

**UNDERSTANDING AND REDUCING ABSENTEISM
IN THE INDIAN CIVIL SERVICE
WITH SPECIAL REFERENCE TO KARNATAKA**

Guo Xu¹

UC Berkeley, Haas School of Business

February 7, 2020

EXECUTIVE SUMMARY

- This report summarizes the findings based on the analysis of the attendance data provided by the Karnataka Evaluation Authority (KEA).
 - The data covers the clock-in and clock-out times of 7,757 individuals between January 2014 and June 2019.
 - *Note:* we analyze variation in clock-in and clock-out patterns recorded at the main premises of the Government Secretariat in Bangalore. To the extent that employees pursue legitimate work-related activities outside of these premises, the recorded hours do not reflect unexcused absence.
- We find substantial variation in attendance, both in the extensive margin (clocking in at work) and the intensive margin (the hours spent at work for those who clock in).
 - Around a third of civil servants are not clocking in on a given day.
 - The mean hours of attendance is 7.5 hours; 5% of the labor days however are shorter than 6.2 hours, and 5% longer than 9.4 hours.

¹ Email: guo.xu@berkeley.edu. We are extremely grateful for Dr. Shalini Rajneesh (IAS) for supporting this research. We also thank Dr. Chaya Degaonkar for the valuable feedback and assistance during the research.

- Variation in attendance *across* individuals explain up to 42% of the overall variation in attendance. While there exists differences in attendance across time and departments, they explain much less of the overall variation.
 - Excluding common holidays, at most 4.3% of the variation in attendance can be explained by time-variation. Absence from the main premises is highest in April and lowest in January; absence is highest on Friday and lowest on Wednesday.
 - Differences in absenteeism and hours present across departments explain at most 6.3% of the overall variation in attendance.

- Senior officers are more likely to be present, and work longer hours per day on average.
 - Compared to junior employees, senior employees are 21% points more likely to be present on any given day, staying on average 43 minutes longer each day.
 - While a third of all employees clock-out between 5-5.30pm, senior officers leave office substantially later.
 - We document a positive relationship between the attendance patterns of the senior officer and the attendance of his/her subordinates.

- Further research can help to better understand how to shape attendance dynamics
 - Data on contractual hours of work is needed to understand if absenteeism or short work hours reflect deviations from the contractual hours. This would allow us to provide a precise assessment on the number of days “lost” due to absenteeism.
 - Longer hours of work need not translate into productivity and work satisfaction. Data on productivity and well-being could help understand the link between hours of work and performance.
 - To benchmark the results, we require similar attendance data for organizations beyond the Government Secretariat (e.g. district administrations).
 - Researchers at the Haas School of Business, UC Berkeley are keen to provide research capacity to foster an international co-generation of knowledge between academia and public sector.

1. BACKGROUND AND DATA

The Karnataka Evaluation Authority (KEA) provided attendance data for all apex state offices on the main premises of the Government Secretariat at Bangalore. The sample covers period January 2014 to June 2019. The dataset covers 7,757 individuals over 4.41 million labor-days.

The attendance data measures the time an employee clocks in and out. The measurement is implemented using the Aadhar-enabled biometric attendance system. Since these measurements are taken at the apex state offices at the Government Secretariat at Bangalore, the data quality is high. We conducted basic quality checks to validate the data.

Interpreting the findings: Before reporting the findings of the empirical analysis, it is important to provide a few important caveats to the interpretation of the results.

The dataset allows us to measure clocking in and clocking out behavior employees. As such, it tells us whether a given employee was working on the government premises or not. Simply being present, however, is not necessarily a measure of productivity. Shorter hours at work thus do not necessarily mean lower performance on the job. Likewise, being present for long hours may not necessarily translate into more output. Similarly, being absent according to the attendance data need not mean that a given employee was shirking and not working: officers could have work-related meetings outside the office, and there are legitimate reasons for being absent from work such as part-time work, home office, or absence due to sickness or vacation. In our analysis, we exclude weekends and national/state holidays.² Nonetheless, understanding the variation in the attendance pattern is an important first step.

² As an approximation, we define common holidays as those where more than 90% of all employees are absent.

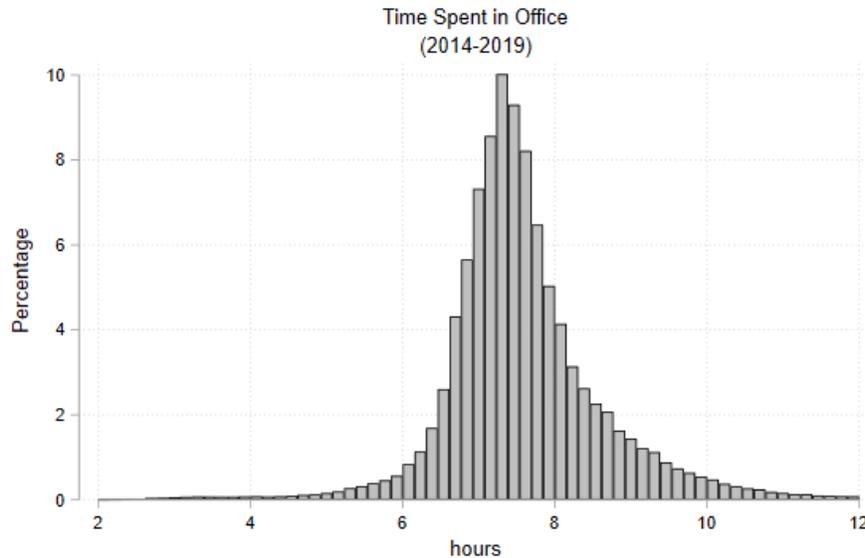
2. HOW MUCH VARIATION IS THERE IN ATTENDANCE?

Over our study period, there exists substantial dispersion in the hours present per day. This dispersion can be driven by two margins: (i) the extensive margin, whereby civil servants are absent altogether and thus have 0 hours of work recorded, and (ii) the intensive margin, whereby civil servants attend work but exhibit variation in the overall duration at work on a given day.

Extensive margin. By our measure, almost a third of the civil servants (32.3%) are not clocking in at work on a given day. While this number excludes weekends and national and state holidays, it does not necessarily mean that these officers are absent from work – we observe their clocking-in and out behavior, but cannot observe whether they are working: as discussed in previous section, it is, for example, possible that a proportion of these officers have business outside of the state office, work at home or part-time or are absent due to sickness or vacation. Absent additional data, we are unable to specify the exact reason of absence.

Intensive margin. Even for officers who do clock in at work, there is variation in the total hours worked on a given day. Figure 1 shows the distribution of hours present for the 7,753 public sector employees, covering a total of the 3.7 million employee-work days. As the figure shows, the average worker is present for about 7.57 hours per day. There exists however substantial variation in the hours present: 5% of the labor days are shorter than 6.23 hours, and 5% of the labor days are longer than 9.46 hours. The interquartile range (measured as the difference in hours between the employee at the 75% percentile and the 25% percentile is 1 hour).

Figure 1. Hours present per day – conditional on attending work



To better understand the drivers of differential attendance, we decompose the overall variation in attendance into three factors: (i) Differences in attendance across employees (ii) Differences in attendance across departments and (iii) Differences in attendance across time.

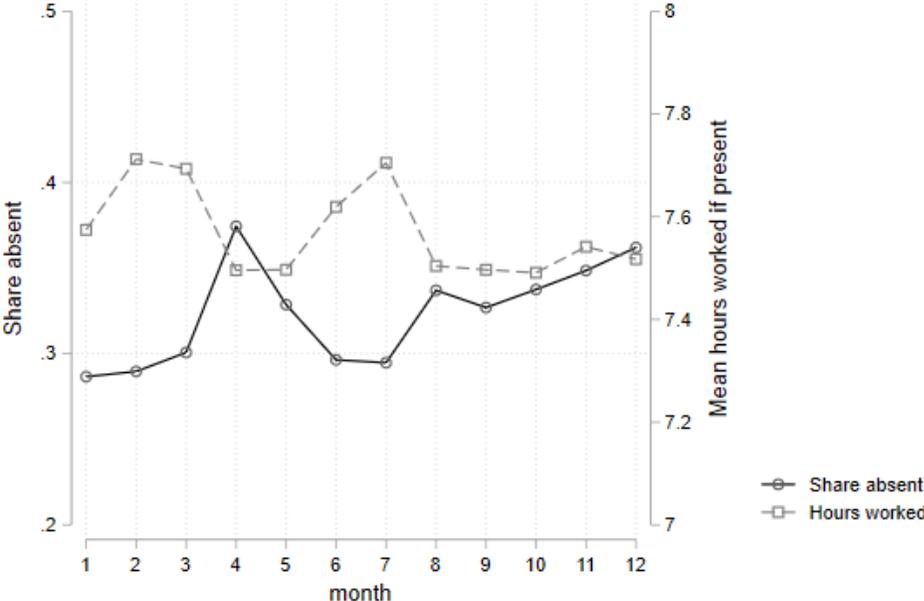
2.1. Differences in attendance across time

Is the variation in attendance rates primarily driven by differences in attendance hours across different days of the week, months, or years? For example, attendance hours may fluctuate depending on the seasons (e.g. summer vs. winter), or with holidays, in which case the observed differences in hours attended could reflect temporal differences.

By week day. Within a week, absenteeism is lowest on Tuesday (30%) and highest on Friday (33%). These are also the days when those present work the longest hours. For those who clock in on Tuesday, the average hours worked are 7.6. In contrast, those who clock in on Friday only work 7.5 hours on average.

By month. We observe seasonality in the attendance patterns. Figure 2 shows the absenteeism rate by month. Absenteeism is highest in April (37.4%) and lowest in January (28.6%). Conditional on attendance, there is variation in hours worked by month: the months with the highest average hours worked are February-March (7.71 and 7.69 hours) followed by July (7.61 hours). The lowest hours are clocked in December (7.51 hours).

Figure 2. Seasonality in absenteeism and hours worked by month



2.2 Differences in attendance across departments

In addition to the temporal variation, is the cross-sectional variation in attendance rates and hours worked primarily driven by differences in attendance hours *across* departments?

For example, employees in different departments (such as revenue department vs. health department) could have different work schedules or workplace culture, in which case the observed differences in hours attended could reflect department-specific (or organization-specific) differences.

As Figures 3a and 3b show, there exists substantial variation both in the absenteeism rate and in the average hours worked across departments.

Figure 3a. Attendance in office by department

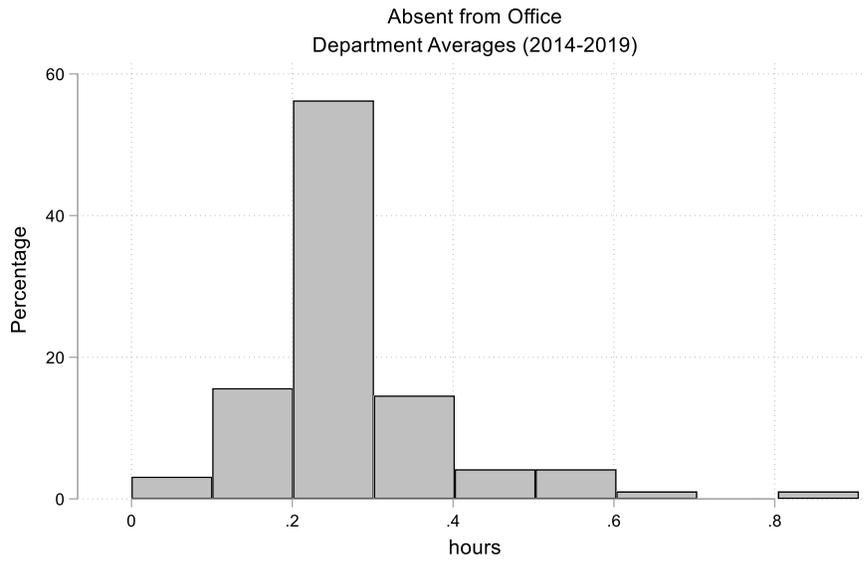
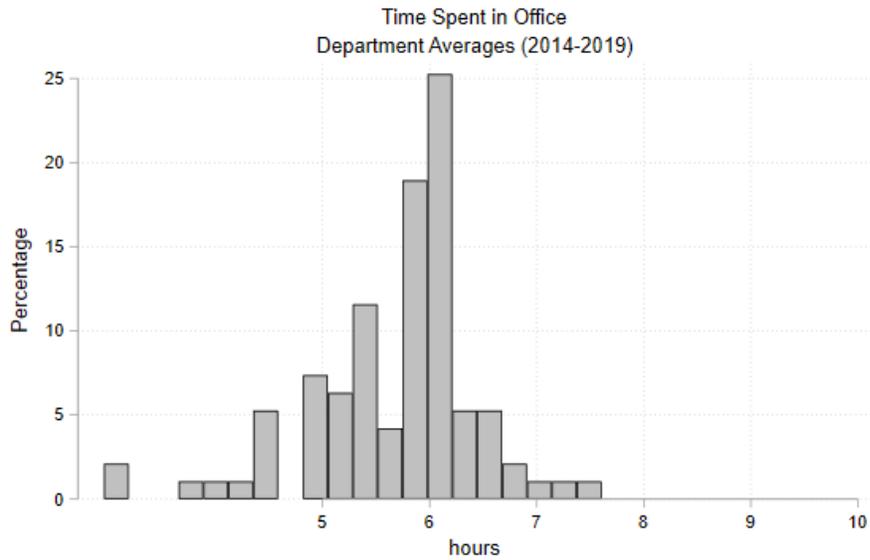


Figure 3b. Average time spent in office by department



2.2 Differences in attendance across employees

Is the variation in attendance rates primarily driven by differences in attendance hours *across* employees? For example, employees may have different contracts (such as part-time vs. full-time contracts) or different work ethics (such as staying late in office vs. leaving early types), in which case the observed differences in hours attended could reflect employee-specific differences.

Interestingly, variation in the extensive margin – that is the propensity of an employee to clock in on a given day (Figure 4a) – is much larger than the variation in the intensive margin – the time spent in office conditional on being present (Figure 4b).

Figure 4a. Attendance in office by individual

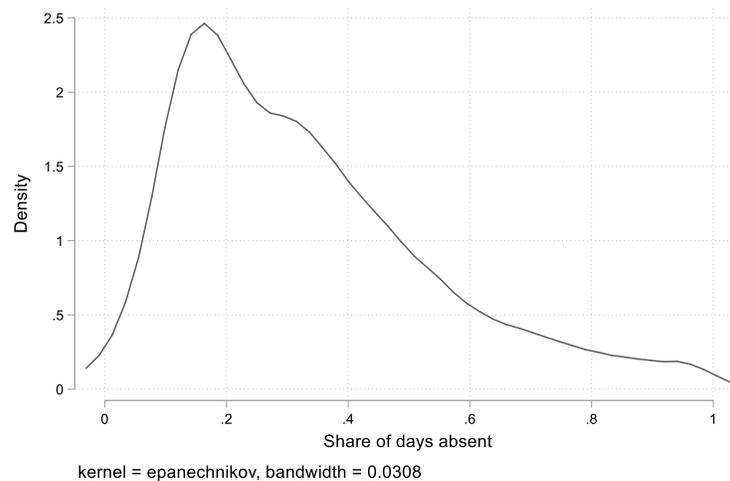
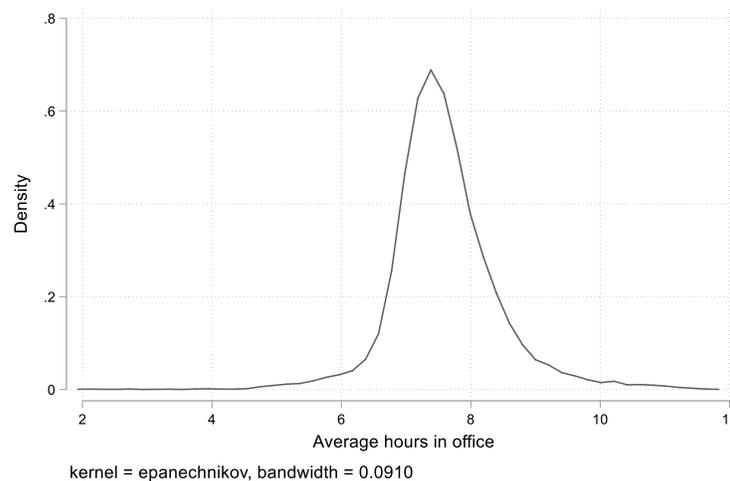


Figure 4b. Average time spent in office by individual



3. DECOMPOSING THE OVERALL VARIATION

To quantify the contribution of each of these factors, we use the fact that we observe the same individuals over time and across different departments. This allows us to statistically decompose the overall attendance pattern.³ The results are summarized in Table 1. We report three different measures of attendance: (i) clocking in on a given day (i.e. not being absent), (ii) the hours spent in the office for those who clocked in (i.e. conditional on not being absent), and (iii) the unconditional hours spent in the office, where those who are absent are imputed a 0 hours attended in the office. Once again, it is important to stress that not being present in the office does not necessarily mean not working.

Among the three factors, differences across employees contribute by far most to explaining variation in attendance hours. Up to 42% of the overall dispersion in attendance (as seen in Figure 1) can be explained purely by differences in attendance across individuals. In contrast, differences in average attendance are much smaller across departments and appear to depend less on seasonality. Differences across departments explain at most 6.3% of the variation in overall attendance; similarly, differences across workdays explain at most 4.3% of the variation.

When combining all explanatory factors – individual, departmental and time – we can account for almost half (46%) of the overall variation in attendance.

Table 1. Share of variation explained by individuals, organizations and seasonality (time)

Share of attendance variation explained (i.e. clocking in)	3.8%	3.5%	16.6%
Share of hourly variation explained (conditional on clocking in)	4.3%	6.3%	42.4%
Share of hourly variation explained (unconditional)	4.2%	3.9%	18.6%
Explanatory factor	Time (Date)	Department	Individual

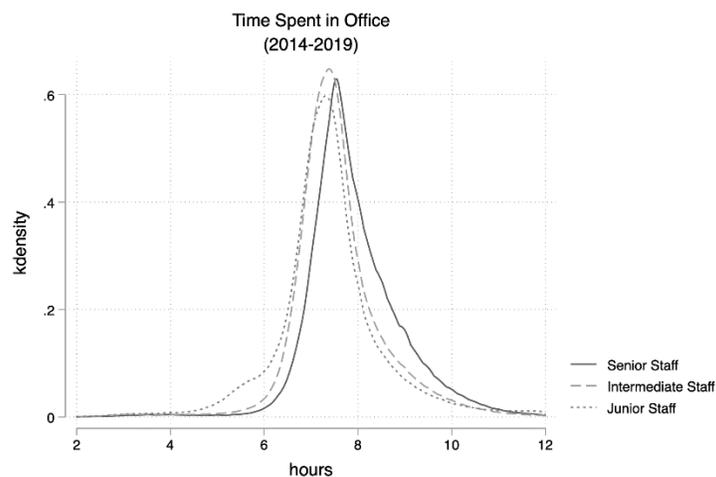
³ Specifically, the method we employ is a fixed effects decomposition where we regress attendance hours on individual, department and day-month-year fixed effects. We then examine the R^2 to assess the extent to which each factor contributes to (descriptively) explaining the variation in attendance.

4. ROLE OF LEADERS

An interesting fact that emerged from the analysis is the importance of individual-level variation in attendance. What drives this variation? Unfortunately, the attendance data itself does not provide many background characteristics that allow us to correlate the observed attendance patterns with employee characteristics. We did, however, construct a measure of seniority based on the job titles all employees. We do so by dividing the employees in every department into seniors (e.g. head of department), intermediate and juniors.

We can thus ask whether attendance patterns vary by seniority. Interestingly, we find stark differences by seniority. On the extensive margin of attending work, we find that seniors are 21.5% points more likely to clock in at the workplace on a given day. With the mean absenteeism rate at 32%, this is a sizeable difference that explains a large part of the variation in attendance. The absenteeism rate on a given day is only 2.9% for seniors. The greater propensity to attend for seniors also shows up in the hours on a given day. In Figure 5, we plot the distribution of hours spent in the office for those who clocked in on a given day, broken down by seniority. As the figure shows, there is a successive shift in the distribution towards higher hours as we move from junior staff to intermediate staff to senior staff: senior staff members spent, on average, more hours in the office than junior staff.

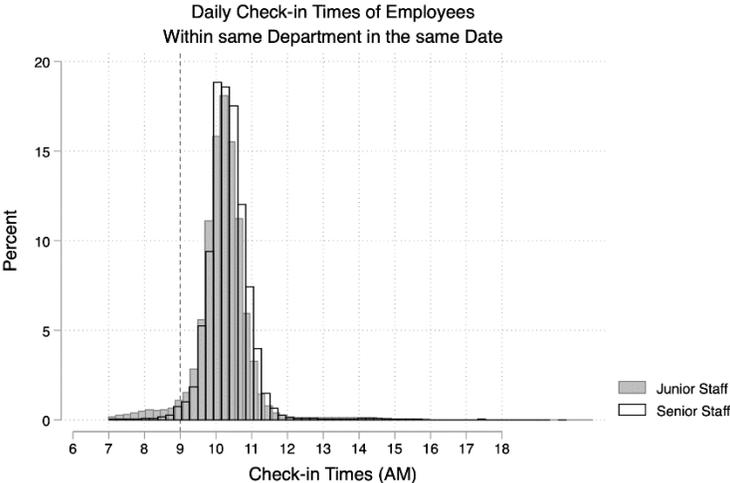
Figure 5. Time spent in office by seniority



The fact that seniors are more present than juniors is surprising. There are multiple interpretations. First of all, it is possible that high attendance employees are more likely to be promoted to the higher ranks. Second, it is likely that more senior roles are more demanding, thus requiring employees to work longer hours. To investigate this, we compared the attendance patterns of employees who were promoted to senior positions. Interestingly, we find that attendance patterns change significantly post-promotion, even when we compare the same employee over time: the same employee who is promoted to senior level is 8% more likely to be present on a given day, stays – conditional on clocking in – on average 9 minutes longer per day.⁴ When counting absent days as 0 hours worked, the difference in average hours present per day increases to a staggering 43 minutes.

We zoom in deeper to understand the mechanisms underlying the difference in attendance between senior and junior employees. To do so, Figures 6a and 6b show the distribution of times when employees clock in and clock out. In terms of clocking in (Figure 6a), most employees clock in shortly after 10am. The median junior officer clocks in 6 minutes earlier (10.13am) than the median senior officer (10.19am).

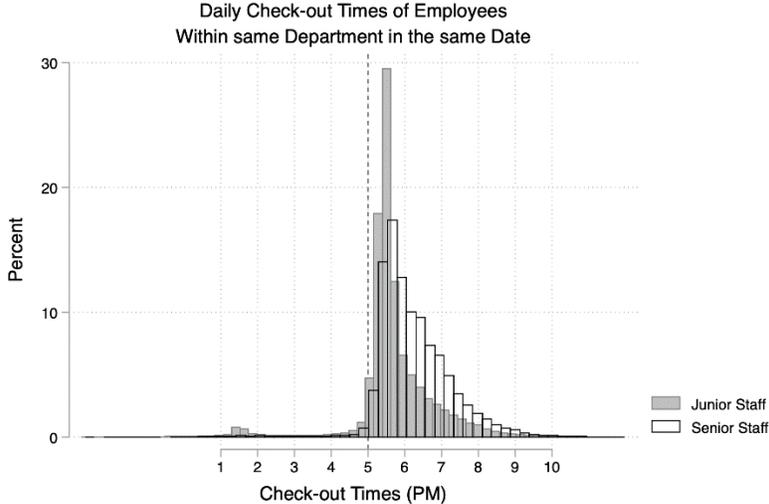
Figure 6a. Distribution of clock-in times by junior and senior staff



⁴ Formally, we regress attendance on a dummy that is 1 if the employee is a senior employee, conditional on employee, department and time fixed effects.

The more interesting pattern concerns the clock-out times (Figure 6b). Almost no employees clock out before 5pm. 34% of the officers clock out between 5pm and 5.30pm. The large increase in clocking out shortly after 5pm, however, is most pronounced for the junior officers. Senior staff, in contrast, tend to clock out significantly later. While 90% of junior officers have already clocked out by 6pm, the share for the senior officers is 80%.

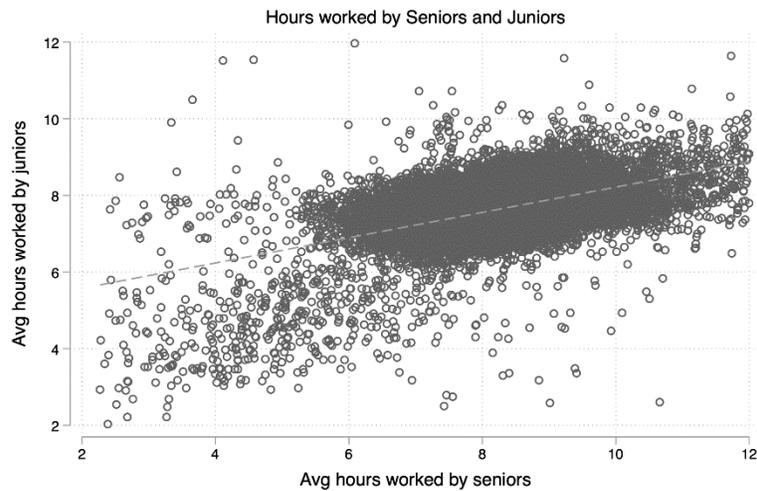
Figure 6b. Distribution of clock-out times by junior and senior staff



Finally, we document a positive correlation between the attendance hours of a senior officer and the attendance of junior officers in the same department (Figure 7). Departments which have senior officers staying longer also tend to have junior officers working longer hours. This pattern is interesting as it could be indicative of a leadership effect, whereby the attendance pattern of the senior employees shape the attendance patterns of their subordinates. These correlations, though suggestive, however should not be viewed as causal.⁵

⁵ There are alternative explanations for this correlation: for example, the association between senior and junior hours could be shaped by common shocks (such as busy workload that affect everyone in the same department).

Figure 7. Relationship between hours worked of seniors vs. juniors



5. CONCLUSIONS

Using large-scale administrative data, we studied the attendance patterns of public sector employees in the state government offices of Bangalore.

The perhaps most interesting finding is that up to 42% of the variation in attendance can be attributed to differences *across* individuals. To the extent that these cross-individual differences reflect differences in work styles, the intriguing question arises about which individual characteristics can predict high vs. low attendance. Indeed, the question on how to attract and select the most motivated workers into the civil service is a question that has received great attention on the academic literature.

Among the characteristics we studied, we found that seniority is a key variable when it comes to explaining variation in individual-level attendance. Senior officers are significantly less likely to be absent, and – conditional on being present in the office – stay significantly longer. Finally, we document a positive relationship between the attendance of senior officers and the attendance of their subordinates.

The study, while revealing novel patterns of attendance using micro-level data, also opens up questions that follow-up studies may be able to address. The current data, for example, does not

shed light on what the contractual working hours are. Such data would be helpful as it would allow us to compute the deviation between the contractual working hours and the actual working hours. This would provide a closer estimate on how many work-days are “lost” due to absenteeism.

It would also be interesting to study the relationship between attendance and productivity: to the extent that high attendance reflects greater productivity and performance, thinking about ways to reduce absenteeism – both in the extensive and intensive margin – would be natural ways to improve public sector effectiveness. In contrast, if longer hours present merely does not translate into productivity, considering policies such as more flexible work schedules may prove effective in increasing performance. Individual-level data on productivity and well-being could help answer these questions.

Finally, the analysis so far is restricted to the departments at the Government Secretariat at Bangalore. This is likely to comprise a set of particularly well managed departments. In order to document attendance patterns more broadly, both beyond the state capital and Karnataka, we require attendance data from a larger number of departmental units that is managed by NIC. The data would be important to benchmark the attendance pattern of the Karnataka civil service against other state civil services, or compare the attendance patterns across different districts of Karnataka.

More broadly, the attendance data provides a useful baseline against which potential policy interventions – e.g. a pilot on more flexible working hours, greater monitoring – could be assessed. As researchers, we are very open to continuing this conversation and assisting the Government of Karnataka with subsequent studies.