Corruption and personnel selection and allocation in the public sector

Sauro Mocetti and Tommaso Orlando ♣

PRELIMINARY DRAFT, DO NOT CIRCULATE

April 2017

Abstract. We examine the impact of corruption on personnel selection and allocation in the public sector. Using difference-in-differences estimation on Italian data, we find that in areas with higher levels of corruption the selection of public employees in terms of observed human capital is worsened in comparison to that of their private-sector counterparts. The effect is concentrated among managers and professionals. Moreover, corruption is associated with a misallocation of human resources and, in particular, with an increase (with respect to the private sector) in the rate of under-qualification among public employees, especially among those in high skill content jobs. These results are robust to various measures of corruption and several robustness checks, including IV estimation that uses historical factors as exogenous source of variation for current corruption.

Keywords: corruption, selection, mismatch, schooling, ability, public employment.

JEL classification: D73, J45.

♣ Sauro Mocetti, Bank of Italy, DG for Economics, Statistics and Research, Structural Economic Analysis Directorate; e-mail: sauro.mocetti@bancaditalia.it; Tommaso Orlando, Bank of Italy, DG for Economics, Statistics and Research, Structural Economic Analysis Directorate; e-mail: tommaso.orlando@bancaditalia.it. We thank Audinga Baltrunaite, Federico Cingano, Silvia Giacomelli and Giuliana Palumbo and seminar participants at the Bank of Italy for useful comments. The views expressed in this paper should be referred only to the authors and not to the Institution with which they are affiliated.

1. Introduction

Corruption is widely believed to entail large economic and social costs. The economic literature has so far explored several channels through which corruption may affect economic outcomes. Some authors highlight its effects in terms of distortion of private decisions, such as investments (Shleifer and Vishny, 1993; Mauro, 1995) and human capital accumulation (Mo, 2001). Others focus on the activities of the public sector, documenting relationships between corruption and inefficiencies in the composition of government expenditure (Mauro, 1998), lower productivity of public investments (Del Monte and Papagni, 2001) and higher shares of goods and services procured by the public administration on non-competitive markets (Hessami, 2014).

In this paper we analyze the impact of corruption on personnel selection and allocation in the public sector. More specifically, we address two issues: first, we examine whether corruption affects the *selection into the public sector* of individuals with different levels of (observable) human capital; secondly, we examine the relationship between corruption and the *allocation within the public sector* of differently educated individuals to jobs with different skill content. Poorer recruitment and misallocation of human resources within public agencies might have significant and long-lasting consequences on the quality of the administration's economic decisions and on the effectiveness of the services provided by the public sector.

The empirical analysis is based on two complementary data sources containing information on Italian public and private employees and exploits several measures of corruption. We use these to examine whether areas characterized by more corruption show peculiar patterns of skill-based selection into and allocation within the public sector. Although we use cross-sectional variation, our empirical strategy mirrors a difference-in-differences approach where the 'treatment' is represented by the intensity of corruption at the local level and the employees in the public and private sectors represent the 'treated' and the 'control' group, respectively. Time-invariant heterogeneity that might be correlated to both corruption and human capital endowments is captured by the inclusion of fixed effects at the local level, thus exploiting public versus private sector sorting of workers within the same geographical area. Moreover, to address reverse causality – the possibility that corruption itself be the consequence of poor selection and allocation of human resources by the public administration – we instrument corruption at the local level with past domination spells and past

economic conditions, i.e. with factors predating the hiring of current public employees.

We find that public employees are, on average, more educated and obtained higher grades at school than their professional counterparts in the private sector. However, in areas characterized by rampant corruption the relationship between educational attainments and the likelihood of joining the public sector is substantially weaker. The negative impact of corruption is concentrated among those whose jobs have higher skill content, such as managers and highly skilled professionals, while the impact is less clear for clerical workers. We also find some evidence that those hired through less formal procedures are, on average, less educated and that this difference is more evident in areas where corruption is higher. This suggests that informal recruitment may be badly used in more corrupt environments. As for the allocation process, we find that a higher level of corruption is associated to a larger likelihood of mismatch between individual educational attainments and the skill content of the job the worker is assigned to. This mismatch is mostly driven by phenomena of under-education – individuals being assigned to jobs which are, on average, undertaken by more qualified personnel – rather than over-education. As above, these effects appear to concern workers at the top of the occupational hierarchy in a stronger fashion. We also show that mismatch is not merely a 'mechanical' consequence of poorer selection processes, nor of inflation in the number of managerial positions.

The literature has already partially tackled the issue of the relationship between corruption and occupational choices. Murphy *et al.* (1991) and Acemoglu and Verdier (1998) argue that corruption magnifies rewards to rent-seeking activities, thus subtracting valuable human resources to entrepreneurship and distorting the allocation of talent across sectors. Concerning selection into the public sector, experimental evidence suggests that more corrupt environments encourage entry by the dishonest into the public sector. Using laboratory data, Hanna and Wang (2013) and Banerjee *et al.* (2015) find negative self-selection into the (supposedly corrupt) Indian public administration, while Barfort *et al.* (2016) find positive self-selection into the (supposedly non-corrupt) Danish public sector.

Our paper innovates over the existing literature along several directions. First, the economic impact of corruption has typically been investigated using cross-country evidence (at a single point in time). However, the cross-sectional relationship might be severely biased because corruption and the other variables of interest are likely to have common correlates that cannot all be credibly controlled for: stated differently, less corrupt societies appear to perform well in

almost any dimension, and the risk of bias due to an omitted variable (e.g. welfare institutions or culture) is large. To address this issue, some papers introduce country-fixed effects by exploiting panel data. However, the reliability of those estimates clearly depends on the longitudinal (within-country) variation of the variables that, in the case of a persistent and structural phenomenon like corruption, is admittedly low. Moreover, panel data alone do not fully address endogeneity concerns, as a variation in corruption and in the outcome variable might reflect common (country-specific) shocks. Finally, the measurement of corruption itself may be problematic either from a cross-section or longitudinal point of view. Indeed, one may question the capacity of international surveys to capture the intensity of corruption equally well in all countries due to differences in culture and social norms or to other perception biases. Similarly, official data on reported crimes might not be comparable across countries due to differences in laws or in the availability of harmonized crime statistics. The extent of these measurement issues can also vary over time. We address all these issues using variation in reported crimes within a single country, thus exploiting (sub-national) territorial variability and using homogenous and comparable measures of corruption. Identification hinges on the differential impact of corruption between the public sector and the private sector, under the assumption that corruption affect personnel choices only – or at least more intensely – in the public sector.

Secondly, previous studies on the relationship between corruption and workforce sorting have used experimental evidence and focused personal propensity to dishonesty rather than on human capital as the main individual attribute. On the contrary, we rely on hard data and drive the attention on skill-based sorting, as measured by educational attainment and grades obtained at school. Indeed, human capital endowments in public agencies – beyond the direct effect of individual attitudes towards corruption – might affect the quality of economic decisions on a number of relevant dimensions such as the level and the composition of public expenditure, the effectiveness of public investments, the quality of public services provided to households and firms, etc.[1]

Third, we focus on corruption's impact on both selection and allocation processes, while previous studies have mainly focused on workforce sorting.

[1] Poor selection and misallocation may also contribute to reinforcing corruption itself. The policy measures against corruption typically rests on (*ex-post*) repressive measures, sometimes accompanied by (*ex-ante*) preventive provisions. According to our findings, one may suspect that the administrations' ability to exert *ex-ante anti-corruptive self-monitoring* might be hindered by corruption itself. Indeed, a less educated workforce may be less conscientious and/or less effective about anti-corruption efforts.

However, human resource misallocation is also relevant: on the one hand, the same group of individuals can produce substantially different results if they are badly matched to jobs requiring different educational qualifications; on the other hand, bad allocation processes and misaligned career rewards might discourage the most skilled individuals from applying for a public job.

The rest of the paper is organized as follows. Section 2 describes the data sources and the main variables of our analysis, including the construction of corruption indicators. Section 3 presents our empirical strategy and clarifies which effects of corruption we are able to identify. Section 4 presents our main findings and some robustness checks. Section 5 concludes.

2. Data and variables

2.1 Individual information on occupation and schooling

Individual data on employment characteristics and observed measures of human capital are drawn from two data sources. The main one is the Italian Labour Force Survey (LFS). The survey is carried out by Istat on a weekly basis and its main aim is to provide accurate and official statistics concerning the employed and unemployed population in Italy. We pool the LFS waves from 2004 to 2010 and we restrict the analysis to non-manual employees (i.e. ISCO major groups 1 to 5). LFS does not provide a clean distinction between the public and the private sector and, therefore, we identify as public employees all those employed in the following three NACE 2-digits groups: public administration and defense, compulsory social security, education and health. We also know the professional qualification of each employee, as measured by the ISCO occupational classification at 3 digits and their education level (in particular, the years of schooling corresponding to the highest educational attainments). Non-manual employees are divided into three groups (from the most to the least qualified): *1*) managers and professionals (ISCO major groups 1 and 2), *2*) technicians and associate professionals (ISCO major group 3) and *3*) clerical workers (ISCO major groups 4 and 5). Among socio-demographic characteristics, we observe age, gender and, most importantly for our goal, the local labor market (henceforth LLM) in which

workers reside.[2] This geographic attribute is used to capture local economic and social conditions that might impact on the likelihood of joining the public sector and on the quality of the match between individual education and job skill content.

The selection process is also investigated through the use of a second data source, the Survey on Household Income and Wealth (SHIW). The survey is carried out by the Bank of Italy and contains information on the socio-economic conditions of a representative sample of the Italian population.[3] We pool the (bi-annual) SHIW waves from 2000 to 2014 and we restrict the analysis to household heads who are non-manual employees, as done with the LFS. The size of the SHIW sample is much smaller than that of the LFS and details on occupations are definitely poorer. When using SHIW data, we divide non-manual employees into two groups only: the first includes senior and junior managers and teachers, the second clerical workers only. However, differently from the LFS, the SHIW allows a clean distinction between public and private sector. More importantly, the SHIW can be used to complement the LFS analysis with a further dimension of individual human capital: the final grade relative to the individual's highest educational attainment. Among socio-demographic characteristics, we include – as before – age, gender and the LLM where the individual resides.

Descriptive characteristics for the (pooled) LFS and SHIW samples are reported in Table 1. Consistently with evidence from previous studies (see for instance Rizzica, 2016), in both samples public sector employees are, on average, older, include a larger share of women and possess relatively richer endowments of human capital, both in terms of education attainments and grades obtained at school. Moreover, the extent of average mismatches (both under-education and lower-education) is similar in the private and the public sector.

2.2 Measures of corruption

There are different definitions of corruption and different approaches to its measurement. While different legal systems may assign this label to different types of offence, in this paper we call "corruption" any instance of distortion of a public employee's behavior, enacted in order to generate a direct or indirect gain to the civil servant and implying the bearing of a cost by a private counterparty. Under

[2] LLMs are geographic units represented by cluster of contiguous municipalities and built on the basis of daily commuting patterns; therefore a LLM represents the area in which most individuals both reside and work.

[3] See Brandolini and Cannari (1994) for more details on the survey.

the Italian law, for instance, such definition encompasses the crimes of *corruption proper* (i.e. bribery), *graft*, and *malfeasance*.[4] These unlawful behaviors all result into additional payoffs accruing to the public employee at the detriment of one or more private agents: the role of such agents may range from being an active part in the enactment of the criminal act (as in bribery) to being the victims of the public servant's prevarication (as in graft).

Measuring corruption is admittedly a challenging task: as all illegal activities, corruption is mostly unobservable and, therefore, difficult to quantify. So far, four basic approaches have been used to measure corruption at some aggregate level. The first approach is based on *subjective* and *direct* estimates or perceptions about the extent of corruption drawn from ad hoc surveys among citizens or a specific group of respondents, sometimes called "experts". Indicators of this type include, for example, the Transparency International Corruption Perceptions Index, the Global Corruption Barometer and the European Quality of Government Index. The second approach relies on *subjective* but *indirect* indicators of corruption. For example, distrust towards (local or national) governments might at least partially reflect whether they are perceived as corrupt entities. The third approach is based on observable consequences of corruption, such as the difference between the public expenditure in certain infrastructures and the corresponding realized outcome. This sort of "missing expenditure" (Olken, 2009) represents an *objective* though *indirect* measure of corruption. The fourth approach relies on *objective* and *direct* measures of corruption such as direct observations, reported crimes or similar evidence arising from governmental audits.[5]

Each of these approaches has advantages and drawbacks. Subjective and perception-based indicators (either direct or indirect) have been widely used, as they are available for a large set of countries and allow to exploit cross-country variation in corruption to examine the latter's relationship with other economic

[4] Crimes perpetrated by public officials are regulated by the Italian criminal law (Codice Penale, articles 314-323, 479-481 and 493): acknowledging oversimplification, *corruption* proper takes place when the public official accepts a bribe from a private counterparty in exchange for the enactment of or the abstention from certain behaviors; *graft* refers to the situation in which the payment is imposed by the civil servant to the private party; *malfeasance* generically defines behaviors enacted by the public employee aiming at earning the latter unlawful benefits: resource embezzlement and document forgery, when perpetrated by Italian public officials or by other providers of public services, may be seen as special cases of malfeasance.

[5] Olken and Barron (2009) designed a study in which surveyors accompanied Indonesian truck drivers on their trips in order to collect direct observations on illegal payments to police, soldiers, and weigh station attendants. Ferraz and Finan (2011) and Brollo et al. (2013) used data on a program of random audits on local governments, with detailed reports on corruption charges. For Italy, Del Monte and Papagni (2001, 2007) and Barone and Mocetti (2014) used official statistics on reported crimes against the public administration.

outcomes.[6] However, the effectiveness of these indicators has been questioned. First, there are significant differences in cultural traits, social norms and laws across countries, so that citizens of one polity may find certain practices more acceptable than citizens of another one, thus leading to different reported perceptions of the extent of corruption. Second, the reliability of survey information has also been questioned as respondents might not report direct experiences but be influenced by what is publicized in the media (Rizzica and Tonello, 2015). The third approach is also intriguing, but missing expenditure – like any other variable measured as a "residual" – is not necessarily attributable to corruption. For example, the effectiveness of public spending in infrastructures might also reflect the efficiency of the local construction industry, unobserved characteristics of the territory or other random elements, thus confounding the interpretation of the computed indicator. Finally, the fourth approach, beyond poor cross-country comparability due to differences in laws and in the organization of the judicial system, might suffer from reporting bias. If crime episodes are collected by police forces or courts, variations in their number might reflect not only the intensity of the criminal activity, but also the efficiency of such institutions and/or their interest in prosecuting that particular crime.

In this paper we adopt two measures of corruption. The first is based on crimes reported by police forces to the judicial authority, extracted from the SDI database[7]. Data at our disposal are collected at the municipality level and cover the period from 2004 to 2011. In particular, we restrict the analysis to crimes intimately linked to corruptive practices: corruption proper, graft and malfeasance. These raw figures are normalized with respect to the employment at the local level (a proxy for the level of economic transactions). This measure is computed at the local labor market (LLM) level and is averaged over the period of observation.[8] To address potential reporting bias we partial out the effect of the local judicial efficiency on crime rates. Namely, we run a regression where we control for the judicial efficiency (as measured by the lengths of penal proceedings in local courts) and we take the residuals. The latter yield a measure of corruption incidence net of local judicial efficiency (C^1 henceforth).[9]

---

[6] See, among the others, Mauro (1995), Knack and Keefer (1995), La Porta et al. (1999), Fisman and Gatti (2002) and Fisman and Miguel (2007).

[7] WHAT IS THE SDI?

[8] We do not exploit within-LLM variation since corruption is a persistent phenomenon and does not show sufficient longitudinal variation.

[9] According to our findings, a variation of one standard deviation of the length of penal proceedings is associated to a 0.14 standard deviations increase in the reported crime rate, thus

The second measure is a synthetic indicator that combines information drawn from different approaches. Namely, we collect four different variables, each echoing one of the four approaches mentioned above (though, for reasons of data availability, they are measured at different geographical levels). The first variable is a subjective assessment of the level of corruption (CPI). Data are drawn from a large European Commission-funded survey (EQI) aimed at measuring the quality of governance within the European Union and they are available at the regional level. More information on the data and related descriptive evidence can be found in Charron et al. (2014). The second variable echoes the subjective and indirect approach to the measurement of corruption. We exploit a survey managed by ISTAT (the so-called "Multiscopo") asking a large set of questions to citizens to various aspects of life, including trust towards local government and other institutions ($TRUST$). Deeming corruption and trust in institutions to be strongly related (Uslaner, 2004; Clausen et al., 2011), we take distrust towards local government as a measure of corruption. These figures are available at the regional level with a further distinction between small municipalities, intermediate municipalities and larger metropolitan areas. The third variable belongs to the group of objective and indirect measures of corruption. Golden and Picci (2005) compute a measure of corruption for Italy as the difference between the value of the public infrastructure and cumulated public expenditure investment in public works (GP). These figures are available at the regional level. Our last variable is reported crime adjusted for judicial efficiency (i.e. the aforementioned C^1). We then rely on a principal component analysis to extract information from these four variables. The first principal component explains about 64 percent of the total variance of the underlying variables and it is positively associated, as expected, to each one of the input variables (Table 2). We call this synthetic indicator C^2.

C^1 and C^2 both have some pros and cons. On the one hand, C^2 might better capture a multidimensional and unobserved phenomenon such as corruption; moreover, the large fraction of variance explained by the first component suggests that the four indicators largely overlap, which is supportive of the measure's rich informational content. On the other hand, C^1 is easier to interpret in economic terms and less subject to arbitrary choices. Moreover, C^1 is available at a finer partition of the territory while C^2 partly reflects indicators that mostly vary at the

suggesting that the latter largely reflects the intensity of the criminal activity at the local level and is only marginally affected by judicial efficiency (as captured by our proxy).

regional level. For these reasons, C^1 is our preferred measure of corruption, though we provide evidence using both indicators throughout the paper.[10]

Summary statistics of the two indicators are reported in Table 3. In order to guarantee comparability between different indicators of corruption, both C^1 and C^2 have been standardized. The two variables display considerable variability across LLMs. A graphical representation of the territorial differences in corruption intensity is reported in Figure 1: both indicators show that corruption is more widespread in Southern Italy, with the North-South divide being more visually evident when C^2 is used; however, in both cases there is also significant variability within each macro-area.

2.3 Descriptive evidence

Corruption and human capital endowments are positively correlated at the LLM level, as shown in Figure 2. This apparently surprising fact is mainly due to other covariates that are correlated to both variables. For example, corruption is positively related to the size of the public sector (Figure 3a), either because large public agencies offer better chances to corruptors or because corruption may hinder the development of more market-oriented activities. Sector composition, however, also affects the incentives to invest in human capital: as it has been widely documented, the public sector tends to attract the most educated workers (Cowley and Smith, 2014; Rizzica, 2016). Moreover, corruption is more widespread in LLMs characterized by a lower level of economic development, as measured by the value added per capita, and by poorer labor market opportunities. The latter, however, might also affect human capital investments reducing, ceteris paribus, the opportunity costs of studying. Indeed, when we control for the above mentioned variables, the correlation between corruption and education disappears. This also suggests that, when attempting to identify a clean effect of corruption on other socio-economic outcomes, we face the challenging task of having to avoid spurious correlation driven by unobserved omitted variables.

The following descriptive statistics are built as a deviation from local mean, thus accounting, though still from a descriptive point of view, for other unobserved local determinants of both corruption and education. Having established that the

[10] Notice that our main results are qualitatively confirmed even if we use raw figures for reported crimes and/or each component of the principal component analysis separately. Results are available from the authors upon request.

share of employees in the public sector is higher in LLMs where the intensity of corruption is higher, we found that the share of managers in those areas is also higher (Figure 3a). In terms of human capital, as measured by the years of schooling, the (positive) gap between the public sector and the private sectors is slightly narrower in more corrupt LLMs, and this is especially true for managers (Figure 3b). Moreover, the relationship between under-education in the private and public sector changes when one moves from less to more corrupt areas. When corruption is low, public-sector employees are *less* likely to be under-educated than private-sector managers by 2.4 percentage points. On the contrary, where corruption is high, public-sector employees are *more* likely to be under-educated than private-sector employees by 1.2 percentage points. For managers only, we observe a shift from a 2.2 percentage points advantage to a 2.6 percentage points disadvantage.

3. Empirical strategy

3.1 Selection of workers into the public sector

The first phenomenon we wish to study is the potential distortionary effect of corruption on the relevance of educational attainments, as well as other measures of individual ability, as predictors of the likelihood of an individual's belonging to the public sector. To this end, we estimate the following linear probability model:

$$Y_i = \alpha + \beta S_i + \delta\left(S_i \cdot C_{LLM(i)}\right) + \gamma' X_i + \rho_{LLM(i),p(i)} + \varepsilon_i \qquad (1)$$

where Y_i is a binary indicator of the occupational status of individual i, taking on the value of 1 whenever i is a public employee and the value of 0 if i is employed in the private sector; S_i is a measure of i's skills endowment, e.g. the completed years of schooling; $C_{LLM(i)}$ is one of the two measures of the incidence of corruptive crimes in the LLM in which individual i resides; X_i are individual controls such as gender and age: these are included as the likelihood of joining the public sector may be affected by gender- or cohort-specific factors; finally, the term $\rho_{LLM(i),p(i)}$ is a group indicator obtained by combining i's LLM and professional area (i.e. managers and professional, technicians, and clerical workers for the LFS sample and manager/professionals and clerical workers for the SHIW sample). Thus, our coefficient of interest δ captures how the impact of schooling on the likelihood of joining a certain professional class $p(i)$ in the public sector (rather than the same

professional class in the public sector) varies across LLMs characterized by different corruption intensity. We might expect $\delta < 0$ as corruption is supposed to decrease the returns to education (displaced by other unobserved soft and relational abilities).

3.2 Allocation of workers within the public sector

Besides affecting, through self-selection and screening, the composition of the available public workforce, corruption may have an impact on how efficiently these human resources are assigned to different jobs and tasks. In particular, we imagine that to each job there corresponds a level of skills or human capital, i.e. the level of an individual which is "just right" for that job. We subsequently test whether corruption shifts the allocation of human resources away from the right matching and, if that happens, whether this prevalently takes the form of under- or over-education, i.e. employees having a much lower or higher, respectively, skill level with respect to that required on average by the jobs they are assigned to. In order to quantify this phenomenon, we estimate the following linear probability model:

$$M_i = \alpha + \beta\left(Y_i \cdot C_{LLM(i)}\right) + \gamma' X_i + \rho_{LLM(i)} + \varphi_{s(i)} + \varepsilon_i \qquad (2)$$

where the dependent variable M_i is a binary indicator for the presence of some form of skills mismatch (under-education or over-education) for individual i. Specifically, an individual is considered to be under-educated if her schooling level falls below the 25^{th} percentile of the distribution of schooling within her profession (defined in terms of the ISCO classification at 3 digits) and, conversely, over-educated if her schooling level exceeds the 75^{th} percentile of that distribution. As before, Y_i denotes whether i is a public employee. [11] LLM-fixed effects ($\rho_{LLM(i)}$) and sector-fixed effects ($\varphi_{s(i)}$) capture local or industry specific variables that might be correlated with mismatch.

Our coefficient of interest is β, which represents how the impact of corruption on mismatch varies across sectors differently exposed to the public sector. We might expect $\beta > 0$ as corruption is supposed to increase the mismatch, particularly in the public sector and/or in private industries that significantly interact with the public sector.

[11] In an alternative specification, we use the continuous measure of exposure described in the previous footnote.

3.3 Identification assumptions

When we examine the impact of corruption on economic outcomes exploiting cross-sectional variation, we should take account of two potential identification threats.

First, unobserved heterogeneity at the local level (e.g. social norms, level of economic development, etc.) might be related to corruption as well as to the accumulation of human capital. These omitted variables are likely to bias the OLS estimates. However, we include LLM-fixed effects aimed at capturing any potential variables varying at the local level. Indeed, our identification strategy exploits the differential sorting between the public and the private sector within each LLM. This strategy mimics a difference-in-differences approach where the treatment is represented by the intensity of corruption at the local level and the employees in the public and private sectors represent the 'treated' and the 'control' group, respectively. The implicit assumption is that corruption affects personnel selection and allocation in the public sector and not – or, at least, not comparably – in the private sector.

Second, we might suspect the presence of reverse causality, as one may argue that skill-biased recruitment processes in the public sector might affect the intensity of corruption. To partially address this problem, we exploit variation in corruption intensity at the local level that is attributable to past dominations and past economic conditions, i.e. to factors predating the hiring of the current public employees. In particular, we use two types of pre-determined indicators as instruments. First, we exploit variation in corruption that is attributable to the following past dominations: Anjou, Austria, Bourbons, Normans, Papal State, Savoy, Spain, Swabians and Venice (with the independent states being the residual category).[12] This analysis is related to a large literature that investigates how history (and historical institutions) may still influence existing institutions and current social behaviors (e.g. Acemoglu and Robinson, 2012). Second, we exploit variation in corruption that is attributable to economic rents at the local level. More specifically, we use data from the Italian 1971 Census to compute the dependence of the private sector from public demand.[13] The idea is that where the

[12] Figures are drawn from Di Liberto and Sideri (2015).

[13] Dependence on the demand of the public administration at the local level is computed in two steps. First, using the input-output matrix, we compute the dependence on the public demand for each sector of economic activity. Second, we translate these figures at the local level using the

latter is higher, the economic rents associated to the discretionary power of the public officials as well as the incentive of the entrepreneurs to influence public spending are also higher. This, in turn, can increase the likelihood of corruption.

4. Results

4.1 The impact of corruption on personnel selection

Table 4 reports the results of the estimation of model (1) for our two main measures of corruption, C^1 and C^2. The sample, drawn from the LFS, include all the employees of the public and private sectors engaging in non-manual activities. Individual human capital endowments are measured by years of schooling.

Higher educational attainments are, as expected, positively associated with the likelihood of having joined the public sector. One additional year of schooling increases the probability of being a public employee by around 2.2 percentage points; the impact is higher among managers and professionals (4.5%) and lower for technicians and clerical workers (1.5%).

More interestingly, corruption reduces the role of education as a predictor of being a public employee. According to the results shown in the top panel of Table 4, moving from a LLM at the 10th percentile of C^1 to one at the 90th percentile (i.e. from a low-corruption to a high-corruption LLM) the impact of one additional years of schooling decreases from 2.3 to 2.0 percentage points; the detrimental effect of corruption is larger for managerial and professional occupations, where the same exercise would lead to a decrease of the impact of education from 5.0 to 3.8 percentage points. On the other hand, we do not detect any significant effect of corruption on education-sorting of workers directed to clerical positions.

The bottom panel of Table 4 replicates the analysis using C^2 as an approximation of corruption intensity at the local level. According to this variable, the skill-biased effect of corruption is even stronger. Moving from a LLM at the 10th percentile to one at the 90th percentile would entail a decrease of the impact of one year of schooling from 2.6 to 1.6 percentage points (from 5.3 to 3.5 for managers and professionals).

In Tables 5 and 6 we rely on the SHIW data rather than on the LFS. Results should be interpreted with some caution given the relatively small number of

past sector composition of the local economy (i.e. the distribution of employees across sectors at the local level as recorded by the Census 1971).

observations and the large number of fixed-effects that we include in the specification in order to control for the relevant unobserved heterogeneity. However, SHIW data allow us to use a second measure of ability, i.e. an index representing the grade obtained by individuals at their highest achieved educational level, which is available only for those with at least secondary education. In Table 5 we consider years of schooling as the only ability measure and the full sample of (non-manual) employees. In Table 6 we restrict the analysis to those having earned at least a diploma and we focus the attention on ability as measured by school grades. Grades are highly correlated with educational attainments, as those who obtain higher grades at secondary level are also those who are more likely to get tertiary education. For this reason, we do not consider the two ability measures jointly interacted with corruption, in order to avoid collinearity.

Table 5 generally confirms previous results, though the impact of education on the probability of joining the public sector is slightly lower on average. Estimates for our parameter of interest, albeit only weakly significant, testify the presence of a detrimental effect of corruption. In Table 6, having obtained one additional grade-point, in a scale ranging from 0 (the lowest grade) to 10 (the highest grade), has a positive but insignificant effect on the overall probability of joining the public sector. However, the impact is significant when we focus on managers and other professionals (for whom an additional grade-point implies a 2 percentage points increase in the likelihood of joining the public sector). Again the impact of grades is differentiated across LLMs characterized by a different intensity of corruption that negatively affect the propensity of more talented students to join the public sector. According to our estimates, the impact of the additional grade-point on the likelihood of joining the public sector for a manager is 3.7 and 0.2 in low- and high-corruption LLMs, respectively (4.5 and 0.9 when we use C^2 instead of C^1 to measure corruption).

4.2 The impact of corruption on the personnel allocation across jobs

In this section we inspect the impact of corruption on the effectiveness of the allocation process of human resources. The latter is examined comparing individual abilities and the skill content of jobs workers are assigned to. A mismatch may happen both in the direction of under-education (an individual is assigned to a task which is on average undertaken by more educated workers) or

over-education (an individual is assigned to a task which is on average undertaken by less educated workers).

As we have described before, under- and over-education are, on average, less frequent in the public sector. The aim of our empirical strategy is, again, to examine differential patterns between low- and high-corruption areas. Tables 7 show the results of the estimation of model (2). The coefficient associated to the interaction term between the dummy for the public sector and the measure of corruption is positive, thus suggesting that corruption increases the correlation between being in the public sector and the likelihood to be under-educated. The effects are again concentrated and display a higher magnitude among professions at the top of the occupational hierarchy. Moreover, the coefficient estimates are fairly similar using either measure of corruption. On the other hand, we do not find any detectable effect in terms of over education and this might be due to an inflation of professions with a higher (formally required) skill content in more corrupt LLMs, thus making over-education less likely by definition.

Skills mismatch might be, at least partially, a mechanical consequence of the negative selection patterns observed in the previous subsection. If corruption tends to make public employment relatively less attractive for the most educated, public agencies in corrupt areas will hire relatively less educated personnel. Assuming that the tasks assigned to each agency do not vary with the level of corruption, under-education will arise as the obvious outcome of having to fill the same job positions with less educated personnel. But under-education may also result from biased management practices which may be more likely to occur where corruption is more intense. As an attempt to disentangle the two channels, in Table 8 we estimate a model which is similar to (2) but includes two additional controls: the average skill content of professions present in each sector-LLM cell (measured as product of the nation-wide average of schooling in each profession and the share of professions in the cell) as well as the average education level in the same sector-LLM cell (measured with the average schooling of the employed in the cell).[14] These should respectively account for different educational endowments (and thus for the effects of selection) as well as for possible inflation in the number of high-level positions managed by public agencies. We indeed find that these additional controls explain, as expected, the level of under- and over-education; for example, under-education is more likely where the average schooling of employees is lower and where the average schooling required by the

[14] This is the LLM-by-sector average of the average education level within professions at the ISCO 3-digit level of disaggregation.

available job positions is higher. Nevertheless, we still find evidence of under-education, due to corruption, among managers and professionals and technicians.

4.3 Robustness

This section contains some robustness checks, motivated by various considerations.

First, one can argue that a sharp comparison between public and private sector may not take into account that the latter is also potentially affected by corruption and possibly differently so depending on the industry. Indeed, corruption typically involves converging interests or – at least – some kind of interaction between the public officials and the private firms whose activity is affected by public decisions. In Table 9 we replicate the previous analysis using a continuous indicator of dependence from public sector instead of the discrete variable. Stated differently, instead of a public sector indicator, we consider a continuous measure of *exposure to the public sector*, whose smallest values correspond to sectors that do not interact with the public sector and operate in competitive markets (e.g. the manufacturing sector), and the largest values to the areas of activity that we identify with the public sector (e.g. public administration). Mid-range values denote industries whose demand partly depends on public spending and/or that rely on regulated market (e.g. electricity, water, waste disposal, construction sectors, social activities, etc.).[15] The results are qualitatively and quantitatively fairly similar to the ones obtained before.[16]

We also replicate our analysis using just the manufacturing sector as control group which is hardly dependent on public spending and, being exposed to international competition, tends to be more-market oriented. Table 10 replicates our baseline results on selection, and under- and over-education using C^1 as our measure of corruption. Our main results are fully confirmed and our estimates are slightly larger, thus implicitly suggesting that misallocation due to corruption is somewhat extended also to the private industries that interact more with the public sector.

---

[15] More specifically, we map economic activities into [0,1], capturing the dependence and/or the proximity between each of economic sector of activity and the public administration, using the input-output matrix. The underlying idea is that corruption is not a phenomenon entirely contained within the public sector and that corruptive crimes typically involve interests of the private sector as well and, in particular, of the private industries that interact more with the public sector.

[16] All results of this subsection are qualitatively similar if we use C^2 instead of C^1. They are not reported for the sake of brevity.

Second, having shown that corruption is more widespread in the South of Italy, we examine to what extent our results are driven by the traditional North-South divide and whether they still hold when we compare more homogenous regions. Table 11 replicates our baseline results restricting the analysis to the LLMs located in the Centre-North, obviously at the cost of losing a significant number of observations and territorial variability. The estimates of the coefficients associated with the interaction term in the selection equation are fairly similar to those of our baseline specification, though the standard errors are higher, thus implying loss of statistical significance in some specifications. As far as misallocation is concerned, corruption continues to be significantly associated to under-education.

Third, our results might also be driven by other omitted variables correlated with corruption and implying differential effects similar to those produced by corruption. More precisely, this concern is not related to potential omitted variables driving the sorting between public and private sectors: those are already controlled for by the introduction of fixed effects at the LLM level. The concern relates to variables having a differential schooling-biased effect similar to that observed for corruption. To address this point, we enrich the specification with other local controls aimed at capturing relevant economic dimensions that are both correlated with corruption and potentially liable to affect individuals occupational choices (Table 12). In the top panel we include the (log of the) value added per employee interacted with schooling as determinants for selection into the public sector. The underlying idea is that higher, on average, economic prospects at the local level (and any other variable correlated with economic development) might affect the education-based sorting between public and private sector. Our main findings are basically unchanged. In the bottom panel we include population density at the LLM level that might affect both corruption and selection patterns, since the scope of public administration can differ between urban and rural areas. Indeed, we find evidence that higher density negatively affects education-based selection, likely due to wider employment opportunities for better educated in the private sector of more agglomerated areas. However, and more importantly for us, our coefficients of interest are unchanged.

One last concern is related to reverse causality. Indeed, corruption might itself be the result of poorly selected public employees while we are interested in the link *from* corruption *to* personnel selection and allocation. To address this issue we rely on an instrumental variable strategy. Our instruments are characterized by the common characteristics of being pre-dated with respect to the

hiring of current public employees. The first instrument is past local dependence from public-sector demand. The underlying idea, implicitly supported by our previous findings, is that corruption episodes are more likely to occur where the role of public spending is higher for the private sector. Indeed, unreported evidence documents that past economic dependence of the private sector on public demand is positively correlated with level of corruption at the LLM level. The second set of instruments considers the impact of past dominations. Again unreported evidence shows broadly consistency with other results on the cultural and institutional legacy of past foreign dominations.[17] Results using the first instrument are reported in Table 13, while those based on past dominations are reported in Table 14. Past economic dependence on public spending appears to be a strong determinant of corruption and the first stage F-statistic of the excluded instrument for the whole sample is above 37. On the contrary the predictive power of past dominations, in our empirically setting, is somewhat weaker and the F-statistic is slightly below 10. Looking at the second stage coefficients, they are qualitatively similar to those of our baseline specifications with both instruments, thus reassuring us on the identification of a link from corruption to personnel selection and allocation in the public sector.

4.4 Corruption and hiring procedures

So far we have showed that corruption has a detrimental impact on education-based selection. In this last section of the paper we examine whether the impact of corruption can be mitigated by different hiring procedures. The public sector is, in several countries including Italy, traditionally characterized by hiring procedures involving a relatively high degree of formalization (with, for instance, strict requirements in terms of age and educational attainments and/or several rounds of written and oral examinations). In Italy, such procedures are known as *concorso*. According to Article 97 of the Italian Constitution, *concorso* are the unique hiring procedure available for entry into the public sector. In practice, the hiring process is implemented very heterogeneously and the procedure might be highly differentiated in terms of the degree of formalization and in the number and types of examinations involved. In this respect, we examine whether corruption affects the differences in ability between those that access public employment

[17] Among past foreign dominations, corruption is positively correlated with the Norman domination and negatively correlated, albeit to a lesser extent, with other spells of foreign domination, except for the Angevine and Swabian ones.

through more or less formalized procedures. Unfortunately the data at our disposal do not allow us to perfectly observe these features and we have to rely on some approximation. To do so, we exploit the panel dimension of the LFS data. About one half of the sample is interviewed in two consecutive years. Among those, we consider workers who have joined the public sector no earlier than one year before the moment they were interviewed. For each of them, we know whether they had undertaken a *concorso* in the year prior to their second interview (or whether they were waiting for the results of a *concorso* in that same year). Our assumption is that only respondents who have undertaken sufficiently formalized procedures identify them as *concorso,* while less formalized procedures (e.g. interviews) are not regarded as *concorso* by the respondent themselves, even though they retain that status from the legal standpoint. This variable should therefore be interpreted with some caution. Moreover, our newly-hired workers may have been hired following a *concorso* taken more than one year before their most recent interview: for this reason, there may be workers who are wrongly recorded as having joined the public sector without taking a *concorso*. These two factors might explain why in our data only 5.3% of newly-hired public employees declare having undertaken a *concorso*[18].

From an empirical point of view, in order to gain insight into the interactions between hiring procedures and the environmental incidence of corruption, we estimate a linear probability model, similar to that discussed in equation (1). The main difference is that the sample has been restricted to newly-hired workers in the public sector and that the dependent variable is a dummy equal to 1 if the workers has taken a *concorso* in the previous years and 0 otherwise. Results, reported in Table 15, show a positive correlation between schooling and the likelihood of having been hired through a *concorso*, conditional on having joined the public sector. Moreover, they show that the educational divide between those who have passed a *concorso* and those who have not increases in the level of corruption. With the caveats previously discussed, we interpret these results as an overall indication that formal examinations may counteract the weakening of education-based selection implied by high levels of corruption. This might occur, for instance, because of a worsening in the selectivity of hiring without *concorso* in areas where corruption is more intense.

[18] A further reason behind the small number of newly-hired public employees who declare to have taken a concorso might be the frequent use of 'stabilization decrees', by which individuals who have been working for the public administration under fixed-term contracts are converted into permanent personnel without undergoing further examinations (see Rizzica, 2015).

3 Conclusions

Our analysis highlights the distortionary effect of corruption on the patterns of selection and allocation of public sector employees. Because of the nature of the tasks assigned to and areas of activity spanned by public agencies and other contracts arrangements, public employees are more educated with respect to their counterparts in the private sector. This gap however is thinner where corruption is higher, and the education-biased induced by corruption is particularly strong for professions at the top of the occupational hierarchy. Similar evidence is found if we consider further dimension of human capital such as grades obtained at school. Besides affecting selection, corruption contributes to deviating the education-based matching between workers and jobs: where corruption is higher, public employees are relatively more likely to be assigned to tasks which are, on average, undertaken by more qualified personnel.

The comparative analysis of our results thus suggests that higher levels of corruption are associated with a poorer capacity of the public sector to select and allocate workers. Hence – if one believes that the workforce's human capital is conducive of better decision making – where corruption is high, the public administration will tend to adopt socially inefficient decisions and to remunerate individual less in terms of schooling ability than of other (unobserved) ability traits such as soft skills, relational capital, craftiness, etc.

The eradication of corruption or, at least, the dampening of its implications have long been a major objective of governmental effort. Actions taken by governmental authorities usually rest on ex-post, repressive measures, which are sometimes accompanied by ex-ante, preventive provisions. The latter often take the form of a requirement for individual agencies to implement "in-house" anti-corruptive programs under governmental supervision. In light of the evidence presented in this paper, one may suspect that the administrations' ability to exert anti-corruptive self-monitoring might be hindered by corruption itself. Indeed, existing levels of crime in the environment may have contributed to the selection of a workforce which will, in general, be more likely to be misallocated as well as less prone to take up action against corruption if called to do so. The risk is that self-regulation aimed at overcoming corruption may work well only where corruption is already rare and fare poorly where it is more intense. Hence our

results suggest caution against over-estimating the additional benefits of ex-ante provisions.[19]

As far as the recruitment process is concerned, our results suggest that formalized procedures may work as a first safeguard against the distortions implied by corruption. Though formal examinations have some drawbacks, in terms of costs and lack of flexibility and though they can also be manipulated to some extent by corrupt officers, they may partially counteract the effects of corruption through the guarantee of substantial homogeneity and objectivity.

[19] Other results indirectly confirm this conclusion. Indeed, in an unreported evidence we run a regression to examine whether corruption was correlated with the score attributed by ANAC (the Italian National Anti-Corruption Authority) to a sample of municipalities in terms of the effectiveness of the anti-corruptive practices implemented at the local level. Our findings show a negative and significant correlation between the two variables that is robust to the inclusions of several controls.

References

Acemoglu, D. and J.A. Robinson (2012), *Why nations fail: the origins of power, prosperity, and poverty*, New York: Random House.

Acemoglu, D. and T. Verdier (1998), Property rights, corruption and the allocation of talent: a general equilibrium approach, *Economic Journal*, vol. 108, pp. 1381-1403.

Banerjee, R., T. Baul and T. Rosenblat (2015), On self selection of the corrupt into the public sector, *Economics Letters*, vol. 127, pp. 43-46.

Barfort, S., N. Harmon, F. Hjorth, A.L. Olsen (2016), Sustaining honesty in public service: the role of selection, working paper.

Barone, G. and S. Mocetti (2014), Natural disasters, growth and institutions: a tale of two earthquakes, *Journal of Urban Economics*, vol. 84, pp. 52-66.

Brandolini, A. and L. Cannari (1994), Methodological appendix: the Bank of Italy Survey of household income and wealth, in A. Ando, L. Guiso and I. Visco (eds.), *Saving and the accumulation of wealth. Essays on Italian households and government saving behavior*, Cambridge: Cambridge University Press.

Brollo, F., T. Nannicini, R. Perotti and G. Tabellini (2013), The political resource curse, *American Economic Review*, vol. 103, pp. 1759‑1796.

Charron, N., V. Lapuente, and L. Dijkstra (2014), Regional governance matters: quality of government within European Union member states, *Regional Studies*, vol. 48, pp. 68-90.

Clausen, B., A. Kraay and Z. Nyiri (2011), Corruption and confidence in public institutions: evidence from a global survey, *World Bank Economic Review*, vol. 25, pp. 212-249.

Cowley, E. and S. Smith (2014), Motivation and mission in the public sector: evidence from the World Values Survey, *Theory and Decision*, vol. 76, pp. 241-263.

Del Monte, A. and E. Papagni (2001), Public expenditure, corruption, and economic growth: the case of Italy, *European Journal of Political Economy*, vol. 17, pp. 1-16.

Del Monte, A. and E. Papagni (2007), The determinants of corruption in Italy: regional panel data analysis, *European Journal of Political Economy*, vol. 23, pp. 379-396.

Di Liberto, A. and M. Sideri (2015), Past dominations, current institutions and the Italian regional economic performance, *European Journal of Political Economy*, vol. 38, pp. 12-41.

Ferraz, C. and F. Finan (2011), Electoral accountability and corruption: evidence from the audits of local governments, *American Economic Review*, vol. 101, pp. 1274‑1311.

Fisman, R. and R. Gatti (2002), Decentralization and corruption: evidence across countries, *Journal of Public Economics*, vol. 83, pp. 325–345.

Fisman, R. and E. Miguel (2007), Corruption, norms, and legal enforcement: evidence from diplomatic parking tickets, *Journal of Political Economy*, vol. 115, pp. 1020–1048.

Giorgiantonio, C., T. Orlando, G. Palumbo and L. Rizzica (2016), Incentives and selection in public employment, Bank of Italy, Occasional Papers no. 342.

Golden, M.A. and L. Picci (2005), Proposal for a new measure of corruption, illustrated with Italian data, *Economics & Politics*, vo. 17, pp. 37-75.

Hanna, R. and S.Y. Wang (2014), Dishonesty and selection into public service: evidence from India, working paper.

Hessami, Z. (2014), Political corruption, public procurement, and budget composition: theory and evidence from OECD countries, *European Journal of Political Economy*, vol. 34, pp. 372–389.

Knack, S. and P. Keefer (1995), Institutions and economic performance: cross-country tests using alternative institutional measures, *Economics and Politics*, vol. 7, pp. 207–227.

La Porta, R., F. Lopez-de Silanes, A. Shleifer and R. Vishny (1999), The quality of government, *Journal of Law, Economics, and Organization*, vol. 15, pp. 222–279.

Mauro, P. (1995), Corruption and growth, *Quarterly Journal of Economics*, vol. 110, pp. 681-712.

Mauro, P. (1998), Corruption and the composition of government expenditure, *Journal of Public Economics*, vol. 69, pp. 263–279.

Mo, P.H. (2001), Corruption and economic growth, *Journal of Comparative Economics*, vol. 29, pp. 66-79.

Murphy, K.M., A. Shleifer and R.W. Vishny (1991), The allocation of talent: implications for growth, *Quarterly Journal of Economics*, vol. 106, pp. 503-530.

Olken, B.A. (2009), Corruption perceptions vs. corruption reality, *Journal of Public Economics*, vol. 93, pp. 950–964.

Olken, B.A. and P. Barron (2009), The simple economics of extortion: evidence from trucking in Aceh, *Journal of Political Economy*, vol. 117, pp. 417‑452.

Rizzica, L. (2015), The use of fixed-term contracts and the (adverse) selection of public sector workers, Bank of Italy working paper n. 1041.

Rizzica, L. (2016), Why go public? A study of the individual determinants of public sector employment choice, Bank of Italy occasional paper n. 343.

Rizzica, L. and M. Tonello (2015), Exposure to media and corruption perceptions, Bank of Italy working paper n. 1043.

Shleifer, A. and R.W. Vishny (1993), Corruption, *Quarterly Journal of Economics*, vol. 108, pp. 599-617.

Uslaner, E.M. (2004), Trust and corruption, in J.G. Lambsdorf, M. Taube and M. Schramm (eds.), *Corruption and the new institutional economics*, London: Routledge.

Tables

Table 1. Descriptive statistics

	Full sample		*Public sector employees*	
	Mean	St. dev.	Mean	St. dev.
	LFS data			
Female	0.561	0.496	0.656	0.475
Young (<35)	0.293	0.455	0.149	0.356
Years of schooling	12.54	3.337	13.54	3.467
Under-education	0.126	0.332	0.124	0.329
Over-education	0.090	0.286	0.103	0.304
# observations	*753,938*		*301,120*	
	SHIW data			
Female	0.365	0.482	0.382	0.486
Young (<35)	0.110	0.313	0.073	0.260
Years of schooling	13.09	3.256	13.43	3.481
Grades obtained at school	0.824	0.136	0.837	0.138
# observations	*11,511*		*5,905*	

Years of schooling are those corresponding to each individual's highest educational attainment. Grades at school is a measure of schooling ability and correspond to the grades (normalized with respect to the maximum obtainable grade) obtained at the highest education attainment of the individual. Under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits).
Sources: authors' elaborations on data drawn from LFS and SHIW.

Table 2. Corruption: principal component analysis

	1st component	2nd component	3rd component	4th component
Eigenvalue	2.573	0.883	0.330	0.214
Proportion	0.643	0.221	0.082	0.054
Cumulative	0.643	0.864	0.946	1.000
	C^1	Trust	GP	CPI
Coefficient 1st component	0,365	0,482	0,382	0,486

Results of the principal component analysis.
Sources: authors' elaborations on ISTAT, Ministry of Justice, Golden-Picci (2015) and EQI data.

Table 3. Descriptive statistics

	Mean	S.D.	10th	25th	50th	75th	90th
Crime rate: C^1	0.000	1.000	-0.570	-0.442	-0.254	0.141	0.765
Principal component: C^2	0.000	1.000	-1.299	-0.739	0.133	0.856	1.225

Corruption indicators are standardized at the LLM level.
Sources: authors' elaborations on ISTAT, Ministry of Justice, Golden-Picci (2015) and EQI data.

Table 4. Selection in the public sector: the impact of schooling (LFS)

Dependent variable:	Employed in the public sector			
Professional area:	All	Managers/ professionals	Technicians	Clerical workers
Years of schooling	0.022***	0.045***	0.016***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.002**	-0.009***	-0.004**	0.001
	(0.001)	(0.002)	(0.002)	(0.001)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.255	0.280	0.260	0.131
Years of schooling	0.021***	0.044***	0.015***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^2	-0.004***	-0.007***	-0.004***	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.256	0.281	0.260	0.131
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS. The dependent variable is equal to 1 for public sector employees and to 0 for private sector employees. Years of schooling are those corresponding to each individual's highest educational attainment. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 5. Selection in the public sector: the impact of schooling (SHIW)

Dependent variable:	Employed in the public sector		
Professional area:	All	Managers/ professionals	Clerical workers
Years of schooling	0.009***	0.033***	-0.003
	(0.002)	(0.003)	(0.003)
Years of schooling × C^1	-0.006*	-0.011*	-0.003
	(0.004)	(0.006)	(0.005)
LLM × professional area FEs	YES	YES	YES
R-squared	0.193	0.255	0.152
Years of schooling	0.008***	0.030***	-0.003
	(0.002)	(0.003)	(0.003)
Years of schooling × C^2	-0.005**	-0.013***	-0.001
	(0.002)	(0.003)	(0.004)
LLM × professional area FEs	YES	YES	YES
R-squared	0.193	0.258	0.152
# observations	*11,477*	*3,858*	*7,617*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the SHIW. The dependent variable is equal to 1 for public sector employees and to 0 for private sector employees. Years of schooling are those corresponding to each individual's highest educational attainment. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider two professional areas for employees: managers and professionals, clerical workers.

Table 6. Selection in the public sector: the impact of grades (SHIW)

Dependent variable:	Employed in the public sector		
Professional area:	All	Managers/ professionals	Clerical workers
Grades at school	0.029	0.203**	-0.046
	(0.069)	(0.095)	(0.076)
Grades at school × C^1	-0.128	-0.286*	-0.046
	(0.110)	(0.157)	(0.134)
LLM × professional area FEs	YES	YES	YES
R-squared	0.229	0.274	0.168
Grades at school	0.018	0.172*	-0.046
	(0.077)	(0.100)	(0.080)
Grades at school × C^2	-0.073	-0.217**	-0.006
	(0.060)	(0.084)	(0.068)
LLM × professional area FEs	YES	YES	YES
R-squared	0.229	0.276	0.168
# observations	*9,107*	*3,636*	*5,471*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the SHIW. The dependent variable is equal to 1 for public sector employees and to 0 for private sector employees. Grades at school is a measure of schooling ability and correspond to the grades (normalized with respect to the maximum obtainable grade) obtained at the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include years of schooling and fixed effects for gender and age cohort. We consider two professional areas for employees: managers and professionals, clerical workers.

Table 7. Under- and over-education in the public sector

Professional area:	All	Managers/ professionals	Technicians	Clerical workers
Dependent variable:		Under-education		
Public sector × C^1	0.009**	0.021***	0.017***	0.007
	(0.004)	(0.007)	(0.006)	(0.004)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.058	0.060	0.092	0.095
Public sector × C^2	0.007***	0.020***	0.011***	0.010***
	(0.002)	(0.003)	(0.003)	(0.003)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.058	0.060	0.092	0.095
Dependent variable:		Over-education		
Public sector × C^1	0.000	-0.000	0.009	-0.002
	(0.003)	(0.003)	(0.006)	(0.004)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.048	0.039	0.071	0.062
Public sector × C^2	-0.002	-0.003	0.004	-0.003
	(0.002)	(0.002)	(0.003)	(0.002)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.048	0.039	0.071	0.062
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS. The dependent variable under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 8. Under- and over-education (education and skill content corrections)

Professional area:	All	Managers/ professionals	Technicians	Clerical workers
Dependent variable:		Under-education		
Public sector × C^1	0.004	0.011*	0.011**	0.002
	(0.003)	(0.006)	(0.005)	(0.004)
Average skill content	0.070***	0.069***	0.078***	0.072***
	(0.002)	(0.005)	(0.004)	(0.003)
Average educational level	-0.067***	-0.070***	-0.079***	-0.056***
	(0.001)	(0.004)	(0.002)	(0.002)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.069	0.069	0.105	0.105
Dependent variable:		Over-education		
Public sector × C^1	0.003	0.004	0.011**	-0.002
	(0.002)	(0.003)	(0.005)	(0.004)
Average skill content	-0.050***	-0.029***	-0.062***	-0.034***
	(0.002)	(0.003)	(0.003)	(0.002)
Average educational level	0.055***	0.027***	0.077***	0.048***
	(0.001)	(0.002)	(0.002)	(0.002)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.058	0.043	0.084	0.073
# observations	*753,048*	*135,127*	*269,490*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS. The dependent variable under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Average skill content is the LLM-by-sector is measured as the product of the nation-wide average of schooling in each profession and the share of each professions in the sector-LLM cell; average educational level is the average schooling of the employed in the sector-LLM cell. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 9. Robustness: continuous measure of exposure

Dependent variable:	Employed in the public sector			
Years of schooling	0.019***	0.035***	0.013***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.002***	-0.008***	-0.004**	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.238	0.263	0.244	0.139
Dependent variable:	Under-education			
Public sector × C^1	0.011**	0.023***	0.022***	0.006
	(0.004)	(0.008)	(0.007)	(0.004)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.058	0.059	0.092	0.095
Dependent variable:	Over-education			
Public sector × C^1	0.001	0.002	0.001	-0.001
	(0.003)	(0.004)	(0.007)	(0.004)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.048	0.039	0.071	0.062
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS; only employees in the public and in the manufacturing sector are included. The dependent variables are the following: employed in the public sector is equal to 1 for public sector employees and to 0 for private sector employees; under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 10. Robustness: only manufacturing sector as control group

Dependent variable:	Employed in the public sector			
Years of schooling	0.016***	0.025***	0.017***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.008***	-0.013***	-0.009***	-0.003**
	(0.001)	(0.002)	(0.002)	(0.001)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.269	0.218	0.321	0.226
Dependent variable:	Under-education			
Public sector × C^1	0.040***	0.041***	0.037***	0.036***
	(0.007)	(0.012)	(0.010)	(0.009)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.045	0.059	0.089	0.085
Dependent variable:	Over-education			
Public sector × C^1	0.002	0.001	0.006	0.001
	(0.005)	(0.008)	(0.009)	(0.007)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.043	0.038	0.073	0.054
# observations	*397,060*	*99,663*	*169,686*	*127,711*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS; only employees in the public and in the manufacturing sector are included. The dependent variables are the following: employed in the public sector is equal to 1 for public sector employees and to 0 for private sector employees; under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 11. Robustness: only Centre-North

Dependent variable:	Employed in the public sector			
Years of schooling	0.022***	0.049***	0.016***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.004	-0.006	-0.009**	-0.001
	(0.003)	(0.004)	(0.005)	(0.003)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.205	0.264	0.224	0.088
Dependent variable:	Under-education			
Public sector × C^1	0.030***	0.041**	0.047***	0.034**
	(0.010)	(0.019)	(0.016)	(0.013)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.063	0.062	0.093	0.098
Dependent variable:	Over-education			
Public sector × C^1	0.014*	0.016	0.009	0.012
	(0.008)	(0.012)	(0.015)	(0.010)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
R-squared	0.050	0.044	0.073	0.061
# observations	*522,716*	*85,589*	*193,105*	*244,022*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS; only employees located in the Centre-North of Italy are included. The dependent variables are the following: employed in the public sector is equal to 1 for public sector employees and to 0 for private sector employees; under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 12. Robustness: adding further controls

Dependent variable:	Employed in the public sector			
Years of schooling	0.022**	0.015	0.030**	0.020*
	(0.010)	(0.016)	(0.012)	(0.011)
Years of schooling × C^1	-0.002**	-0.009***	-0.004**	0.001
	(0.001)	(0.002)	(0.002)	(0.001)
Years of schooling × VA	0.000	0.008*	-0.003	-0.001
	(0.003)	(0.004)	(0.003)	(0.003)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.255	0.281	0.260	0.131
Years of schooling	0.0233***	0.046***	0.018***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.002**	-0.009***	-0.004**	0.001
	(0.001)	(0.002)	(0.002)	(0.001)
Years of schooling × Density	-0.002***	-0.002**	-0.003***	-0.001**
	(0.000)	(0.001)	(0.001)	(0.001)
LLM × professional area FEs	YES	YES	YES	YES
R-squared	0.255	0.280	0.260	0.131
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS. The dependent variable is equal to 1 for public sector employees and to 0 for private sector employees. Years of schooling are those corresponding to each individual's highest educational attainment. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). We also include controls that might affect the public-private sector sorting between individuals with different educational level: the log of value added per employee at the LLM level (VA) and the LLM population density (Density). Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 13. Robustness: IV estimates using past public dependence

Dependent variable:	Employed in the public sector			
Years of schooling	0.022***	0.045***	0.016***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.012**	-0.040***	-0.024***	0.007*
	(0.005)	(0.007)	(0.006)	(0.004)
LLM × professional area FEs	YES	YES	YES	YES
F-stat of excluded instruments	98.40	33.23	37.51	37.56
R-squared	0.255	0.271	0.258	0.130
Dependent variable:	Under-education			
Public sector × C^1	0.054***	0.093***	0.065***	0.057***
	(0.013)	(0.025)	(0.018)	(0.020)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
F-stat of excluded instruments	37.4	31.6	37.0	36.3
R-squared	0.057	0.057	0.091	0.094
Dependent variable:	Over-education			
Public sector × C^1	-0.004	-0.008	0.003	-0.009
	(0.011)	(0.014)	(0.019)	(0.014)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
F-stat of excluded instruments	37.4	31.6	37.0	36.3
R-squared	0.048	0.039	0.071	0.062
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS; only employees located in the Centre-North of Italy are included. The dependent variables are the following: employed in the public sector is equal to 1 for public sector employees and to 0 for private sector employees; under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider reported crimes net of judicial efficiency (C^1) instrumented with past public sector dependence. Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

Table 14. Robustness: IV estimates using past dominations

Dependent variable:	Employed in the public sector			
Years of schooling	0.022***	0.045***	0.016***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of schooling × C^1	-0.010***	-0.024***	-0.013***	-0.002
	(0.003)	(0.004)	(0.004)	(0.002)
LLM × professional area FEs	YES	YES	YES	YES
F-stat of excluded instruments	22.59	7.481	7.643	9.567
R-squared	0.255	0.278	0.260	0.131
Dependent variable:	Under-education			
Public sector × C^1	0.028***	0.068***	0.041***	0.029***
	(0.007)	(0.014)	(0.010)	(0.011)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
F-stat of excluded instruments	8.8	7.9	7.9	9.4
R-squared	0.058	0.059	0.092	0.095
Dependent variable:	Over-education			
Public sector × C^1	-0.008	-0.009	0.015	-0.013*
	(0.005)	(0.007)	(0.011)	(0.007)
LLM FEs	YES	YES	YES	YES
Sector of activity FEs	YES	YES	YES	YES
F-stat of excluded instruments	8.8	7.9	7.9	9.4
R-squared	0.048	0.039	0.071	0.062
# observations	*753,043*	*135,127*	*269,491*	*348,425*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual employees, drawn from the LFS; only employees located in the Centre-North of Italy are included. The dependent variables are the following: employed in the public sector is equal to 1 for public sector employees and to 0 for private sector employees; under-education (over-education) is equal to 1 if the employee has a number of years of schooling below the 25th percentile (above the 75th percentile) of the years of schooling distribution of the jobs he/she is assigned to (ISCO occupational at 3 digits). Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider reported crimes net of judicial efficiency (C^1) instrumented with length of past dominations. Other controls include fixed effects for gender and age cohort. We consider three professional areas for employees: managers and professionals, technicians, clerical workers.

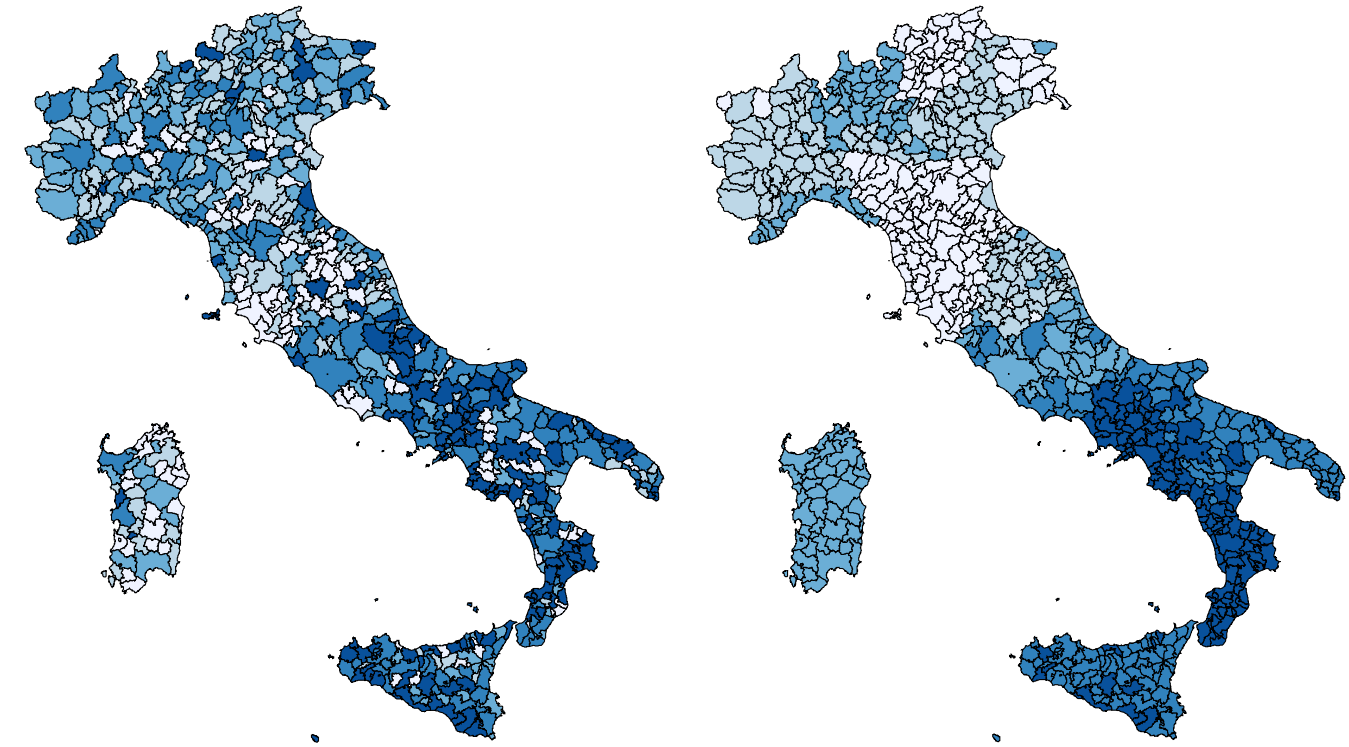
Table 15. Selection in the public sector: the role of *concorso*

Dependent variable:	Having taken a *concorso*	
Years of schooling	0.008***	0.009***
	(0.001)	(0.001)
Years of schooling × C^1	0.004*	
	(0.002)	
Years of schooling × C^2		0.003***
		(0.001)
LLM × professional area FEs	YES	YES
# observations	*8,669*	*8,669*

Standard errors are clustered at the LLM level (* p<0.1, ** p<0.05, *** p<0.01). The sample includes non-manual newly-hired public sector employees, drawn from the LFS. The dependent variable is equal to 1 for those who have taken a *concorso* in the previous year and to 0 otherwise. Years of schooling are those corresponding to the highest education attainment of the individual. Corruption is measured at the LLM level and we consider two measures: reported crimes net of judicial efficiency (C^1) and the principal component of CPI, $TRUST$, GP and C^1 (C^2). Other controls include fixed effects for gender and age cohort.

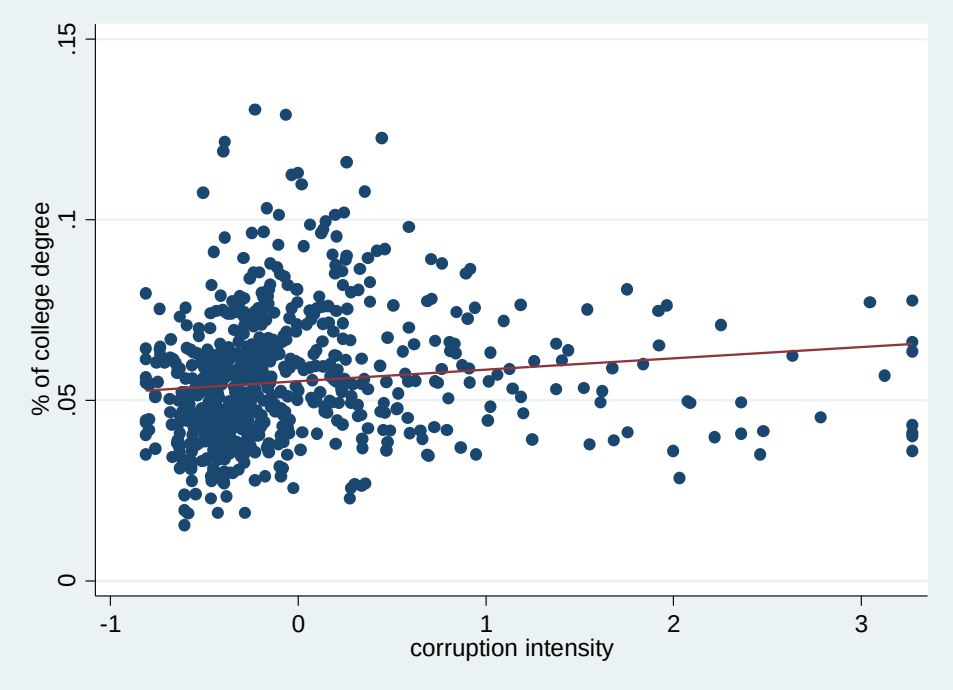
Figures

Figure 1. Map of corruption



Sources: authors' elaborations on ISTAT, Ministry of Justice, Golden-Picci (2015) and EQI data.

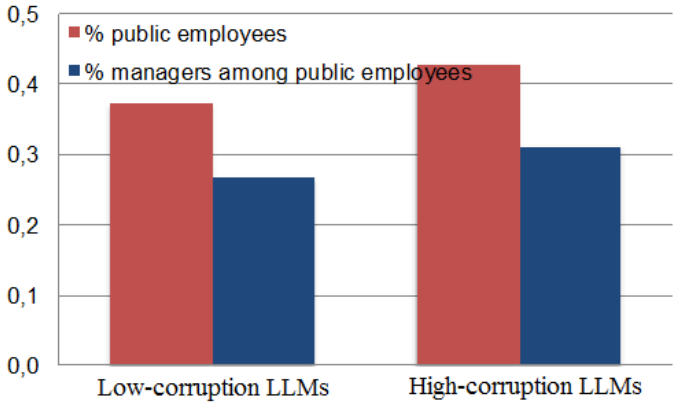
Figure 2. Corruption and education across LLMs



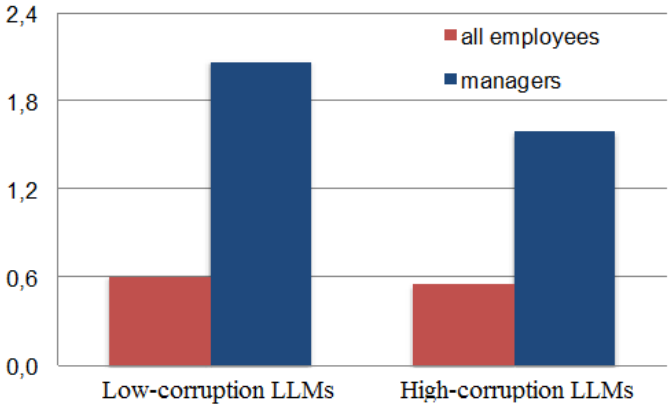
Corruption intensity is measured with reported crimes net of judicial efficiency; the share of population with a college degree is drawn from Census 2001.
Sources: authors' elaborations on ISTAT and Ministry of Justice data.

Figure 3. Corruption and public employment

Share of public employees



Public-private difference in schooling



Sources: authors' elaborations on ISTAT and Ministry of Justice data.