Managers and Productivity in the Public Sector

Alessandra Fenizia*

May 4, 2020

Abstract

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

JEL: H11, O10, D73, M50.

*Corresponding author: Alessandra Fenizia, George Washington University, Department of Economic, G Street NW, Room #364, Washington, DC 20052, (e-mail: afenizia@gwu.edu). This paper is based on the main chapter of my dissertation that I completed at UC Berkeley. I am particularly indebted to my advisors David Card, Patrick Kline, and Hilary Hoynes for their invaluable support and guidance throughout the process. I am grateful to Eric Avis, Tito Boeri, Giulia Bovini, Luca Citino, Ernesto Dal Bó, Fred Finan, Michela Giorcelli, Daniel Haanwinckel, Erin Kelly, Nicholas Li, John Loeser, Juliana Londoño-Vélez, Isabela Manelici, Edward Miguel, Matteo Paradisi, Francisco Perez-Gonzalez, Evan Rose, Raffaele Saggio, Yotam Shem-tov, David Silver, Bryan Stuart, Kevin Todd, Jose Vasquez, Christopher Walters, Guo Xu, Danny Yagan, Zarek Brot-Goldberg for helpful suggestions. The views expressed in paper are those of the author and don't necessarily reflect the position of the Istituto Nazionale di Previdenza Sociale.

I Introduction

Public sector managers are the cornerstone of modern bureaucracies. They oversee day-to-day operations of complex public organizations and supervise policy implementation. Their effectiveness may have large consequences for citizens' welfare. For instance, delayed unemployment insurance benefit payments can aggravate hardship of the newly unemployed, and longer processing times for disability insurance claims can directly reduce employment and earnings for multiple years (Autor et al., 2015). 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, promotions, incentive-pay schemes). In most countries, public sector workers enjoy strong job security, and often receive promotions and pay raises that depend only on seniority as opposed to individual performance. On the other hand, public sector managers may play a particularly 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 is notoriously hard to measure the performance of government agencies. A set of recent studies has made progress by measuring individual managerial practices, qualitative policies and procedures that are thought to be associated with well-run organizations, and has established that these measures positively correlate 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, pensions, subsidies to the poor and other welfare and insurance programs. A key innovation is the use of an output-based measure of productivity of public offices constructed using detailed administrative quarterly 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. 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 a homogeneous product, and there are virtually no differences in physical capital across offices. Important for the study of government productivity, I do not need to rely on wages to infer productivity (Eeckhout and Kircher, 2011).

I begin my analysis by documenting large variation in productivity across offices within INPS, dispersion that is not fully explained by regional differences that typically characterize the Italian economy. 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 increase in managerial talent is associated with a 10% increase in office productivity. I also find that the (bias-corrected) covariance between manager and office fixed effects is negative, suggesting that INPS currently allocates the best managers to the least productive sites. This result is consistent with INPS trying to reduce inequality in productivity across sites.

In the second part of the paper, I exploit the rotation of managers as a natural experiment to 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 output may be particularly challenging. Instead, good managers may keep production high while reducing costs. I extend the current literature by documenting what makes for a productive manager in the public sector. I find that the productivity gains are driven primarily by the exit of older workers who retire when a productive manager takes charge. Better managers maintain production without resorting to hiring or assigning overtime hours to compensate for the reduction in full-time equivalent employment. This is consistent with more productive managers being able to better match workers with tasks. One might be worried that higher output per worker comes at the cost of quality of service provided. INPS also measures a quality index that captures both timeliness in processing claims and the rate of errors in subsequent random audits. I use this quality index to assess whether there is any a trade-off between productivity and quality of service, and I find that higher output per worker does not come at the cost of lower quality in this setting.

In the final section of the paper, I discuss how governments could use these findings to improve public service provision by evaluating the efficiency gains from alternative managerial allocation schemes. The estimates from my productivity model imply that an optimal social allocation assigns the best managers to the largest and most productive offices. I find that if managers were reassigned on this basis, the productivity of the agency would increase by at least 6.9%. This result suggests that there may be large social returns to carefully modelling public sector productivity and the impacts of managerial talent.

This paper contributes to three strands of the literature. First, it contributes to the research that documents the impact of management and managerial practices on firm-level outcomes (Bertrand and Schoar, 2003; Perez-Gonzales, 2006; Bloom and Van Reenen, 2007; Bloom et al., 2013; Lazear et al., 2015; Bloom et al., 2018; Bruhn et al., 2018; Hoffman and Tadelis, 2018; Giorcelli, 2019; Bandiera et al., 2020). The effects of managerial practices are not confined to the private sector: better management practices are correlated with improved public service delivery in hospitals (Tsai et al., 2015), schools (Bloom et al., 2015), and civil service organizations

(Rasul and Rogger, 2018; Rasul et al., 2019). My paper extends this literature in two ways. First, I study a domain – the public sector – where we know very little about the productivity impacts of managers.[1] Second, to the best of my knowledge, this is the first paper to document *how* managers matter in a very constrained environment where they have limited ability in hiring and firing workers and no discretion over investment.

Second, my work relates to the literature that studies the impact of civil servants on 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). The scarcity of reliable output measures for the public sector has severely limited the extent to which researchers could study the performance of government agencies, and my work attempts to fill this gap.[2] My paper is most closely related to the work of Best et al. (2019) and Janke et al. (2019), who find conflicting evidence on whether public officials matter: Best et al. (2019) find that more effective bureaucrats are able to substantially lower procurement costs in Russia, while Janke et al. (2019) show that CEOs of large public hospitals do not affect hospital performance. The measurement challenges that papers in this space have faced is reflected in their sophisticated approaches to measuring performance: Best et al. (2019) develop a quality-adjusted price index using a machine learning classifier, while Janke et al. (2019) aggregate their production variables into a single performance index to proxy for multifaced hospital performance. 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 the rich INPS administrative data to recover the channels through which managers improve productivity and examine what makes

[1]Choudhury et al. (2019) find that better alignment between the CEO and middle-level managers boosts research productivity in India's public R&D labs. Janke et al. (2019) find that CEOs do not affect the performance of large public sector hospitals in the UK.

[2]A broader literature has focused on the performance of relatively small institutions and the consensus is that better teachers and principals (Chetty et al., 2014a,b; Kane and Staiger, 2008; Lavy and Boiko, 2017; Branch et al., 2012; Bloom et al., 2015; Bohlmark et al., 2016; Coelli and Green, 2012; Dhuey and Smith, 2014; Grissom et al., 2015), and more effective bureaucrats (Best et al., 2019; Bertrand et al., 2019) improve the quality of the service provided. One notable exception is the work by Janke et al. (2019) discussed below.

for a productive manager.

Third, this paper also fits in the broad literature on productivity differentials between workplaces. Several papers have documented large and persistent differences in productivity across firms, even in narrowly defined industries (Syverson, 2004, 2011; Chandra et al., 2016). My paper contributes to this literature by providing compelling evidence that this phenomenon is not limited to the private sector and that it arises even *within* a large centralized public agency.

II 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, subsidies to the poor, and audits to firms and workers. 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. The overall demand facing a given office largely reflects the demographic characteristics of residents and macroeconomic conditions. I study the offices that conduct the routine work associated with reviewing and processing claims. My sample includes 111 main satellite offices and 383 local branches.[3] Within each office, a single manager (*figura apicale*) oversees production workers who assess whether to accept or reject claims.

At INPS, managers assign workloads and responsibilities, coordinate work inside the office, and ensure resources are used effectively. Their tasks include monitoring the production process and devising solutions whenever office performance falls short of production targets. However, managers are constrained from improving productivity through payroll decisions. Firing is uncommon in Italian government positions; a hiring freeze was instituted in 2008 (*blocco del turnover*) and covers the period of my analysis.[4] Thus, managers have to make the best out

[3]These offices employ the vast majority of INPS workers. Refer to Online Appendix C for more details on the sample construction. Figure J.I in the online appendix shows the distribution INPS offices across the country.

[4]The hiring freeze was introduced in 2008 and was aimed at progressively downsizing the public sector. This

of their assigned set of workers. Anecdotally, more productive managers make their mark: they reassign workloads and responsibilities, change workplace practices, enforce break times, directly oversee employees' performance, and evaluate office operations using quantitative data.

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 (and therefore contribute to the estimation of the two-way fixed effects model discussed below). In the full sample, the average manager is 54 years old and has 24 years of civil service experience, commensurate with most managers having spent their entire career in civil service. Close to 60% were born in Southern Italy or the Islands, potentially reflecting the relative attractiveness of civil service jobs to people from those areas. About one-third of managers have a university degree in Law, and another 13% have a degree in Business, Administration, or Economics. Interestingly, over 20% have no university-level education. In comparison to the overall sample, managers who move across offices are younger, more likely to be male, and more likely to hold a university degree.

Next, I briefly describe manager rotation. INPS posts manager vacancies and their corresponding eligibility criteria on an internal website that is visible to all employees. As there are no official rules or unofficial guidelines on how to choose among qualified candidates, human resources officers select managers by making a case-by-case assessment (Online Appendix A). Managers stationed in main offices are forced to rotate every five years as part of anti-corruption law *190-2012*, which aims to prevent managers from becoming susceptible to corruption as a result of becoming too entrenched. Because of staggered tenures, relatively few vacancies are open in any given year, limiting the extent to which managers can sort.[5] However, this law does not apply to managers serving in local branches (*responsabili di agenzia*). 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

reform allows government agencies to hire one worker for four employees who leave.

[5]While ideally, I would like to limit my sample to managers stationed in main offices as their moves are plausibly more exogenous, I can not do so due to the limited sample size.

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. Specifically, if INPS decides to reallocate managers, the HR department faces a complicated problem. The same hiring and firing constraints that apply to production workers also apply to managers, so the HR department has to fill a given number of managerial positions by reshuffling a given set of managers. I further corroborate this argument by testing for endogenous mobility in Section V.

Workers and managers' salaries have a fixed component (*retribuzione tabellare*) and a bonus (*retribuzione accessoria*). The former is tied to job title and the latter is a strictly increasing function of the levels of productivity and quality of service as well as the improvements of these two indicators relative to the previous year (refer to Online Appendix B for details on the bonus structure and how it can induce managers to sort into different office types). While bonuses represent a small share of overall employee compensation, they amount to 15-30% of managers' salary.

III Conceptual Framework

In this section, I present a simple conceptual framework that serves two purposes. First, it highlights where the existing empirical and theoretical work have focused. Second, it lays out the basis for my empirical analysis and for the counterfactual exercises I develop in Section VIII.

I begin by considering a constant returns to scale production function of a homogeneous product (Y_{it}) that takes capital (K_{it}), labor ($e_{it}L_{it}$), and managerial talent (M_{it}) as inputs:

$$Y_{it} = A_{it} f(K_{it}, e_{it}L_{it}, M_{it}),$$

where i and t index office and time, respectively. A_{it} represents total factor productivity and it is the product of an office-specific and a transitory component (i.e., $A_{it} = \tilde{A}_i v_{it}$). M_{it} is the portable component of managerial talent and managers are heterogeneous in their ability. Several papers have argued that increasing managerial talent increases output (Lazear et al., 2015; Frederiksen et al., 2017) and ultimately improves firm performance (Bertrand and Schoar, 2003; Perez-Gonzales, 2006; Bennedsen et al., 2010, 2011). Previous research finds that effective managers in the private sector 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 output may be particularly challenging. Good managers may be those who are able to keep production high while reducing costs.

The Italian context suggests some simplifying assumptions that allow me to disentangle the channels through which managers impact office productivity. First, all production employees work on the same software and labor is the main input of production. There are virtually no differences in per-worker physical capital across sites and little scope for manager input. However, larger offices have more workers and more workstations. Thus, I specify physical capital as $K_{it} = k_t \times L_{it}$, the product between a per-worker capital component k_t, which does not vary across offices, and labor L_{it}.

Managers are likely to matter the most with respect to workers. To formalize it, I specify L_{it} as a labor aggregate $h(L_1, L_2, ...,)$ where I assume that there are ℓ worker types who differ in their innate productivity. In principle, managers can either affect worker composition by affecting each L_ℓ individually (Hoffman and Tadelis, 2018) or office size $\sum_\ell L_\ell$ (Lucas, 1978). Finally, managers can also affect $e_{it} = m_{it}^{\lambda}$, effort, which I specify as an increasing function of per-worker managerial input (m_{it}). The vast majority of papers that study the public sector focus on $e_{it}L_{it}$ by examining the role of bureaucrats and front-line providers (Finan et al., 2017; Xu, 2018; Bertrand et al., 2019; Best et al., 2019; Khan et al., 2019). A key innovation of this paper is that I study the direct impact of the portable component of managerial talent on office performance (M_{it}) as well as the effects mediated through personnel decisions ($e_{it}L_{it}$).

Letting office production take a Cobb-Douglas form $Y_{it} = A_{it}K_{it}^{a}(e_{it}L_{it})^{b}M_{it}^{1-a-b}$, the log output per workers $\ln P_{it}$ can be expressed as:

$$\ln P_{it} = \left[\ln \tilde{A}_i\right] + [a \ln k_t] + [(1 - a - b(1 - \lambda)) \ln m_{it}] + \ln v_{it}. \tag{1}$$

I approximate (1) with a combination of office, time, and manager effects. I discuss the indentifying assumptions in Subsection V.B.

The facts that technology is constant across offices, the same rules apply to all sites, and offices produce a homogeneous product make this the ideal setting to study the impact of managers on office-level outcomes. In other words, many of the confounding factors that typically bias the estimates of the impact of managers on firm outcomes are held constant across sites.

IV Data

This section details the quarterly office level data that form the basis of 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.

IV.A Office-Level Productivity Measures

INPS has computerized quality control and internal monitoring system aimed at tracking every step of the production process.[6] I use their internal monitoring data from Q1 2011 to Q2 2017. These data report inputs including the number of full-time equivalent workers devoted to production (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. INPS also constructs a (complexity-weighted) standardized index of claims processed by each office as a measure for office output (Y_{it}). Specifically, the number of claims ($c_{v,it}$) of different types ($v = 1, ...V$) processed by office i in quarter t are aggregated into a single output measure by weighting them by their

[6] INPS ensures that the information is accurate by recording the data centrally and supervising the data collection in real-time.

complexity ($w_{v,t}$).

$$Y_{it} = \sum_{v=1}^{V} c_{v,it} \times w_{v,t}.$$

The weights represent the time employees *should* take to process each type of claim (refer to Online Appendix C for more details on the weighting system).[7] Importantly, appropriately weighting claims by their complexity controls for differences in tasks across offices (Autor et al., 2006; Autor, 2013; Stinebrickner et al., 2018). Although INPS employees' main task consists in processing paperwork, they also take turns working at the front-office where they assist beneficiaries. I provide more information on front office operations and show that measurement error in this variable is not driving my results in the Online Appendix I.

I combine the measures of office output and FTE employment to construct my measure of productivity (P_{it}) as output per worker:

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

P_{it} represents the monthly worker productivity at office i averaged over quarter t, namely the number of complexity-adjusted claims processed per worker. As INPS incentivizes managers on the basis of monthly output per worker averaged over each quarter, I multiply the denominator by three to normalize my measure of productivity in the same fashion.[8] In the empirical analysis, I approximate the labor aggregate L_{it} as $FTE_{it} \times 3$.

A concern is that dispersion in productivity may be driven by demand volatility and that when demand is low workers are left idle. This is not the case due to two reasons: first, offices have a large backlog (refer to Table II and Online Appendix C). Second, managers who are in charge of offices facing low demand are instructed to contact a high-demand office and ask to transfer claims to equalize workloads across sites (Online Appendix D). Another concern that is sometimes raised in the analysis of productivity is that employees are selectively allocating

[7]For example, processing the paperwork associated with overdue pension benefits should take on average six minutes while evaluating a mortgage (i.e., *mutuo ipotecario*) should take on average four hours.

[8]As I use log productivity in my analysis, this normalization does not affect any of the estimated coefficients.

their effort to tasks that are rewarded by the employer while neglecting other activities that are valuable but not directly rewarded (i.e., multitasking) (Holmstrom and Milgrom, 1991). In my context, this worry is mitigated by the fact that I focus on workers whose tasks mainly consist in processing paperwork, and INPS invests a lot of effort into carefully measuring all steps of the production process. One may also worry that, if weights do not correctly reflect task complexity, managers could try to game the system by shifting production toward overvalued claims and neglect undervalued ones. I address these concerns in Section VII.

Throughout the analysis, I primarily focus on productivity as it relates to quantity of output. However, the data also allow me to test whether managers matter in affecting 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)[9] and the "error rate" (the fraction of claims that has to be processed more than once because of an error in initial processing).[10] My data does not contain these two sub-components, therefore I can not analyse them separately.

IV.B Office-Employee Data

In addition to office-level productivity data, I have access to a personnel file that allows me to track employees over time within INPS (2005-2017). This dataset includes office location, job title, hiring, firing, and promotions.[11] Anecdotally, most employees are hired through a competitive examination (i.e., *concorso pubblico*) or from other government agencies. Workers rarely quit a public sector job, and the vast majority of them leave the INPS when they retire. Since I do not observe retirement directly, I use this anecdotal evidence to construct a proxy for it. I define retirements as coming from two sources: (1) voluntary separations of workers over age 60 and (2) automatic separations due to worker age limits (refer to Online Appendix

[9]While INPS generally sets the timeliness threshold at thirty days, there is a handful of products for which the threshold is set at sixty and ninety days.

[10]INPS audits 5% of each office production twice per year, and most mistakes are detected during these audits. Some are also found when denied beneficiaries file an appeal (Online Appendix C). My data does not contain information on the audits, the number of mistakes, and the number of appeals filed so I can not analyze these components separately.

[11]The personnel data does not include information on wages and earnings.

C for details).

IV.C Descriptive Statistics and Stylized Facts

In this subsection, I present an overview of Social Security Offices and I document two stylized facts related to productivity.

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 typical main office employs on average 115 workers, while a local branch has on average only 16 employees. As labor is the main input of the production process, larger office size translates to higher output. Offices have a large backlog which amounts to 80% of the average quarterly inflow of new claims. While all offices have large backlogs, this phenomenon is more pronounced in main offices. Interestingly, main offices are 12% more productive than local branches on average. Despite these stark differences between main offices and local branches, the quality index and absenteeism rates do not seem to differ substantially across these two types of production sites. Overall, employees devote a very small fraction of their time to training and overtime work (column 1). Hiring is extremely limited in this context, 0.5 workers per office separate from INPS on average every quarter (62% of which are due to retirement), and 0.9 workers transfer to another office within the Social Security Agency (Online Appendix C).

In the remaining part of this Section, I document two stylized facts. First, there is a surprising amount of variation in productivity across offices over time, even within a large centralized agency. Offices located at the 90-th percentile of the productivity distribution are 2.6 times more productive than those at the 10-th percentile (Figure J.IIa). Although comparing productivity differentials across industries is notoriously hard, I benchmark my estimates with previous studies. I compare the distribution of log productivity in my sample (Panel A of Table J.I) 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).

Second, there are large productivity differentials not only across but also within regions. Figure J.III plots the average productivity in each province over my sample period. This figure shows that while the North is more productive on average, there is a substantial variation within each geographical region (for details regarding workload composition across regions and dispersion in sub-components of productivity refer to Online Appendix D).

V Do managers matter in the public sector?

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 that arise in this context and estimate the model. I then perform a series of diagnostic checks which evaluate the performance of my specification in this setting. I conclude this section by summarizing the implications of this model in a variance decomposition exercise and comparing my estimates to the literature.

V.A Preliminary Evidence

If the variation in productivity across offices was entirely driven by time-invariant differences across sites, changes in manager ability should have no impact on office productivity. If instead managers directly affect office-level outcomes, we would expect productivity gains to be positively correlated with changes in managerial ability. To test this claim, I correlate changes in log productivity with changes in manager ability.

I construct the former as the difference between average log productivity in the four quarters after a change in management and average log productivity in the four quarters prior to a change. I proxy for the ability of manager m at office i with the (trend adjusted) leave-i-out mean of log productivity.[12] This exercise does not take sorting into account; in particular, the leave-office-out mean overstates (understates) manager ability for officers who work in highly (un)productive offices. Figure I shows a positive correlation between the two variables of interest (slope 0.211 and SE 0.101), suggesting that improvements in managerial talent are associated with productivity gains.[13]

The pattern looks remarkably linear; this implies that a unit change in managerial talent produces symmetric gains and losses. I classify these switches by the outgoing manager leave-out-mean quartile (group of origin), and the slopes for these four groups look roughly the same (Figure I). In other words, productivity gains associated with a change in leadership do not appear to depend on the talent of the outgoing manager.

V.B Econometric Model

I develop an econometric model that allows me to separately identify the impact of managers and office heterogeneity on productivity. In particular, I posit that log productivity ($\ln P_{it}$) is the sum of a manager component ($\theta_{m(i,t)}$), where $m(i,t)$ is the identity of the specific manager of office i in quarter t, a permanent office component (α_i), a time effect (τ_t), and an error term (u_{it}):

$$\ln P_{it} = \alpha_i + \tau_t + \theta_{m(i,t)} + u_{it}, \tag{2}$$

i and t index office and time, respectively. I interpret $\theta_{m(i,t)}$ as the portable component of managers' ability. I refer to it is as manager quality or managerial talent interchangeably. The office effects (α_i) proxy for the time-invariant characteristics of the office (e.g., geographical location, average quality of the workers at office i, 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

[12] Namely, the average quarterly productivity of other offices during the quarters they were managed by m.

[13] If there was no selection of managers into offices and no measurement error, we would expect a slope of 1. Measurement error biases the coefficient toward zero, while selection can bias it in either direction. In this case, the slope is less than one due to a combination of measurement error in the proxy for managerial talent and sorting. I discuss the implications of sorting for my empirical strategy and document this phenomenon in Subsections V.B and V.D, respectively.

effects to absorb seasonality and macroeconomic shocks.

Model (2) postulates that productivity changes discreetly as a new manager takes over. However, in reality, managers may take some time to change work practices. As I do not want the estimated manager effects to be confounded by switching costs or measurement error in manager identity[14], I estimate (2) excluding the first quarter in which the new manager is in charge.

I can re-write (2) in matrix notation as:

$$\ln(P) = D\alpha + G\theta + T\tau + u, \tag{3}$$

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:

$$E[d_i' u] = 0 \ \forall i, \tag{4}$$

$$E[g_m' u] = 0 \ \forall m, \tag{5}$$

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

It is well known that the office and manager effects in (2) are identified by movers. As I can separately identify manager from office effects only within a connected set (Abowd et al., 1999), I can meaningfully compare the estimated fixed effects only within and not across connected sets (Online Appendix E). The identifying assumptions (4) and (5) impose that manager mobility is as-good-as random, conditional on office and time fixed effects.

Loosely speaking, these orthogonality conditions are satisfied if the assignment of managers to offices depends only on the permanent component of office productivity (α_i) and/or the permanent component of managerial ability ($\theta_{m(i,t)}$). For example, better managers sorting into more productive offices would not violate the identifying assumptions. By the same token,

[14] If the switch does not occur on the first day of the quarter, I assign the quarter of the switch to the manager with the longest spell in that quarter.

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. As Best et al. (2019) note, (4) and (5) allow for rich patterns in the sorting of managers to offices. Violations of the exogenous mobility assumption occur when managers sort on the error term.

I follow Card et al. (2013) and consider three forms of endogenous mobility that depend on any office-manager match component of productivity ($\eta_{im(i)}$), on any office-specific trend in productivity (ζ_{it}), or on any transitory component of office productivity (ϵ_{it}). In particular, I specify the following composite structure of the error term:

$$u_{it} = \eta_{im(i)} + \zeta_{it} + \epsilon_{it}, \tag{6}$$

I assume that $\eta_{im(i)}$ has mean zero for all offices and all managers in the sample. ζ_{it} is a drift component, which captures offices improving or deteriorating over time; I assume this component has a mean zero for each office but contains a unit root. ϵ_{it} is an idiosyncratic error term and represents transitory shocks; I assume that ϵ_{it} has mean zero for each office.

Given the error structure posited in equation (6), the assumptions in equations (4) and (5) rule out three types of sorting. First, managers can not sort into offices on the basis of their comparative advantage. Second, better managers cannot be systematically sent to offices whose performance is worsening over time (i.e., assortative matching based on underlying trends in productivity). Third, a better manager cannot join an office in response to a negative transitory productivity shock (Brown, 1982).

I use the estimated manager effects and the non-parametric proxy for managerial talent described in Subsection V.A to conduct a series of tests for the presence of endogenous mobility in Subsection V.D.

V.C Results

Table III describes the structure of my sample of quarterly level observations on office-level productivity. The first column reports the statistics for the full sample, while the second column restricts attention to the balanced-analysis sample. 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. 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 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).

In order to assess the amount of dispersion in public sector productivity attributed to managers, I follow Bertrand and Schoar (2003). I compare the adjusted R^2 estimated from a regression of the logarithm of productivity on office and time fixed effects, model (7), to the one from model (2) which also includes manager fixed effects.

$$\ln P_{it} = \alpha_i + \tau_t + \tilde{u}_{it}. \tag{7}$$

Model (2) nests (7) under the assumption that managers have no impact on office productivity. Table IV reports the estimates from (7) and (2) in columns 2 and 3, 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).

To test more formally whether managers affect productivity, I perform an F-test for the null hypothesis that the manager effects are jointly zero. I reject the null at any standard significance level (p-value=0.000). Notice that the adjusted R^2 of columns 2 and 4 are high relative to the one reported in column 3, which suggests that manager and office fixed effects are highly correlated in this setting. These R^2's are lower than those of two-way fixed-effect models that decompose wages (Card et al., 2013). The reason is that productivity is intrinsically more volatile than wages.[15]

V.D Diagnostic Checks

In Section II I argued that the institutional framework severely limits the ability of managers to sort into offices. I now test for detectable evidence of this phenomenon focusing on patterns that might be related to the components of the error specified in equation (6). First, I discuss sorting on the drift component (i.e., ζ_{it} in equation 6). Second, I address concerns related to managers being assigned to offices on the base of unobservable factors determining their comparative advantage (i.e., the term $\eta_{im(i)}$ in equation 6). Third, I consider sorting on the transitory component of the error term (i.e., the term ϵ_{it} in equation 6).

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. Intuitively, managers cannot impact office performance before they take charge, hence any correlation between future manager ability and baseline office characteristics is indicative of sorting. As a benchmark, if managers were randomly assigned one would expect manager productivity to be uncorrelated with observable pre-determined characteristics of the office.

The results in column 1 of Table V show that, more productive managers are less likely to serve in main offices and more likely to be assigned to Northern or Central Italy. Importantly, future manager effects do not appear to be correlated with office growth rates. I also test

[15] For completeness, I report the same exercise using quarterly data in Online Appendix J (Table J.II). Although quarterly productivity is a somewhat noisier outcome, the results are largely unchanged.

whether the explanatory variables are jointly statistically significant and whether growth rates can jointly predict future manager fixed effects. I can reject the latter but not the former. Overall, there is some evidence of managerial sorting on geography and office type, but not on growth rates. As I discussed earlier, manager assignment being correlated with time-invariant characteristics of the office does not pose a threat to my empirical strategy. 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. These findings show that there is no evidence of managers sorting on the drift component. They also further motivate the use of office fixed effects in my main specification to control for sorting based on time-invariant characteristics of the office.[16]

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 (2) associated with specific forms of manager changes. When an office goes through a change in management, it can experience three types of transitions: an overall increase in manager ability ($\widehat{\Delta M}_i > 0$, where $\widehat{\Delta M}_i$ represents the change in the estimated manager fixed effects), a decrease in management quality ($\widehat{\Delta M}_i < 0$), or no significant change ($\widehat{\Delta M}_i \approx 0$). Figure II reports the event study for (trend-adjusted) office productivity for these types of transitions (i.e., tertiles of $\widehat{\Delta M}_i$).[17] Log productivity remains relatively flat in the four quarters before the change in management and jumps discontinuously at the time of the event. 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 VI, and I do not find evidence of this phenomenon. Importantly, the fact that productivity appears to be slightly lower in the quarter of the switch than in the following three quarters, motivates my choice of excluding the quarter in which the takeover takes place from my two-way fixed-effect model.

Next, I test for sorting on the match component of productivity by examining gains and losses as managers move from office to office. As already noted, Figures I and II display a

[16] These results are robust to the exclusion of the lagged variables (Table J.III in the Online Appendix).

[17] Figure II is constructed averaging (trend adjusted) log productivity by event time and transition type.

remarkably symmetric pattern, which is not consistent with managers sorting on their idiosyncratic match component (i.e., sorting based on comparative advantage of specific managers at specific offices). I also compare the fit of model (2) with a fully saturated model that includes manager-by-office dummies. In the presence of match components, the latter should fit substantially better than the former. The adjusted R^2 of the fully saturated model is only marginally higher than the two-way fixed effect specification (0.764 in column 5 vs. 0.762 in 3 of Table IV), suggesting that match components are not quantitatively relevant in this context.

Finally, manager rotation could be correlated with the transitory component of the error term. This could be the case if managers were to relocate to a less productive office after a particularly bad ϵ_{it} draw. Once again, this is not consistent with the lack of pre-trends reported in Figure II and in Section VI.[18]

A final set of concerns about model (2) 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. I compute the mean residual for each cell. Figure III reports these statistics. All values are rather low, and the highest mean residual is equal to 0.01.[19] 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.

V.E Variance-Covariance Decomposition

We might expect social norms and workforce composition (proxied by office fixed effects) to be important drivers of productivity. However, it is less obvious whether public sector managers,

[18]Refer to Online Appendix B for a discussion of how the bonus structure can induce managers to sort into different office types and whether this is a threat to the empirical strategy.

[19]The mean residuals have been computed using all the connected sets in which there are at least four offices and four managers. Managers and offices are ranked within a connected set. Figure J.IV reports the same exercise on the largest connected set.

who operate in a severely constrained environment, can have an impact on productivity. I propose a variance decomposition exercise that allows me to assess whether these two dimensions matter and their relative importance. 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. I use (2) to decompose the variance of productivity into (Abowd et al., 1999):

$$\begin{aligned} Var(\ln P_{it}) &= Var(\alpha_i) + Var(\theta_{m(i,t)}) + Var(\tau_t) + Var(u_{it}) \\ &\quad + 2Cov(\theta_{m(i,t)}, \alpha_i) + 2Cov(\theta_{m(i,t)}, \tau_t) + 2Cov(\alpha_i, \tau_t). \end{aligned} \tag{8}$$

Table VI reports the bias-corrected variances and covariances estimated on the largest connected set (Andrews et al., 2008; Gaure, 2014). This procedure allows me to obtain consistent estimators of the variances and covariances of interest in the presence of limited mobility bias. 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, which is mainly driven by seasonality in productivity and the overall improvement in the Social Security Agency performance over time.

Interestingly, the bias-corrected covariance between manager and office effects is negative, namely more productive managers currently work at less productive offices (i.e., negative assortative matching). This finding is crucial for the interpretation of the counterfactual exercises I develop in Section VIII. It is worth emphasizing that this result is somewhat unusual. Most economic models predict positive assortative matching and the recent literature on wage determination suggests that higher-wage workers tend to sort to firms that offer higher wage premiums (Card et al., 2013; Song et al., 2015). This result is consistent with INPS trying to reduce inequality in productivity across sites.[20]

[20]Several papers have found evidence of negative assortative matching using the two-way fixed effects framework. However, these findings may be tainted by limited mobility bias. I refer to Andrews et al. (2012) for a complete treatment of this issue. A recent study by Adhvaryu et al. (2020) documents negative assortative

V.F Manager Effects and Observable Characteristics

I conclude this Section by discussing the magnitude of the estimated manager effects and showing how they correlate with observable characteristics of top-level officers.

Managers have a large impact on office performance. A two standard deviation increase in managerial talent leads to a 20% increase in office productivity (Table VI). Although it is challenging to directly compare point estimates across industries and countries, I benchmark the magnitude of this effect with the work of Bloom et al. (2013). The authors find that the adoption of management practices induces a 17% increase in productivity in textile firms in India. My results suggest that a very good manager has a comparable effect on productivity in the Italian government.

To study how estimated managerial talent correlates with observable characteristics, I regress the estimated manager fixed effects from (2) on gender, experience, experience squared, a set of dummies for region of birth, and a set of indicators for the highest educational attainment, as well as connected set fixed effects (Table VII). Female managers appear to be on average more productive than their male counterparts. Not surprisingly, managerial talent is strongly correlated with experience although it exhibits decreasing returns to experience. There is some suggestive evidence that managers born in Southern Italy or the Islands are more productive than those from the North and that those who never attended college are better than those who studied law or STEM. Importantly, these coefficients should not be interpreted causally; these correlations can be explained by differential selection patterns into public sector jobs and managerial career. In particular, 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 there is substantial variation in managerial talent within this large centralized agency and managers have a quantitatively meaningful impact on office

matching between managers and production lines in an Indian garment factory and argue that it is driven by the strong incentives in reducing delays on any particular order. The work of Limodio (2019) also suggests that high performing World Bank managers often work in poorly performing countries and that the negative assortative matching strengthens in the aftermath of a natural disaster.

productivity. Next, I open the black box of manager fixed effects and try to analyze the specific mechanisms that drive the effects of more and less productive managers.

VI What makes for a productive manager?

Better managers could affect office productivity through a variety of mechanisms that include better personnel decisions, more competent management of office operations, and eliciting more effort from workers by motivating and monitoring them appropriately. Given the institutional constraints discussed in Section II, managers are unlikely to have an impact on hiring and firing. However, they can in principle affect office operations by changing workers' time allocation (e.g., training, overtime hours, absenteeism rate) and the assignment of tasks to workers. As employees enjoy strong job security, soft skills may be particularly important when eliciting effort from workers. In this section, I utilize manager rotations as a quasi-experimental analog of random assignment of managers to offices to 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 construct a covariate index that allows me to estimate the extent to which the productivity gains can be explained by observable characteristics. Fourth, I evaluate the productivity-quality trade-off. Finally, I test for heterogeneous treatment effects.

VI.A Event Study Strategy

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

$$y_{it} = \alpha_i + \sum_{k \neq 1} \left[\pi_0^k D_{it}^k + \pi_1^k D_{it}^k \Delta M_i \right] + h_t \left(X_{it} \right) + \varepsilon_{it} \tag{9}$$

where k indexes quarters relative to a change in management (positive values of k refer to quarters after the event and negative values to those prior) and the π_0^k 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 the extent to which effects differ depending on the change in quality of incoming managers relative to the managers they replace (i.e., ΔM).[21] Permanent differences in office productivity are captured by α_i and $h_t(X_{it})$ controls flexibly for time trends. The identifying assumption is that changes in management quality are not coincident with the evolution of other unobservable factors. Event study frameworks of the type described in equation (9) are typically estimated via OLS, and the identifying assumption is tested by evaluating the pre-trends.

If manager effectiveness were observable to the econometrician, equation (9) could be estimated directly. Managerial talent is fundamentally unobservable but the two-way fixed effects model enables me to estimate it by exploiting the rotation of managers across sites. However, using the first-step estimates as covariates in (9) could bias π_1^k. Idiosyncratic productivity shocks could affect both my estimates of manager effectiveness and the outcome of interest, creating a spurious correlation even in the absence of a causal relationship. A natural solution is to purge idiosyncratic shocks by estimating the first step leaving out data where correlations may arise. I overcome this challenge by modifying the standard event study specification in two ways. First, I subtract from equation (9) in each event time the corresponding values in event time $k = -1$. Second, to purge the regressor of potential mechanical correlations, I generate the change in manager productivity by using estimates from separate two-way fixed effect models, *excluding* data from event times -1, 0, and k.[22] Formally,

$$\Delta y_i^k = \pi_0^k + \pi_1^k \widehat{\Delta M}_i^{L,k} + \Gamma^k X_i + \Delta \epsilon_i^k \tag{10}$$

[21]Equation (9) implicitly assumes that there is at most one change in leadership per office and it can be easily modified to include multiple events.

[22]As described in Section V, managers may take some time to change work practices. Hence, I never include the quarter of the switch (i.e., $k = 0$) when estimating manager effects via (2).

where

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

and the $\hat{\theta}^{L}_{i,\cdot}$'s are the leave-out estimated manager effect of the incoming and outgoing managers, respectively (where L superscript stands for "leave-out").[23] 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 (10) suggests estimating separate regression models for each event time. I focus on the $[-4, 6]$ window, and I limit my sample to the subset of events which are balanced over this time horizon (i.e., balanced-analysis sample).[24] I bootstrap the standard errors to account for the presence of a generated regressor.

The separate regression models and leave-out procedure ensure that $\widehat{\pi_1^k}$'s are not spuriously driven by contemporaneous, idiosyncratic shocks that affect both the estimated manager effects and the outcomes of interest. The underlying parameters of the differenced model and the standard model in equation (9) are nonetheless the same, allowing me to directly test the over-identifying restrictions of parallel trends. I interpret violations of the parallel pre-trend assumption as evidence of time-varying unobservable confounding factors.

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. While leaving out more data in the first-step mitigates concerns relative to mechanical correlation, it also increases measurement error in my estimates of manager effectiveness. Given the limited number of years in my sample, I do not have enough data to pursue a leave-office-out estimation strategy.[25] Reassuringly, serially correlated productivity shocks do not appear to be a concern

[23] $\widehat{\Delta M}_i^{L}$ ranges from -0.4 to 0.55 and looks approximately normally distributed (Figure J.V). Crucially, $\widehat{\Delta M}_i^{L}$ does not depend on the normalization I choose (Online Appendix E).

[24] I can not expand my window further as the balanced sample becomes excessively thin.

[25] A leave-office-out strategy allows me to examine only switches where both the incoming and outgoing managers are movers. Given the nature of my data, this requirement excessively reduces my sample size.

in my setting. First, the autocorrelation coefficient from fitting an AR(1) model to the residuals from equation (2) is extremely small ($\hat{\rho} = 0.04$). Second, in simulation analyses, I hold constant the mobility structure from my sample and generate simulated outcomes suffering from different degrees of autocorrelation. I show that a substantial degree of autocorrelation is needed to represent a serious threat to my empirical strategy (Online Appendix F).

VI.B Decomposition of Productivity Impacts

The two-way fixed effect model documents that managers matter and parsimoniously summarizes their contribution in a single measure. However, this measure does not give any insight into the mechanisms and their timing. I begin the analysis of the mechanisms by examining the timing of the productivity gains; next, I decompose them into their effect on output and full-time equivalent employment.

Figure IVa reports the estimated impact of an increase in managerial quality on office-level productivity.[26] This figure collects the estimated coefficients and their 95% bootstrapped confidence intervals obtained by running a separate regression for each time horizon (k).[27] Reassuringly, changes in managerial talent do not predict changes in productivity before the event takes place (placebo tests), which alleviates concerns regarding endogenous mobility. Productivity starts increasing when the better manager takes office and it stabilizes to its new level one quarter after the change. These results show that improving management quality increases productivity, but that the effect does not fully materialize in the first quarter. This is consistent with the presence of adjustment costs associated with manager rotation.

Previous research finds that effective managers in the private sector 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 output may be particularly challenging. Instead, good managers may keep production high while reducing costs. Since productivity is constructed as the ratio

[26] Table J.IV reports the results of Figures IVa, IVb, and IVc in a table format.

[27] I bootstrap standard errors and confidence intervals over 1,000 replications.

between output and full-time equivalent employment, I can decompose the impact of managerial talent on office productivity into its effect on output and FTE. When a better manager takes over there is a modest (although not statistically significant) increase in output (Figure IVb). This pattern is consistent with better managers using resources more effectively, motivating their employees, and monitoring them more closely. 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 IVc). Most of the productivity gains are driven by the reduction in the number of workers assigned to the office. In particular, a 10% increase in managerial talent (i.e., $\widehat{\Delta M}_i^L = 0.1$) increases office output by 1.7% and generates a 4.9% reduction in full-time equivalent employment six quarters after the event (Table J.IV , column 2 and 3, respectively). This finding speaks directly to the consequences of downsizing the public sector and suggests that reducing the number of public sector employees does not necessarily lower the volume of services provided.

VI.C Reduced Form Impacts of Managerial Talent on Workers

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.

Better managers induce older workers to retire (column 1).[28] This effect is concentrated in the first two quarters after the change in leadership. Importantly, managers and higher level officials cannot force older workers to retire and they can not negotiate exit packages or golden parachutes to persuade them to leave (see Online Appendix C for a description of the retirement system). In light of these constraints, it seems most plausible that the experience of working under a more productive manager causes some older employees who are near to or past standard retirement age to retire rather than continue working under the new regieme.

[28] Inverse hyperbolic sine is defined at zero. I choose this functional form over the logarithmic one whenever my outcome variable can take value zero. This is an alternative specification to $\ln(x+1)$ and I abbreviate $Asinh(x)$ as $A(x)$.

As INPS is characterized by a rather senior workforce, inducing workers past retirement age to retire may be a sustainable way to increase productivity in this context.[29] Unsurprisingly given the limited ability of managers to hire and fire workers, an increase in management quality does not have any 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), although the latter is only marginally statistically significant.

I also investigate how time allocation changes with the takeover of a more productive manager. Changes in managerial quality do not appear to produce strong and persistent effects on training, overtime work, and total hours (Online Appendix G), while there is some suggestive evidence that better managers may be able to reduce abstenteeism rate. Remarkably, more productive managers succeed in keeping up production without resorting to more overtime hours to compensate for the reduction in FTE (column 3 of Table G.I). This is consistent with more productive managers being able to better match workers with tasks.

VI.D Share of Productivity Gains Explained by Observable Characteristics

I construct a covariate index to estimate the share of productivity gains explained by observable characteristics and evaluate the relative importance of worker composition versus time allocation. I regress $\ln P_{it}$ on office and time fixed effects and a set of covariates[30] and I construct my covariate index as the predicted values of this model.

To evaluate what components of the covariate index are more responsive to changes in manager quality, I generate predicted values using only office and time fixed effects in column 2 ($\widehat{\ln P1_{it}}$) and include progressively larger subsets of estimates. Column 3 adds covariates that are a function of the office age structure ($\widehat{\ln P2_{it}}$), and column 4 includes the gender composition

[29] The average INPS employee was 55 years old in 2017.

[30] The set of covariates includes share of employees in each of the 10 deciles of the age distribution, average office age, fraction female, fraction of female among top-officials, ln FTE, asinh(absences), asinh(over-time), and asinh(training). I include a linear and a quadratic term for each of these covariates as well as their two-way interactions with time fixed effects and a dummy for main offices.

of the office ($\widehat{\ln P3_{it}}$).

Changes in the observable characteristics of the office explain 56% of productivity gains six quarters after the event (column 1 of Table IX).[31] Although these changes jointly explain a significant share of the productivity gains, there are some mechanisms through which managerial talent operates I can not account for. These include motivating and monitoring workers as well as reassigning tasks within the office. Column 1 shows no evidence of pre-trends and displays a sharp increase in the first quarter the new manager takes office. These patterns suggest that changes in the observable characteristics of the office were not ongoing before the event took place and productivity gains are indeed driven by the change in management. Changes in the demographic composition of the office explain 13.7% of the overall effect attributable to observable characteristics (column 3 of Table IX), while changes in the number of workers and in the time allocation explain the remaining 86.3% of the variation (column 4 of Table IX).[32]

VI.E Reduced Form Impacts of Managerial Talent on Quality and Backlog

One might fear that there is some trade-off between productivity and quality of service provided. Column 1 of Table X shows that the arrival of a more productive manager does not negatively impact quality. This result is likely to be driven by the fact that the incentive-pay scheme provides a direct incentive for managers to increase productivity without letting quality deteriorate. There is also some suggestive evidence that effective managers lower backlog (column 2 of Table X). In light of the work by Autor et al. (2015), this result points toward potentially large benefits to the claim beneficiaries driven by a reduction in processing time.

[31] I construct the share of productivity gains explained by observable characteristics taking the ratio of the predicted change in log productivity (column 1 of Table IX) and the change in ln productivity (column 1 of Table J.IV) six quarters after the event. Namely, 0.371/0.661=56%.

[32] 13.7% is obtained as the ratio of 0.047 (column 3) and 0.371 (column 1). 86.3% is its complement to 1.

VI.F Heterogeneity

A productive manager might be able to have a larger impact on an unproductive rather than on a very productive office. Likewise, she could be more effective in a smaller than in a larger site, or in some specific geographical areas. Productivity gains do not appear to differ by geographical location, office size, office type, baseline productivity and social capital, although I have admittedly limited power to detect heterogeneous treatement effects (see Online Appendix H).[33]

VII Robustness Checks

In this section, I address some concerns regarding my empirical strategy. I first show that manager effects are not confounded by demand shocks. Second, I test whether managers appear to game the system. Third, I relax the assumption that manager effects translate linearly in office level outcomes.

VII.A Demand

A concern might be that high-fixed effect managers are classified as "productive" because they happen to be working at offices that received a site-specific positive demand shock. Demand for services is measured as the number of claims that originate from the office catchment area and it is exogenous to the office, as it depends on the demographic characteristics of those living in the catchment area as well as its economic condition. Unlike managers in the private sector, top-level bureaucrats cannot advertise their products or take actions to affect the local demand for public services.

I use model (10) to test whether demand is higher on average when a more productive manager is in charge. Column 3 of Table X shows that more productive managers are not in

[33]These findings are in line with the absence of heterogeneous treatment effects postulated by my two-way fixed effects model (see 2).

charge during periods of high demand. Demand appears to be extremely volatile even when controlling for time fixed effects, which is consistent with the fact that it is mostly driven by local idiosyncratic shocks. As an additional robustness check, I correlate my estimated manager fixed effects with those estimated by controlling for the logarithm of the number of claims originated in each quarter. The correlation is extremely high (98.8%). I conclude that the estimated manager effects do not appear to be confounded by demand shocks.

VII.B Do Managers Game the System?

As bonuses are a strictly increasing function of productivity, one could worry that the managers I classified as productive are, in fact, those who are able to game the system. In particular, if the weights are mismeasured, managers might try to shift production toward the overvalued products and shift away from the undervalued ones.[34]

First, manipulation should be mitigated by the fact that if managers decided not to process undervalued products in a timely fashion, this would reflect negatively on the quality index and, in turn, on their bonuses. Quality is not negatively impacted when a more productive manager takes over (column 2 of Table X), which attenuates these concerns. Second, 5% of all claims processed by each office are audited twice per year; the purpose of this cross-check is to monitor the production process and detect anomalies or illicit behavior. Managers are responsible for the claims processed under their watch and are held personally accountable. Third, if the backlog of a specific type of claim increases across multiple offices, INPS reassesses the weight associated with that product.

In order to test more formally whether managers game the system, I divide all the products into nine categories, and I estimate whether the number of claims processed in each of these categories changes differentially when a better manager takes office.[35] I interpret shifts toward a product category or away from it as evidence of gaming.

[34] Notice that if the weights are measured correctly, there is no scope for gaming.

[35] 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.

More specifically, I test the presence of such behavior using a generalized difference-in-difference approach which leverages the staggered timing of manager rotation:

$$y_{it} = \mu_i + \nu_t + \sum_{k=-\underline{K}}^{\bar{K}} \beta^k D_{it}^k + \sum_{k=-\underline{K}}^{\bar{K}} \delta^k D_{it}^k \times \widehat{\Delta M}_i + \iota X_{it} + u_{it}, \tag{11}$$

where y_{it} represents the outcome of interest for office i at quarter t, μ_i and ν_t are office and time fixed effects, respectively. u_{it} is an idiosyncratic error term. D_{it}^k is a dummy variable which indexes event time it is defined as $D_{it}^k = \mathbb{1}(e_i + k)$, where e_i is the quarter in which the change in leadership takes place. I measure treatment intensity associated with the switch as the difference between the incoming and outgoing estimated manager effects, i.e., $\widehat{\Delta M}_i = \hat{\theta}_{incoming} - \hat{\theta}_{outgoing}$. X_{it} includes time fixed effects interacted with a dummy for Center-North, time fixed effects interacted with a dummy for main office, and demand for the nine broad product categories. Controlling for demand is important in this context because it controls for the shifts toward some products dictated by external factors that are not under the control of the manager. δ^k represents my main parameter of interest. It captures the change in y_{it} associated with changes in managerial ability (i.e., $\widehat{\Delta M}_i$) k periods after the event.

Figures Va and Vb show that there is no evidence of productive managers shifting the production mix. Overall, I find no evidence consistent with managers gaming the system.

VII.C Quartile Specification

Model (10) assumes that productivity gains are a linear function of changes in managerial talent. In this subsection, I relax this assumption and propose an alternative exercise where I divide my measure of changes in manager ability (i.e., $\widehat{\Delta M}_i^L$) into four quartiles. I estimate the impact of the change in management for each of these groups:

$$\Delta y_{it}^k = \beta_0^k + \sum_{v=2}^{4} \beta_v^k \times Q_{iv} + \Delta\tau + \psi^k \Delta X_{it} + \Delta\epsilon_{it}^k. \tag{12}$$

Let Q_{iv} be a dummy that takes a value of one if office i belongs to the v-th quartile of the $\widehat{\Delta M}_i^L$ distribution. All the other variables are defined as above. Since I omit the first quartile, β_v^k identifies the difference between offices belonging to the v-th and the first quartile at event time k. I iterate over the values of k as described in Section VI.

Figure VIa shows that the higher the treatment intensity (i.e. $\widehat{\Delta M}_i^L$), the larger the treatment effect. Figures VIb and VIc display the same pattern. These findings suggest that the linear specification is not a poor approximation of the data.

VIII Counterfactual Exercises

In this section, I discuss how governments could use these findings to improve public service provision. I use my estimates to construct counterfactual exercises that illustrate the magnitude of manager effects and evaluate the efficiency gains from alternative managerial allocation schemes.

I assume that the social planner maximizes the aggregate agency productivity and that she cannot directly influence the permanent office component of productivity (α_i) and the number of workers assigned to the office (w_i). She can, however, 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 some time-invariant office characteristics (α_i) and on the ability of the manager ($\theta_{m(i)}$). Let w_i represent the full-time equivalent number of workers assigned to the office. Assuming $\ln P_i\left(\alpha_i, \theta_{m(i)}\right) = \alpha_i + \theta_{m(i)}$, the planner's objective function is:

$$\max_{\underline{\theta}} \sum_i \gamma_i e^{\theta_{m(i)}}, \tag{13}$$

where $\gamma_i = e^{\alpha_i} w_i$. I choose an additively separable model as I have shown that it approximates the data fairly well and it fits naturally in the framework I have developed.

I consider four counterfactual policies that the social planner can implement: 1) She can maximize (13) by reassigning existing managers to offices; 2) She can fire the bottom 20% of managers and substitute them with the median manager (but allocate them as in the current environment); 3) She can implement both policies at once; and 4) She can randomly assign managers to offices.[36] Being able to reassign managers within but not across connected sets puts additional constraints on (13).[37] As the optimal solution to the constrained problem can never exceed that of the unconstrained one, the estimated impact of the first and third intervention represent a lower bound on the true effect.

$P_i w_i$ is twice differentiable and supermodular, therefore the optimal allocation is an assortative matching equilibrium where the best managers are sent to offices which are both productive and large (Becker, 1974), where γ_i implicitly weights these two dimensions. Such an allocation exacerbates productivity inequality across sites. Traditionally, the argument about equalizing quality of services across offices relies on the idea that beneficiaries should not receive a different treatment depending on where they live. As claims can be easily redistributed across sites and processed anywhere, it is unclear why productivity should be equalized across offices in this setting.

Table XI reports the efficiency gains from alternative managerial allocation schemes. If the social planner reassigns managers using the optimal allocation rule, aggregate productivity increases by at least 6.9%. If instead, she fires the bottom 20%, aggregate productivity raises only by 2.9%. In this setting, the first policy is more effective than the second because there are strong complementarities between managerial talent and the permanent component of office productivity and, as the most productive managers are currently allocated to the least productive offices (Table VI), there is quite some scope for reallocation. Implementing both these policies simultaneously increases aggregate productivity by at least 7.4%, corroborating the finding that managerial allocation is key in this context. Figure VIIa, VIIb, and VIIc illustrate these concepts graphically. Last but not least, I randomly reassign managers to offices

[36] As the second policy requires the social planner to fire the bottom 20% of managers, I drop all connected sets with less than five managers in all the counterfactual exercises.

[37] As my sample contains multiple connected sets, I implement each policy within each connected set. Connected sets reflect the mobility patterns in my sample and often overlap with broad geographical regions (Appendix A).

and I take the average of the implied productivity gains and losses over 1,000 iterations. If managers were randomly reassigned, aggregate productivity would increase by 2%. This is because random assignment moves the allocation closer to the socially optimal one by undoing the negative assortative matching equilibrium.[38] In practice, reallocating managers across sites is feasible and INPS has experimented with it in 2019 when Tridico was appointed to INPS President. However, hiring and firing top-level officials is extremely challenging in the Italian legal framework and it is unlikely to be a viable policy option.

IX Conclusion

Managerial practices are positively correlated with public good provision across a variety of settings (Bloom et al., 2015; Tsai et al., 2015; Rasul and Rogger, 2018; Rasul et al., 2019), yet there is very little evidence on whether managers affect the performance of public sector organizations.

This paper studies the productivity impacts of managers in the public sector 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 and that the rise in productivity associated with the arrival of a more productive manager is mainly driven by the exit of older workers. Importantly, a good manager sustains production levels without resorting to hiring or overtime to compensate for the decrease in full-time equivalent employment. These results speak directly to the debate on downsizing the public sector and they suggest that managerial talent can go a long way in sustaining adequate public service provision in a context where the workforce is shrinking. I also discuss how governments could use these findings to improve public service provision by evaluating alternative managerial allocation schemes. I find that complementarities between managerial talent and the permanent component of office productivity play a key role in this

[38] Table J.V and Figures J.VIa, J.VIb, and J.VIc report the estimates computed on the largest connected set (Online Appendix J). Reassuringly, the pattern of results is unchanged.

setting. These results suggest that there may be large social returns to rigorously modelling public sector productivity and the impacts of managerial abilities. Governments should design policies to attract, retain, and properly allocate managerial talent within the public sector.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Adhvaryu, Achyuta, Vittorio Bassi, Anant Nyshadham, and Jorge A Tamayo, "No Line Left Behind: Assortative Matching Inside the Firm," Working Paper 27006, National Bureau of Economic Research April 2020.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Economics Letters*, 2012, *117* (3), 824 – 827.

_ , _ , Thorsten Schank, and Richard Upward, "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society Series A*, 2008, *171* (3), 673–697.

Autor, David H., "The "task approach" to labor markets: an overview," *Journal for Labour Market Research*, Sep 2013, *46* (3), 185–199.

_ , Lawrence F. Katz, and Melissa S. Kearney, "The Polarization of the U.S. Labor Market," *American Economic Review*, May 2006, *96* (2), 189–194.

_ , Nicole Maestas, Kathleen J Mullen, and Alexander Strand, "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants," Working Paper 20840, National Bureau of Economic Research January 2015.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun, "CEO Behavior and Firm Performance," *Journal of Political Economy*, 2020, *128* (4), 1325–1369.

Becker, Gary S., *A Theory of Marriage*, University of Chicago Press,

Bennedsen, Morten, Francisco Perez-Gonzalez, and Daniel Wolfenzon, "Do CEOs Matter?," Working Paper September 2010.

_ , _ , and _ , "Estimating the Value of The Boss: Evidence from CEO Hospitalization Events," Working Papers August 2011.

Bertrand, Marianne and Antoinette Schoar, "Managing with Style: The Effect of Managers on Firm Policies," *The Quarterly Journal of Economics*, 2003, *118* (4), 1169–1208.

_ , Robin Burgess, Arunish Chawla, and Guo Xu, "The Glittering Prizes: Career Incentives and Bureaucrat Performance," *The Review of Economic Studies*, 05 2019, *87* (2), 626–655.

Best, Michael, Jonas Hjort, and David Szakonyi, "Individuals and Organizations as Sources of State Effectiveness," Working paper October 2019.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ , Aprajit Mahajan, David McKenzie, and John Roberts, "Do Management Interventions Last? Evidence from India," Working Paper 24249, National Bureau of Economic Research January 2018.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

_ , Renata Lemos, Raffaella Sadun, and John Van Reenen, "Does Management Matter in schools?," *The Economic Journal*, 2015, *125* (584), 647–674.

Bohlmark, Anders, Erik Gronqvist, and Jonas Vlachos, "The Headmaster Ritual: The Importance of Management for School Outcomes," *The Scandinavian Journal of Economics*, 2016, *118* (4), 912–940.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin, "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals," Working Paper 17803, National Bureau of Economic Research February 2012.

Brown, M. Craig, "Administrative Succession and Organizational Performance: The Succession Effect," *Administrative Science Quarterly*, 1982, *27* (1), 1–16.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar, "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico," *Journal of Political Economy*, 2018, *126* (2), 635–687.

Card, David, Jorg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Cartocci, Roberto, *Mappe del tesoro*, Bologna: Il Mulino, 2007.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson, "Productivity Dispersion in Medicine and Manufacturing," *American Economic Review*, May 2016, *106* (5), 99–103.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, September 2014, *104* (9), 2593–2632.

_ , _ , and _ , "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, September 2014, *104* (9), 2633–79.

Choudhury, Prithwiraj, Tarun Khanna, and Christos A Makridis, "Do Managers Matter? A Natural Experiment from 42 RandD Labs in India," *The Journal of Law, Economics, and Organization*, 12 2019, *36* (1), 47–83.

Coelli, Michael and David A. Green, "Leadership effects: school principals and student outcomes," *Economics of Education Review*, 2012, *31* (1), 92 – 109.

Dhuey, Elizabeth and Justin Smith, "How important are school principals in the production of student achievement?," *Canadian Journal of Economics/Revue canadienne d'Acconomique*, 2014, *47* (2), 634–663.

Eeckhout, Jan and Philipp Kircher, "Identifying Sorting—In Theory," *The Review of Economic Studies*, 03 2011, *78* (3), 872–906.

Finan, F., B.A. Olken, and R. Pande, "Chapter 6 - The Personnel Economics of the Developing State," in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 2 of *Handbook of Economic Field Experiments*, North-Holland, 2017, pp. 467 – 514.

Frederiksen, Anders, Lisa B Kahn, and Fabian Lange, "Supervisors and Performance Management Systems," Working Paper 23351, National Bureau of Economic Research April 2017.

Gaure, Simen, "Practical Correlation Bias Correction in Two-way Fixed Effects Linear Regression," Memorandum 21/2014, Oslo University, Department of Economics 2014.

Giorcelli, Michela, "The Long-Term Effects of Management and Technology Transfers," *American Economic Review*, January 2019, *109* (1), 121–52.

Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb, "Using Student Test Scores to Measure Principal Performance," *Educational Evaluation and Policy Analysis*, 2015, *37* (1), 3–28.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, June 2004, *94* (3), 526–556.

Hoffman, Mitchell and Steven Tadelis, "People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis," Working Paper 24360, National Bureau of Economic Research February 2018.

Holmstrom, Bengt and Paul Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law Economics and Organization*, 1991, *7*, 24–52.

Janke, Katharina, Carol Propper, and Raffaella Sadun, "The Impact of CEOs in the Public Sector: Evidence from the English NHS," Working Paper May 2019.

Kane, Thomas J. and Douglas O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," Working Paper 14607, National Bureau of Economic Research December 2008.

Khan, Adnan Q., Asim Ijaz Khwaja, and Benjamin A. Olken, "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings," *American Economic Review*, January 2019, *109* (1), 237–70.

Lavy, Victor and Adi Boiko, "Management Quality in Public Education: Superintendent Value-Added, Student Outcomes and Mechanisms," Working Paper 24028, National Bureau of Economic Research November 2017.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton, "The Value of Bosses," *Journal of Labor Economics*, 2015, *33* (4), 823–861.

Limodio, Nicola, "Bureaucrat Allocation in the Public Sector: Evidence from the World Bank," Working Paper October 2019.

Lucas, Robert E., "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 1978, *9* (2), 508–523.

Perez-Gonzales, Francisco, "Inherited Control and Firm Performance," *American Economic Review*, December 2006, *96* (5), 1559–1588.

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

_ , _ , and Martin Williams, "Management and Bureaucratic Effectiveness: Evidence from the Ghanaian Civil Service," Working Paper July 2019.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," Working Paper 21199, National Bureau of Economic Research May 2015.

Stinebrickner, Todd R, Ralph Stinebrickner, and Paul J Sullivan, "Job Tasks and the Gender Wage Gap among College Graduates," Working Paper 24790, National Bureau of Economic Research July 2018.

Syverson, Chad, "Product Substitutability and Productivity Dispersion," *The Review of Economics and Statistics*, 2004, *86* (2), 534–550.

_ , "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–65.

Tsai, Thomas C., Ashish K. Jha, Atul A. Gawande, Robert S. Huckman, Nicholas Bloom, and Raffaella Sadun, "Hospital Board And Management Practices Are Strongly Related To Hospital Performance On Clinical Quality Metrics," *Health Affairs*, 2015, *34* (8), 1304–1311. PMID: 26240243.

Xu, Guo, "The Costs of Patronage: Evidence from the British Empire," *Forthcoming American Economic Review*, 2018.

Tables

Table I: Manager Characteristics

	Full Sample	Movers
Male	0.63	0.70
Age	54.98	52.96
Experience Publ. Sec.	27.43	24.60
North-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
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
Observations	851	207

Note: Full sample, 2011q1-2017q2. All statistics are computed over the full sample of managers in the first column, and over the sub-sample of movers in column 2. Movers are defined as those managers who have at least two appointments over my sample period.

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	0.06	0.18	0.02
Separations	0.50	1.53	0.17
Fires	0.01	0.01	0.00
Inbound Transfers	0.87	2.64	0.32
Outbound Transfers	0.41	0.98	0.23
Retirement	0.31	0.97	0.10
Divorce	0.87	0.88	0.87
Blood Donations	0.03	0.03	0.03
Age	52.57	52.69	52.53
Female	0.58	0.57	0.58
Female Top-Officials	0.39	0.40	0.39
Number of office-quarter obs.	13212	3142	10070
Number of managers	851	221	638
Number of offices	494	111	383

Note: The full sample includes all main offices and local branches, 2011q1-2017q2. All statistics are calculated across office-quarter observations. The number of office-quarter observations for the full sample, main offices and local branches for quality are 10943, 2658, and 8285, respectively; for divorce these statistics are 11052, 2622, and 8430, and for blood donations 11104, 2648, and 8456. Output, demand, and backlog are measured in thousands of hours, while FTE, training, hours, and overtime are measured in full-time equivalent units.

Table III: Sample Characteristics

	(1) Full Sample	(2) Balanced-Analysis Sample
# Managers	851	601
# Offices	494	282
# Managers >1 Office	207	184
# Offices >1 Manager	404	282
# Connected Sets	276	143
# Switches	635	318
# Switches in Main Offices	226	80
# Switches in Local Branches	409	238

Note: Column 1 reports the summary statistics computed over the full sample (2011q1-2017q2, N=13,212). Column 2 reports the same statistics over the balanced-analysis sample (2011q1-2017q2, N=8165). 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. All statistics are calculated over office-quarter observations. "# Managers >1 Office" is the number of managers who serve in at least two sites over my sample period. "# Offices >1 Manager" represents the number of offices that experience at least one change in leadership over my sample period. Switches are defined as changes in leadership.

Table IV: Analysis of Variance of Yearly Measures of Productivity per Worker at INPS Offices

	(1) Ln(P)	(2) Ln(P)	(3) Ln(P)	(4) Ln(P)	(5) Ln(P)
N	3316	3316	3316	3316	3316
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: Full sample at the yearly level, 2011-2017. I perform an analysis of variance of log productivity to study how much of its variation is explained by the office, manager, and time components. The p-value at the bottom of the table tests the hypotheses that manager effects are jointly zero.

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 + 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	
Adj. R sq.	0.210		0.397	
CS FE	Yes		Yes	
P-value (All)	0.000		0.000	
P-value (Growth Rates)	0.807		0.864	

Note: The sample includes all events balanced on $[-4, 0]$. The dependent variable is the manager effect estimated using (2) in column 1, while it is the difference between the estimated effect of the incoming and outgoing manager effects in column 2. P, Y, and FTE stand for productivity, output and full-time equivalent employment respectively. t indexes the time of the event, and P_{2011} represents 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 -(x+1) and -1. "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 (CS) fixed effects. Standard errors are clustered at the office level and are reported in parenthesis.

Table VI: Biased-Corrected Variance-Covariance decomposition

	Var. Component	Sh. 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: The sample includes the largest connected set, 2011q1-2017q2. The model includes dummies for individual managers, for individual offices, and quarter fixed effects.

Table VII: Manager Effects and Observable Characteristics

	(1) Manager FE
Male	-0.06
	(0.03)
Experience Publ. Sec.	0.02
	(0.01)
Experience Publ. Sec. 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
CS FE	Yes

Note: Full sample of managers. The dependent variable is the manager effect estimated from (2). 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 parenthesis.

Table VIII: Estimated Effects of Changes in Managerial Talent on Office Composition

	(1)	(2)	(3)	(4)	(5)
k	A(Retirement)	A(Hires)	A(Fires)	A(Inbound T)	A(Outbound T)
-4	0.044	0.178	0.003	0.077	-0.055
	(0.125)	(0.110)	(0.004)	(0.106)	(0.144)
-3	-0.037	0.093	0.002	0.027	-0.108
	(0.089)	(0.090)	(0.004)	(0.085)	(0.131)
-2	-0.048	0.034	0.003	0.031	-0.090
	(0.059)	(0.073)	(0.004)	(0.074)	(0.112)
0	0.299	0.024	-0.008	0.000	-0.043
	(0.087)	(0.018)	(0.010)	(0.154)	(0.053)
1	0.401	0.027	-0.056	-0.098	-0.019
	(0.102)	(0.033)	(0.031)	(0.158)	(0.066)
2	0.380	0.024	-0.049	-0.274	-0.167
	(0.108)	(0.033)	(0.040)	(0.167)	(0.096)
3	0.396	0.006	-0.063	-0.405	-0.245
	(0.119)	(0.038)	(0.045)	(0.169)	(0.113)
4	0.458	-0.015	-0.040	-0.454	-0.243
	(0.121)	(0.039)	(0.038)	(0.166)	(0.124)
5	0.422	0.005	-0.061	-0.471	-0.303
	(0.123)	(0.039)	(0.041)	(0.170)	(0.124)
6	0.376	-0.077	-0.059	-0.581	-0.402
	(0.132)	(0.056)	(0.042)	(0.180)	(0.135)
N	318	318	318	318	318
Time FE	Yes	Yes	Yes	Yes	Yes
Mean	0.415	0.038	0.019	0.900	0.374

Note: Balanced-analysis sample (2011q1-2017q2). This subsample includes only events which are balanced on $[-4, 6]$. The dependent variable is Δy_{it}^{k} (see text) and cumulative y_{it} is reported at the top of each column. $A(.)$ stands for asinh. 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 parenthesis.

Table IX: Estimated Effects of Changes in Managerial Talent on Predicted Productivity

k	(1) Pred Ln(P)	(2) Pred Ln(P1)	(3) Pred Ln(P2)	(4) Pred Ln(P3)
-4	0.018	0.036	0.031	0.023
	(0.056)	(0.016)	(0.039)	(0.043)
-3	0.010	0.026	0.011	0.000
	(0.058)	(0.014)	(0.040)	(0.042)
-2	-0.015	0.021	0.013	0.009
	(0.059)	(0.013)	(0.035)	(0.040)
0	0.204	0.007	0.040	0.030
	(0.067)	(0.016)	(0.029)	(0.033)
1	0.121	0.009	0.047	0.036
	(0.050)	(0.016)	(0.033)	(0.036)
2	0.215	-0.001	0.061	0.055
	(0.066)	(0.016)	(0.040)	(0.044)
3	0.234	0.007	0.065	0.050
	(0.048)	(0.011)	(0.037)	(0.037)
4	0.207	0.014	0.024	0.011
	(0.080)	(0.014)	(0.039)	(0.043)
5	0.290	0.015	0.034	0.017
	(0.059)	(0.014)	(0.037)	(0.039)
6	0.371	-0.013	0.047	0.051
	(0.072)	(0.013)	(0.040)	(0.044)
N	318	318	318	318
Time FE	Yes	Yes	Yes	Yes
Mean	0.058	-0.212	-0.057	-0.053

Note: Balanced-analysis sample (2011q1-2017q2). This subsample includes only events which are balanced on $[-4, 6]$. The dependent variable is Δy_{it}^k (see text) and y_{it} is reported at the top of each column. The dependent variable in column 1, $\widehat{\ln P_{it}}$, is the fitted value from a regression of $\ln P_{it}$ on office and time fixed effects and a set of covariates. These covariates include share of employees in each of the 10 deciles of the age distribution, average office age, fraction female, fraction of female among top-officials, ln FTE, asinh(absences), asinh(over-time), asinh(training). This model also includes a linear and a quadratic term for each of these covariates, as well as their two-way interactions with time fixed effects and a dummy for main offices. Dependent variables in columns 2, 3, and 4 are fitted values constructed using a subset of variables. I use office and time effects in Column 2, add age in column 3, and gender composition in column 4.
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 parenthesis.

Table X: 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
Mean Dep. Var.	0.232	0.114	

Note: Balanced-analysis sample (2011q1-2017q2). This subsample includes only events which are balanced on $[-4, 6]$. The dependent variable is Δy_{it}^k (see text) and 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 parenthesis.

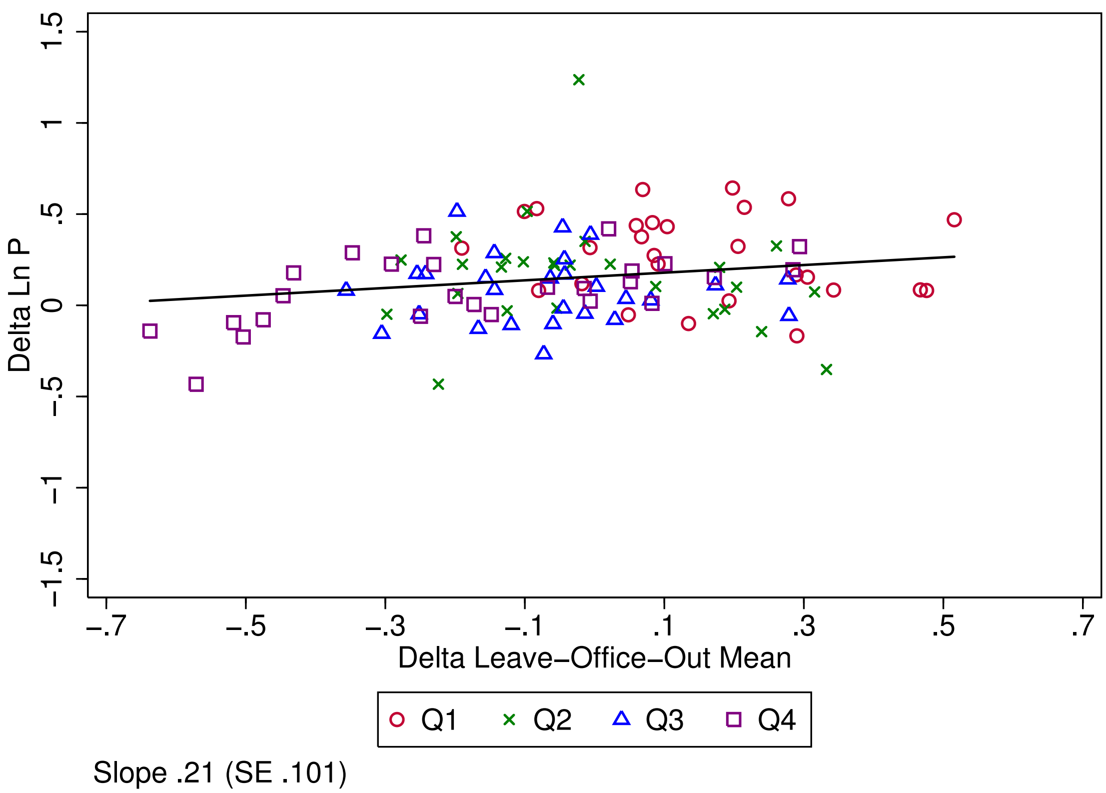
Table XI: Counterfactual Exercises

		ΔP
Policy 1:	Reassign	6.9%
Policy 2:	Replace bottom 20%	2.9%
Policy 3:	Replace bottom 20% + Reassign	7.4%
Policy 4:	Random allocation	2%

Note: The sample includes all the connected sets with at least five managers. I consider four counterfactual policies that the social planner can implement. Policy 1: she can rallocate existing managers according to the optimal rule. Policy 2: she can fire the bottom 20% of top-level bureaucrats and substitute them with the median manager (but allocate them as in the current environment). Policy 3: she can implement both Policy and 2. Policy 4: she can randomly assign existing managers to offices (1,000 iterations).
As my sample contains multiple connected sets, I implement each policy within connected set, estimate the implied counterfactual for each office, and then aggregate over offices using w_i as weights.

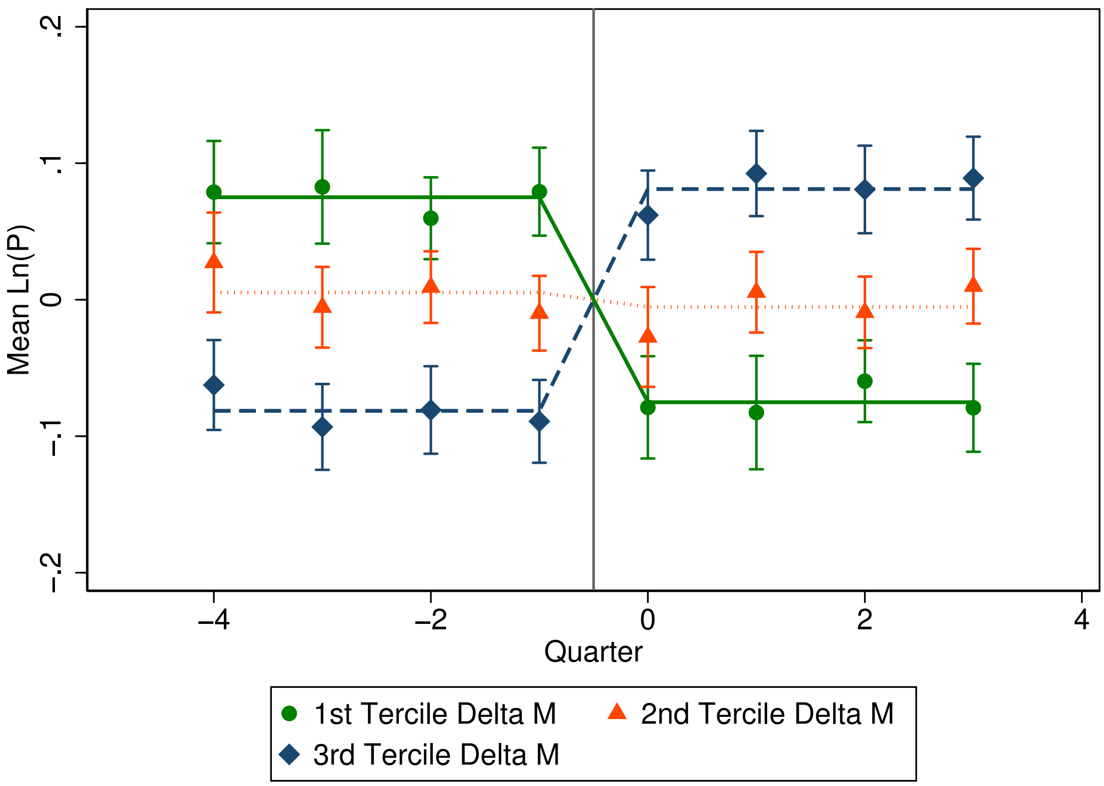
Figures

Figure I: Productivity gains and losses associated with changes in managerial talent by group of origin



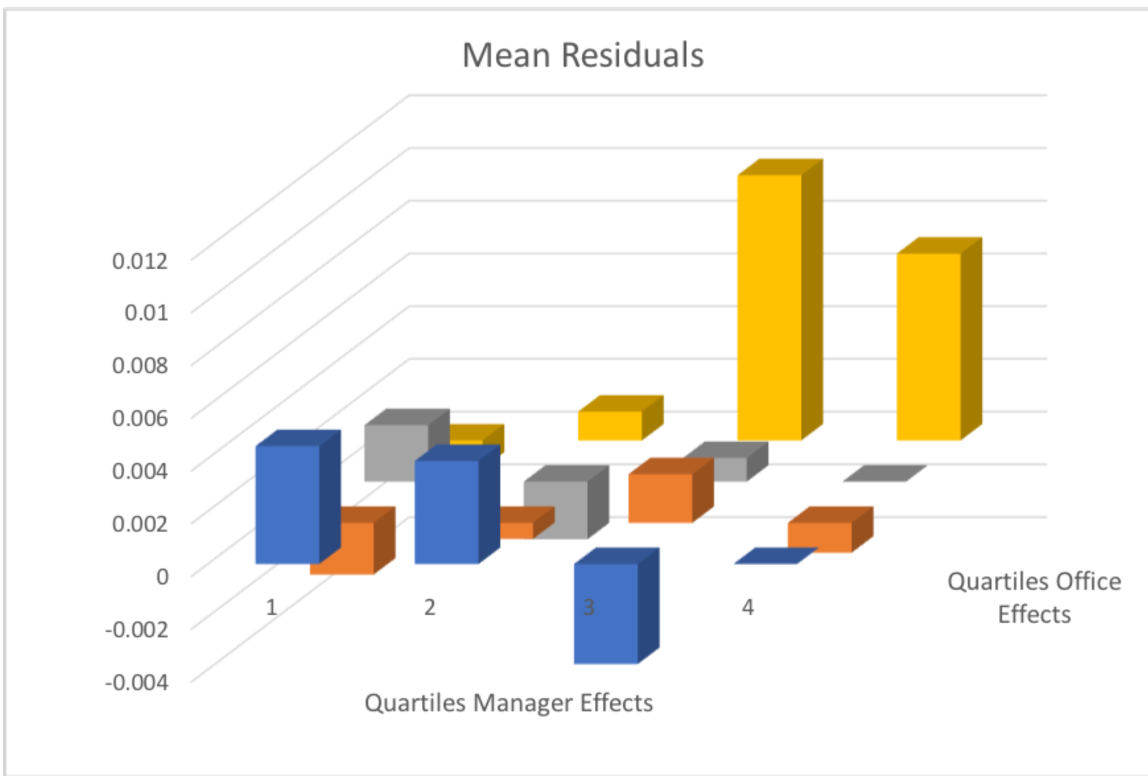
Note: Delta Ln(P) is the difference between agerage log productivity in the four quarters after a change in management and average log productivity in the four quarters prior to a change. Delta Leave-Office-Out Mean is the change in the (trend adjusted) leave-i-out mean log productivity induced by the change in leadership. The leave-i-out mean proxies for manager ability. Q_j with $j = \{1, 2, 3, 4\}$ indexes the group of origin. The group of origin is computed as the quartile of the outgoing manager leave-i-out mean and proxies for the ability of the outgoing manager.

Figure II: Mean Productivity for Offices which Experience a Change in Leadership Classified by Tercile of Changes in Manager Effects



Note: The dependent variable is the mean of (trend-adjusted) log productivity. Whiskers represent 95% confidence intervals. $\widehat{\Delta M}_i$ represents the change in the estimated manager fixed effects. When an office goes through a change in management it can experience three types of transitions: an overall increase in manager ability (third tercile of $\widehat{\Delta M}_i$), a decrease in management quality (first tercile of $\widehat{\Delta M}_i$) or no significant change (second tercile of $\widehat{\Delta M}_i$). X-axis indexes event time.

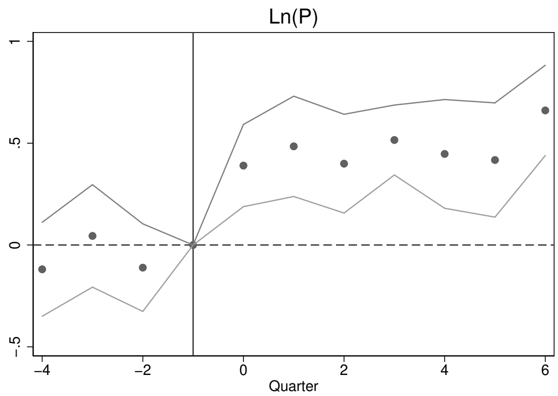
Figure III: Mean Residual by Manager/Office Quartiles, 2011q1-2017q2



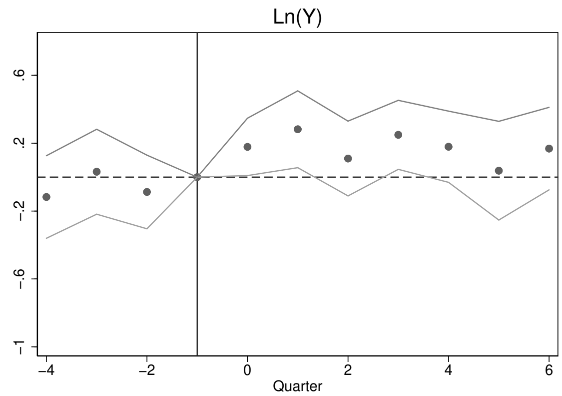
Note: This Figure shows mean residuals from model (2) with cells defined by quartiles of estimated manager effect, interacted with quartiles of estimated office effect.

Figure IV: Decomposition of Productivity Effects

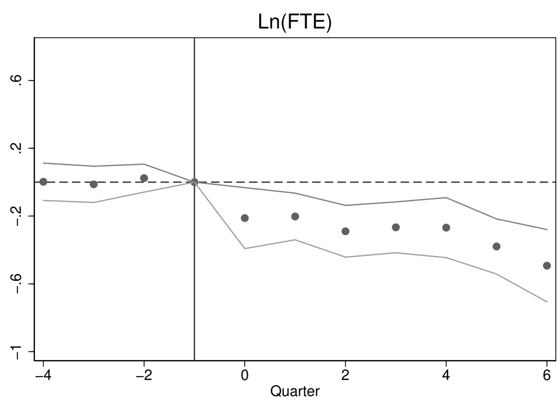
(a)



(b)



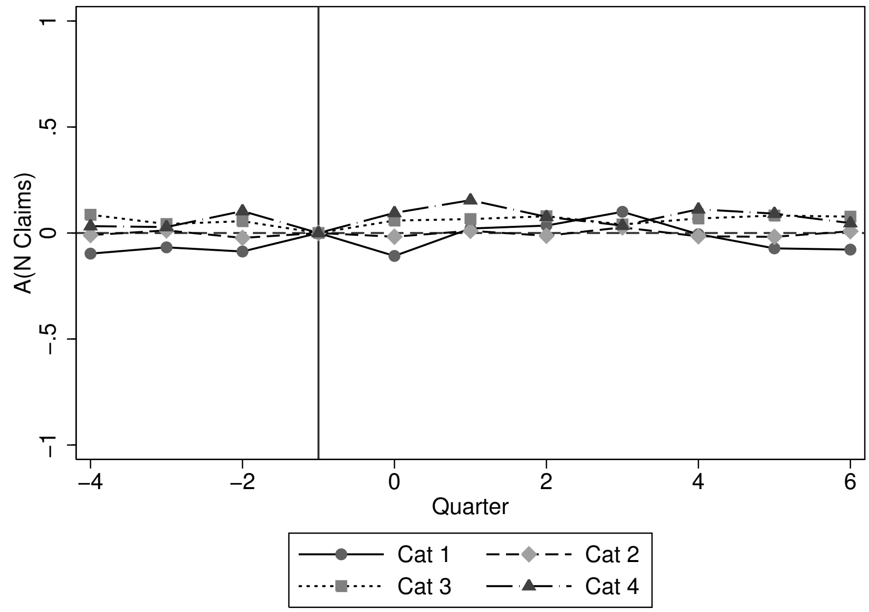
(c)



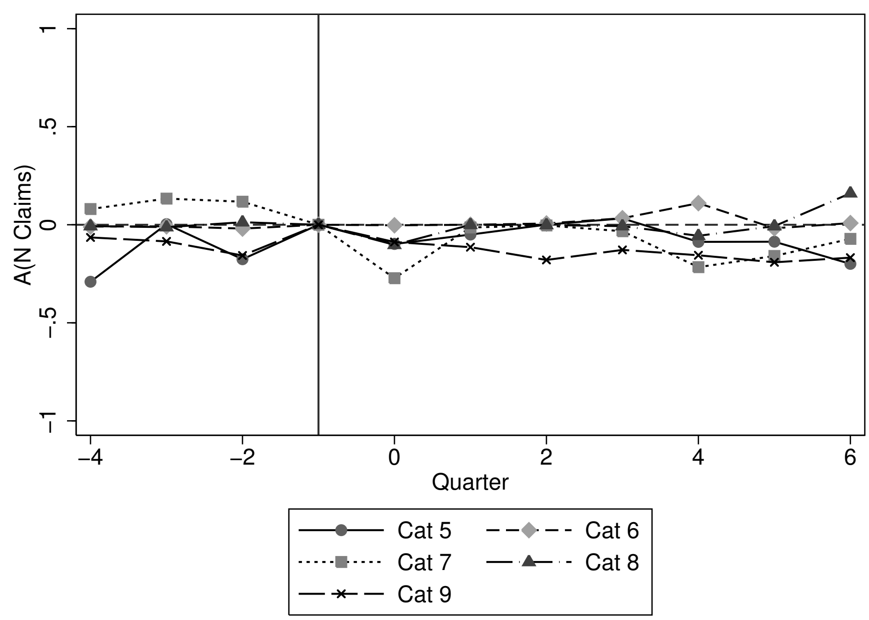
Note: Panels (a)-(c) report point estimates ($\hat{\pi}_1^k$) and 95% bootstrapped confidence intervals from (10). The dependent variables are log productivity (Ln(P)), log output (Ln(Y)), and log full-time equivalent employment (Ln(FTE)), respectively. X-axis indexes event time. Refer to Table J.IV for these results in table format.

Figure V: More Productive Managers Do Not Shift Production

(a)



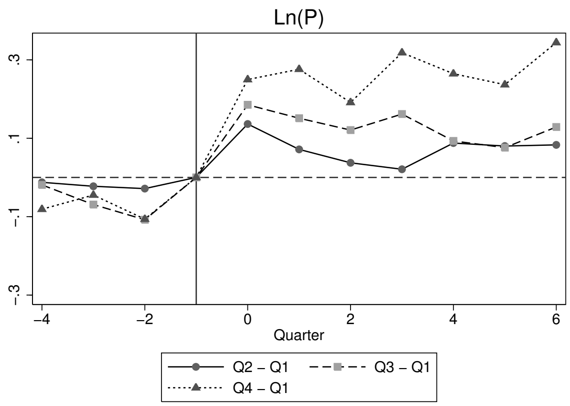
(b)



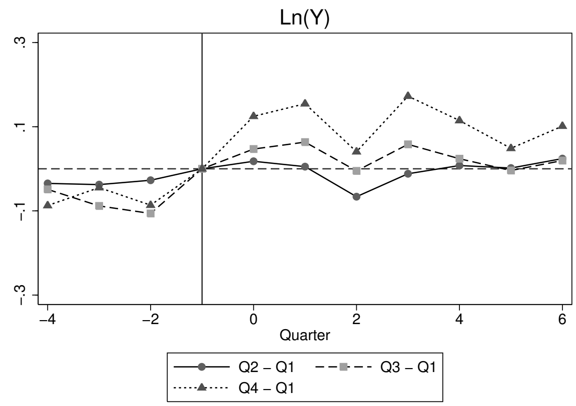
Note: The dependent variable is the number of claims belonging to product category n processed by each office. n=1, 2, ..., 9. Panels report the point estimates for nine product categories. The categories are defined as follows: 1: Insurance and pensions, 2: Subsidies to the poor, 3: Services to contributors, 4: Social and medical services, 5: Specialized products, 6: Archives and data management, 7: Administrative cross-checks, 8: Checks on benefits, 9: Appeals. X-axis indexes event time.

Figure VI: Estimated Effect of Leadership Changes by Quartile of Changes in Manager Effects

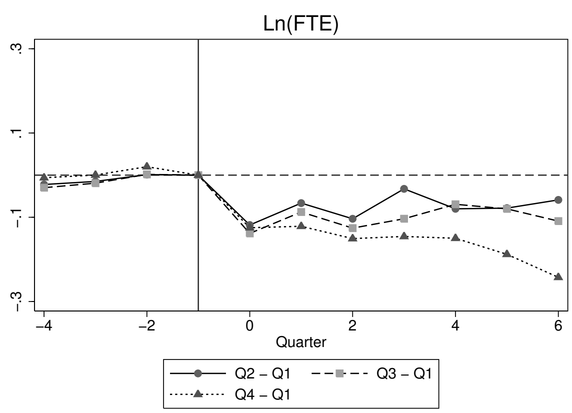
(a)



(b)



(c)

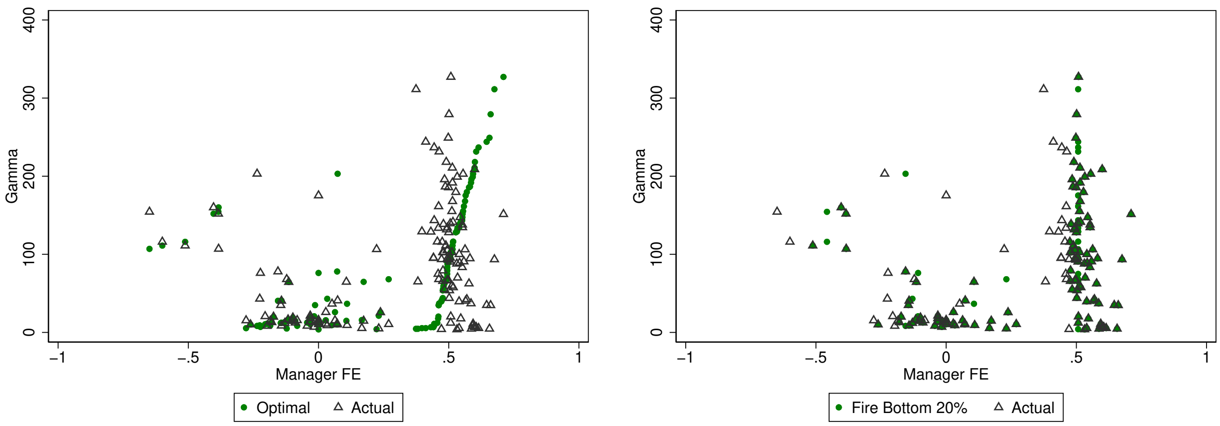


Note: These Figures report point estimates from (12). The dependent variables are log productivity (Ln(P)), log output (Ln(Y)), and log full-time equivalent employment (Ln(FTE)), respectively. As the first quartile (Q1) of ΔM_i^L is the omitted category, all coefficients identify the difference between the j-th and the first quartile $(Q_j - Q_1)$. X-axis indexes event time.

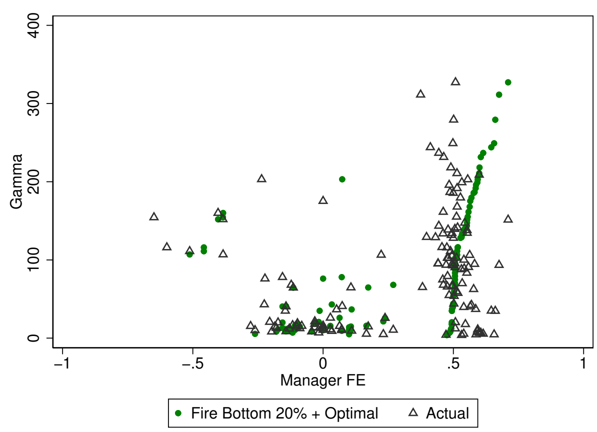
Figure VII: Counterfactual Exercises

(a) Policy 1: Reassign

(b) Policy 2: Fire bottom 20%



(c) Policy 3: Replace bottom 20% and Reassign



Note: This figure illustrates how the allocation of managers to offices is impacted by three counterfactual exercises. Policy 1: the social planner can reallocate existing managers according to the optimal rule. Policy 2: the social planner can fire the bottom 20% of top-level bureaucrats and substitute them with the median manager (but allocate them as in the current environment). Policy 3: the social planner can implement both Policy 1 and 2. As my sample contains multiple connected sets, I implement each policy within the connected set. The hollow triangles represent the current allocation of managers to offices while the green circles represent the one implied by the counterfactual exercises.