*** (0.03)	0.10*** (0.03)	0.10*** (0.03)
Observations	22,583	22,583	22,583
R^2	0.44	0.52	0.51
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.61	0.61	0.61
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Model Controls	✓	✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

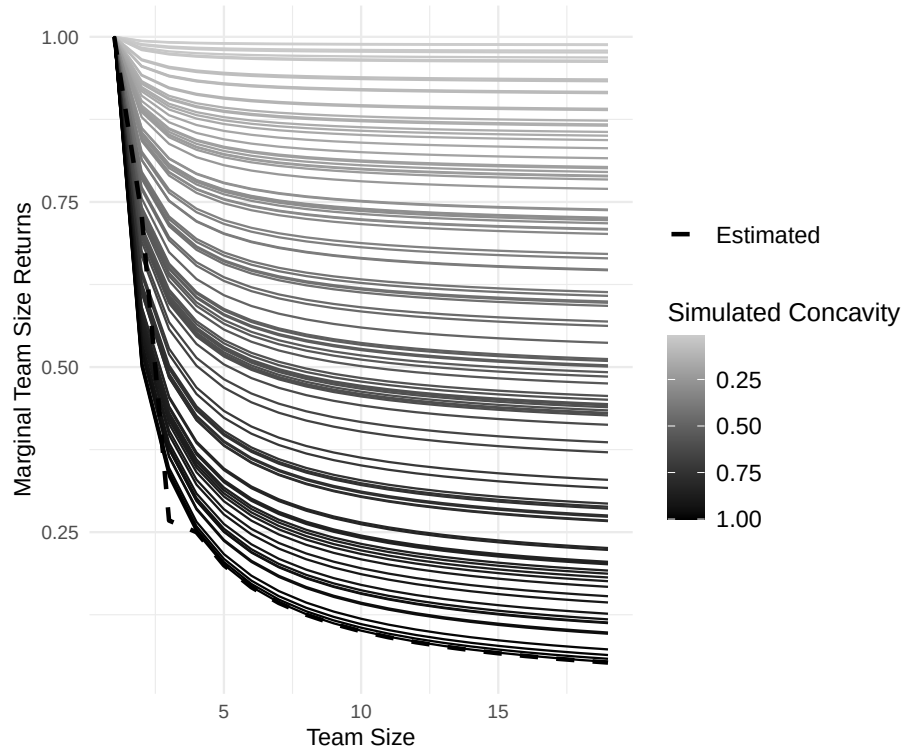
Note: The table presents results of estimating equation (1), using estimated returns to team size as described in Appendix Section A.3. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.11: Unpaid and Estimated Inputs (Position Level, Bonhomme, 2025)

	Officer Ability		
	(1)	(2)	(3)
Unpaid	0.22*** (0.05)	0.21*** (0.05)	0.26*** (0.08)
Observations	433	433	433
R^2	0.05	0.07	0.42
Boroughs	219	219	219
Outcome Mean	0.00	0.00	0.00
Cluster	Borough	Borough	Borough
Controls		✓	✓
Borough fixed effects			✓

Note: Table presents results from regressing estimated team performance inputs using Bonhomme (2025) on unpaid status. Column (3) includes borough fixed effects. Observations are at the level of a position. The number of positions in each sample is indicated in the table. Standard errors are clustered at the borough level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is the estimated position input $\hat{\alpha}_{bp}$ from the team production model in Appendix Section A.1.

Figure A.2: Simulated Values of $\lambda_{N_{bt}}$ Across Team Sizes



Note: The figure displays the values of $\lambda_{N_{bt}}$ used in our robustness analysis, as a function of team size N_{bt} , for selected values of ζ . Team sizes range from 1 to 45.

A.4 Comparable Control Groups

We use the model from equation (10) to structure our discussion of potential omitted variable bias.

To estimate γ consistently via OLS, we require $\mathbb{E}[\epsilon_{bt} \mid N_{bt}, U_{bt}] = 0$. This requires:

$$\mathbb{E}[\epsilon_{bt} \mid N_{bt}, U_{bt}] = 0 \quad \text{and} \quad \mathbb{E}[\epsilon_{bpt} \mid N_{bt}, U_{bt}] = 0 \text{ for all } bp \in P_{bt}.$$

Fixed Effects The inclusion of fixed effects in equation (1) implies a set of within-transformations that remove variation at the borough, task, and position level. Most importantly, borough and task fixed effects remove all components of ϵ_{bt} that do not vary across tasks or boroughs, respectively. These transformations thus help isolate variation that is orthogonal to systematic differences across boroughs, tasks, and position patterns. Below, we highlight the terms subtracted from the main variables in equation (1) as a consequence of the specific fixed effect.

- **Borough Fixed Effects (δ_b)**

– Y_{bt} :

$$\frac{1}{O_b \cdot T} \sum_{p=1}^{O_b} \sum_{t=1}^T Y_{bt} = \frac{1}{T} \sum_{t=1}^T Y_{bt}$$

Captures borough-level productivity differences across tasks (i.e., ‘some boroughs do better than others’).

– w_{bpt} :

$$\frac{1}{O_b \cdot T} \sum_{p=1}^{O_b} \sum_{t=1}^T w_{bpt}$$

Captures borough-level variation in staffing intensity (i.e., ‘some boroughs have more positions’ or ‘some positions cover more tasks’).

– $w_{bpt}u_{bp}$:

$$\frac{1}{O_b \cdot T} \sum_{p=1}^{O_b} \sum_{t=1}^T u_{bp} w_{bpt}$$

Captures borough-level variation in unpaid staffing (i.e., ‘some boroughs have more unpaid positions’ or ‘some boroughs assign them to more tasks’).

- **Task Fixed Effects (δ_t)**

– Y_{bt} :

$$\frac{1}{B \cdot O_b} \sum_{b=1}^B \sum_{p=1}^{O_b} Y_{bt} = \frac{1}{B} \sum_{b=1}^B Y_{bt}$$

Captures task-specific difficulty (i.e., ‘some tasks are easier’).

– w_{bpt} :

$$\frac{1}{B \cdot O_b} \sum_{b=1}^B \sum_{p=1}^{O_b} w_{bpt}$$

Captures task implementation and staffing (i.e., ‘some tasks are more common across boroughs’ or ‘some tasks have larger teams’).

– $w_{bpt}u_{bp}$:

$$\frac{1}{B \cdot O_b} \sum_{b=1}^B \sum_{p=1}^{O_b} u_{bp}w_{bpt}$$

Captures unpaid staffing across tasks (i.e., ‘some tasks are more often staffed with unpaid positions’ or ‘unpaid teams are larger in certain tasks’).

• **Position Fixed Effects** (δ_p)

– Y_{bt} :

$$\frac{1}{\underline{P}} \sum_{b=1}^B \sum_{p=1}^{O_b} Y_{bt} \mathbb{1}(p = \underline{p})$$

Captures the average task performance of positions with the same position title $\underline{p}$, with $\underline{P} = \sum \mathbb{1}(p = \underline{p})$.

– w_{bpt} :

$$\frac{1}{\underline{P}} \sum_{b=1}^B \sum_{p=1}^{O_b} w_{bpt} \mathbb{1}(p = \underline{p})$$

Measures the scope of typical position task assignments (i.e., ‘this position covers more tasks’).

– $w_{bpt}u_{bp}$:

$$\frac{1}{\underline{P}} \sum_{b=1}^B \sum_{p=1}^{O_b} w_{bpt}u_{bp} \mathbb{1}(p = \underline{p})$$

Measures the breadth of unpaid assignments (i.e., ‘this unpaid position is assigned to more tasks’).

In addition, we can vary the set of fixed effects included in the regression specification. Appendix Table A.12 presents results from a specification that includes task scope fixed effects, position-by-task fixed effects, and county-by-task fixed effects. The estimated correlation with unpaid status remains stable, indicating that our main findings are not sensitive to task-specific confounders that vary at the county level.

Controls Appendix Table A.13 reports results from equation (1), including position-level controls. This addresses potential omitted variable bias from components in ε_{bpt} . Specifically, we interact team assignment (w_{bpt}) with indicator variables capturing whether a position is a magistrate, a member of the governing body or the common council, a Justice of the Peace, subject to Crown approval, or part-time. Across all specifications, our main coefficient of interest remains stable and comparable to the baseline estimates. Moreover, the additional controls are not significantly associated with team performance, suggesting that these observed characteristics do not confound the main relationship.

Appendix Table A.14 reports results from equation (1), allowing borough report length to impact tasks differentially. We include the page length of the borough report, chunked into deciles, and interacted with task indicators. This addresses potential omitted variable bias from components in ε_{bpt} from reporting bias. Across all specifications, our main coefficient of interest remains stable and comparable to the baseline estimates, again suggesting that these observed characteristics do not confound the main relationship.

Matching To more comprehensively account for potential differences at the borough and position levels, we implement a propensity score matching approach. Specifically, we estimate a probit model to predict a position’s unpaid status using the full set of borough and position characteristics. Based on the estimated propensity scores, we use a nearest neighbor matching algorithm to pair each unpaid position with the most similar paid counterpart. We then re-estimate regression equation (1) on this matched sample. Results, reported in Appendix Table A.15, are qualitatively consistent with our main specification, indicating that the observed relationship is not driven by systematic differences in observables.

Outliers Next, we assess whether our results are driven by specific tasks. Appendix Figure A.3 presents estimates from regression equation (1), sequentially omitting one task at a time from the sample. Across all leave-one-out specifications, the estimated effect of unpaid status remains statistically significant and qualitatively consistent with our main findings, suggesting that no single task is disproportionately influencing the results.

Binarizing Performance We also assess robustness to the binarization of task performance. Our main measure introduced in Section 3.1 binarizes each ordinal or count measure at the median. Instead, we consider alternative thresholds in a neighborhood of the median: for example, the cutoff for ‘number of courts held’ is three. We then consider cutoffs in (2, 3, 4). We apply this neighborhood cutoff rule to all performance measures. Then, we choose a random draw from the cross-product of all possible cutoff rules

and re-estimate regression equation (1). Appendix Figure A.4 shows results: throughout all permutations, coefficients remain significant and positive.

Alternative Performance Measures We also assess robustness to the definition of task performance. In addition to our main measures introduced in Section 3.1, we consider alternatives to *each* of our measures. Five of these measures are from sources fully independent of the main report. One is derived from a later analytical index to the municipal report, and the one that is from the report proper (the number of submitted documents) cannot be ascribed to the biases of the commissioners. We describe all seven in turn below.

For *governing*, we use an indicator for whether the borough had moved to adopt the British Imperial measurement system in 1826 (House of Commons, 1826). Historically, many places in England held customary and varying definitions of units of mass, volume, and distance; the Weights and Measures Act of 1824 sought overall conversion to the British Imperial standard. However, some places were slower to adopt the reform. Here, we exploit variation in whether boroughs had checked their standards against the central models provided. For *judicial*, we count the number of days borough courts of record—civil courts for fast resolution of small claims and debt cases—sat in 1836 from Parliamentary reports on borough courts (House of Commons, 1839b; 1839c). This measure, by contrast with the measure used in the paper (the number of courts), indicates the actual activity of the town’s legal system. For *infrastructure*, we scrape Historic England’s catalog of historic sites to construct an indicator for whether a borough had a market cross, a structure representing the site of the town’s market (Historic England, 2026). For *clerical*, we use an indicator for whether the commissioners for the 1835 Municipal Corporations Report were supplied with documents by the borough’s town clerk.² For *financial*, we consult the Analytical Index to the Municipal Report to construct an indicator for whether the town’s accounts were audited (House of Commons, 1839a). For *prison*, we use the number of cells per prisoner from an 1833 Parliamentary survey (House of Commons, 1833a). For *law enforcement*, we take the negative of the number of criminal convictions from an 1825 Parliamentary report tabulating all categories of crime (property and violent) in that year by borough (House of Commons, 1825).

Then, we choose a random draw from the cross-product of all possible variable combinations and re-estimate regression equation (1). Appendix Figure A.5 shows results: throughout all permutations, coefficients remain significant and positive.

²For instance, some towns offered the commissioners comprehensive documentation, e.g., complete financial accounts, lists of burgesses, boundary maps, ancient charters, judicial records, etc., while others provided few or no documents at all.

Extensive and Intensive Margins A further concern is the distinction between the *extensive* and *intensive* margins of performance: whether a public good is implemented at all versus how well it is implemented. Our performance measures introduced in Section 3 allow us to separate these margins. With the exception of the ‘governing’ category, a zero in the raw (pre-binarization) data indicates that a public good is not implemented at all. This motivates a set of robustness checks. Appendix Table A.16 applies a binarization at zero. For the ‘governing’ category, we instead define a binary variable that indicates whether any governing position is listed in the report as appointing subordinate positions. Coefficients are positive and statistically significant: unpaid status is associated with greater public goods provision. The effects are economically meaningful: adding an unpaid position to a team is associated with a three percentage point higher likelihood of successful implementation. These results are robust and stronger when excluding the ‘governing’ task, at five percentage points (Appendix Table A.17).

Appendix Table A.18 examines the intensive margin by excluding tasks that are never implemented and using the logged raw performance measures as continuous outcomes. Unpaid status is again associated with better performance.

Finally, our main specification is, by construction, conditional on a position being assigned to a task; we do not observe nonexistent positions or unassigned tasks. To address this limitation, we adopt a borough-level correlational approach. We construct a variable measuring the fraction of tasks implemented, equal to zero if a borough implements no public goods and one if it implements all. We then estimate equation (13) at the borough level. Appendix Table A.19 shows that unpaid status is associated with greater public goods provision, suggesting that our position-task-level estimates of the extensive-margin contribution of the Embedded State may be conservative.

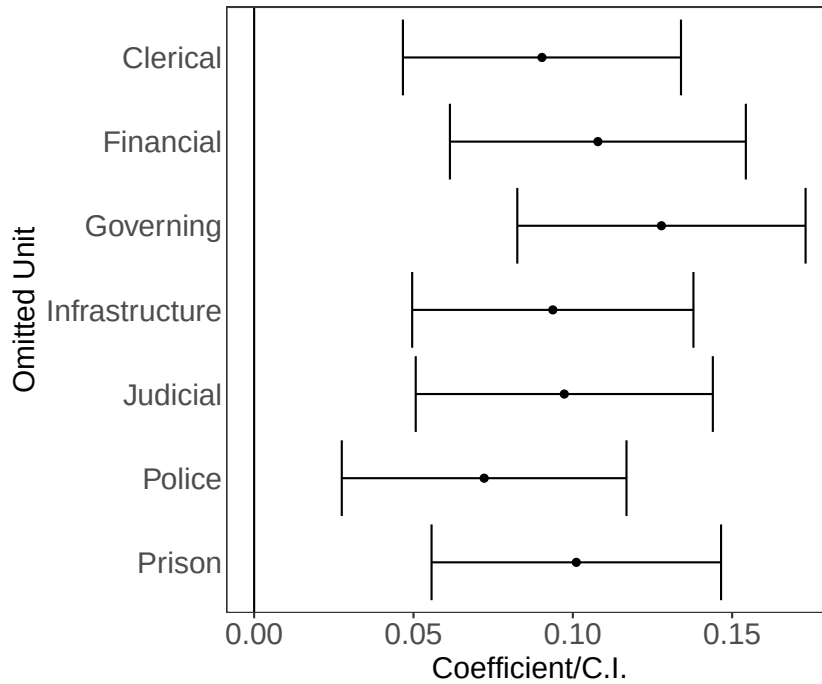
Any Wages Versus Market Wages Finally, we assess robustness to the definition of paid status. We omit all positions from our data that are paid a small wage, where paid or unpaid status might be ambiguous. We omit all positions paid less than 20 pounds a year (the median wage among paid positions is 16 pounds) and show that results are robust to omitting positions with low wages in Appendix Table A.20. Appendix Figure A.6 shows the unpaid advantage holds within objective outcomes (recorded facts and counts) and within subjective outcomes (commissioner ratings) estimated separately, which addresses the concern that reliance on a single reporting source drives the result. Appendix Figure O.C.1 provides a complementary check on the constructed inputs: permuting borough incomes shows that the relationship between unpaid and unaffordable status is not driven by our treatment of market wages or the salary prioritization ranking.

Table A.12: Unpaid and Performance (Fixed Effects)

	Performance			
	(1)	(2)	(3)	(4)
Unpaid Position on Team	0.06*** (0.02)	0.07*** (0.02)	0.08*** (0.02)	0.06*** (0.02)
Observations	39,970	39,970	39,970	39,970
R^2	0.54	0.54	0.58	0.73
Outcome Mean	0.61	0.61	0.61	0.61
Controls	✓	✓	✓	✓
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Borough fixed effects	✓	✓	✓	✓
Task fixed effects	✓	✓		
Position Title fixed effects	✓	✓		
Task Scope fixed effects		✓	✓	✓
Position Title $\times$ Task fixed effects			✓	✓
County $\times$ Task fixed effects				✓

Note: The table presents results of estimating equation (1), using restrictive fixed effects as indicated in the table. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Figure A.3: Unpaid and Performance (Leave-Out)



Note: The figure presents results of estimating equation (1), leaving out one task at a time. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.13: Unpaid and Performance (Position Characteristics)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.10*** (0.02)	0.12*** (0.02)	0.12*** (0.02)
Magistrate $\times$ On Task	0.06 (0.04)	0.06 (0.04)	0.06 (0.04)
Part of Governing Body $\times$ On Task	0.01 (0.04)	0.004 (0.04)	0.003 (0.04)
Part of Common Council $\times$ On Task	-0.02 (0.04)	0.010 (0.04)	0.008 (0.04)
Justice of the Peace $\times$ On Task	-0.05 (0.04)	-0.04 (0.04)	-0.04 (0.04)
Approval by Crown $\times$ On Task	0.07 (0.07)	-0.01 (0.06)	-0.01 (0.06)
Part-Time Office $\times$ On Task	-0.009 (0.02)	-0.02 (0.02)	-0.02 (0.02)
Observations	39,970	39,970	39,970
R^2	0.33	0.39	0.39
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.61	0.61	0.61
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), controlling for position characteristics as indicated in the table. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.14: Unpaid and Performance (Report Length)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.09*** (0.02)	0.07*** (0.02)	0.08*** (0.02)
Observations	39,970	39,970	39,970
R^2	0.33	0.45	0.45
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.61	0.61	0.61
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

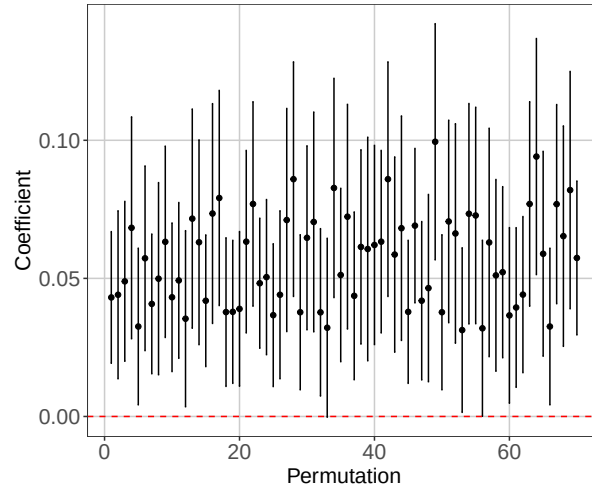
Note: The table presents results of estimating equation (1), additionally controlling for deciles of borough report page length interacted with task indicators. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.15: Unpaid and Performance (Matching)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.11*** (0.03)	0.13*** (0.03)	0.15*** (0.04)
Observations	25,788	25,788	25,788
R^2	0.33	0.39	0.39
Boroughs	237	237	237
Positions	3684	3684	3684
Tasks	7	7	7
Outcome Mean	0.62	0.62	0.62
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

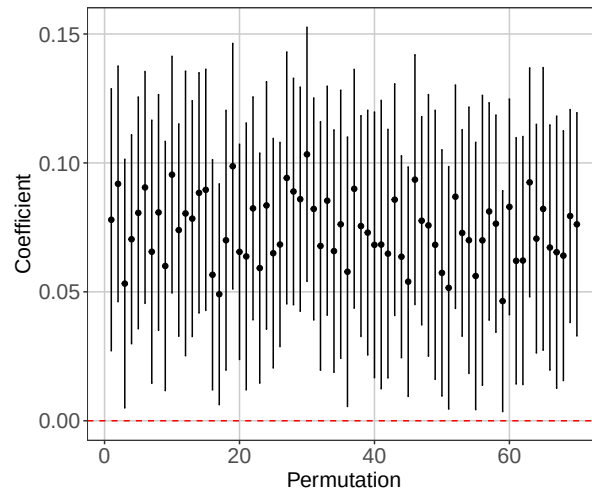
Note: The table presents results of estimating equation (1), in a matched sample. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Figure A.4: Unpaid and Performance (Binarization Variation)



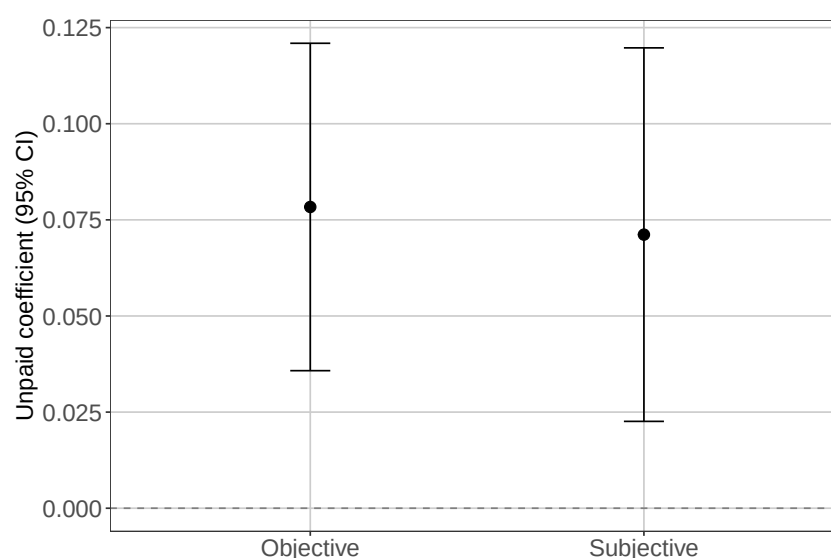
Note: The figure presents results of estimating equation (1), varying the cutoffs for borough-task performance binarization as detailed in Appendix Section A. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Figure A.5: Unpaid and Performance (Definition Variation)



Note: The figure presents results of estimating equation (1), varying task performance definitions as detailed in Appendix Section A. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Figure A.6: Unpaid and Performance (Objective vs Subjective Outcomes)



Note: The figure presents results of estimating equation (1), estimated separately within objective outcomes (recorded facts and counts: documents, courts, accounts, paving) and within subjective outcomes (commissioner quality ratings: governance, police, gaol). Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.

Table A.16: Unpaid and Performance (Extensive Margin)

	Performance (Extensive)		
	(1)	(2)	(3)
Unpaid Position on Team	0.02** (0.01)	0.03*** (0.01)	0.03*** (0.01)
Observations	39,970	39,970	39,970
R^2	0.66	0.68	0.68
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	7	7	7
Outcome Mean	0.75	0.75	0.75
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), with extensive-margin performance as an outcome. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is extensive-margin borough-task performance, as defined in Appendix Section A.4.

Table A.17: Unpaid and Performance (Extensive Margin, Excluding Governing)

	Performance (Extensive)		
	(1)	(2)	(3)
Unpaid Position on Team	0.04*** (0.01)	0.05*** (0.01)	0.05*** (0.01)
Observations	34,260	34,260	34,260
R^2	0.38	0.42	0.42
Boroughs	258	258	258
Positions	5710	5710	5710
Tasks	6	6	6
Outcome Mean	0.87	0.87	0.87
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), with extensive-margin performance as an outcome, excluding the 'governing' task. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is extensive-margin borough-task performance, as defined in Appendix Section A.4.

Table A.18: Unpaid and Performance (Intensive Margin)

	Performance (Intensive)		
	(1)	(2)	(3)
Unpaid Position on Team	0.12*** (0.03)	0.11*** (0.02)	0.12*** (0.03)
Observations	29,829	29,829	29,829
R^2	0.88	0.89	0.89
Boroughs	254	254	254
Positions	5689	5689	5689
Tasks	6	6	6
Outcome Mean	-1.08	-1.08	-1.08
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), with intensive-margin performance as an outcome. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is intensive-margin borough-task performance, as defined in Appendix Section A.4.

Table A.19: Unpaid and Extensive Margin Performance (Borough Level)

	Performance (Extensive, Borough)		
	(1)	(2)	(3)
Unpaid Positions	0.01*** (0.005)	0.01*** (0.005)	0.01** (0.004)
Observations	258	258	258
R^2	0.30	0.51	0.60
Boroughs	258	258	258
Outcome Mean	0.76	0.76	0.76
Cluster	County	County	County
Controls			✓
County fixed effects		✓	✓

Note: The table presents results of estimating equation (13), with treatment and controls as defined in the estimating equation, and performance defined as 'share of tasks implemented.' Observations are at the level of a borough. The number of boroughs in each sample is indicated in the table. Standard errors are clustered at the county level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table A.20: Unpaid and Performance (Omitting Low Wages)

	Performance		
	(1)	(2)	(3)
Unpaid Position on Team	0.08*** (0.02)	0.09*** (0.02)	0.10*** (0.02)
Observations	32,193	32,193	32,193
R^2	0.33	0.39	0.39
Boroughs	254	254	254
Positions	4599	4599	4599
Tasks	7	7	7
Outcome Mean	0.62	0.62	0.62
Cluster	Borough $\times$ Task	Borough $\times$ Task	Borough $\times$ Task
Task Control	✓	✓	✓
Controls		✓	✓
Borough fixed effects	✓	✓	✓
Task fixed effects	✓	✓	✓
Position Title fixed effects			✓

Note: The table presents results of estimating equation (1), omitting positions with low wages. Observations are at the level of a position-task. The number of position-tasks and positions in each sample is indicated in the table. Standard errors are clustered at the borough-task level. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively. The dependent variable is borough-task performance, as introduced in Section 4.