Submitted to *Econometrica*

MANAGERS AND PRODUCTIVITY IN THE PUBLIC SECTOR

ALESSANDRA FENIZIA
Department of Economics, The George Washington University

This paper studies the impacts of managers in the administrative public sector using novel Italian administrative data containing an output-based measure of productivity. Exploiting the rotation of managers across sites, I find that a one standard deviation increase in managerial talent raises office productivity by 10%. These gains are driven primarily by the exit of older workers who retire when more productive managers take over. I use these estimates to evaluate the optimal allocation of managers to offices. I find that assigning better managers to the largest and most productive offices would increase output by at least 6.9%.

KEYWORDS: Productivity, Managers, Public Sector.

1. INTRODUCTION

The public sector represents a large share of modern economies and is not disciplined by economic forces of competition. Public sector managers are the cornerstone of modern bureaucracies. They oversee day-to-day operations of complex public organizations and supervise policy implementation. As such, manager effectiveness may have important consequences for the performance of government agencies and ultimately citizens' welfare. However, we know little about the extent to which differences in manager quality ultimately impact public service provision.

On the one hand, managers may not be able to affect the performance of their organizations because they lack many of the tools available to private sector firms (e.g., firing and promotions). In most countries, public sector workers enjoy strong job security and receive promotions and pay raises that depend on seniority rather than individual performance. On the other hand, public sector managers may play an important role precisely because of the lack of other tools to motivate their workers.

One reason why little is known about the effectiveness of public sector managers is that it is notoriously hard to measure the performance of government agencies. A set of recent studies has made progress by measuring managerial practices, qualitative policies and procedures that are thought to be associated with well-run organizations, and has established that these measures correlate positively with public service delivery (Tsai et al., 2015, Bloom et al., 2015, Rasul and Rogger, 2018). Yet, it is unclear how to translate these correlations into quantitative measures of the causal impact of managers.

This paper studies the productivity impacts of managers in the public sector using novel data from the Italian Social Security Agency (*Istituto Nazionale di Previdenza Sociale* —INPS, hereafter). INPS administers applications for unemployment insurance, disability insurance,

Alessandra Fenizia: afenizia@gwu.edu

I thank David Card, Hilary Hoynes, and Patrick Kline for their invaluable support and guidance. I also thank two anonymous referees for the thoughtful comments that greatly improved my paper. I am grateful to Eric Avis, Tito Boeri, Giulia Bovini, Luca Citino, Ernesto Dal Bó, Fred Finan, Michela Giorcelli, Zarek Brot-Goldberg, Daniel Haanwinckel, Erin Kelly, Nicholas Li, John Loeser, Juliana Londoño-Vélez, Isabela Manelici, Matteo Paradisi, Francisco Perez-Gonzalez, Daniel Rogger, Evan Rose, Raffaele Saggio, Yotam Shem-Tov, David Silver, Bryan Stuart, Jose Vasquez, Mauricio Ulate, Christopher Walters, Guo Xu, and Danny Yagan for helpful suggestions. My special thanks goes to all the members of the administrative and technical staff of the INPS who provided invaluable support and help. The findings and conclusions expressed are solely those of the author and do not represent the views of INPS.

pensions, subsidies to the poor, and other welfare and insurance programs. A key innovation of the paper is measurement. First, INPS uses an objective output-based measure of productivity constructed using detailed administrative data on both output—measured by a (complexity-weighted) standardized index of claims processed by the office—and on full-time equivalent workers assigned to the office. Second, INPS monitors service quality, reflected in claim processing time and error rates. I assess whether changes in productivity come at the cost of quality. This is an ideal setting to isolate the contribution of managers to office performance because all sites are subject to the same rules, workers produce homogeneous products, and there are virtually no differences in physical capital across offices.

I begin my analysis by documenting large variation in productivity across offices, which is comparable to the within-industry plant-level distribution of productivity documented by Syverson (2004). I then use a two-way fixed effects model to decompose log productivity into the components due to office characteristics, manager effects, and time effects. A simple model with additive office and manager components may raise two concerns. First, managers could be assigned to offices on the basis of unobserved factors that determine their comparative advantage. I test for match-driven sorting and find no evidence of comparative advantage-based mobility. Second, manager rotation might be correlated with office-specific trends. I find no evidence of sorting based on trends.

Using bias-corrected measures of the variance components, I find that manager fixed effects explain 9% of the total variation in productivity at the office level—about one third as much as the permanent component of productivity associated with different offices. Overall, a one standard deviation better manager increases office productivity by 10%. I also find that the (bias-corrected) covariance between manager and office fixed effects is negative, suggesting the best managers work at the least productive sites.

In the second part of the paper, I study the mechanisms through which better managers achieve higher productivity. Previous research finds that effective private sector managers increase productivity by making better personnel (Hoffman and Tadelis, 2018) and investment decisions (Bennedsen et al., 2010, 2011). Public sector managers do not have the tools of their private sector counterparts, so increasing productivity may be particularly challenging. Instead, good managers may be better at matching workers with tasks, may use indirect strategies to encourage worker effort, or may make unproductive workers quit or retire. I find that the productivity gains are driven primarily by the exit of older workers who retire when a productive manager takes charge. These findings are consistent with anecdotes suggesting that senior workers leave when better managers reassign more prestigious and better compensated tasks to junior employees. Despite the decrease in labor input, productive managers are not associated with a decrease in output or increase in overtime hours. Importantly, higher output per worker does not come at the cost of lower quality.

In the final section of the paper, I evaluate the efficiency gains from reassigning managers. The estimates from my productivity model imply that an optimal social allocation would assign the best managers to the largest and most productive offices. I find that if managers were reassigned optimally, agency output would increase by at least 6.9% – more than twice as much as firing the worse 20% of managers.

This paper contributes to three strands of the literature. The first strand documents the impact of (1) managerial practices associated with better organizations and (2) managers themselves and their decisions. Both practices and managers matter for private sector firms (Bertrand and Schoar, 2003, Bloom and Van Reenen, 2007, Bloom et al., 2013, Lazear et al., 2015, Giorcelli, 2019, Bandiera et al., 2020, Baltrunaite et al., 2020). In the public sector, studies have found that better organizational practices correlate with better outcomes in hospitals (Tsai et al., 2015), schools (Bloom et al., 2015), and civil service organizations (Rasul and Rogger, 2018, Rasul et al., 2019). But, it is not a given that managers' effectiveness in private sector firms would necessarily carry over to the public sector. Most government operations are very bureaucratic, fulfilling broad mandates with no semblance of competition or prices for individual services. My paper is the first to document managers' effectiveness within public sector bureaucracy and study *how* they improve office performance.

Second, my work relates to the literature that studies the impact of civil servants on organizational culture and the performance of public sector institutions (Finan et al., 2017, Xu, 2018, Bertrand et al., 2019, Best et al., 2019, Choudhury et al., 2019, Khan et al., 2019, Janke et al., 2019, Azulai et al., 2020). The primary challenge in this literature is credibly measuring output. For instance, two related papers use different methods and arrive at different conclusions. Best et al. (2019) study public procurement. They conclude that bureaucrats vary in their ability to procure low cost goods, adjusting for observable quality differences using a machine learning classifier. Janke et al. (2019) study large public sector hospitals and conclude that CEOs are not able to improve multifaceted hospital output. These sophisticated approaches highlight the measurement challenges, especially because the economic agents multitask over the production of many heterogeneous outputs. My setting provides several important advantages. First, I construct a comprehensive output-based measure of office productivity that is not subject to the concerns relative to multitasking that characterize previous studies. Second, I assess the productivity-quality trade-off using a measure of quality that is unavailable to the existing literature. Third, I use rich administrative data to recover the channels through which managers improve productivity and examine what makes for a productive manager.

Finally, this paper also fits in the literature on productivity differentials between workplaces (Syverson, 2004, 2011, Chandra et al., 2016). My paper contributes to this literature by providing compelling evidence that large and persistent differences across units are not limited to the private sector and that they arise even *within* a centralized public agency.

2. INSTITUTIONAL BACKGROUND

The *Istituto Nazionale di Previdenza Sociale* (INPS) employs 30,000 workers and administers applications for virtually all social welfare and insurance programs in the country including unemployment insurance, disability insurance, social security transfers, maternity leave, and subsidies to the poor. Even though INPS is a large, centralized government agency, claim processing is decentralized. Every office has a catchment area and processes all claims that originate from it. I study the 111 main satellite offices and 383 local branches that review and process routine claims. Within each office, a single manager oversees production workers who assess whether to accept or reject claims.

Anecdotally, more productive managers make their mark: they reassign workloads and responsibilities, change workplace practices, enforce break times, and closely oversee employees. However, managers cannot make payroll decisions to improve productivity. Firing is almost non-existent in Italian government positions, and a hiring freeze was instituted in 2008, covering the period of my analysis. Managers have to make the best out of their assigned set of workers.

Table I presents summary statistics of managers' characteristics. The first column includes all managers observed in my sample, while column 2 presents the characteristics of managers who are observed in at least two different offices. In the full sample, the average manager is 55years old and has 27 years of civil service experience. Close to 60% were born in Southern Italy or the Islands, potentially reflecting the relative attractiveness of civil service jobs in poorer regions. About one-third have a university degree in Law, and another 13% have a degree in Business, Administration, or Economics. Still, over 20% have no university-level education.

The managers who move across offices are younger and are more likely to be male and college-educated.

INPS posts manager vacancies and their corresponding eligibility criteria on an internal website that is visible to all employees. There are no official rules or unofficial guidelines on how to choose among qualified candidates. Human resources officers select managers on a case-by-case basis. Managers stationed in main offices rotate every five years as part of anti-corruption law *190-2012*, which prevents managers from becoming too entrenched and susceptible to corruption. Relatively few vacancies are open in any given year because of staggered tenures, limiting the extent to which managers can sort. However, this law does not apply to managers serving in local branches. As such, one may be concerned that managers may switch due to both plausibly exogenous reasons (e.g., retirement) and potentially endogenous choices (e.g., work closer to home). Nonetheless, the limited pool of candidates eligible to fill these positions, the lack of guidelines, and the many constraints related to the manager rotation limit the ability of managers to sort into offices. I further corroborate this argument by testing for endogenous mobility in Section 4.

TABLE I

MANAGER CHARACTERISTICS

	(1) Full Sample	(2) Movers
Demographic Characteristics		
Male	0.63	0.70
Age	54.98	52.96
Experience Publ. Sec.	27.43	24.60
Region of Birth		
Nort-East	0.13	0.09
North-West	0.11	0.14
Center	0.16	0.12
South or Islands	0.59	0.65
Abroad	0.01	0.00
Highest Educational Attainment		
High-School Diploma	0.26	0.10
Econ, Business, and Admin	0.13	0.11
Sci, Engen, Math, and Stat	0.04	0.06
Social Sciences and Humanities	0.20	0.21
Law	0.30	0.42
Missing Educ	0.07	0.10
N	851	207

Note: The table reports the summary statistics of manager characteristics. The statistics are computed over the full sample of managers in column (1) and over the subsample of movers in column (2). Movers are the managers that oversee at least two offices over my sample period. Experience in the public sector is defined as the number of years since the manager was first hired in any public sector institution.

To some approximation INPS maximizes output subject to quality constraints. Workers and managers' salaries have a fixed component and a bonus. The former is tied to job title, and the latter depends on both the level and the year-over-year change of productivity and quality of service. (See Online Appendix C for details). On average, managers earn more than the median Italian household income. While bonuses represent a small share of overall employee compensation, they amount to 15–30% of managers' salary.

3. DATA

This section details the quarterly office level data used in my analysis. These data are comprised of two main elements: data on office-level inputs and output and a personnel file that allows me to observe individual worker assignments to offices.

3.1. *Office-Level Productivity Measures*

I use INPS data from the internal monitoring system for Q1 2011 to Q2 2017. These data report inputs including the number of full-time equivalent employment (FTE_{it}) at office i in quarter t as well as information on absences, overtime hours, and hours devoted to training by workers in each office. For output, I use INPS's standardized index of claims Y_{it}, a sum of claims processed c_{vit} weighted by their complexity w_{vt},

$$Y_{it} = \sum_{v} c_{vit} \times w_{vt}.$$

The claim categories are very fine—there are more than 1,000 types. The weights reflect the time employees *should* take to process each type of claim (see Online Appendix D for details on how the weights are measured). Importantly, objective complexity weights control for differences in tasks across offices.[1] Although INPS employees' main task consists in processing paperwork, they also take turns working at the front-office where they assist beneficiaries.[2] I combine the measures of office output and FTE employment to construct my measure of productivity (P_{it}) as output per worker:

$$P_{it} = \frac{Y_{it}}{FTE_{it} \times 3} = \frac{\sum_{v=1}^{V} c_{vit} \times w_{vt}}{FTE_{it} \times 3}.$$

One advantage of my setting is that workers' tasks are well-defined, and INPS devotes significant effort into measuring each individual task. So, contrasting with other studies of productivity, multitasking (Holmstrom and Milgrom, 1991) is much less of a concern in my setting.

The data also allow me to test whether more productive managers decrease quality. In particular, INPS constructs an index of service quality, a weighted average of "timeliness" (the fraction of claims processed within the first thirty days) and the "error rate" (the fraction of claims that has to be processed more than once because of an error in initial processing).[3]

3.2. *Office-Employee Data*

I track INPS employees over time using the personnel file (2005-2017). This dataset includes office location, job title, hiring, firing, separations, and promotions.

Anecdotally, most employees are hired through a competitive examination or transfer from other government agencies. Workers rarely quit a public sector job, and the vast majority leave INPS when they retire. Because I do not observe retirement directly, I proxy for retirement using voluntary separations of workers over age 60.

---

[1] If weights do not correctly reflect task complexity, one may worry that managers shift production toward overvalued claims. I address these concerns in Section 6.

[2] My results are robust to the exclusion of front-office operations (see Online Appendix E).

[3] INPS audits 5% of each office production twice per year, and many errors are detected during these audits (Online Appendix D).

3.3. *Descriptive Statistics and Stylized Facts*

Table II reports the summary statistics for the full sample in column 1; columns 2 and 3 display the statistics for main offices and local branches, respectively. Main offices are substantially larger than local branches. A main office employs on average 115 workers. Local branches have only 16 employees on average. Naturally, larger offices produce more output because labor is the main input of the production process. Offices have a large backlog, which amounts to 80% of the average quarterly inflow of new claims. Overall, training and overtime work is a very small fraction of employees' time. Hiring is extremely limited. Only 0.5 workers per office separate from INPS every quarter (48% are due to retirement), and 0.3 workers transfer to another office.

Comparing columns 2 and 3, main offices are 12% more productive than local branches on average. However, quality and absenteeism do not seem to differ substantially across the two types of production sites.

TABLE II

CHARACTERISTICS OF SOCIAL SECURITY OFFICES

	(1) Full Sample	(2) Main Offices	(3) Local Branches
Productivity	94.56	103.65	91.72
Output (Thousands)	10.24	29.18	4.33
FTE	39.95	115.39	16.41
Hours	31.66	91.76	12.91
Training	0.62	1.73	0.28
Overtime	0.70	2.10	0.26
Absenteeism Rate	0.21	0.21	0.21
Quality	100.52	101.31	100.27
Backlog (Thousands)	54.24	197.68	9.48
Demand (Thousands)	68.02	220.55	20.42
Hires	0.05	0.13	0.02
Separations	0.50	1.53	0.17
Fires	0.00	0.01	0.00
Inbound Transfers	0.72	2.08	0.29
Outbound Transfers	0.30	0.63	0.20
Retirement	0.24	0.72	0.09
N	13212	3142	10070
Number of Managers	851	221	638
Number of Offices	494	111	383

Note: The table reports the summary statistics for social security offices. All statistics are calculated across office-quarter observations. The statistics are computed over the full sample of offices in column (1), and over the subsample of main offices and local branches in column (2) and (3), respectively. For all rows except divorce and blood donations, there are 10943, 2658, and 8285 observations, for the full sample, restricting to main offices, and restricting to local branches, respectively. For divorce, these statistics are 11052, 2622, and 8430, and for blood donations they are 11104, 2648, and 8456, respectively. Output, demand, and backlog are measured in thousands of hours, while FTE, training, hours, and overtime are measured in full-time equivalent units.

Although comparing productivity differentials across industries is notoriously hard, I benchmark my estimates with previous studies. Appendix Table A.I compares the distribution of log productivity in my sample (Panel A) with the within-industry plant-level distribution moments in Syverson (2004) (Panel B). There might be reasons for believing the dispersion in productivity across offices that belong to the same centralized agency is substantially smaller than the one across plants within the same industry. Yet, my estimates are somewhat smaller, but comparable, to those in Syverson (2004).

Nevertheless, a concern is that dispersion in productivity may be driven by the fact that when demand is low workers are idle. This is not the case. First, offices have a large backlog (Table II). Second, if offices face low demand, they are supposed to process claims of high-demand offices (see Online Appendix D for details).

4. DO MANAGERS MATTER IN THE PUBLIC SECTOR?

Here, I develop a framework which exploits manager rotation across sites to decompose productivity into a manager and an office component. I discuss the identification challenges and perform a series of diagnostic checks that evaluate the model specification. I conclude this section by summarizing the implications of this model in a variance decomposition exercise.

4.1. *Identification*

I model log productivity ($\ln P_{it}$) as the sum of a manager component ($\theta_{m(i,t)}$), a permanent office component (α_i), a time effect (τ_t), and an error term (u_{it}):[4]

$$\ln P_{it} = \alpha_i + \tau_t + \theta_{m(i,t)} + u_{it}. \tag{1}$$

I interpret $\theta_{m(i,t)}$ as the portable component of managers' ability, which I interchangeably refer to as "manager quality" or "managerial talent." The office effects (α_i) proxy for the time-invariant characteristics of the office (e.g., geographical location, average quality of the workers, main office vs local branch) and for size/composition of the workforce to the extent that these variables do not change over time. I include time fixed effects τ_t to absorb seasonality and macroeconomic shocks.

Model (1) postulates that productivity changes discretely as a new manager takes over. However, in reality, managers may take some time to change work practices. I therefore estimate (1) excluding the first quarter in which the new manager is in charge. I re-write (1) in matrix notation as $\ln(P) = D\alpha + G\theta + T\tau + u$, where D, G, and T collect all the office, manager, and time dummies, respectively. OLS identifies the parameters of interest under the following identifying assumptions:

$$E[d_i' u] = 0 \ \forall i \qquad E[g_m' u] = 0 \ \forall m, \tag{2}$$

where d_i is the i-th row of the matrix D and g_m is the m-th row of the matrix G.

Identification requires that manager mobility is as-good-as random, conditional on office and time fixed effects. These orthogonality conditions permit managers to be assigned to offices on the basis of the permanent component of office productivity (α_i) or the permanent component of managerial ability ($\theta_{m(i,t)}$). That is, better managers sorting into more productive offices *would not* violate the identifying assumptions. By the same token, if productive managers were systematically sent to local branches or to a specific geographical area, this *would not* represent a threat to the identification strategy. I follow Card et al. (2013) and consider three forms of endogenous mobility. Let the error term take the form:

$$u_{it} = \eta_{im(i,t)} + \zeta_{it} + \epsilon_{it}, \tag{3}$$

[4]Online Appendix B derives this estimating equation from claim-specific production functions.

where $\eta_{im(i,t)}$ reflects a office-manager match component of productivity, ζ_{it} are office-specific trends in productivity, and ϵ_{it} are transitory shocks to office productivity. I assume that $\eta_{im(i,t)}$ has mean zero for all offices and all managers in the sample. I also assume the drift component ζ_{it} is mean zero for each office but contains a unit root. Finally, I assume that ϵ_{it} has mean zero for each office.

Each term reflects a type of sorting that would violate the exogenous mobility identifying assumption: (1) if managers were to sort into offices on the basis of their comparative advantage; (2) if better managers were to be systematically sent to offices whose performance is worsening over time; (3) or if INPS were to reallocate managers to offices in reaction to a negative transitory shock. I tests for the presence of endogenous mobility in Subsection 4.3.

4.2. *Results*

Table III describes the structure of my sample. Column 1 reports the statistics for the full sample. Column 2 restricts attention to the balanced-analysis sample. Here, I require outgoing managers to be observed for at least four quarters before and incoming managers to be observed for at least six quarters after a change in leadership. The full sample contains 851 managers, 494 offices, and 276 connected sets (Table III, column 1). Roughly one-fourth of these managers move across sites and almost 80% of offices experience a change in management between 2011 and 2017 (column 1). The remaining 20% of the offices (employing 10% of the managers) do not contribute to the estimation of the manager effects. All offices experience a change in leadership in the balanced-analysis sample by construction, and 30% of managers move across sites (column 2). To assess the amount of dispersion in public sector productivity attributed

TABLE III

SAMPLE CHARACTERISTICS

	(1) Full Sample	(2) Balanced-Analysis Sample
# Managers	851	601
# Offices	494	282
# Managers >1 Office over the Sample Period	207	184
# Offices >1 Manager over the Sample Period	404	282
# Connected Sets	276	143
# Events	635	318
# Events in Main Offices	226	80
# Events in Local Branches	409	238
N	13,212	8,165

Note: The table reports the sample characteristics for the full sample of offices in column (1) and for the balanced-analysis sample in column (2). The latter includes the subset of offices for which I observe the outgoing manager being in charge for at least four quarters before the change in leadership and the incoming manager being assigned to the office for at least six quarters after that. "# Managers >1 Office over the Sample Period" represents the number of managers who serve in at least two sites over my sample period. "# Offices >1 Manager over the Sample Period" represents the number of offices that experience at least one change in leadership over my sample period. Each office has only one manager at each point in time but may have multiple managers over my sample period. Events are defined as changes in leadership. "N" represents the number of office-quarter observations.

to managers, I follow Bertrand and Schoar (2003). Specifically, I compare the adjusted R^2 estimated from a regression of the logarithm of productivity on office and time fixed effects to the full model (1) with manager fixed effects.[5] Columns 2 and 3 of Table IV report the model

[5]The models coincide when managers' effects are all identically equal to zero.

estimates excluding and including manager effects, respectively. The adjusted R^2 increases from 0.69 in column 2 to 0.76 in column 3, suggesting that managers explain a non-trivial amount of the variation in productivity across sites. Although the increase in the adjusted R^2 might seem small, its magnitude is very similar to that reported in Bertrand and Schoar (2003). More formally, an F-test strongly rejects the null hypothesis that all the manager effects are zero (p-value=0.000).

TABLE IV

ANALYSIS OF VARIANCE OF OFFICE PRODUCTIVITY

	(1) Ln(P)	(2) Ln(P)	(3) Ln(P)	(4) Ln(P)	(5) Ln(P)
N	3,316	3,316	3,316	3,316	3,316
R sq.	0.325	0.727	0.835	0.789	0.839
Adj. R sq.	0.324	0.679	0.762	0.720	0.765
Time FE	Yes	Yes	Yes	Yes	Yes
Office FE	No	Yes	Yes	No	No
Manager FE	No	No	Yes	Yes	No
Manag-by-Office FE	No	No	No	No	Yes
Pvalue			0.000	0.000	

Note: This table investigates how much of the variance in log productivity is explained by the office, manager, and time components in the full sample. N represents the number of office-year observations. The p-value at the bottom of the table tests the null hypothesis that manager effects are jointly zero.

4.3. *Diagnostic Checks*

One might be wary of endogenous mobility related to office-specific trends in productivity. As a concrete example, if good managers were able to systematically move to offices which are improving over time, my model would overestimate their managerial quality. I investigate this concern in Table V by evaluating the correlation between baseline office characteristics and estimated fixed effects of future managers. The results in column 1 of Table V show that there is no evidence of managers sorting on trends in productivity, which would bias my results.[6] I repeat the same exercise using the change in the estimated fixed effects as the dependent variable (column 2 of Table V). The overall pattern of results is largely unchanged.

Another way to test whether the sorting of managers to offices is driven by serially correlated error components in office or manager productivity is to examine the residuals from (1) associated with specific forms of manager changes. Figure 1 plots average (trend-adjusted) office productivity relative to quarters where the office experiences a change in management, classifying offices by terciles of the change in managerial quality $\widehat{\Delta M}_i = \hat{\theta}_{i,incoming} - \hat{\theta}_{i,outgoing}$. The lack of pre-trends corroborates the claim that officers do not sort into sites based on the drift component. I test more formally for the presence of pre-trends in Section 5, and I do not find evidence of this phenomenon.

Second, absent match-driven sorting, a good manager has the same effect (on log productivity) regardless of the office. Thus, the effect of moving from a low- to a high-productivity

[6]I find some evidence of better managers sorting according to geography and office type. This type of sorting is not a threat to my empirical strategy.

TABLE V

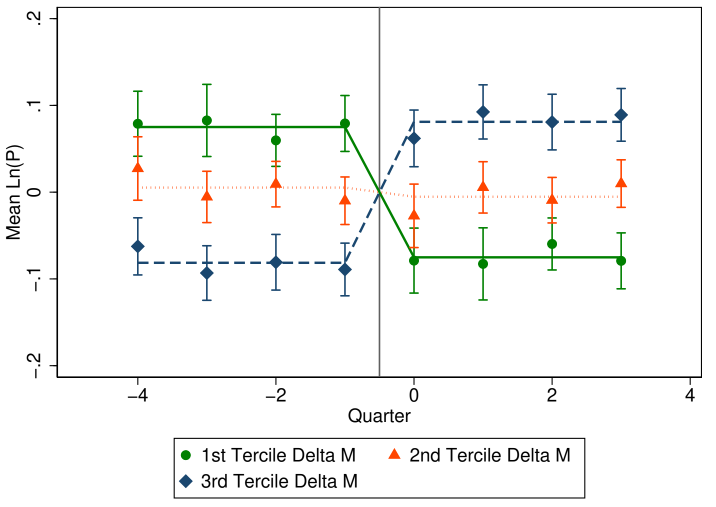
CAN OBSERVABLES PREDICT INCOMING MANAGER FE?

	(1) Manager FE		(2) Change in Manager FE	
Main Office	-0.612	(0.095)	0.029	(0.040)
North or Center	0.161	(0.071)	-0.017	(0.027)
P_{2011}	0.000	(0.002)	-0.002	(0.001)
Y_{2011}	0.000	(0.000)	0.000	(0.000)
FTE_{2011}	-0.000	(0.003)	-0.001	(0.001)
P_{t-1}	0.001	(0.002)	-0.001	(0.001)
P_{t-2}	-0.001	(0.001)	-0.001	(0.001)
P_{t-3}	0.000	(0.001)	-0.000	(0.001)
P_{t-4}	0.000	(0.001)	-0.002	(0.001)
Y_{t-1}	-0.000	(0.000)	-0.000	(0.000)
Y_{t-2}	0.000	(0.000)	-0.000	(0.000)
Y_{t-3}	0.000	(0.000)	0.000	(0.000)
Y_{t-4}	-0.000	(0.000)	0.000	(0.000)
FTE_{t-1}	0.002	(0.002)	-0.002	(0.001)
FTE_{t-2}	0.001	(0.002)	0.001	(0.002)
FTE_{t-3}	0.001	(0.001)	-0.000	(0.001)
FTE_{t-4}	-0.003	(0.002)	0.001	(0.002)
Growth Rate P - 3 q	0.058	(0.078)	-0.032	(0.065)
Growth Rate P - 2 q	-0.008	(0.100)	0.061	(0.095)
Growth Rate P - 1 q	-0.125	(0.115)	-0.007	(0.078)
Growth Rate Y - 3 q	0.021	(0.077)	0.086	(0.059)
Growth Rate Y - 2 q	-0.002	(0.040)	-0.033	(0.043)
Growth Rate Y - 1 q	0.075	(0.101)	-0.014	(0.075)
Growth Rate FTE - 3 q	-0.084	(0.156)	-0.033	(0.149)
Growth Rate FTE - 2 q	0.138	(0.175)	0.074	(0.146)
Growth Rate FTE - 1 q	-0.115	(0.207)	0.093	(0.186)
N	521		521	
R sq.	0.482		0.605	
Connected Set FE	Yes		Yes	
P-value (All)	0.000		0.000	
P-value (Growth Rates)	0.807		0.864	

Note: This table investigates the extent to which office characteristics predict the incoming manager FE or the change in manager FE. The sample includes all events balanced on $[-4, 0]$. The dependent variable in column (1) is the manager effect estimated using (1), while in column (2) it is the difference between the estimated effect of the incoming and outgoing manager. P, Y, and FTE represent productivity, output, and full-time equivalent employment, respectively. t indexes the time of the event, and P_{2011} represents the office productivity at baseline. Y_{2011} and FTE_{2011} are defined accordingly. “Growth Rate P - x q" is defined as the productivity growth rate of office i between $t - (x + 1)$ and $t - 1$. “N" represents the number of office-quarter observations. “P-value (All)" and “P-value (Growth Rates)" are the p-values for the null hypothesis that all regressors of interest are jointly statistically significant and that the growth rates are jointly significant, respectively. All regressions include connected set fixed effects. Standard errors are clustered at the office level and are reported in parentheses.

manager is equal and opposite to the effect of moving from a high- to a low-productivity manager. I present two tests for this threat to identification. Figure 1 does not show evidence of managers sorting according to comparative advantage. The first and third terciles are remarkably symmetric. Moreover, if match components were present, a more flexible saturated model with manager-by-office dummies should fit better than the additively-separable specification of equation (1). The fit improves only marginally, moving from an adjusted R^2 of 0.762 (column 3) to 0.764 (column 5), suggesting that match components are not quantitatively relevant in this context.

FIGURE 1.—Mean Productivity for Offices which Experience a Change in Leadership Classified by Tercile of Changes in Manager Effects



Note: This figure plots the mean (trend-adjusted) log office productivity and the associated 95% confidence intervals relative to change in leadership events. The figure plots three types of leadership transitions, classified by terciles of the change of managerial ability: (1) an overall increase (blue diamonds), (2) an overall decrease (green circles), and (3) no significant change (orange triangles). $\widehat{\Delta M}_i$ represents the change in the estimated manager fixed effects. The x-axis indexes event time.

Third, manager rotations could coincide with transitory shocks. For instance, INPS may reallocate good managers as a reaction to a bad ϵ_{it} draw. Once again, this is not consistent with the lack of pre-trends reported in Figure 1 and in Section 5.

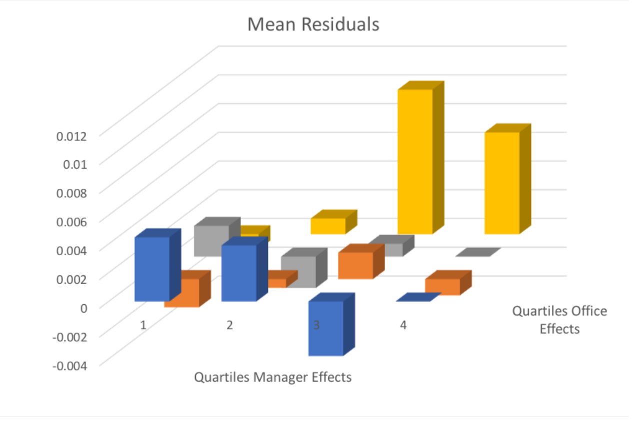
A final set of concerns about model (1) regards the assumption of additive separability between the permanent office component and managerial ability. A violation of the additive separability assumption would result in abnormally large/small mean residuals for some office-manager pairs. To assess whether this is the case, I divide the estimated manager and office effects into quartiles and compute the mean residual for each cell. Figure 2 reports these statistics. All values are rather low, and the highest mean residual is equal to 0.01. Overall, this finding suggests that match effects, if present, are not quantitatively relevant in this context. The analysis presented in this subsection supports the claim that the two-way fixed effect model approximates the data fairly well in this setting.

4.4. *Variance Decomposition and Observable Characteristics*

I perform a variance decomposition to assess the relative importance of permanent office characteristics (e.g., local social norms, workforce composition) and managers' effects. If manager ability and office characteristics are important determinants of productivity, then $Var(\alpha_i)$ and $Var(\theta_{m(i,t)})$ should explain a large share of the variation in observed productivity. Applying equation (1), the variance of log productivity decomposes into (Abowd et al., 1999):

$$\begin{aligned} Var(\ln P_{it}) =& Var(\alpha_i) + Var(\theta_{m(i,t)}) + Var(\tau_t) + Var(u_{it}) \\ &+ 2Cov(\theta_{m(i,t)}, \alpha_i) + 2Cov(\theta_{m(i,t)}, \tau_t) + 2Cov(\alpha_i, \tau_t). \end{aligned} \tag{4}$$

FIGURE 2.—Mean Residual by Manager/Office Quartiles



Note: This Figure shows mean residuals from model (1) with cells defined by quartiles of estimated manager effect, interacted with quartiles of estimated office effect.

TABLE VI

BIASED-CORRECTED VARIANCE-COVARIANCE DECOMPOSITION

	(1) Component	(2) Share of Total
Var(Ln(P))	0.1106	100 %
Var(Manager)	0.0102	9.22%
Var(Office)	0.0319	28.84 %
Var(Time)	0.0408	36.89%
Cov(Manager, Office)	-0.0096	-8.68%
Cov(Time, Manag. + Office)	0.0015	1.39%
N	2,735	

Note: This table presents the bias-corrected variance-covariance decomposition of log productivity in the largest connected set. The model includes dummies for manager, office, and quarter fixed effects.

Table VI reports the bias-corrected variances and covariances estimated on the largest connected set (Andrews et al., 2008).[7] Manager fixed effects explain roughly 9% of the variance in log productivity, about one third as much as the permanent component of productivity associated with different offices. Time fixed effects explain a non-trivial share of the variation in productivity, reflecting seasonality in productivity and overall improvement in the Social Security Agency performance over time. The bias-corrected covariance between manager and office effects is negative: more productive managers work at less productive offices. The negative assortative matching result is consistent with INPS trying to reduce inequality in productivity across sites, and with the most productive managers having a preference for the least productive sites.[8]

[7] Andrews et al. (2008) show that if there are too few movers in the data (i.e., limited mobility bias) sample variances tend to be biased upward and the sample covariance between manager and office tends to be biased downward. They develop a correction for the bias.

[8] Two recent studies documented negative assortative matching between managers and production lines (Adhvaryu et al., 2020) and World Bank managers and countries (Limodio, 2019).

TABLE VII

MANAGER EFFECTS AND OBSERVABLE CHARACTERISTICS

	(1) Manager FE	
Male	-0.06	(0.03)
Experience in the Public Sector	0.02	(0.01)
Experience in the Public Sector Squared	-0.00	(0.00)
Center	0.07	(0.06)
South or Islands	0.02	(0.04)
North-West	0.00	(0.05)
Abroad	0.00	(0.06)
Econ, Business, and Admin	0.04	(0.05)
Sci, Engen, Math, and Stat	-0.08	(0.06)
Social Sciences and Humanities	0.02	(0.04)
Law	-0.05	(0.04)
Missing Educ	-0.09	(0.07)
N	851	
R sq.	0.45	

Note: This table presents the correlation between the manager effects estimated from equation (1) and manager characteristics. These characteristics include gender, experience, the region of birth, and highest educational attainment. N represents the number of managers in my sample. “Experience in the public sector" is defined as the number of years since the manager was first hired in any public sector institution. The omitted categories are “Female", “North-East", and “No college". Controls include connected set fixed effects. Robust SE in parentheses.

To better understand what predicts more productive managers, I regress the estimated manager fixed effects from (1) on observable characteristics in Table VII. These characteristics include: gender, experience, experience squared, a set of dummies for region of birth, and a set of indicators for the highest educational attainment. Female managers are on average more productive than their male counterparts. Furthermore, managerial talent is strongly correlated with experience, but it exhibits decreasing marginal returns. The regression analysis also suggests that managers born in Southern Italy or the Islands are more productive than those from the North and that managers who never attended college are more productive than those who studied law or STEM. These findings are consistent with negative selection into government jobs for men, those born in the North, and those who have a STEM major.

In this section, I have shown that managers have a quantitatively meaningful impact on office productivity. Next, I analyze the specific mechanisms that drive the effects of more productive managers.

5. WHAT MAKES FOR A PRODUCTIVE MANAGER?

In this section, I characterize *how* managers matter. I first decompose the productivity gains induced by a change in leadership into its effects on output and full-time equivalent employment. Second, I explore how changes in managerial talent impact workers through personnel decisions and changes in their time allocation. Third, I evaluate the productivity-quality tradeoff and mangers’ impacts on backlog.

5.1. *Event Study Strategy*

I begin by specifying a basic event study regression that relates outcome y (e.g., productivity, output, FTE, new hires etc.) to changes in managerial talent ΔM_i:

$$y_{it} = \alpha_i + \sum_{k \neq -1} \left[\pi_0^k D_{it}^k + \pi_1^k D_{it}^k \Delta M_i\right] + g_t\left(X_{it}\right) + \xi_{it} \tag{5}$$

where k indexes quarters relative to a change in management and the π_0^k's coefficients capture dynamics related to a change in leadership that are common across all offices. The main objects of interest are π_1^k's, which capture effect heterogeneity that depends on how incoming managers' talent differs from the managers they replace (i.e., ΔM). Permanent differences in office productivity are captured by α_i, and $g_t\left(X_{it}\right)$ controls flexibly for time trends.

If manager effectiveness were observable to the econometrician, equation (5) could be estimated directly. Fundamentally, managerial talent $\theta_{m(i,t)}$ cannot be directly observed, but the two-way fixed effects model enables me to estimate it. However, using the first-step estimates $\widehat{\theta}_{m(i,t)}$ as covariates in (5) could bias the $\hat{\pi}_1^k$'s. Idiosyncratic productivity shocks could affect both the estimates of manager effectiveness and the outcome of interest, creating a spurious correlation even in the absence of a causal relationship. To see this, consider productivity as the dependent variable in equation (5). In finite samples, idiosyncratic productivity shocks are not completely averaged away in the estimates of manager effectiveness. The same idiosyncratic productivity shocks in the residual of the event study ξ_{it} also appear as finite-sample estimation error in $\hat{\theta}_{m(i,t)}$.[9]

I develop a procedure that estimates manager effectiveness, removing data that generates the bias. Specifically, I take advantage of the fact that because $\widehat{\theta}_{m(i,t)}$ averages over many periods, I can purge the mechanical correlation by leaving out data where the correlations may arise. To this end, I modify the standard event study specification in two ways. First, I subtract from equation (5) in each event time the corresponding values in event time $k = -1$ (i.e., $\Delta y_i^k = y_i^k - y_i^{-1}$). Second, to purge the regressor of potential mechanical correlations, I generate the change in manager productivity $\widehat{\Delta M}_i^{L,k}$ by using estimates from separate two-way fixed effect models *excluding* data from event times -1, 0, and k.

Formally for a generic outcome y,

$$\Delta y_i^k = \pi_0^k + \pi_1^k \widehat{\Delta M}_i^{L,k} + \Gamma^k X_i + \Delta \xi_i^k \tag{6}$$

$$\text{where} \qquad \widehat{\Delta M}_i^{L,k} = \hat{\theta}_{i,incoming}^{L,k} - \hat{\theta}_{i,outgoing}^{L,k}$$

and the $\hat{\theta}_{i,\cdot}^{L}$'s are the leave-out estimated manager effect of the incoming and outgoing managers, respectively. X_i includes indicators for being in the Center-North of Italy, for being main offices, for quartiles of baseline productivity, and two-way interactions between each of these. I also flexibly control for trends by including time dummies and time dummies interacted with the dummy for being in the Center-North of Italy. The specification in (6) suggests

[9] More explicitly, in this case $\xi_{it} = u_{it}$, where u_{it} is the same composite error from the two-way fixed effects model in equation (3). Even if unsystematic, idiosyncratic shocks in period t, ϵ_{it}, appear as estimation error in the estimated manager effect and tend to zero with more observations per manager. In finite samples, idiosyncratic productivity shocks generate a mechanical correlation between estimated managerial productivity $\hat{\theta}_{m(i,t)}$ and the event study residual ξ_{it}.

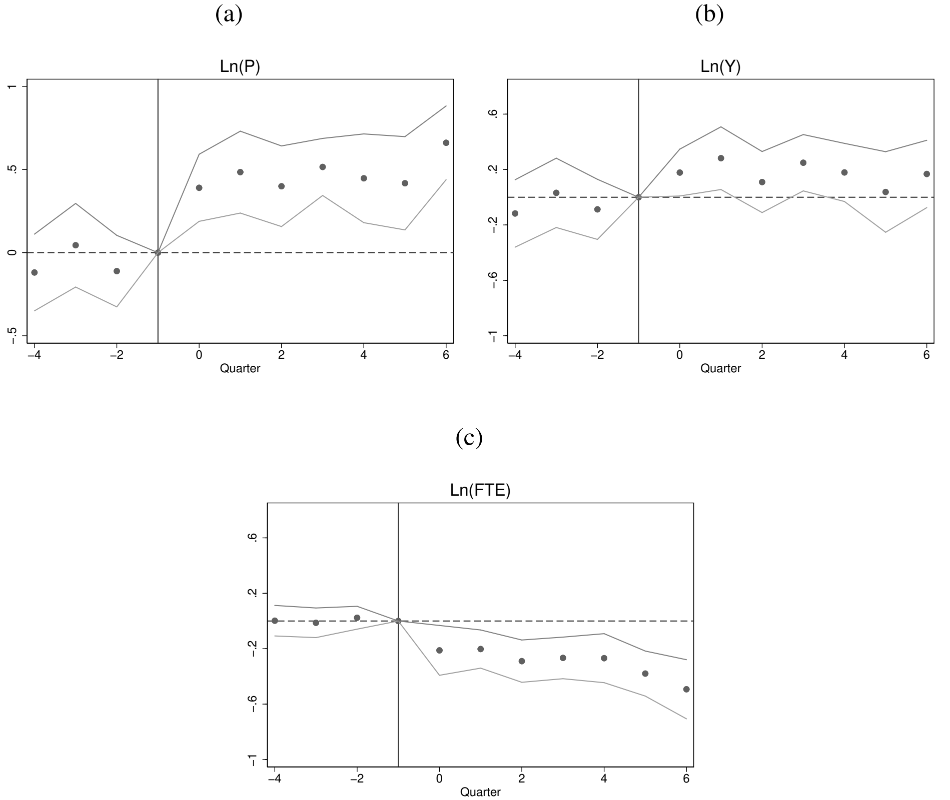
estimating separate regression models for each event time. I focus on the balanced-analysis sample. I bootstrap the standard errors to account for the presence of a generated regressor.

This procedure ensures that the outcomes of interest are not directly related to my measures of manager ability. However, if unobserved productivity shocks u_{it} are serially correlated, then my leave-out measure may still be spuriously, yet indirectly correlated with outcomes. However, serially correlated productivity shocks do not appear to be a concern in my setting as the autocorrelation coefficient from fitting an AR(1) model to the residuals from equation (1) is extremely small ($\hat{\rho} = 0.04$).

5.2. *Decomposition of Productivity Impacts*

Figure 3 reports the estimated impact of an increase in managerial quality on three log-transformed outcomes: productivity (Panel a), output (Panel b), and full-time equivalent labor (Panel c). Reassuringly, changes in managerial talent do not predict changes in productivity before the event takes place, which alleviates concerns regarding sorting on trends. Productivity gains materialize when the better manager takes office.

FIGURE 3.—Decomposition of Productivity Effects



Note: This figure plots coefficients measuring the impact of managerial talent $\hat{\pi}_1^k$ from equation (6) and their associated 95% bootstrapped confidence intervals. The dependent variables are log productivity (Panel a), log output (Panel b), and log full-time equivalent employment (Panel c). The x-axis indexes event time.

The figure decomposes productivity impacts into separate effects on output (numerator) and FTE (denominator). Holding office composition constant, managers may increase output by eliciting more effort from workers or by better matching employees to tasks. Because public sector managers have limited discretion over workers' promotions and compensation, increasing output may be particularly challenging. Managers also have limited ability to make de jure personnel decisions: firing is virtually impossible and the hiring freeze prevents them from bringing in additional workers. In this setting, good managers may be those who are better at matching workers with tasks and use indirect strategies to elicit workers' effort and make unproductive workers quit or retire.

When a better manager takes over, there is a positive (although not statistically significant) effect on output (Figure 3b). This pattern is consistent with more effective managers better matching workers to tasks or motivating their employees better (Burgess et al., 2010). This finding is striking in light of the sharp decrease in the number of employees assigned to the office after a better manager takes charge (Figure 3c).

5.3. *Reduced Form Impacts of Managerial Talent*

How do managers reduce office size in such a constrained environment? Table VIII explores the channels through which managers impact the composition of workers they supervise. The dependent variables represent *cumulative* flows.

TABLE VIII

ESTIMATED EFFECTS OF CHANGES IN MANAGERIAL TALENT ON OFFICE COMPOSITION

k	(1) Retirement	(2) Separations	(3) Hires	(4) Fires	(5) Inbound T	(6) Outbound T
-4	0.041	0.099	0.178	0.003	0.120	-0.039
	(0.125)	(0.111)	(0.110)	(0.004)	(0.121)	(0.146)
-3	-0.044	-0.011	0.093	0.002	0.069	-0.111
	(0.088)	(0.078)	(0.090)	(0.004)	(0.090)	(0.133)
-2	-0.055	0.033	0.034	0.003	0.068	-0.122
	(0.059)	(0.053)	(0.073)	(0.004)	(0.082)	(0.112)
0	0.301	0.038	0.024	-0.008	0.031	-0.023
	(0.085)	(0.081)	(0.018)	(0.010)	(0.159)	(0.053)
1	0.393	-0.011	0.027	-0.056	-0.033	-0.010
	(0.100)	(0.099)	(0.033)	(0.031)	(0.163)	(0.067)
2	0.381	-0.057	0.024	-0.049	-0.196	-0.142
	(0.105)	(0.117)	(0.033)	(0.040)	(0.172)	(0.097)
3	0.392	-0.012	0.006	-0.063	-0.327	-0.234
	(0.117)	(0.126)	(0.038)	(0.045)	(0.174)	(0.114)
4	0.438	-0.0722	-0.015	-0.040	-0.385	-0.231
	(0.116)	(0.128)	(0.039)	(0.038)	(0.171)	(0.123)
5	0.413	0.015	0.005	-0.061	-0.403	-0.303
	(0.120)	(0.135)	(0.039)	(0.041)	(0.174)	(0.125)
6	0.399	-0.121	-0.082	-0.059	-0.537	-0.405
	(0.125)	(0.155)	(0.057)	(0.042)	(0.184)	(0.135)
N	318	318	318	318	318	318
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the reduced-form impacts of managerial talent on office composition. More specifically, it reports the coefficients π_1^k obtained estimating (6) on the balanced-analysis sample. N represents the number of office-quarter observations. The dependent variable, cumulative Asinh(y_{it}), is reported at the top of each column. All models include time fixed effects, main effects and two-way interactions between a dummy for Center-North, a dummy for main offices, a set of dummies for quartiles of baseline productivity, as well as time effects interacted with the dummy for Center-North. k indexes event time. Each coefficient is obtained from a separate regression. Bootstrapped standard errors are reported in parentheses.

Better managers drive older workers to retire (column 1). The bulk of the manager-induced retirements occur in the first two quarters after the change in leadership. Managers and higher level officials cannot force older workers to retire, and they cannot negotiate severance packages to persuade them to leave. So, why do retirement-age workers only leave when better managers arrive? One plausible explanation comes from the norm that more senior workers are usually given more prestigious tasks that come both with additional responsibilities and importantly, additional compensation. Anecdotal evidence suggests that more productive managers reallocate these tasks when they take charge. Either because they are slighted or because they lose the extra compensation, senior employees retire. Although it is hard to evaluate how productive managers impact the satisfaction of the remaining workers, the fact that younger emplyees do not quit pushes against the interpretation that these workers are significantly worse-off (Table VIII, column 2). Given institutional constraints, an increase in management quality unsurprisingly does not have a statistically significant impact on hiring and firing (columns 2 and 3). The arrival of a more productive manager is associated with fewer inbound (column 4) and outbound transfers (column 5).

I also investigate how time allocation changes with the takeover of a more productive manager. Changes in managerial quality do not appear to come from training, overtime work, or the total number of hours (Table G.I in the Online Appendix). There is some suggestive evidence that better managers may be able to reduce absenteeism rate.

Furthermore, the increase in productivity does not come at the cost of quality (column 1 of Table IX). There is also some suggestive evidence that effective managers lower claim backlogs (column 2 of Table IX). Overall, these findings suggest that good public sector managers are those that affect culture and organization (Azulai et al., 2020).

6. ROBUSTNESS CHECKS

In this section, I show that manager effects are not confounded by contemporaneous "demand shocks" and that managers are not "gaming the system" by targeting specific claim types.

6.1. *Demand*

Column 3 of Table IX shows no consistent pattern of demand increasing at the same time more productive managers enter. Moreover, estimating managerial productivity controlling directly for local demand yields largely similar results. The correlation of the estimated manager effects with and without the control is 0.988. Altogether, there is little evidence that classifying managers as high or low productivity is confounded by idiosyncratic local demand shocks.

6.2. *Claim Targeting*

INPS did not prioritize any subset of claims during the period of study. However, different managers may place more or less weight on different tasks, reflecting differing social values or their attempt to take advantage of "over-priced" tasks. Accordingly, managers have the latitude to allocate resources to tasks that they value highly.

However, if managers prioritized some claim types at the expense of others, the quality index would suffer. Empirically, quality is not negatively impacted when a more productive manager takes over (column 1 of Table IX). Moreover, 5% of all claims processed by each office are audited twice per year. The purpose of this cross-check is to detect anomalies or illicit behavior.

To test more formally whether managers change the product mix, I divide all the products into nine categories, and I estimate the effect of better managers on claims processed in each of

TABLE IX

ESTIMATED EFFECTS OF CHANGES IN MANAGERIAL TALENT ON QUALITY, BACKLOG, AND DEMAND

k	(1) Ln(Quality)	(2) Ln(Backlog)	(3) Ln(Demand)
-4	-0.036	0.150	0.124
	(0.025)	(0.117)	(0.114)
-3	-0.029	0.140	0.113
	(0.024)	(0.100)	(0.129)
-2	-0.049	0.053	0.071
	(0.026)	(0.081)	(0.105)
0	-0.058	-0.129	0.100
	(0.041)	(0.090)	(0.124)
1	0.064	-0.077	0.282
	(0.067)	(0.099)	(0.129)
2	-0.091	-0.248	-0.275
	(0.054)	(0.116)	(0.189)
3	0.049	-0.345	-0.075
	(0.041)	(0.120)	(0.130)
4	0.055	-0.258	0.061
	(0.035)	(0.172)	(0.165)
5	0.010	-0.071	0.176
	(0.031)	(0.160)	(0.175)
6	-0.008	-0.145	0.171
	(0.068)	(0.178)	(0.143)
N	300	318	313
Time FE	Yes	Yes	Yes

Note: This table reports the reduced-form impacts of managerial talent on quality, backlog, and demand. More specifically, it reports the coefficients π_1^k obtained estimating (6) on the balanced-analysis sample. N represents the number of office-quarter observations. The dependent variable, y_{it}, is reported at the top of each column. All models include time fixed effects, main effects and two-way interactions between a dummy for Center-North, a dummy for main offices, a set of dummies for quartiles of baseline productivity, as well as time effects interacted with the dummy for Center-North. k indexes event time. Each coefficient is obtained from a separate regression. Bootstrapped standard errors are reported in parentheses.

these categories. I interpret shifts toward a product category as evidence of targeting. Specifically, I estimate event studies described by equation (5), controlling for demand for the nine broad product categories. Figures 4a and 4b show no evidence of productive managers shifting the production mix. Overall, I find no evidence of claim targeting.

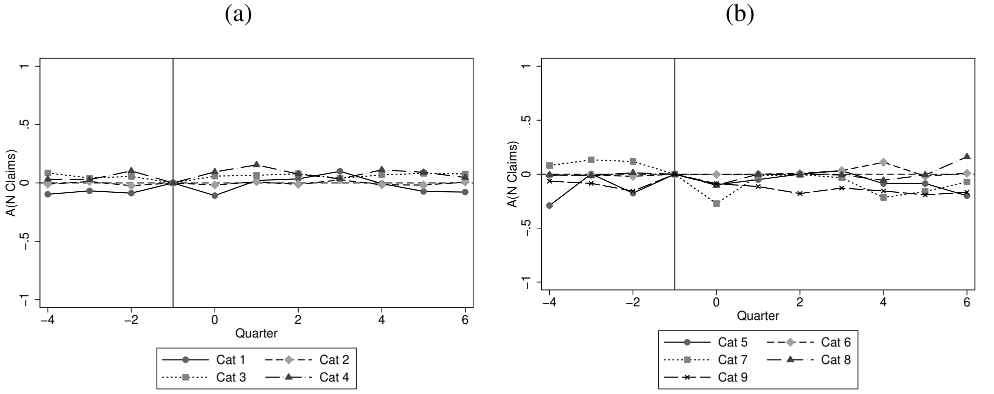
7. COUNTERFACTUAL EXERCISES

In this section, I evaluate the efficiency gains from alternative managerial allocation schemes. I consider a static allocation problem and subsequently omit time indices t. I assume that the social planner maximizes the aggregate agency output but cannot directly influence office productivity (α_i) or the full-time equivalent number of workers assigned to the office (L_i). However, she can hire and fire managers and freely assign them to offices. Let $P_i\left(\alpha_i, \theta_{m(i)}\right)$ be the average productivity at office i, which depends on α_i and the ability of the manager $\theta_{m(i)}$. Assuming $\ln P_i\left(\alpha_i, \theta_{m(i)}\right) = \alpha_i + \theta_{m(i)}$, the planner's objective function is:

$$\max_{m(\cdot)} \sum_i \gamma_i e^{\theta_{m(i)}}, \tag{7}$$

where $\gamma_i = e^{\alpha_i} L_i$ and the social planer chooses how to allocate managers to offices, $m(\cdot)$.

FIGURE 4.—More Productive Managers Do Not Shift Production



Note: Panels (a) and (b) report estimates of the effect of managerial talent on each of nine separate claim categories. The categories are defined as follows: 1: Insurance and pensions, 2: Subsidies to the poor, 3: Services to contributors, 4: Social and medical services, 5: Specialized products, 6: Archives and data management, 7: Administrative cross-checks, 8: Checks on benefits, 9: Appeals. The x-axis indexes event time.

I consider four counterfactual policies: 1) reassigning existing managers to offices optimally; 2) firing the worse 20% of managers and substituting them with a manager of median quality without otherwise changing the assignment of managers to offices; 3) implementing both policies at once; and 4) random assignment of managers to offices.

P_iL_i is twice differentiable and supermodular, therefore the optimal allocation is an assortative matching equilibrium where the best managers are sent to the largest and most productive offices with the highest γ_i.[10]

If the social planner reassigns managers using the optimal allocation rule, aggregate productivity increases by at least 6.9%. Firing the bottom 20% of managers raises aggregate productivity by only 2.9%. The first policy is more effective than the second because the most productive managers are currently allocated to the least productive offices (Table VI). Implementing both these policies simultaneously increases aggregate productivity by at least 7.4%. Managerial allocation is key in this context. Finally, random assignment increases aggregate productivity by 2% by undoing the negative assortative matching equilibrium.

8. CONCLUSION

This paper is the first to estimate the productivity impacts of managers in the public sector bureaucracy. Bureaucracy represents a large part of modern economies. But because prices, profit-motive, and competition are not driving forces in the public sector, simply measuring productivity when government has many objectives has limited previous research.

I overcome the contextual challenges by using novel administrative data containing an output-based measure of productivity of public offices. I find that public sector managers have a quantitatively meaningful impact on the productivity of the offices they oversee, despite their

[10]Such an allocation widens the spread of productivity across sites. Traditionally, agencies made efforts to equalize quality of services across sites on horizontal equity grounds. Beneficiaries should not receive a different treatment depending on where they live. As claims can be electronically redistributed across sites at virtually no cost and processed anywhere, it is unclear why productivity should be equalized across offices in this setting.

limited ability to make personnel decisions. These productivity gains are mainly driven by the exit of older workers. A good manager sustains production levels without resorting to hiring or overtime to compensate for the decrease in full-time equivalent employment. These findings are consistent with anecdotes suggesting that senior workers leave when better managers reassign more prestigious and better-compensated tasks to more junior employees. Assessing alternative managerial allocation schemes, I find that reallocating managers across sites would increase the agency output by more than twice as much as firing the worst 20% of managers. This is important given how challenging it is to substantially improve the talent pool among public sector employees through civil service reforms.

I view these findings as broadly relevant for agencies where offices primarily administer claims, a large share of modern bureaucracies. These include the Social Security Administration and other ministries that administer old age pension programs; national and local agencies that dispense social welfare transfers; and licensing authorities such as the departments of motor vehicles. These findings may not apply to very different public sector organizations like hospitals, schools, and police departments, a promising avenue for future research. Moreover, while this paper is relevant for top-level bureaucrats managing small- and medium-size teams, it may not generalize to the management of large teams; to settings where the agency mission is not clearly defined; or where the output is not observable. Finally, the fact that managers matter the most with respect to personnel selection is not unique to this setting; it is echoed by findings in the private- and public-sector literature (Branch et al., 2012, Hoffman and Tadelis, 2018). The precise mechanisms through which managers matter are likely to depend on the institutional constraints that limit managers' actions.

REFERENCES

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. A. (2020). No line left behind: Assortative matching inside the firm. Working Paper 27006, National Bureau of Economic Research.

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A*, 171(3):673–697.

Azulai, M., Imran, R., and Daniel, R. (2020). Can Training Improve Organizational Culture? Experimental Evidence from Ghana's Civil Service. Working paper.

Baltrunaite, A., Bovini, G., and Mocetti, S. (2020). Opening the Black-Box of Managerial Talent. Working paper.

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

Bennedsen, M., Perez-Gonzalez, F., and Wolfenzon, D. (2010). Do CEOs Matter? Working paper.

Bennedsen, M., Perez-Gonzalez, F., and Wolfenzon, D. (2011). Estimating the Value of The Boss: Evidence from CEO Hospitalization Events. Working papers.

Bertrand, M., Burgess, R., Chawla, A., and Xu, G. (2019). The Glittering Prizes: Career Incentives and Bureaucrat Performance. *The Review of Economic Studies*, 87(2):626–655.

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

Best, M., Hjort, J., and Szakonyi, D. (2019). Individuals and Organizations as Sources of State Effectiveness. Working paper.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*, 128(1):1–51.

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

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

Branch, G. F., Hanushek, E. A., and Rivkin, S. G. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals. Working Paper 17803, National Bureau of Economic Research.

Burgess, S., Propper, C., Ratto, M., Scholder, S. v. H. K., and Tominey, E. (2010). Smarter task assignment or greater effort: The impact of incentives on team performance. *The Economic Journal*, 120(547):968–989.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Chandra, A., Finkelstein, A., Sacarny, A., and Syverson, C. (2016). Productivity dispersion in medicine and manufacturing. *American Economic Review*, 106(5):99–103.

Choudhury, P., Khanna, T., and Makridis, C. A. (2019). Do Managers Matter? A Natural Experiment from 42 RandD Labs in India. *The Journal of Law, Economics, and Organization*, 36(1):47–83.

Finan, F., Olken, B., and Pande, R. (2017). Chapter 6 - the personnel economics of the developing state. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 2 of *Handbook of Economic Field Experiments*, pages 467 – 514. North-Holland.

Giorcelli, M. (2019). The long-term effects of management and technology transfers. *American Economic Review*, 109(1):121–52.

Hoffman, M. and Tadelis, S. (2018). People management skills, employee attrition, and manager rewards: An empirical analysis. Working Paper 24360, National Bureau of Economic Research.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law Economics and Organization*, 7:24–52.

Janke, K., Propper, C., and Sadun, R. (2019). The impact of ceos in the public sector: Evidence from the english nhs. Working paper.

Khan, A. Q., Khwaja, A. I., and Olken, B. A. (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review*, 109(1):237–70.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The value of bosses. *Journal of Labor Economics*, 33(4):823–861.

Limodio, N. (2019). Bureaucrat Allocation in the Public Sector: Evidence from the World Bank. Working paper.

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

Rasul, I., Rogger, D., and Williams, M. (2019). Management and bureaucratic effectiveness: Evidence from the ghanaian civil service. Working paper.

Syverson, C. (2004). Product substitutability and productivity dispersion. *The Review of Economics and Statistics*, 86(2):534–550.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Tsai, T. C., Jha, A. K., Gawande, A. A., Huckman, R. S., Bloom, N., and Sadun, R. (2015). Hospital board and management practices are strongly related to hospital performance on clinical quality metrics. *Health Affairs*, 34(8):1304–1311. PMID: 26240243.

Xu, G. (2018). The costs of patronage: Evidence from the british empire. *Forthcoming American Economic Review*.

APPENDIX: A

TABLE A.I

DISPERSION IN PRODUCTIVITY

Productivity Measure	Within-Industry Productivity Moment	
Panel A: My Measure		
Labor productivity:	Median	4.524
log(weighted claims/employee)	IQ range	0.426
	90-10 percentile range	0.860
	95-5 percentile range	1.161
	St. deviation	0.366
	N	13,212
Panel B: Syverson (2004)		
Labor productivity:	Median	3.174
log(value added/employee)	IQ range	0.662
	90-10 percentile range	1.417
	95-5 percentile range	2.014

Note: Panel A reports the statistics of interest for my productivity measure calculated over the full sample. N represents office-quarter observations. Panel B is taken from Table 1 of Syverson (2004) and reports plant-level productivity distribution moments across 433 (four-digit SIC) manufacturing industries.