Why Are Some Regions So Much More Productive than Others?

Chiara Lacava*

June 2020

Abstract

Differentials in aggregate labor productivity across regions could be due to differences (1) in workers' skills, (2) in firms' technologies, (3) in how efficiently skills and technologies are matched, and (4) in institutional factors specific to each region. I introduce a framework to separately identify each determinant using matched employer-employee data. I estimate skills by comparing the wages of each worker with co-workers in the same firm and in the same region; technologies by comparing the productivity of firms with the same workforce's skills and in the same region; and positive complementarities between skill and technology by measuring how workers and firms jointly contribute to productivity in the same region through a model of the aggregate production function. Finally, I disentangle region-specific factors from technologies since some firms have plants in more than one region. In an application to the Italian regions, I find that differences in firms' technologies and in region-specific factors account respectively for 65% and 27% of the large productivity differentials. In contrast, little contribution is due to differences in the distribution and allocation of skills. Also, optimal reassignment of workers does not shrink productivity differences, but productivity levels do rise by around 16%.

Keywords: cross-regional differences, productivity, mismatch, sorting, wage dispersion, matched employer-employee data.

JEL codes: O47, R11, E24, J24, J31.

---

*Goethe University, Theodor-W.-Adorno-Platz 3, 60629 Frankfurt. Email: lacava@econ.uni-frankfurt.de. I am grateful to Claudio Michelacci for his invaluable and patient guidance. I thank Nicola Fuchs-Schündeln, Leo Kaas, Thibaut Lamadon, Rafael Lopes de Melo, Alexander Ludwig, Iourii Manovskii, Guido Menzio, Eleonora Patacchini, Facundo Piguillem, Víctor Ríos Rull and Fabiano Schivardi for the helpful conversations, as well as seminars and conferences participants at University of Rome Tor Vergata, UPenn Macro Club, EIEF, IFN-Stockholm, Goethe University Frankfurt, Bank of Italy-CEPR-EIEF conference on Firm Dynamics, CASD-IAB Advances in Social Sciences with Administrative Data, Search and Matching BI Oslo, Midwest Macro Georgia, SED Washington St.Louis, SIEPI RomaTre, RES PhD Meetings Westminster, XXXII Jornadas de Economìa Industrial Navarra, IWEEE Milano Bicocca. I gratefully acknowledge financial support by NORFACE Dynamics of Inequality across the Life Course (TRISP) grant 462-16-120.

1 Introduction

Huge and persistent differences in output per worker across economies are extensively documented. In 2019, a single worker in the United States produced the same output of 4 workers in China, 6 in India and 30 in Zimbabwe. Notwithstanding the many studies investigating this issue, there is still a big debate on what drives such differentials in labor productivity.[1] They might be due to differentials in the skills of the workers or in the technologies of the firms. Moreover, in the presence of production complementarities they might depend on how efficiently workers and firms are matched (Shimer and Smith, 2000). Finally, the literature agrees that differentials in productivity are also explained by a technological component specific to the economy and disembodied from workers and firms, defined as social infrastructure in Hall and Jones (1999) or as institutions in Acemoglu et al. (2005).

This paper proposes a strategy to identify separately each of these channels by using matched employer-employee data. I apply this novel approach to disentangle productivity differentials between regions in a single country. The Italian case is of particular interest since dispersion in labor productivity across regions is especially high compared to other countries. I find that productivity differentials are mainly explained by differences in firms' technologies (accounting for over 65% among most different regions) and in region-specific factors (around 27%). The distribution of skills is very similar across regions and has a limited contribution to explaining differences in productivity (3% to 9%). Similarly, the allocation of workers to firms is homogeneous across regions and plays a minor role in determining heterogeneity in productivity. I also verify whether productivity differentials are amplified or reduced by the migration of workers and by firms relocating across regions: mobility has a negligible impact on these differentials.

To isolate the skill of each worker, I compare employees of the same firm in the same region. If the wage is increasing in a worker's skill, it is a sufficient statistic of the skill ranking of a worker within an establishment. Following the methodology proposed by Hagedorn et al. (2017), I compare workers in different firms and regions by observing how they are ranked with respect to their co-workers who move across firms and regions. Then, after having retrieved the skills of the workers, I estimate technologies by comparing the productivity of firms with similar workforce skills and in the same region. I specify a model of the aggregate production function to estimate how skills and technologies contribute jointly to productivity in the same region. By choosing a CES specification, the elasticity parameter captures the complementarity between skill and technology at the match level, and its value determines how much the similarity between the productivity of the worker and the firm increases the productivity of the match. Finally, I disentangle region-specific factors by observing firms with plants in more than one region and by assuming that the technology of a firm

[1] Syverson (2011) reviews the approach adopted by different literatures. Beyond being a crucial inquiry in growth economics, this question shaped the agenda also in labor economics, international economics and industrial organization.

is constant across regions.

By using a dataset of large manufacturing firms, containing administrative records on the working history of every employee in each firm, I find large differences in the distributions of the technologies of firms across Italian regions, while the distributions of skills of employed workers are homogeneous across regions. This is also driven by the low variation of skills between firms, accounting only for 13% of the total variance of skills. Coherently, the correlation between the ranking of the worker and the ranking of the employer firm is close to zero, suggesting that workers do not sort systematically into firms of similar productivity. However, direct estimation of the complementarities in the production function of the economy detects an elasticity of 0.46 between a worker's skill and a firm's technology and reveals the presence of positive assortative matching, a result that is robust to several specifications. This indicates that the allocation of workers across firms do not exploit the complementarities between skills and technologies, and the mismatch of workers across firms generates aggregate losses.

To assess the contribution of each channel to differentials in productivity, I compute the counterfactual output obtained by assigning one at a time the skill distribution, the technology distribution, the allocation and the regional factor of the most productive region, Northeast, to the other regions. The most important channel of productivity differentials is related to the technologies of the firms: out of a total difference ranging between 9% and 17%, when assigning the distribution of technologies observed in the Northeast, labor productivity increases by 1% in the Northwest, by 4% in Central Italy, by 10% in the South and by 13% in the Islands. Also, differences in the region-specific factors account for an important part of the differentials in productivity, notwithstanding that some sources of institutional differences traditionally considered relevant by the literature (like the legislative and fiscal framework) are common to all the regions. On the contrary, skills and allocation of the workers do not drive productivity differentials. Previous work by Hellerstein and Neumark (2007) for the U.S. and Fox and Smeets (2011) for Denmark also find only a modest role for skills of the workforce in explaining productivity differences between firms in the manufacturing industry.

Mismatch is homogeneous across regions, as also confirmed by similar values of correlation measures across regions, and it does not significantly explain productivity differentials. However, the loss in productivity due to suboptimal allocation of workers to firms is high: under optimal assignment of better workers to better firms within each region, labor productivity would rise by a factor between 11% and 21%, depending on the region. Finally, this framework allows for computing variations in productivity in a scenario where workers cannot move across economies and firms cannot relocate to other economies. Mobility seems to play little role in explaining labor productivity differentials: there are small changes in productivity by assigning to each region the distribution of skills of workers born there or the distribution of technologies of firms that have their headquarter in that region.

This paper contributes to the empirical literature on growth and inequality across economies (reviewed in Caselli, 2005) by proposing a novel approach to measure what explains differentials in productivity using matched employer-employee data, currently available for many countries. Following the seminal contribution of Barro (1991), countless works used cross-country regressions to detect empirical linkages between the dispersion in output per capita and a variety of indicators related to human capital, physical capital and institutional indicators (see Durlauf et al., 2005 for a review). While aggregate time series provide insufficient information to quantify the role due to different determinants (Levine and Renelt, 1992; Sala-I-Martin, 1997), exploring disaggregate data can provide new insights.

I implement the estimation techniques developed in the empirical labor literature for the identification of worker and firm unobservable productivity (e.g., Abowd et al., 1999; Card et al., 2013; Hagedorn et al., 2017) to separate the contribution to productivity attributable to sole workers' skills and sole firms' technologies. Then, I integrate this estimation approach into a model of the production function to infer complementarities between skill and technology and to assess the importance of differences in skills, technologies, in their allocation and in other factors unrelated to workers and firms.

Moreover, this work is related to the literature on misallocation and productivity, documenting that the allocation of inputs across heterogeneous production units is an important determinant of aggregate productivity (see Restuccia and Rogerson, 2013 for a survey). Several papers study the consequences of frictions affecting the firm's choice of labor and capital on allocation and aggregate output, focusing on the role of either implicit wedges (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Bartelsman et al., 2013) or credit market imperfections (Buera et al., 2011; Midrigan and Xu, 2014). While these papers focus on deviations from the optimal amount of resources assigned to heterogeneous firms or entrepreneurs, I investigate the effect of deviations from the optimal quality of resources (i.e. skill levels) assigned across heterogeneous firms. More specifically, I evaluate mismatch with an indirect approach similar to the one in Hsieh and Klenow (2009): the cost of suboptimal allocation is measured as the difference between the output observed in the data and the potential output simulated in the absence of allocative distortions –in their model, by imposing equal revenue productivity across firms; here, by imposing an optimal assignment of workers. Previous quantifications of the loss from the mismatch of workers to firms are proposed in the context of equilibrium models of the labor market with wage dispersion by Lise et al. (2016), Hagedorn et al. (2017) and Bagger and Lentz (2019).

Finally, this work contributes to the literature evaluating assortative matching in the labor market by providing an alternative approach to identify the sign and strength of sorting. Previous literature tries to identify sorting using the wages and transitions of workers alone. I complement this information with data on productivity at the firm level, which allows me to estimate the elasticity of substitution in the production function, relying on a small set of assumptions, namely the production function specification and the strict monotonicity of wages in the skill type within each establishment. The use of firm-level data was previously exploited by Bartolucci et al. (2018), who rank firms by their profits.

The rest of the paper is organized as follows. Section 2 describes the model of production and section 3 explains the identification strategy. In section 4 I show that the Italian case is of particular interest due to large differentials across regions and I describe the data. Section 5 reports the estimates of the determinants of productivity, section 6 describes the counterfactual exercises and section 8 concludes.

2 Model

The economy is a collection of regions ($r = 1, \dots, R$), firms ($f = 1, \dots, F$) and workers ($i = 1, \dots, I$). Denote with $\mathcal{F}_{rt}$ the set of firms in region r at time t and with F_{rt} the number of firms in region r, $F_{rt} \equiv \sum_{f=1}^{F} \mathbf{1}(f \in \mathcal{F}_{rt})$. A firm can hire workers in more than one region. The collection of workers assigned to a firm f in region r is called a plant. Let $\mathcal{E}_{frt}$ and E_{frt} be respectively the set and the number of job matches in firm f in region r, $E_{frt} \equiv \sum_{i=1}^{I} \mathbf{1}(i \in \mathcal{E}_{frt})$. Employment E_{rt} is the total number of workers employed by plants in region r.

Agents are heterogeneous in productivity. Each worker i is characterized by a skill x_i and each firm f by a technology y_f. Skills and technologies are known by all agents and constant in time. Workers do not enjoy any learning-by-doing or human capital accumulation; workers' schooling decisions are taken as exogenous since they happen before entering the labor market; and experience is assumed to have no impact on the worker's ability. Firms choose their capital optimally and the level of capital is embodied in the unidimensional measure y_f.[2] Moreover, skills and technologies are independent of location. The location of a job in a region affects the output of the match directly through a region-specific factor A_r, which varies between regions but is constant for any worker-firm match within the same region and in time.[3]

The output of a match is a function of the skill of the worker, the technology of the firm and the regional factor, $Y_{ifr} = g(x_i, y_f, A_r)$. Firms hire multiple workers. The output of a plant is the sum of the output of the job matches of that plant and the observable output of a firm f at time t is

$$Y_{ft} = \sum_{r=1}^{R} \mathbb{1}_{(E_{frt}>0)} g(x_i, y_f, A_r). \tag{1}$$

[2]The empirical literature shows that the productivity of a firm changes very slowly. Then, the assumption of constant technology is especially plausible when the analysis spans an interval of time of a few years.

[3]The assumption of technology independent of the location implies that the production process is similar across plants of the same firm, and excludes strategic allocation of plants across regions. In the application presented here, this is a plausible assumption since plants with different production chains are usually classified under different identifiers, notwithstanding the ownership by the same group.

Variations of the firm output in time are due to changes is the set of workers E_{frt}. From accounting, the output of a region (i.e. the GDP of region r) is the sum of the output of all matches of all plants in region r. Assume that the match production function is specified as a CES function,

$$g(x_i, y_f, A_r) = A_r \cdot \left((1-\lambda)x_i^{\rho} + \lambda y_f^{\rho}\right)^{\frac{1}{\rho}}, \tag{2}$$

where the region-specific factor A_r enters as a Hicks-neutral component, $\lambda \in [0,1]$ sets the weight of skill relative to technology in production and ρ determines the degree of substitutability between skill and technology, with $1/(1-\rho)$ being the elasticity of substitution. Since skills and technologies are heterogeneous, the value of ρ governs whether more aggregate output is produced when high skills are matched to high technologies and low skills to low technologies, or when workers are assigned to firms with technologies different from their skills. If $\rho = 1$, the production function exhibits no complementarities between the skill of the worker and the technology of the firm. In this case, the marginal productivity of a firm is independent on the skill of the worker with whom it is matched. Instead, if the production function exhibits supermodularity, $\rho < 1$ (submodularity, $\rho > 1$), the marginal productivity of a worker or firm rises more when the type of the partner is more similar (diverse) to their type.[4] Under the CES specification, the output of region r is

$$Y_{rt} = \sum_{f \in \mathcal{F}_r} Y_{frt} = A_r \cdot \sum_{f \in \mathcal{F}_{rt}} \left(\sum_{i \in \mathcal{E}_{frt}} \left(\lambda {x_i}^{\rho} + (1-\lambda){y_f}^{\rho}\right)^{\frac{1}{\rho}} \right). \tag{3}$$

Output per worker in region r is obtained by dividing the previous equation by number of employed workers in the region, E_{rt}.

This model allows for separating the contribution to differentials in productivity of the four considered determinants. Productivity crucially depends on the pool of skills and technologies available in that region (respectively the skills x_i of the workers in set $\mathcal{E}_{fr}$ and the technologies y_f of the firms in set $\mathcal{F}_{rt}$). Also, productivity increases when the allocation of skills to technologies exploits potential complementarities in the production function. With a supermodular (submodular) production function, positive (negative) assortative matching arises: assigning high skills to high (low) technologies and low skills to low (high) technologies yields to greater aggregate output than that produced under random allocation. Finally, the residual productivity disembodied from workers and firms is due to the location of the match in region r. Interpretation of this region-specific factor is broad and follows the literature on growth. The parameter A_r captures differences in laws, other social infrastructures in the fashion of Hall and Jones (1999), physical infrastructures and geographical characteristics. Also, the presence of networks of firms or districts is not explicitly modeled: the average effect of the agglomeration of more or larger firms in a region is absorbed

[4]The condition for supermodularity is preserved under the sum of multiple workers within a firm, see Eeckhout (2018).

in the common regional coefficient A_r, while the marginal effect on a single firm is included in the technology of that firm. Investigating the determinants of the region-specific factors is out of the scope of this work, which focuses on measuring the intrinsic contribution of workers and firms to productivity by considering their interactions but assuming that they cannot affect the aggregate factors due to their atomistic power.

Finally, I describe the wage determination rule. Firms choose their wage setting policy. This heterogeneity implies non-monotonicity of wages with respect to the skills of the workers at the aggregate level. However, within a firm, the wage of a worker is strictly increasing in his skill type. This is a reasonable assumption, in particular under the case of perfect information. Consider two workers in firm f in region r with skills $x_1 > x_2$: if the output of the match with worker 1 is greater than the output of the match with worker 2, then worker 1 earns a higher wage than worker 2, $w_1(x_1, y_f, A_r) > w_2(x_2, y_f, A_r)$. Hagedorn et al. (2017) show that this assumption is consistent with a search model where the value of a vacancy is increasing in the technology of a firm. On the contrary, this assumption excludes any role of the history of matches and wages of a worker in determining her future matches and wages, like in models with on-the-job search à la Postel-Vinay and Robin (2002) and Cahuc et al. (2006).

3 Identification Strategy

The model of regional output in equation (3) relates the labor productivity of a region to the regional distributions of skills and technologies, the allocation of workers to firms and a region-specific component. This work uses this framework to retrieve the contribution to productivity of each input from measures observable in the data. More specifically, I identify the skills and technologies in the economy and the parameters of the aggregate production function from matched employer-employee data with information on the value added of firms, on the working history and wages of every worker in each firm, provided that some of the firms have plants in multiple regions.

The identification strategy is organized into three steps. First, I recover a global ranking of the workers in the economy. I follow the methodology proposed by Hagedorn et al. (2017) to infer from wages and job transitions the ordering of the workers with respect to their skills. Then, I estimate the region-specific factors by considering firms with plants in more than one region. Since the technology of a firm is constant in time, variations of productivity over time at the firm level are only explained by the variation in the skills and in the location of the employees. By controlling for the change in the estimated rank of the employed workers, I estimate the contribution of locating the production in a region by observing how output varies with the reallocation of the workforce across regions. Finally, I use the model of firm output in equation (1) to estimate simultaneously a global ranking of firms with respect to technologies, and how skills and technologies jointly contribute to production. Given the production function parameters and controlling for the workers' ranks and allocation, the rank of a firm corresponds to the fixed component that explains the residual heterogeneity in observed value added per worker between firms. Therefore, any set of production function parameters (λ and ρ) provides an estimate of the global ordering of firms' technologies. I estimate the production function parameters by minimizing the distance between the value added of firms predicted by the model and the value added of firms observed in the data.

3.1 Ranking workers

I estimate the ranks of the workers by implementing the ranking procedure proposed by Hagedorn et al. (2017). The only variation to their estimation procedure is that I rank workers within firm-region sets rather than just within firms, since, in this model, the geographic location generates heterogeneity in the match output. Production is increasing in both skill and technology and, in the same firm and region, wage is increasing in the skill of the worker. Then, the order of co-workers according to their skills can be recovered directly from the information on the wages of the workers. In the model, the wage of a match is assumed constant over time and unrelated to the time-varying characteristics of the worker (like age and experience) that exist in the real data. Longitudinal data on the wage history allows for controlling for these observable characteristics and for isolating a permanent measure of wage based on workers' compensation over different job spells, namely the average residual wage.

Once workers are ranked within plants, since the skill of the worker is constant over time, the transition of a worker from one plant to another is informative on the ranking of his co-workers in the two plants. The transition of a worker allows for establishing that his co-workers ranked lower (higher) than him in the first plant are also ranked lower (higher) than all his co-workers ranked higher (lower) than him in the second plant. In other words, since the global ranking of workers satisfies the transitive property (such that if $x_1 > x_2$ and $x_2 > x_3$ then $x_1 > x_3$), then the transitions of workers across different firms and regions allow for comparing workers that were never employed by the same plant. With a sufficient number of transitions of workers across plants, the global ranking of workers is recovered in an accurate way.

However, in practice, the rank aggregation problem involves two major issues. First, inconsistencies between rankings are possible, e.g., $w(x_1, y_1, A_r) > w(x_2, y_1, A_r)$ and $w(x_1, y_2, A_r) < w(x_2, y_2, A_r)$. In this case, the ranking retrieved from plant 1 is $x_1 \succ x_2$, while the ranking obtained from plant 2 is $x_1 \prec x_2$. The Kemeny-Young approach (Kemeny, 1959 and Kemeny and Snell, 1963) defines optimal global ranking as the ranking that minimizes the number of inconsistencies among within-plant rankings. The Kendall-tau rank distance between two rankings π' and π'', $\mathcal{K}(\pi', \pi'')$, is computed as the sum of pairwise inconsistencies between the rankings π' and π''. Given P within-plant rankings $\pi_1, \ldots, \pi_F$, the objective global ranking is $\pi_0^* = \text{argmin}_{\pi_0} \frac{1}{P} \sum_{k=1}^{P} \mathcal{K}(\pi_k, \pi_0)$. Second, computing the pairwise relation between workers is a non-deterministic polynomial-time hard (NP-hard) problem: the exact algorithm requires considering all possible permutations of workers, which is computationally infeasible. In fact, an approximate algorithm is implemented. Given an initial global ranking, it computes for each worker i the wage difference with any co-worker j, corrected by potential measurement error in the wages and weighted for the number of spells of each worker in the considered firm. From these wage differences, it is possible to specify the probability that worker i is ranked higher than co-worker j. The algorithm considers the rank with the highest probability as the best ranking for worker i and updates the starting global ranking by giving to i the position of b in the ranking, leaving unchanged all other relative rankings. The optimal ranking is retrieved upon iteration.

The optimal ranking is then discretized in bins of similar skills both for tractability and because this substantially decreases the measurement error. Monte Carlo simulations performed by Hagedorn et al. (2017) with 50 bins show that the correlation between the true binning and the one estimated with the procedure is 0.995 or higher.

3.2 Estimating region-specific factors

The estimation of region-specific factors $\{A_r\}_{\forall r \in R}$ is possible thanks to the longitudinal dimension of the data with respect to firms. I consider the subsample of firms with plants in multiple regions, as this subsample is the only one that has the information required to identify differences in region-specific factors. The output per worker of a firm is typically observed for several periods of time.[5] Since firm technology is constant over time, changes in the output per worker of a firm reflect changes in the set of workers and in the allocation of workers across regions. By controlling for the deviation of the average rank of workers employed by the firm, the deviation from the average value added per worker is due to the changes in the fraction of workers in region r. Indeed, workers in different regions enjoy a different region-factor A_r. Therefore, regional factors are retrieved by estimating the model

$$\frac{Y_{ft}}{E_{ft}} = \overline{\left(\frac{Y_f}{E_f}\right)} \times \left(\sum_{r=1}^{R} A_r \frac{E_{frt}}{E_{ft}} + \zeta(s_{ft} - \bar{s}_f)\right), \tag{4}$$

where s_{ft} is the average rank of workers in a firm at t and the overline notation indicates the averages over time. The estimation of the regional factors exploits the within-firm variation in time and then is then viable without estimating firms' technologies.

3.3 Ranking firms and estimating production function parameters

Once the ranking of workers and the region-specific factors are estimated, the model of the firm output per worker allows for identifying the global ranking of firms and the parameters of the production function. The latter are common across firms. Subject to these parameters, the output of a firm is fully determined by its geographical location, by the skills of its employees and by

[5] Because information on output is reported in the balance sheet here each period coincides with a year.

its technology. If two firms in the same region have similar workers, the difference in the value added per worker across firms is only explained by the difference in their technologies. Given a specific set of production function parameters, and controlling for the observed variation of workers across plants (observed in the data) and for the estimated ranks of employees, the model allows for estimating a residual firm-specific component fixed over time. The order of the estimated firm-specific components reflects the ranks of the firms.[6] Similar to what is done for workers, firms are binned in groups of similar rank. Each different set of parameters yields a different ranking of firms.

The production function parameters and the ranking of the firms are estimated jointly by minimizing the distance

$$\min \quad \sum_{ft} \left(\frac{Y_{ft}}{E_{ft}} - \widehat{\left(\frac{Y_{ft}}{E_{ft}} \right)} \right)^2. \tag{5}$$

The parameter λ is pinned down by the sensitivity of the output to variations in skills relative to the variation in technologies, and the parameter ρ by comparing how much having workers of productivity more similar to their technology increases the output of firms with the same technology. Notice that the estimated rankings only specify the ordering of the workers and firms without providing information on the relative distance of the skill (technology) of two consecutive workers (firms). Given a rank χ_i for each worker and a rank γ_f for each firm, the skill of a worker i and the technology of a firm f are retrieved respectively as

$$x_i = \Phi^{-1}(\chi_i) \quad \text{and} \quad y_f = \Psi^{-1}(\gamma_f).$$

For the sake of simplicity, in the baseline estimation I assume that the skill (technology) types with consecutive rank are equally distant. In section 7, I estimate the production function and the ranking of firms under the alternative assumption of log-normal distribution of skills and technologies.

From equation (1), the value added per worker of a firm is specified as

$$\frac{Y_{ft}}{E_{ft}} = \eta \sum_{r=1}^{R} \left(\mathbb{1}_r \cdot A_r \cdot \frac{1}{E_{frt}} \cdot \sum_{i \in \mathcal{E}_{fr}} (\lambda \cdot \Phi^{-1}(\chi_i)^{\rho} + (1-\lambda) \cdot \Psi^{-1}(\gamma_f)^{\rho})^{\frac{1}{\rho}} \right), \tag{6}$$

where the number of workers in each plant E_{frt} is observed in the data and η is a positive scale factor converting the output in monetary terms. Since η is positive and common to all the firms, it does not require being estimated directly, but it is a common multiplier of the region-specific

[6] This procedure uses the information on value added of firms and follows closely the theoretical framework. Notice that in Hagedorn et al. (2017) firms are ranked by the order of the out-of-unemployment wage premium they pay to newly hired workers. Given the monotonicity of the production function, a firm with higher technology will produce higher output than a firm with lower technology with a worker of skill x. Then, the wage offered to that worker by the firm with higher technology will also be higher than the wage offered by the firm with lower technology. However, if the standard deviation of the wage premium across firms is small, like it happens to be in the data under analysis, this approach might lead to mismeasurement.

factors.

4 The Italian case

4.1 Explaining large differentials in labor productivity

The theoretical framework introduced above disciplines a decomposition analysis of productivity differentials across economies. However, an application to differentials in productivity across countries would raise empirical difficulties because the identification might be weak with few transitions of workers or relocations of firms across countries, and the linked employer-employee data sources are not yet structured to trace workers and firms across countries. In order to still get some insights on the potential determinants of labor productivity differentials between countries, I perform the decomposition exercise within a country with large differentials in output per worker across its regions. More specifically, this paper uses the presented framework to study differences in output per worker across Italian regions.

The Italian case is of particular interest since cross-regional differentials are especially high in comparison to other countries. Figure 1 displays the standard deviation of the output per worker across regions within countries after equalizing the mean value of the country to 100. Relative to other countries, Italy exhibits high dispersion in productivity across regions. This dispersion is not driven by specific outliers (like the Paris region in France, District of Columbia in the U.S. and the London region in the U.K.). As opposed to Germany, whose regions present high differentials but experienced a convergence process after the reunification of the country, differentials between Italian regions are very persistent over time. Moreover, since Italy is not a federal republic, the institutional framework is homogeneous across regions and the interpretation of the differentials in the region-specific factors can assume away a major role of differences in the legislative setting.

By examining the Italian case, it is possible to study a span in labor productivity equivalent to a third of the difference in productivity between the U.S. and Mexico. The range of GDP per worker of Italian regions covers an important portion of the difference in GDP per worker among OECD countries. Figure 2 compares output per worker across countries and across Italian regions, by combining data on real GDP measured at PPP from the Penn World Table 9.1 and data from the Italian National Institute for Statistics. The output per worker in the least productive region in Italy is about 30% lower than the output in the most productive region. In this work, I use matched employer-employee data representative of the manufacturing sector. GDP differentials among Italian regions in manufacturing and in the total economy are highly correlated (0.88).

The current analysis considers regions according to the NUTS 1 Eurostat classification. Indeed, identification is only guaranteed if, in each region, there are enough firms with plants in multiple regions. It would not be the case while considering the NUTS 2 level of regional classification. Notice that the chosen definition of region is not based on a spatial normalization (i.e. some regions are bigger than others), but reflects historical conventions that recognize a region as an area characterized by common institutions or culture, in line with the widespread approach in the growth literature.

4.2 Data

I use a matched employer-employee data set built using INVIND, an annual survey of manufacturing firms conducted by the Bank of Italy. This longitudinal sample includes around 1,200 firms per year and is representative of manufacturing firms with at least 50 employees. It contains detailed information on firms' characteristics, including sector of activity, year of creation, headquarters, number of employees and some financial accounts (e.g., revenues, costs, investments, debt).[7] The Italian Social Security Institute (INPS) extracted individual employment records for all the workers ever employed in an INVIND firm. The information consists of the complete working history of these workers for 18 years (1980-1997), and is comprehensive of spells in private firms not included in the INVIND sample. Each observation reports the social security information (O1/M mandatory contribution form) that the employer is required to transmit for each employee, with details on age, gender, birthplace, location of the job, annual gross earnings, number of weeks worked and occupational status (blue collar, white collar, manager).

The matched employer-employee data INPS-INVIND covers the years from 1991 to 1997, since information on the costs used to compute value added is not reported in waves before 1991 and I cannot access workers' data for the years after 1997. Additional information on the structure of the data is available in Iranzo et al. (2008), which I also followed for the data cleaning. This data offers the big advantage of observing the entire workforce of each firm in the INVIND sample, which is crucial in order to identify the technology of the firm from the skills of the employed workers. Importantly, in contrast to establishment-level administrative datasets, it provides information on the location of all the workers employed by the firm, allowing to exploit the existence of firms with plants in different regions and identify the region-specific factor. Also, earnings are not top-coded.

Each INVIND wave provides weights, ω_{jt}, allowing to reproduce the Italian manufacturing industry. This analysis also requires assigning a weight to each plant and to each individual spell observed. The weight of a plant in a given region is computed by multiplying the survey weight by the share of total weeks worked by employees in that region. Similarly, the weight for a worker in a plant is the number of weeks she works over the total number of weeks worked by all employees of the plant.[8]

[7] As each firm is identified by a unique fiscal code, I complement the firm information with the administrative balance sheet information from the Company Accounts Data Service (CADS). Double checks between the INVIND and CADS data prove the high accuracy of the information provided in the survey.

[8] The weight of the plant of firm j in region r in year t is

$$\omega_{jrt} = \frac{\text{total weeks worked in plant } jr \text{ in year } t}{\text{total weeks worked in all plants of firm } j \text{ in year } t} \cdot \omega_{jt}.$$

4.3 Descriptive statistics

Table 1 reports summary statistics. I use the full sample of employment records to estimate the skills of the workers as described in Section 3.1. I observe around 19 million employment spells of 1.3 million workers. The majority of them (49%) works in the Northwest, around 28% in the Northeast and the remaining share is distributed in Central Italy and the South, with only 4% in the Islands. Around 23% of workers are women and this share decreases by around 10 percentage points when considering the South and the Islands. It is important to notice that almost all workers have multiple spells: 77% of them are employed in more than one firm over their working history (on average by three different employers) and 13% of them works in more than one region. All these features ensure that the data offers sufficient variability for a reliable identification of the ranking of workers' skills.

The decomposition analysis uses the INPS-INVIND sample. The latter includes around 3.2 million employment spells of almost 0.9 million workers, whose distribution across regions is similar to that of the full sample. They are employed by 1,464 firms, observed over seven years. One third of the firms have plants in more than one region and they account for 65% of the workforce. More than a half of the firms have a plant in the Northwest, almost 40% in the Northeast, 33% in Central Italy, and a lower share of firms have plants in the South and the Islands (21% and 12%, respectively). On average, a firm has 200 employees, while the median and the 95th percentile are 93 and 503 employees, respectively. Figure 3 displays the distribution of firm productivity across regions, the average value added per employee is 45,578 euro at a 1995 fixed price. Value added per worker at the national level in Italy estimated by using the sample corresponds to the value added per worker reported in the official estimates of the Italian National Institute of Statistics (ISTAT, 2005).[9]

To obtain a reliable global ranking of workers, it is necessary to have enough transitions of workers among firms so that the firms are connected between each others. In the data, the largest connected set includes 88% of the firms and 98% of the workers in the sample. Identification of workers' rank also requires within-plant variation of wages. In addition to the contracted salary, workers receive irregular payments such as overtime, shift work and bonuses, which increase the wage variability within the plant. Notwithstanding the presence of the collective bargaining agreements that define a minimum pay for the workers, the combination of minimum levels specific by sector, occupation and location, and additional payments at the individual level, produces suffi-

. The weight for the worker i in plant jr in year t is

$$\omega_{jrti} = \frac{\text{total weeks worked in plant } jr \text{ in year } t \text{ by worker } i}{\text{total weeks worked in plant } jr \text{ by all workers in year } t} \cdot \omega_{jrt}.$$

[9]Unfortunately, the official statistics of the aggregate value added per worker at the regional level are not fully comparable since they are based on a sample of firms with more than 20 employees and they assign the value added of the firm to the region of the headquarters without considering the location of the employee.

cient heterogeneity in wages. In particular, the percentage premium over the base salary is 6.6% at the 10th percentile and 13% at the median. I consider weekly wages, deflated for yearly national inflation. I exclude spells with a monthly wage less than 100 euro.

Table 2 summarizes the residual wage by region and job category after a regression with sector and gender fixed effects. Median and dispersion in wages across regions indicate how large the gain of moving into another region is: the median worker moving to the Islands to the Northwest gains a 7% premium in wage. The difference in the median wage is similar for blue-collar (7.1%) and white-collar workers (4.4%), while median wages of managers are higher in Central Italy and lower in the northern regions (albeit this should be discounted for the lower number of managers in the South). In addition, wage dispersion is similar across regions: the between-region component explains only 1.2% of the total wage variability and this pattern is persistent in time. The heterogeneity in productivity across regions is only mildly reflected in wage dispersion across regions. This feature is explored in Boeri et al. (2019) with more recent data. They document a mean wage difference between North and South of 4.2% and negative correlation between real wages and productivity. By comparing wage differentials in Italy and Germany, they highlight the importance of centralized wage bargaining as a potential driver of this asymmetry between productivity and wage differentials.

Even if the wage is an imperfect indicator of the skill, the dispersion of wages across firms is indicative of the direction of sorting. In the period 1980-1997, within-firm variability accounts for 55.6% of total wage variability, in line with what was reported by Lazear and Shaw (2009).[10] This suggests that, on average, firms have a broad skill mix and it excludes strong sorting of workers across firms: workers increase their salary mostly by achieving promotions within the firm.[11] In Appendix A, I present a decomposition of the variance of the wages with respect to the fixed components of wage due to workers, firms and regions.

5 Estimates

5.1 Skills

Workers are ordered within each plant by a permanent residual wage. Wages might partially reflect individual time-varying characteristics such as age, experience and tenure, or the aggregate performance of the economy at the time of the spell. To compare workers employed at different dates, I compute residual wages from a regression of deflated weekly log-wage on a quadratic specification for experience and tenure, and fixed effects for the year and the number of weeks of the contract. Table 3 shows that the wage has a concave profile in experience, with similar results

[10]They report a ratio of the average within-firm standard deviation of wages to the standard deviation of wages close to 0.6 for Italy, the US, the Netherlands and France. The ratio is close to 0.8 for Denmark, Finland, Norway and Sweden. They also find high variance in wage growth rates across individuals within the firm.

[11]The between-firm variability of wages is pretty stable during the period, seeming not to be influenced by eventual changes that occurred in technology during these years.

when considering age instead of experience. Once controlling for experience, seniority in a firm has a small negative effect, as detected also by a similar wage regression in Iranzo et al. (2008). In column 2, I include a binary variable for male workers to take into consideration that there can be systematic gender differences in wages, which is confirmed by the high positive coefficient estimated. It is especially important to control for that to avoid potential biases due to the heterogeneous share of female employees across regions (lower in the South and Islands) and across firms (lower in firms with higher value added per worker).[12] For each worker-plant pair, the permanent residual wage is computed as the average of the residual from this regression.

The ranking of workers is obtained by aggregating within-plant rankings, as described in Section 3.1, while adjusting for the number of spells spent in each firm and for measurement error in the data. Since these are administrative data, the accuracy of the reported wages is high. Measurement error estimated by assuming error independence across years and workers accounts for around only 1.5% of the variance of the wages. As an initial ranking for the algorithm, I use the ordering based on the workers' fixed effects from the model in Abowd et al. (1999), henceforth AKM.[13] Once the ranking is estimated, I assign the workers to 100 bins of equal size increasing in rank. To cross-validate the estimated ranking, in Figure 4, I plot the distribution of workers' ranks by occupational status. As expected, white-collar workers are, on average, more skilled than blue-collar workers and managers are concentrated in the highest bins.

Table 4 reports summary statistics of the ranks of workers across regions. Each of the first four moments are homogeneous across regions; the mean rank is higher by around 3 percentiles in the northern regions. In order to assess how different the distributions of skills are beyond these point statistics, I compute two metrics that summarize the distance between the distribution of skills in each region and the distribution of skills in the Northeast: the Jensen-Shannon divergence and the Bhattacharyya coefficient.[14] They both confirm that the distribution of skills are nearly equal across regions. Differences remain small when considering the distribution of workers by their birthplace

[12] Table 1 reports the share of women in each region. Instead, if firms are grouped into quartiles with respect to the value added, women are around 40% of the employees of firms in the first quartile, around 30% of firms in the second quartile, and around 20% of firms in the third and fourth quartiles. The correlation between the global rankings of workers if the residual wage is estimated controlling and not controlling for gender is 0.79.

[13] Regression estimates are reported in the third column of Table 3. The correlation between the initial and the optimized ranking is 0.62.

[14] The Jensen-Shannon divergence, introduced by Lin (1991), is a symmetrized version of the Kullback-Leibler divergence. Given two discrete probability distributions $P(x)$ and $Q(x)$,

$$\begin{aligned} JSD(P,Q) &\equiv S\left(P, \frac{P+Q}{2}\right) + S\left(Q, \frac{P+Q}{2}\right) \\ &= H_S\left(\frac{P+Q}{2}\right) - \frac{1}{2}H_S(P) - \frac{1}{2}H_S(Q), \end{aligned}$$

where $S(P,Q)$ is the Kullback-Leibler divergence, $\sum_x P(x)\log\frac{P(x)}{Q(x)}$, and $H_S(P) = -\sum_x P(x)\log P(x)$ is the Shannon entropy. The Bhattacharyya coefficient (Bhattacharyya, 1943) is an approximate measurement of the amount of overlap between two statistical samples. Given n partitions of the samples P and Q, it is defined as $BC(P,Q) = \sum_{i=1}^{n}\sqrt{P_i Q_i}$, $0 < BC(P,Q)$.

rather than by their job-place. While this evidence is surprising in light of the studies using proxies for the skills based on education, it should be remarked that this data excludes unemployed and self-employed workers, thus the skills of the considered workers are strictly tied to the skill needed by the specific production process. Indeed, I use the information on education in the Italian Labor Force Survey data in 1995 to plot the average regional deviations from the average schooling of workers employed in the manufacturing sector. Figure 5 shows that the difference is extremely low.

5.2 Technologies

I classify firms into 100 bins of equal size and increasing rank. Notice that the joint estimate of the firm rank and of regional factors only considers the permanent component of the skills of the workforce, since potential differences across firms depending on the workforce composition by experience and tenure are cleaned out in the residual wage estimation. Figure 6 shows that the distributions of the experience and the tenure of workers are similar both across regions (panel A) and across firms in different quartiles of value added per worker (panel B). This suggests that these differences do not play an important role in explaining differentials across regions and firms. Instead, when these differences are large, the ranks of workers should be explicitly augmented by parametrically modeling experience and tenure.

To understand how the estimated rank of firms is related to observable characteristics, I run a linear regression of the rank of the firm on the latter. Table 5 shows that, on average, firms with plants in more than one region have higher ranks, while firms with plants in the South and the Islands have lower ranks. Also, rank and size of the firm are positively correlated. Even if the sample mostly includes firms that have been active for many years (the median of years since foundation is 29), I observe that firms founded in the last 5 years usually have a lower rank. Interestingly, the rank of the firm is positively correlated with the share of employees in the first quartile of the rank distribution, but negatively with the share of employees in the highest quartile. This might suggest that firms of the highest technology do not find it optimal to hire only the best workers. It is plausible that the rank of a firm is highly related to the sector of activity. Figure 7 plots the distribution of ranks by sector. Firms operating in the chemical, mineral and food industries tend to have a higher rank.

Table 6 reports summary statistics of the distribution of technologies across regions, highlighting relevant heterogeneity. The average firm rank in the South and in the Islands is respectively 5% and 10% lower than the average firm rank in the Northeast. Also, the higher skewness of the distribution of technologies in the South and in the Islands reveals that, in these regions, the mass of the distribution of firms is concentrated in low ranks rather than high ranks, which is in contrast to what happens in the northern regions and in Central Italy. This difference in the tails of the rank distribution only arises for technologies, whereas the empirical density of the skills is more uniform within and across regions. Also, the statistics of distance of the distribution of firms' ranks of each region from the one in the Northeast are much higher than for the workers' ranks. While the maximum value of the Jensen Shannon divergence is 0.20 for the ranking of workers, it rises to 0.65 for the ranking of firms. Symmetrically, the minimum Bhattacharyya coefficient of overlap with the distribution of ranks in the Northeast is 0.95 for workers, but 0.87 for firms. These differences between the ranks of firms across regions are even larger if we consider the ranks of firms according to the region where their headquarters is located (panel B of Table 6).

5.3 Allocation

Looking at the distributions of the ranks of workers and firms offers an incomplete perspective of their impact on aggregate productivity. Indeed, in the presence of complementarities in production between skills and technologies, a different assignment of workers to firms changes the aggregate value added produced. As a first step to understand which worker is assigned to which firm, I compute the variance of workers' ranks between firms. This is based on the measure proposed by Kremer and Maskin (1996) for wages and equals the ratio of the between-firms variance of ranks and the sum of between-firms variance and within-firms variance of ranks.[15] As reported in Table 7, the between-firms component accounts for 13% of the variance of ranks, indicating low segregation of workers by skill at the firm level. A potential explanation is that firms might prefer a diverse skill mix rather than hiring workers of similar skill. This value is close to the one documented by Iranzo et al. (2008), which uses the AKM worker fixed effect as a measure of skill. They interpret the predominance of the within-firms skill dispersion as a result of vertically integrated organizational structures. Interestingly, the between-firm dispersion of skills is slightly higher in less productive regions.[16] Lopes de Melo (2018) proposes the correlation between the fixed effects of workers and their co-workers as an alternative measure of the segregation of skills across firms. I estimate an average co-worker correlation of 0.36, which is also higher in the least productive regions. This is close to the values he reports for Germany, Denmark and Brazil. Iranzo et al. (2008) find a value of 0.17, which is close to what I obtain when ranking the workers without controlling for gender in the residual wage regression.

The literature on assortative matching uses the correlation between the rankings of workers and firms as the standard indicator of sorting. A positive value indicates that high-skill workers are usually assigned to high-technology firms and low-skill workers to low-technology firms, while a negative value suggests a higher probability of finding a worker with low skill in a firm with high technology. The sorting is stronger the closer the rank correlation is to its bounds $[-1, 1]$, while a rank correlation close to zero suggests random assignment of workers to firms.

In the Italian manufacturing sector with at least 50 employees, rank correlation is 0.04, pointing

[15]In Appendix A, I show that the between-firms variance of wages yields a lower value, suggesting that part of the variance in wages is due to heterogeneity in wage-setting policies across firms.

[16]With respect to the variation across regions, the between-region variance of skills explains less than 1% of the total dispersion, consistent with the almost null between-region share of wages in Appendix A.

to high substitutability of workers' skills, in contrast to higher values reported by the literature for other countries.[17] This might be due to the specific nature of the sample, i.e. large manufacturing firms. Considering that these firms employ mostly blue- and white- collar workers, and that managers are just around 2% of the workers, a high-technology firm might have little benefit in employing many workers of high skill. This is in line with the hierarchical structure of the firm, as described in Garicano and Rossi-Hansberg (2006), where the presence of knowledge hierarchies makes optimal for the firm to employ highly skilled workers with high decisional power beside less skilled workers, since what matters for productivity is the complementarity between the technology of the firm and the skills of the workers with high decisional power. High substitutability between skills is also supported by high interchangeability across occupational statuses: around 2/3 of white-collar workers were also employed as blue-collar workers during their working history.[18] Notice that low rank correlation can also be induced by frictions in the wage bargaining: Card et al. (2013) document a rank correlation between 0.03 and 0.10 for Germany before 1996, while it rises up to 0.25 in more recent years simultaneously with institutional changes in wage structure. Rank correlation is close to zero in all regions and around 0.08 in the Northwest and in the Islands.

To have a graphical representation of the sorting patterns in the economy, I reclassify workers into 6 classes of increasing skill and firms into 10 classes of increasing technology, and then, for each technology class, I plot the fraction of workers in each skill class.[19] The left panel of Figure 8 is a graphical representation of the rank correlation: the composition of the worker classes is very similar in different firm classes. The right panel of Figure 8 plots the mean log weekly wage by firm class. As expected, workers of similar skill employed by firms in a higher class earn a higher wage. Worker heterogeneity in each class seems relevant. The gain of working in a high class does not seem to depend on the worker class, suggesting again that there is not a high segregation of workers across firms. Finally, this figure also shows that a large share of wage heterogeneity is within firms in the same class, in line with the low contribution of firm fixed effects to the variance of wages documented in Appendix A.[20]

[17] My estimate is similar to the value reported by Iranzo et al. (2008) with the same data, while Bartolucci et al. (2018) find a value of 0.52 by using data of all sectors in Veneto, a county in the Northeast. Hagedorn et al. (2017) estimate a correlation of 0.75 for Germany, Bagger and Lentz (2019) find 0.37 for Denmark, Bonhomme et al. (2019) document a range between 0.30 and 0.50 for Sweden and Lopes de Melo (2018) a value of 0.52 for Brazil. However, it is not clear to which extent these differences are driven by true differences in the similarity of types rather than by differences in the estimation approach.

[18] See the companion table in Figure 4.

[19] This replicates Figure 2 in Bonhomme et al. (2019). By using Swedish data, they find a similar behavior of the mean log-earnings, but a stronger sorting pattern.

[20] Notice also that, within a worker class the mean wage is not constrained to be monotonic in the firm type, which is instead assumed by construction in the AKM fixed effects model.

5.4 Production function

Table 8 reports the estimates of the production function parameters. The region-specific factors normalized with respect to the Northeast are estimated by equation 4 on the sample of firms with plants in more than one region. Then, these are used with the estimated ranks of workers to jointly estimate the remaining production function parameters and the ranks of the firms. The complementarity parameter ρ is negative, indicating the presence of positive complementarities between the skill of the worker and the technology of the firm. In a frictionless assignment, like in Becker (1973), skill and technology are perfectly substitutable when ρ assumes a value of 1. Shimer and Smith (2000) prove that, in a model of search and matching with frictions, the condition is stricter, $\rho < 0$. The distance of the parameter ρ from the threshold value is informative about the strength of complementarities between worker and firm: the elasticity of substitution between skill and technology is 0.46. By estimating a CES production function in an empirical equilibrium model with sorting, Bagger and Lentz (2019) find a lower elasticity of 0.12 between worker and firm in Denmark, while Lise et al. (2016) report an elasticity of 0.53 for college graduates and perfect substitutability for workers with lower schooling in the United States. The parameter λ is informative about the relative share of value added per worker explained by the skills of the employed workers in a firm and is close to the canonical value of labor share retrieved by the productivity literature. Finally, thanks to additivity of A_r in the production function, the estimates of the regional specific factors capture the differentials in value added per worker directly imputable to residual factors disembodied from workers' and firms' characteristics and allocations. Since the value added observed in the data is expressed in monetary terms, estimates of the regional factors can only be interpreted after discounting for a common scale factor (η). I report their relative value with respect to the region-specific factor of the Northeast. These coefficients should be read in comparison with the productivity of regions summarized in Table 1.

6 Counterfactual exercises

6.1 Importance of different channels

Once the parameters of the production function are estimated, I use the model to establish how much the difference in each of the four considered channels contributes to explaining the gap between regions. In particular, I focus on the difference between the output per worker of a region and the most productive region, which is the Northeast. To this aim, I estimate the output per worker of a region after equalizing one channel at a time to the corresponding value for the Northeast, and I evaluate how much the counterfactual output reduces the aggregate gap between the regions. Since the region-specific components enter the production function as Hicksian multipliers, it is sufficient to compute the ratio between the region-specific factor of the regions in order to assess how value added per worker would change whether all regions would have the same region-specific factor as the Northeast. Instead, simulating changes in the distribution of skills and technologies requires mimicking the shape of the distribution in the most productive region. In order to do that, I approximate the within region distributions of the ranks of workers and firms by using a Beta probability density function. This specification is convenient because it is both flexible and parsimonious, since it is defined by only two shape parameters. The estimated parameters are reported in Table 9 and describe the differences in the skills and technologies across regions. The counterfactual output per worker if a region had the skill distribution of the Northeast is computed by assigning to the workers in that region a skill level such that the within region distribution reproduces the distribution of skills in the target region, while keeping unchanged the within-region distribution of the firms' technologies, the region-specific factor and the observed allocation of workers across plants. I implement an analogous procedure to evaluate the changes due to assigning the distribution of plants' technologies of the Northeast. Finally, to assess the changes due to the allocation of workers, I compare the aggregate output per worker in the region with the aggregate output per worker in the Northeast when jointly imputing the distributions of skills and technologies and the regional factor of the region of interest. The percentage variations in labor productivity are reported in panel A of Table 10. Despite the heterogeneity in the effects across regions, differentials in labor productivity seem to be crucially driven by the technologies of the firms available in each region. Indeed, assigning to plants in the other regions the distribution of technologies of plants in the Northeast generates increases in labor productivity of around 3% in Central Italy, 15% in the South and 23% in the Islands. The increase in the Northwest is very low, as the distribution of skills in Northeast is only mildly dominating and the mean firm rank is even slightly higher in the Northwest. Region-specific factors are relevant too in order to explain productivity differentials with the Northeast. In particular, the lowest the productivity gap between regions, the higher is the relative importance of region-specific factors.

To quantify the relative importance of the change in each channel, I compute how much it accounts for the total gap in aggregate productivity recovered by the model. The ratio between the two is reported in panel B of Table 10. As the Northwest has a very similar aggregate productivity and structure of the economy to the Northeast, the regional factor almost completely explains the differentials. This is not the case for the other regions. On average, the skills of the workers and the allocation of workers to firms play a limited role, they account for 3% to 8% of the total gap between Central Italy and the South and the Northeast, and they do not drive the gap with the Islands.[21] The main channel explaining total differentials in labor productivity is the difference in technologies of the firms, accounting for around 65% of the gap between the most and the least productive regions, and even more in the case of the Islands. An important role is also played by

[21]The finding of the limited role of skills dispersion is not new to the empirical literature. By using Danish employer-employee data, Fox and Smeets (2011) document that the role of labor input quality in productivity dispersion across firms is low and is twice as important in services than in manufacturing.

the region-specific factors, which explains around 27% of the total differentials. Notice that this exercise considers only direct effects of a single channel on labor productivity, abstracting from plausible effects of the change in one channel on other channels. In Lacava (2019), I investigate the general equilibrium effects of the change in the distribution of workers due to migration across regions by using a search and matching model with heterogeneous agents and firms.

6.2 Cost of suboptimal allocation

Since the production function is estimated to be supermodular, according to the result in Becker (1973), the optimal assignment for a given population of jobs is the one with the highest skilled worker matched with the highest technology job, the second highest skilled worker with the second highest technology job, and so forth. In this spirit, I reassign workers to firms in the order of their ranking within each region, keeping the firm size unchanged, and I measure how much the output per worker produced under perfect Beckerian allocation increases with respect to that produced under the real allocation observed in the data. This exercise provides a measure of the cost of mismatch of workers to firms for aggregate productivity. Gains are relevant across all regions: the relocation of workers would generate percentage increases in labor productivity of between 17% and 27%. At the national level, a within-region reallocation of workers would increase productivity by 19%. The high value of this gain reflects that the complementarity estimated in the production function is not exploited, as documented by the low rank correlation statistics. Mismatch seems to be similar across regions, in line with both the homogeneous rank correlations and with the small role played by the allocation channel in explaining differentials across regions. Remaining differences across regions might be probably driven by the sectoral composition of the regional economy. Indeed, mismatch varies significantly across sectors, consistently with the idea that skill-technology complementarity is more important in some sectors than in others.

Since, in the model, workers are characterized by a unidimensional skill, this reassignment ignores the possibility that the worker ability might be specific to a certain task. Unfortunately, the data do not provide any information on the task within the firm and I cannot control empirically for that dimension. To consider at least the presence of sector-specific abilities, I repeated the exercise by reassigning workers only within sector. The aggregate gain decreases only by around one sixth, indicating that most of the mismatch is within sector, as it is also suggested by the high variance of skills within firms documented in Section 5.4.

In an analogous exercise with Danish data, Bagger and Lentz (2019) find an output gain of 7.8%, while Lise et al. (2016) estimate an output gain of around 3.8%. In contrast, Hagedorn et al. (2017) assess that imposing the matching to be on the main diagonal implies a 0.23% decline in output relative to the observed allocation.[22] The magnitude of the estimated effect can be

[22]When analyzed in the context of a general job search model, Gautier and Teulings (2015) estimate that, without search frictions, output would increase by 9% to 16%; Hagedorn et al. (2017) find that the output of the counterfactual

interpreted also in the light of the counterfactual output reported by Hsieh and Klenow (2009) when eliminating misallocation between capital and labor in a model where firm-specific wedges generate heterogeneity in capital-labor ratios across firms. They find that assigning the U.S. dispersion of marginal products increases Total Factor Productivity by 30%-50% in China and by 40%-60% in India.

6.3 Impact of mobility

The distribution of skills and technologies across regions can be an equilibrium outcome of the labor market. To assess how the relocation of workers and firms across regions impacts labor productivity, I use a Beta probability density function to estimate the distribution of skills of workers born in a given region, independently of the location of the job. Then, I impose the estimated distribution of skills to the workers in the region and I measure the change in labor productivity. As expected, since the distribution of skills are very similar across regions, the relocation of workers through migration does not have a big effect on productivity differentials. To dig deeper into this result, in Figure 9, I plot the distribution of the within-region ranking of workers moving in and out of the region. In each region, the shape of emigrants and immigrants is similar, meaning that the migration of workers from the region of birth does not reshape in a significant way the distribution of skills. Workers with higher skill are also the ones that experiment higher mobility. Notice also that there are migrants at all levels of the rank distribution. This helps to build confidence in the ranking procedure, since the transitions across regions are not concentrated in a single range of the within-region rankings.

Then, I repeat the exercise by forcing the distribution of technologies to be that of the firms with headquarters in the considered region. The effect is negligible also in the case of firm relocation, explained by the large share of single plant firms, with the exception of the Islands. Table 13 reports the technology level of plants that relocated their production in other regions. The share of multi-plant firms in the South and in the Islands is higher than in the most productive region and around 10% of the firms are relocated from the Northwest. The difference in the gain by relocation between the South and the Islands is explained by the average rank of these firms: lower in the South (56) and much higher in the Islands (72).

7 Log-normal distribution of skills and technologies

The non-parametric ranking algorithm used to estimate the global ranking of workers does not provide any cardinal information on the difference in skill between two consecutive workers. In the baseline estimate, I bin the workers in 100 groups of equal size and increasing rank and I assume a fixed difference in skill between each group–equal to 0.01 since $x \in (0, 1)$. However, the

allocation in the absence of frictions and frictional unemployment would be 8.47% higher.

shape of the distribution of skills is potentially relevant for estimating aggregate productivity and firm technologies. Indeed, suppose that the majority of the workers has a low and similar level of skill, while a small group of workers has high skill: substitutability might be very high in the first large group even between workers with a very distant rank because their real level of skill is close. On the contrary, workers with close rank but on the right tail of the distribution can have a very different skill and be hardly substitutable. Similar reasoning holds true for the technologies of firms of consecutive rank, that so far I assumed being equally distanced, but whose shape might be potentially different. To verify the sensitivity of the baseline results to the assumption of equally distanced types of consecutive rank, I repeat the estimation by assuming an alternative rank-to-type mapping. In particular, I assume that skills follow a log-normal distribution replicating the shape of the distribution of wages, and technologies follow a log-normal distribution replicating the shape of the distribution of firms' value added per worker.

Table 14 reports three variations to the baseline assumption on the rank-to-type mapping. In column 1, I assign a lognormal shape to the distribution of skills only, in column 2 a lognormal shape of the distribution of technologies only, and in column 3 I assume that both distributions of skills and technologies are lognormal. The estimates of the production function parameters do not depart dramatically from the baseline. The elasticity of substitution seems to be lower under a lognormal shape of technologies. Results are similar when the lognormal distribution of skills replicates the shape of the residual wages or the residual wages controlling also for firm fixed effects.

Table 15 shows the variation in rank correlation and in the results of the counterfactual exercises in Section 6 under the three alternative sets of assumptions. In all cases, the Spearman rank correlation between workers and firms is slightly higher than in the baseline estimate, but still the allocation is not characterized by a strong sorting pattern. Productivity differentials are still prominently driven by differences in firms' technologies and in regional factors. In particular, in the South the differences in technologies play a bigger role (around 75%) than in the baseline case (65%). On the contrary, in the Northwest differentials are entirely driven by a region-specific multiplier which is lower than in the Northeast, while assigning the skills, technologies and allocation of the Northeast to the Northwest would always yield to lower productivity. Reallocating the best workers to the best firms produces a gain of around 10% of the productivity at the national level. The effects are lower than in the baseline estimation, especially fo the South (12% instead of 27%), but still considerable. Mismatch seems to be similar across regions, and mainly within-sector. Finally, Table 15 confirms that migration of workers and relocation of firms have a minor impact on productivity, as highlighted by the baseline results. All in all, alternative assumptions about the distributions of types do not seem to have a major impact on the findings.

8 Conclusion

I introduced a strategy to measure how much skills of the workers, technologies of the firms, efficiency in the matching of workers and firms, and region-specific factors contribute to differentials in labor productivity across regions by using information on job transitions, wages and value added contained in increasingly available matched employer-employee data. Skills are identified by comparing the wages of co-workers in the same firm and same region; technologies by comparing output per worker of firms with the same workforce's skills and in the same region; complementarities between worker and firm by estimating the joint contribution of skill and technology to productivity in the same region; and region-specific factors as a residual by observing firms with plants in more than one region.

I present an application to Italy, a country with large differentials of productivity between regions. Differentials in productivity are explained for around 65% by differences in technologies of the firms and for 27% by differences in region-specific factors. On the contrary, the skills and allocative efficiency of workers are similar across regions, and the gains from reallocation of workers between regions are small. This has a strong policy implication, suggesting that, in Italy, the historical stagnation of the South and the Islands is crucially linked to the difficulty to gather firms of high technology in those regions. I estimate positive complementarity between skill and technology in the production function, implying that it is optimal to assign most skilled workers to firms with the highest technology. Optimally reallocating workers would increase productivity by around 16%, but since the mismatch is rather homogeneous across regions, it would not significantly affect productivity differentials. Moreover, the migration of workers does not impact labor productivity.

This analysis can be extended in several ways by relaxing some assumptions of the model. First, by considering the output of the firm as the sum of the output of each job in the firm, I assume the absence of any interaction between jobs in the same firm. On the contrary, empirical evidence implies the presence of complementarities between skills in the same firm and the possibility of increasing returns to scale. In this direction, Eeckhout and Kircher (2018) analyze sorting in relation to firms with multiple workers: they introduce a theory of production with the firm simultaneously optimizing its size and the skill composition of its workforce. Second, I impose constant returns to scale in the aggregation of the output of firms at the regional level. This is at odds with the regional and networks literature, which find that the agglomeration effect might play an important role in determining productivity. In my framework, both the eventual complementarity between workers and the existence of an agglomeration effect will be captured in the estimates of the firms' technologies and the average agglomeration effect will also contribute to the estimate of the regional factor. Third, I assume that skills and technologies are permanent in time, while it is well established in the literature that human capital accumulation and on-the-job training play a strong role in reshaping the skills of the workers and, over longer horizons, firms might invest large resources to the reorganization of production and other policies affecting their technology. Finally, in counterfactual exercises, I analyzed variations of productivity by changing one channel at a time in a static setting. As changes in one channel could plausibly generate readjustments in the others, a model of general equilibrium can shed more light on the joint changes.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*, 251–333.

Acemoglu, D., S. Johnson, and J. Robinson (2005). Institutions as a Fundamental Cause of Long-Run Growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*. Amsterdam: Elsevier.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? *Journal of the Royal Statistical Society. Series A (Statistics in Society) 171*(3), 673–697.

Bagger, J. and R. Lentz (2019). An Empirical Model of Wage Dispersion with Sorting. *The Review of Economic Studies 86*(1), 153–190.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics 106*(2), 407–443.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013). Cross-Country Differences in Productivity: The Role of Allocation and Selection. *American Economic Review 103*(1), 305–34.

Bartolucci, C., F. Devicienti, and I. Monzón (2018). Identifying Sorting in Practice. *American Economic Journal: Applied Economics 10*(4), 408–38.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy 81*(4), 813–846.

Bhattacharyya, A. (1943). On a Measure of Divergence between Two Statistical Populations defined by their Probability Distributions. *Bull. Calcutta Math. Soc. 35*, 99–109.

Boeri, T., A. Ichino, E. Moretti, and J. Posch (2019). Wage Equalization and Regional Misallocation: Evidence from Italian and German Provinces. *NBER Working Papers 25612*.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A Distributional Framework for Matched Employer Employee Data. *Econometrica 87*(3), 699–783.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review 101*(5), 1964–2002.

Cahuc, P., F. Postel-Vinay, and J. Robin (2006). Bargaining with On-the-Job Search: Theory and Evidence. *Econometrica 74 (2)*, 323–364.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics 128 (3)*, 976–1015.

Carneiro, A., P. Guimarães, and P. Portugal (2012). Real Wages and the Business Cycle: Accounting for Worker, Firm, and Job Title Heterogeneity. *American Economic Journal: Macroeconomics 4*, 133–152.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. Volume 1 of *Handbook of Economic Growth*, pp. 679 – 741. Elsevier.

Correia, S. (2016). A Feasible Estimator for Linear Models with Multi-Way Fixed Effects. *mimeo*.

Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005). Growth Econometrics. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 8, pp. 555–677. Elsevier.

Eeckhout, J. (2018). Sorting in the Labor Market. *Annual Review of Economics 10*(1), 1–29.

Eeckhout, J. and P. Kircher (2011). Identifying Sorting - In Theory. *The Review of Economic Studies 78*(3), 872–906.

Eeckhout, J. and P. Kircher (2018). Assortative Matching with Large Firms. *Econometrica 86*, 85–132.

Fox, J. T. and V. Smeets (2011). Does Input Quality drive Measured Differences in Firm Productivity? *International Economic Review 52*(4), 961–989.

Garicano, L. and E. Rossi-Hansberg (2006). Organization and Inequality in a Knowledge Economy. *Quarterly Journal of Economics 121*(4), 1383–1435.

Gautier, P. A. and C. N. Teulings (2006). How Large Are Search Frictions? *Journal of the European Economic Association 4*(6), 1193–1225.

Gautier, P. A. and C. N. Teulings (2015). Sorting and the Output Loss due to Search Frictions. *Journal of the European Economic Association 13*(6), 1136–1166.

Hagedorn, M., T. H. Law, and I. Manovskii (2017). Identifying Equilibrium Models of Labor Market Sorting. *Econometrica 85(1)*, 29–65.

Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*, 83–116.

Hellerstein, J. K. and D. Neumark (2007). Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-Employee Data Set. In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, pp. 31–71. University of Chicago Press.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403.

Iranzo, S., F. Schivardi, and E. Tosetti (2008). Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data. *Journal of Labor Economics 26*(2), 247–285.

ISTAT (2005). Conti economici delle imprese (1980-2004). *Italian National Institute of Statistics*.

Kemeny, J. (1959). Mathematics without Numbers. *Daedalus 88*(4), 577–591.

Kemeny, J. and J. Snell (1963). Mathematical Models in the Social Sciences. *Blaisdell Publ 2*.

Kline, P., R. Saggio, and M. Sølvsten (2019). Leave-out Estimation of Variance Components. *NBER Working Paper 26244*.

Kremer, M. and E. Maskin (1996). Wage Inequality and Segregation by Skill. *NBER Working Paper 5718.*

Lacava, C. (2019). Brain Drain, Employment and Productivity. *Mimeo*.

Lazear, E. P. and K. L. Shaw (2009). Wage structure, Raises and Mobility. In E. P. Lazear and K. L. Shaw (Eds.), *The Structure of Wages: An International Comparison*, pp. 1–57. Chicago: University of Chicago Press.

Levine, R. and D. Renelt (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review 82*(4), 942–963.

Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inf. Theor. 37*(1), 145–151.

Lise, J., C. Meghir, and J. M. Robin (2016). Matching, Sorting and Wages. *Review of Economic Dynamics 1*(19), 63–87.

Lopes de Melo, R. (2018). Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence. *Journal of Political Economy 126*(1), 313 – 346.

Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review 104*(2), 422–58.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*, 2295–2350.

Restuccia, D. and R. Rogerson (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics 11*(4), 707–720.

Restuccia, D. and R. Rogerson (2013). Misallocation and Productivity. *Review of Economic Dynamics 16*(1), 1 – 10.

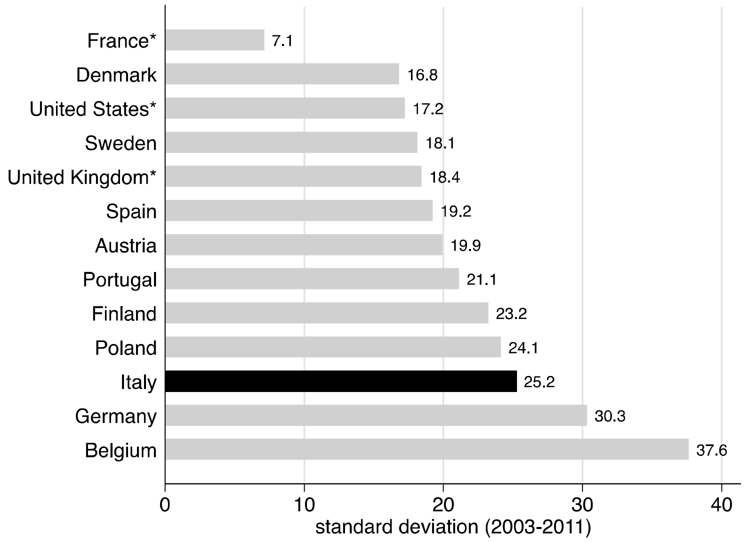
Sala-I-Martin, X. (1997). I Just Ran Two Million Regressions. *The American Economic Review 87*(2), 178–183.

Shimer, R. and L. Smith (2000). Assortative Matching and Search. *Econometrica 68*, 343–369.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature 49(2)*, 326–65.

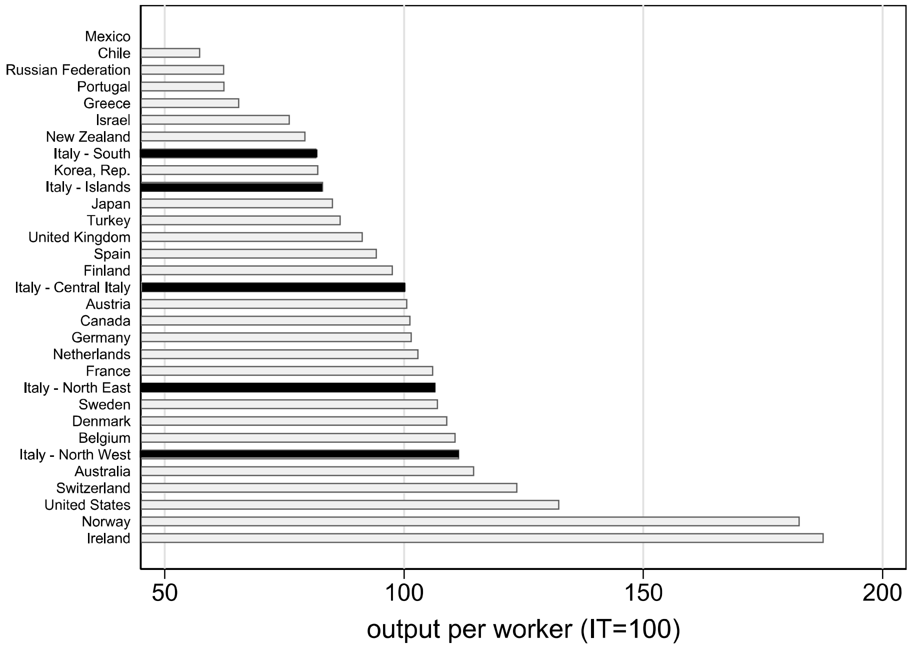
9 Figures

Figure 1: Output per worker dispersion across regions by country



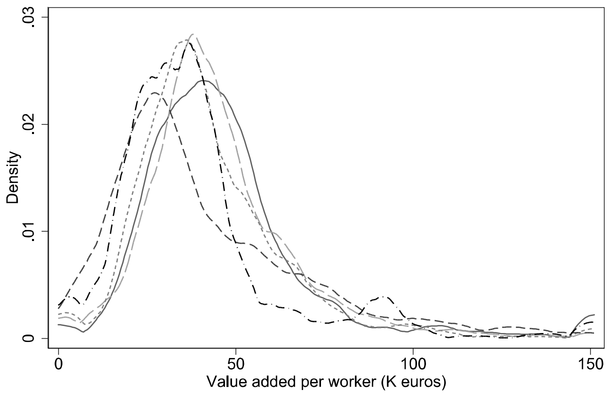
Notes. Data sources: Eurostat and Bureau of Economic Analysis. Standard deviation is computed after rescaling the regional output per worker with respect to the mean value of the country (set to 100). Outliers are excluded: Île de France in France, DC in the US, and Inner London in the UK.

Figure 2: Value added per worker across countries and Italian regions



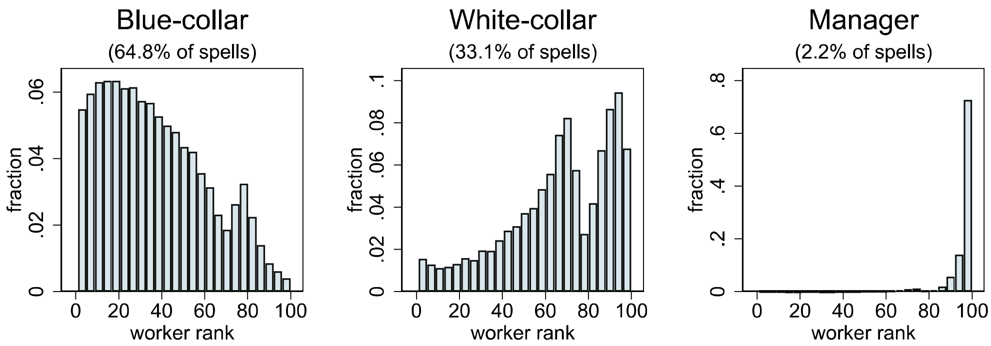
Notes. GDP per person employed (constant 2011 PPP$) across countries and Italian regions (in black) in 2016. The sample includes EU15 and OECD countries with more than one million employees. Data sources: Penn Table 9.1 and Italian National Statistics Institute. Values reported as a percentage on the value for Italy (IT=100: 61,128 PPP$).

Figure 3: Firm value added per worker by region



Notes. Density of value added per worker of firms by region. Output per worker is reported in euro (1995 fixed prices) and the 1-st and 99-th percentiles are winsorized.

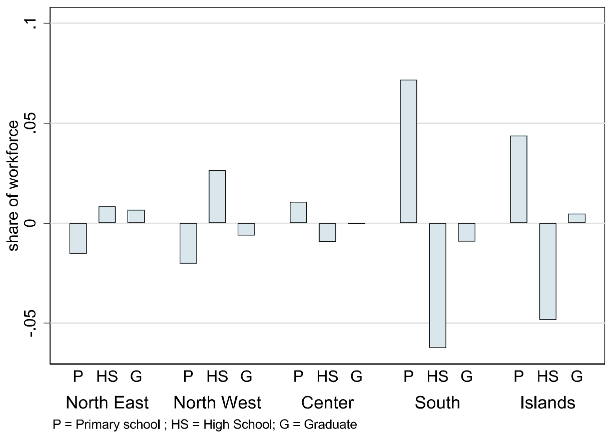
Figure 4: Ranks of workers by job category



Mobility across job categories (%)	Blue-collar	White-collar	Manager	All
Blue-collar	58.1			
White-collar	15.3	22.7		
Manager	0.1	2.2	0.7	
All				0.2

Notes. The figure reports the distribution of the ranks of workers by the job category of any spell in the full sample of workers. Ranks are reported as the corresponding percentiles of the global ranking of workers. The fraction of workers with spells in more than one category is reported in the companion table. The distributions of the ranks computed by considering only the workers with spells in a single category have similar profiles.

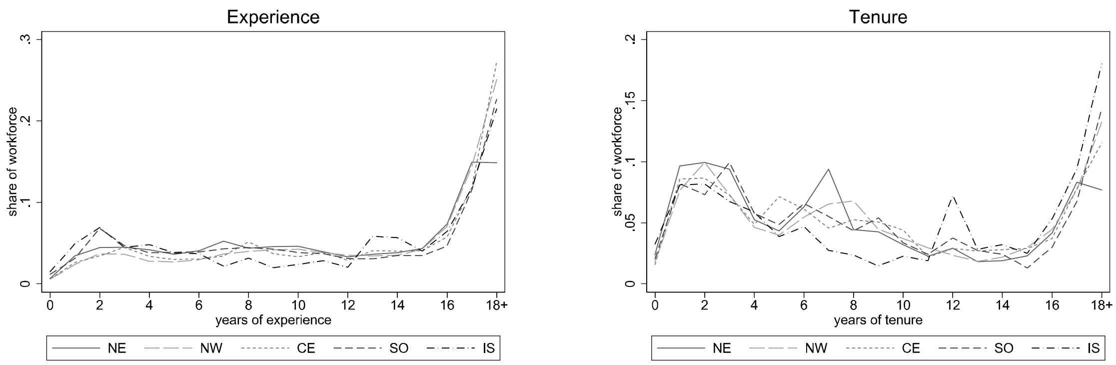
Figure 5: Schooling levels of employees in the manufacturing sector by region



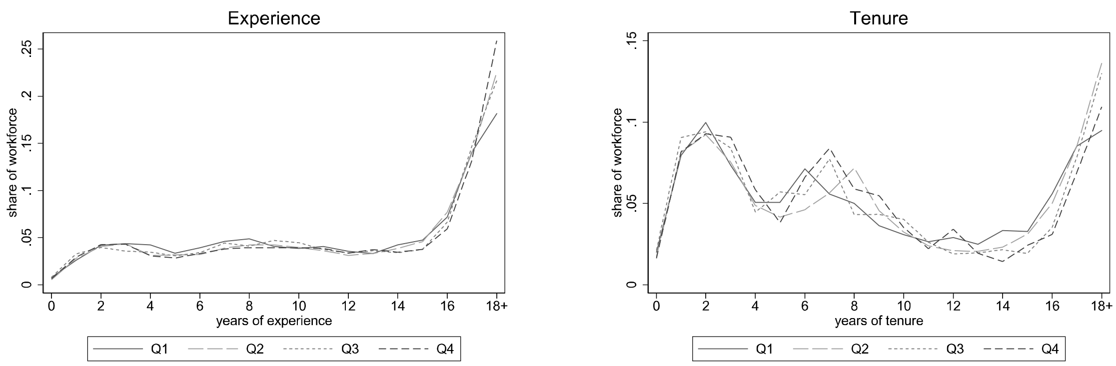
Notes. The figure reports the regional deviation from the average composition of workforce schooling in the manufacturing sector in 1995. Computation using the longitudinal data from the Labour Force Survey by the Italian National Institute of Statistics. National average shares by group of completed schooling level: Primary school (P)=68.13%, High School (HS)=28.85%, Graduate (G)=3.02%.

Figure 6: Distribution of workforce by experience and tenure

Panel A: Across regions

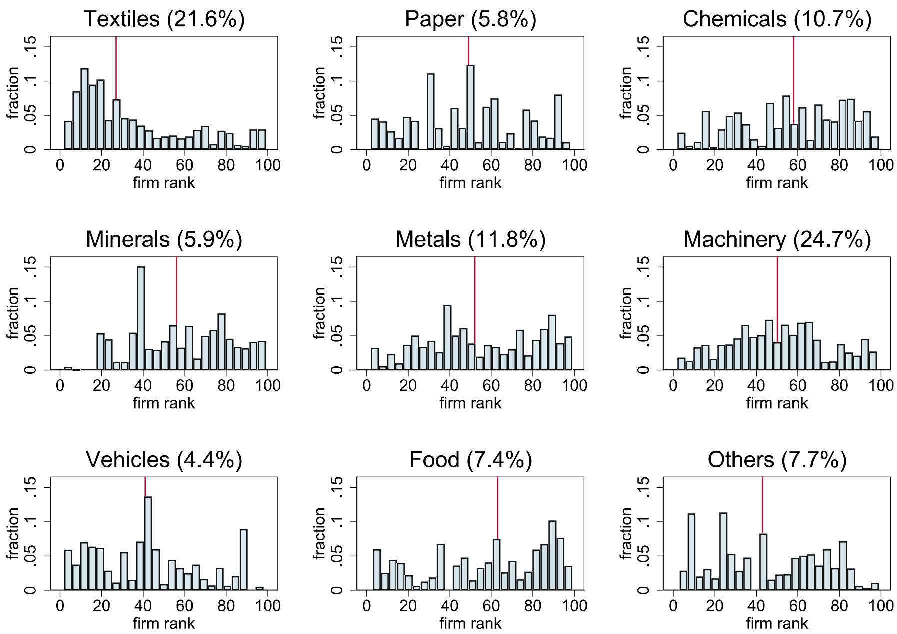


Panel B: Across quartiles of firm's value added per worker



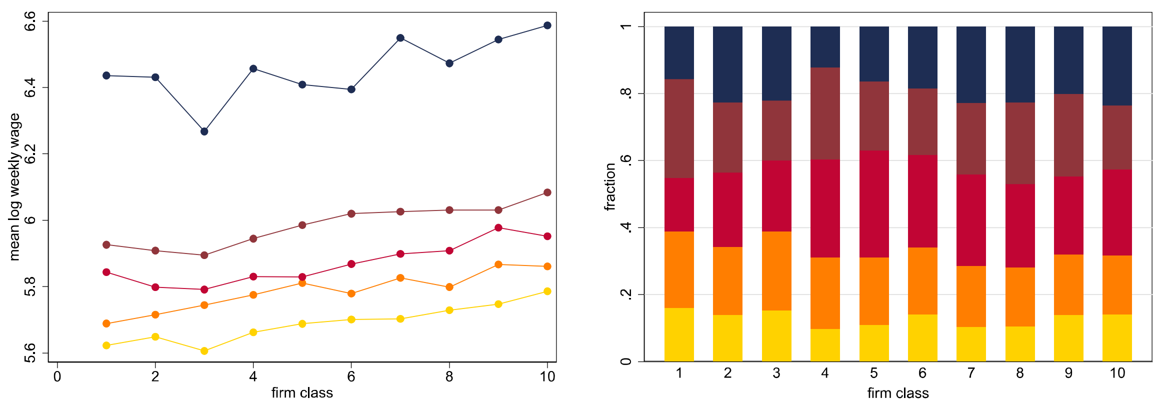
Notes. Share of workers in 1997 by years of experience and tenure across regions (panel A) and across firms in different quartiles of value added per worker (panel B). A year is defined as 52 weeks of work as employed worker.

Figure 7: Ranks of firms by sector of activity



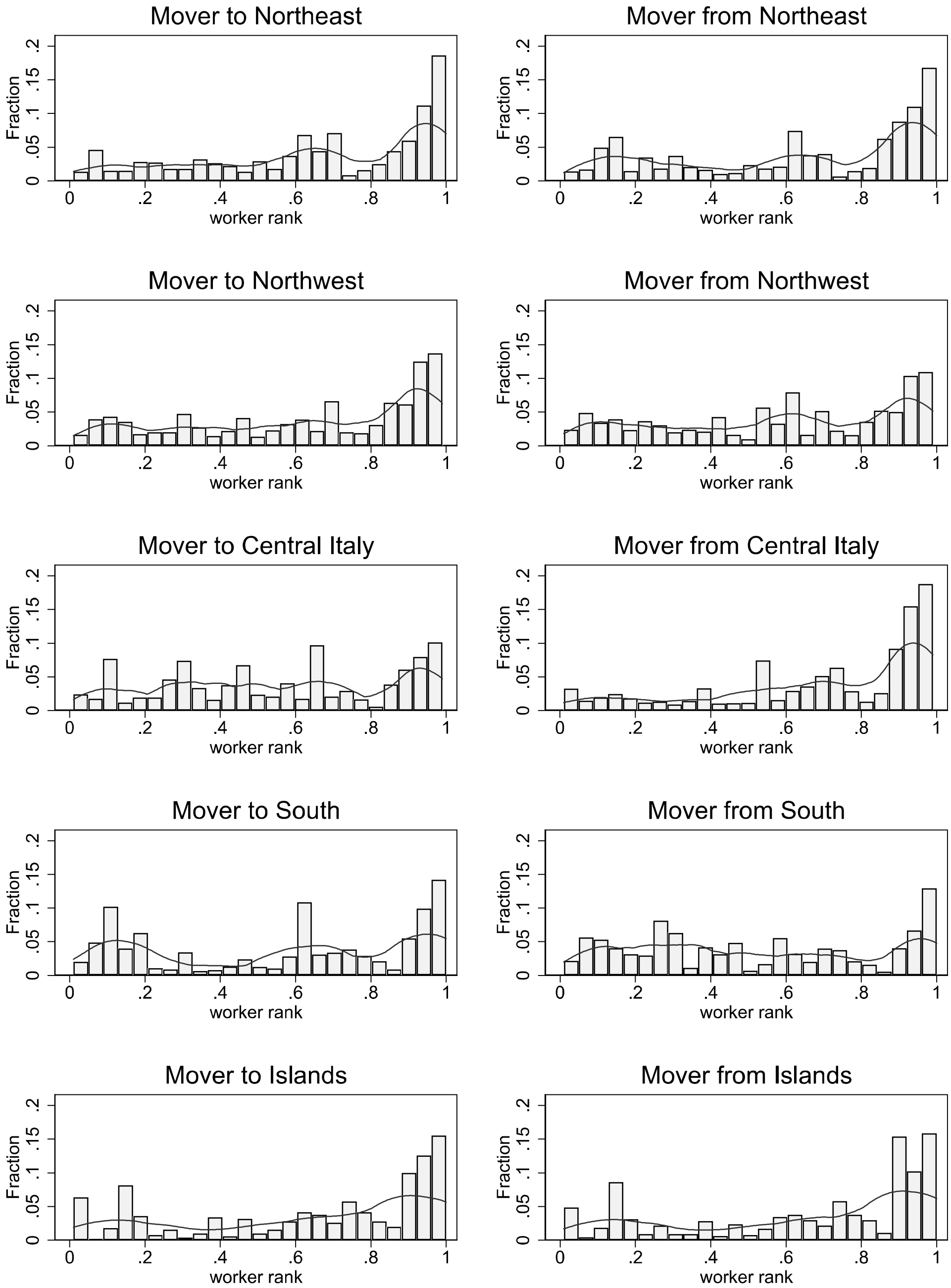
Notes. Distribution of firms by their rank in each sector of activity. Ranks are reported as the corresponding percentiles of the global ranking of firms. The share of firms in each sector is reported in parentheses. The red line indicates the median worker rank in each sector.

Figure 8: Mean log wages and proportions of workers across firms



Notes. Workers are categorized in 6 classes of increasing rank (distinguished by color) and firms are ordered on the x-axis in 10 classes of increasing rank. The left panel plots means of log weekly wages of male employees by worker class and firm class. The right panel shows estimates of the proportions of worker ranks in each firm class.

Figure 9: Rank of movers by region



Notes. Distribution of movers to and from each region according to the within-region ranking of workers. A worker is defined as a mover to (from) a region if he had at least one job spell in a different (that) region and then moved to that (a different) region.

10 Tables

Table 1: Descriptive statistics

	Italy	Northeast	Northwest	Central Italy	South	Islands
Workers: INPS (1980-1997)						
Employment spells	18,806,792	4,480,119	8,389,651	2,730,587	2,526,960	676,532
		(23.82%)	(44.61%)	(14.52%)	(13.44%)	(3.60%)
Workers	1,335,161	369,582	656,607	235,012	210,836	58,795
		(27.68%)	(49.18%)	(17.60%)	(15.79%)	(4.40%)
Share of women	22.95	24.74	23.17	20.68	14.47	10.35
Share of workers with spells in:						
- more than one year	97.43	97.32	98.03	97.44	98.01	97.27
- more than one firm	76.87	81.12	79.78	78.65	77.36	73.53
- more than one region	12.71	20.90	19.78	29.23	31.76	38.43
Average n. years in the sample:	12.65	11.89	13.10	12.55	13.05	12.33
Average n. years in the sample:	2.93	3.46	3.11	3.15	2.92	3.54
Workers: INPS-INVIND (1991-1997)						
Employment spells	3,188,743	772,968	1,428,486	463,631	425,274	98,384
		(24.24%)	(44.80%)	(14.54%)	(13.34%)	(3.09%)
Workers	884,427	210,509	387,337	133,729	134,898	30,948
		(23.80%)	(43.80%)	(15.12%)	(15.25%)	(3.50%)
Share of women	23.43	27.80	25.09	21.71	15.00	13.33
Share of workers with spells in:						
- more than one year	73.94	76.66	74.24	73.63	71.78	72.87
- more than one firm	5.68	3.52	9.22	4.20	2.70	1.98
- more than one region	1.45	2.11	2.31	4.54	4.03	2.90
Average n. years in the sample:	3.35	3.45	3.48	3.37	3.03	3.11
Average n. different employers:	1.06	1.04	1.10	1.04	1.03	1.02
Firms: INPS-INVIND (1991-1997)						
Observations	6,195	2,413	3,542	2,071	1,278	718
		(38.95%)	(57.18%)	(33.43%)	(20.63%)	(11.59%)
Firms	1,464	586	841	516	343	196
		(40.03%)	(57.45%)	(35.25%)	(23.43%)	(13.39%)
Share of firms active in:						
- more than one year	85.31	82.94	83.59	80.43	78.43	78.57
- more than one region	32.72	56.14	49.35	68.22	71.43	79.08
Average n. years in the sample:	4.23	4.41	4.38	4.39	4.19	4.31
Average n. different regions:	1.59	2.22	1.99	2.52	2.91	3.52
Regions: INPS-INVIND (1991-1997)						
Value added per worker (euro 1995)	45,578	48,966	44,957	42,552	42,200	41,725

Table 2: Weekly wage by region and occupational status

Region	All workers		Blue-collar			White-collar			Managers		
	median	sd	median	sd	(%)	median	sd	(%)	median	sd	(%)
Northeast	436.8	249.3	407.5	90.0	(66)	512.1	189.8	(31)	1267.9	779.0	(3)
Northwest	439.0	248.8	407.6	82.4	(68)	524.5	166.2	(28)	1314.0	628.8	(3)
Central Italy	426.6	296.8	404.1	88.5	(73)	505.1	159.0	(24)	1686.2	784.4	(3)
South	417.5	234.5	400.4	83.1	(76)	502.6	298.6	(22)	1356.0	553.3	(2)
Islands	408.1	225.4	381.0	91.4	(70)	502.4	179.1	(28)	1338.5	638.2	(2)
Italy	434.1	255.3	405.5	86.2	(69)	516.2	183.1	(28)	1321.5	711.5	(3)

Notes. Median and standard deviation of residual weekly wages (in euro, 1995 fixed prices) of workers in the INPS-INVIND sample after controlling for year fixed effect and sector and gender interactions. Shares of workers by occupational status are reported in parentheses. Baseline variables: gender: woman, year: 1991, sector: paper.

Table 3: Residual wage estimation

Dep. Var.: Log weekly wage	(1)	(2)	(3)
Experience (years)	0.056***	0.052***	0.037***
	(0.000)	(0.000)	(0.000)
Experience squared	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Tenure (years)	-0.009***	-0.005***	-0.001***
	(0.000)	(0.000)	(0.000)
Tenure squared	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Man		0.225***	
		(0.000)	
Year FE	yes	yes	yes
Week FE	yes	yes	yes
Worker FE			yes
Plant FE			yes
R-sq.	0.184	0.226	0.798
N. Obs.	18,806,792	18,806,792	18,806,792

Notes. OLS estimates with robust standard errors (in parentheses) are reported. The sample includes all employment spells in the period 1980-1997 of individuals who worked in an INVIND firm in the period 1991-1997. Experience and tenure count the number of years worked by the individual and the years worked in the same firm, respectively. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 4: Ranks of workers across regions

A. REGION OF JOB	mean	sd	skewness	kurtosis	p25	p75	JSD_{NE}	BC_{NE}
Northeast	51.04	29.09	-0.07	1.80	25	76	0	1
Northwest	51.31	28.87	-0.03	1.81	27	77	0.05	0.97
Central Italy	48.02	28.56	0.18	1.79	24	74	0.08	0.98
South	48.74	28.99	0.08	1.78	24	75	0.11	0.96
Islands	48.14	27.14	0.10	1.92	24	71	0.20	0.95
B. REGION OF BIRTH	**mean**	**sd**	**skewness**	**kurtosis**	**p25**	**p75**	JSD_{NE}	BC_{NE}
Northeast	50.79	28.95	-0.07	1.82	25	75	0	1
Northwest	51.85	28.90	-0.06	1.79	27	77	0.04	0.98
Central Italy	48.55	28.77	0.17	1.78	25	75	0.08	0.98
South	48.49	28.81	0.10	1.84	24	74	0.07	0.98
Islands	46.16	27.28	0.15	1.95	22	68	0.10	0.98

Notes. Statistics of the percentiles of the global ranking of workers according to their skills. The sample includes all the workers in the INPS-INVIND matched sample. $0 < JS_{NE} < 2$ is the Jensen-Shannon divergence between the distribution of ranks in a region and the distribution of ranks in the Northeast. $0 < BC_{NE} < 1$ is the Bhattacharyya coefficient of overlap between the distribution of ranks in a region and in the Northeast.

Table 5: Rank of firms explained by observables

Dep. Var.: Firm rank	(1)	(2)	(3)
Firm has plants in more than one region	11.728***	10.819***	11.099***
	(1.169)	(1.906)	(1.607)
Share of employees by region:			
– Northwest	0.564	0.812	2.466**
	(1.210)	(1.181)	(0.969)
– Central Italy	-2.269	-2.179	0.758
	(1.463)	(1.455)	(1.374)
– South	-7.111***	-7.279***	-3.562*
	(2.157)	(2.125)	(1.952)
– Islands	-11.330***	-12.712***	-7.983***
	(2.399)	(2.444)	(2.287)
Number of employees:			
– 100 - 250	1.314	0.943	1.542*
	(1.023)	(1.003)	(0.802)
– 250 - 1000	2.236**	2.002*	1.875**
	(1.084)	(1.065)	(0.845)
– >1000	3.164**	2.623	4.228***
	(1.588)	(1.675)	(1.351)
Firm was founded in last 5 years	-6.044*	-6.597*	-3.483
	(3.595)	(3.496)	(2.919)
Share of employees by worker rank:			
– second quartile		2.566	0.336
		(1.796)	(1.685)
– third quartile		8.917***	4.531**
		(1.313)	(1.897)
– fourth quartile		-10.217***	-12.419***
		(1.339)	(1.152)
Value added per worker			0.401***
			(0.128)
R-squared	0.034	0.060	0.321
N. Obs.	6,195	6,195	6,195

Notes. OLS estimates with robust standard errors (in parentheses) are reported. The sample includes all firms in the matched INPS-INVIND sample. Baseline variables are: share of employees in the Northeast, number of employees lower than 100, and share of employees in the first quartile of the ranking of workers. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 6: Ranks of firms across regions

A. REGION OF JOB	mean	sd	skewness	kurtosis	p25	p75	JSD_{NE}	BC_{NE}
Northeast	49.27	26.52	0.16	1.99	27	68	0	1
Northwest	49.72	26.43	0.04	1.90	28	72	0.17	0.98
Central Italy	47.67	27.59	0.20	1.79	23	71	0.24	0.97
South	44.01	32.78	0.32	1.56	12	80	0.42	0.94
Islands	39.45	27.04	0.56	2.28	17	61	0.65	0.87
B. REGION OF HEADQUARTER	mean	sd	skewness	kurtosis	p25	p75	JSD_{NE}	BC_{NE}
Northeast	48.95	26.64	0.17	1.98	27	67	0	1
Northwest	50.06	26.29	0.03	1.91	29	72	0.16	0.95
Central Italy	47.50	27.69	0.20	1.78	23	71	0.26	0.92
South	43.89	33.16	0.33	1.55	12	80	0.45	0.83
Islands	37.92	26.43	0.63	2.41	17	56	0.70	0.62

Notes. Statistics of the percentiles of the global ranking of firms according to their technologies. The sample includes all the firms in the INPS-INVIND matched sample. $0 < JS_{NE} < 2$ is the Jensen-Shannon divergence between the distribution of ranks in a region and the distribution of ranks in the Northeast. $0 < BC_{NE} < 1$ is the Bhattacharyya coefficient of overlap between the distribution of ranks in a region and in the Northeast. The information on the location of the headquarter is missing for 14% of the firms.

Table 7: Allocation of workers across firms

	IT	NE	NW	CE	SO	IS
Between firm variance	0.13	0.12	0.14	0.17	0.18	0.21
Co-worker correlation	0.36	0.33	0.37	0.40	0.41	0.44
Spearman rank correlation	0.04	0.01	0.08	0.01	0.01	0.09

Notes. The table reports alternative indicators of allocation at the national level and within each region.

Table 8: Production function

Step 1	Parameter	Estimate	Std.Err.
	A_{NW}/A_{NE}	0.961	0.038
	A_{CE}/A_{NE}	0.948	0.052
	A_{SO}/A_{NE}	0.941	0.066
	A_{IS}/A_{NE}	0.923	0.118
	N. Obs.	2331	
	R-Squared	0.376	

Step 2	Parameter	Estimate	Std.Err.
	λ	0.700	0.029
	ρ	-1.178	0.212
	η	494.213	68.640
	N. Obs.	6195	
	R-Squared	0.314	

Table 9: Estimated parameters of $Beta(\alpha, \beta)$ probability density function

	SKILLS				TECHNOLOGIES			
By region of:	job		birth		plant		headquarter	
	α	β	α	β	α	β	α	β
Northeast	0.934	0.858	0.954	0.900	1.290	1.276	1.272	1.272
Northwest	0.951	0.852	0.967	0.848	1.337	1.350	1.358	1.354
Central Italy	0.962	0.943	0.959	0.922	1.202	1.267	1.191	1.265
South	0.953	0.960	0.906	0.896	0.753	0.888	0.732	0.862
Islands	1.083	1.118	1.012	1.140	0.990	1.415	1.005	1.530

Notes. Fitted parameters of $Beta(\alpha, \beta)$ probability density functions of the ranks of workers by region of job and by region of birth, and of the ranks of firms by region of the job and by region of the headquarter. These parameters are used to replicate the distributions of skills and technologies in the counterfactual exercises in Section 6.

Table 10: Accounting for labor productivity differentials

Panel A				
Increase in productivity (%)	Northwest	Central Italy	South	Islands
Regional factor	4,08	5,51	6,22	8,29
Skills	-0,78	0,96	1,89	0,95
Technologies	0,40	3,05	14,81	22,98
Allocation	-0,43	1,28	-0,17	-1,26
Panel B				
Contribution to differentials (%)	Northwest	Central Italy	South	Islands
Regional factor	124,49	51,04	27,34	26,78
Skills	-23,65	8,88	8,29	3,06
Technologies	12,26	28,26	65,13	74,22
Allocation	-13,10	11,82	-0,76	-4,06

Notes. Panel A reports the percentage change in the aggregate output per worker of each region (columns) after imposing the value of a single determinant (rows) equal to the value of the corresponding determinant in the most productive region. Panel B reports the ratio between the change in panel A and the difference in aggregate output per worker in the considered region and in the most productive region.

Table 11: Changes in labor productivity if firms and workers are matched optimally

	$\%\Delta P_{IT}$	$\%\Delta P_{NE}$	$\%\Delta P_{NW}$	$\%\Delta P_{CE}$	$\%\Delta P_{SO}$	$\%\Delta P_{IS}$
All sectors pooled	18,74	18,90	17,41	19,73	27,37	17,34
Within sector	15,71	14,98	15,74	16,27	21,22	10,66
Textiles	20,83	15,52	21,59	25,97	22,15	-1,24
Paper	11,21	14,19	9,40	10,54	45,00	-11,67
Chemicals	11,87	11,35	12,53	7,75	19,74	14,49
Minerals	7,58	4,30	6,14	14,09	2,40	12,26
Metals	18,84	20,76	16,36	13,41	53,76	9,18
Machinery	16,60	13,20	20,15	15,68	6,36	4,33
Vehicles	6,82	13,83	4,14	5,26	-10,82	34,47
Food	14,22	19,97	6,81	9,43	20,07	12,51
Others	20,35	23,19	17,45	15,58	31,19	0,94

Notes. The first row displays the percentage change in the national (IT) and regional (other columns) aggregate output per worker (P) after reassigning workers to firms in the order of their ranking within the area. The second row reports the change when the reassignment is done within area and sector. The remaining rows report the change by sector of activity.

Table 12: Impact of worker migration and firm relocation on productivity

	$\%\Delta P_{IT}$	$\%\Delta P_{NE}$	$\%\Delta P_{NW}$	$\%\Delta P_{CE}$	$\%\Delta P_{SO}$	$\%\Delta P_{IS}$
No migration	-0.28	0.48	-0.31	0.31	-0.20	-2.31
No firm relocation	-0.66	0.47	-0.38	-0.41	-0.30	-3.53

Notes. Percentage change in the national (IT) and regional (other columns) aggregate output per worker (P) after assigning to the area the distribution of skills of workers born in the area (first row) and the distribution of technologies of firms with headquarter in the area (second row), respectively.

Table 13: Mean rank of movers and multi-plant firms by region

Panel A. Movers		Job									
		Northeast		Northwest		Central Italy		South		Islands	
Birth	Northeast	50.56	(88.9%)	52.75	(5.4%)	57.62	(1%)	53.83	(0.5%)	79.36	(0.9%)
	Northwest	57.12	(4.3%)	51.50	(80.1%)	65.47	(1.9%)	62.78	(1.1%)	64.48	(1.2%)
	Central Italy	60.66	(1.8%)	65.29	(1.4%)	46.82	(89.5%)	55.46	(2.8%)	76.68	(1.4%)
	South	49.94	(3.7%)	48.00	(9.0%)	52.23	(5.7%)	47.81	(94.8%)	62.87	(1%)
	Islands	48.18	(1.3%)	44.04	(4.1%)	44.34	(1.9%)	63.80	(0.8%)	47.89	(95.4%)
Panel B. Multi-plant firms		Plant									
		Northeast		Northwest		Central Italy		South		Islands	
Headquarter	Northeast	49.10	(94.6%)	33.19	(1.4%)	65.45	(1.9%)	53.85	(2.1%)	49.48	(2.1%)
	Northwest	61.26	(4.6%)	49.82	(97.2%)	57.20	(7.1%)	56.46	(9.1%)	72.03	(10%)
	Central Italy	58.20	(0.7%)	59.25	(1%)	47.34	(90.1%)	47.29	(3.1%)	66.27	(2.3%)
	South	77.92	(0.1%)	69.98	(0.4%)	75.77	(0.5%)	43.39	(85.5%)	58.14	(0.8%)
	Islands	25.82	(0.1%)	32.00	(0%)	19.83	(0.5%)	40.07	(0.2%)	38.21	(84.9%)

Notes. Panel A displays the mean rank of movers by region of birth (row) and region of job (column). Panel B displays the mean rank of plant by region of headquarter (row) and region of subsidiary (column). Rank is multiplied by 100. In parentheses, the fraction of workers or firms in each category out of the total in the region of the job or plant is reported.

Table 14: Production function under log-normal distribution of types

Parameter	*Log-normal skills*	*Log-normal technologies*	*Log-normal skills and technologies*
	(1)	(2)	(3)
λ	0.841	0.687	0.627
	(0.000)	(0.000)	(0.000)
ρ	-0.795	-3.291	-1.350
	(0.077)	(0.000)	(0.000)
η	927.760	600.240	574.020
	(67.596)	(0.018)	(0.046)
N. Obs.	6,195	6,195	6,195
R-Squared	0.300	0.396	0.337

Notes. Estimates of the production function parameters by assuming log-normal distribution of skills (column 1), technologies (column 2), and both skills and technologies (column 3). Standard errors in parentheses. Regional parameters estimated in the first step are left unchanged (see baseline estimate in Table 8).

Table 15: Counterfactual exercises under log-normal distribution of types

Panel A. Log-normal skills	IT	NE	NW	CE	SO	IS
Spearman rank correlation	0.08	0.03	0.11	0.06	0.03	0.13
Contribution to productivity differentials (%):						
- Regional factors			151.09	52.62	23.42	30.19
- Skills			-5.01	4.80	1.79	2.24
- Technologies			-28.16	37.91	75.58	68.45
- Allocation			-17.93	4.68	-0.78	-0.87
ΔP under reallocation (%), all sectors	13.27	13.62	13.44	10.53	17.14	12.21
ΔP under reallocation (%), within sector	10.33	11.33	9.56	9.27	14.70	0.59
ΔP under no migration (%)	-0.06	-0.10	0.14	0.12	-0.05	-0.51
ΔP under no firm relocation (%)	-0.58	-0.37	0.51	-0.18	-1.17	-2.39

Panel B. Log-normal technologies	IT	NE	NW	CE	SO	IS
Spearman rank correlation	0.08	0.04	0.11	0.08	0.03	0.11
Contribution to productivity differentials (%):						
- Regional factors			162.58	65.38	28.47	33.66
- Skills			-16.11	0.21	1.51	-2.33
- Technologies			-30.13	38.27	74.30	75.41
- Allocation			-16.35	-3.86	-4.28	-6.74
ΔP under reallocation (%), all sectors	8.41	6.53	10.24	9.12	12.19	5.83
ΔP under reallocation (%), within sector	7.50	6.51	8.41	7.52	12.28	3.10
ΔP under no migration (%)	-0.14	0.13	0.25	0.03	-0.50	-0.86
ΔP under no firm relocation (%)	-0.45	-0.66	0.54	-0.14	-0.58	-1.74

Panel C. Log-normal skills and technologies	IT	NE	NW	CE	SO	IS
Spearman rank correlation	0.09	0.05	0.12	0.09	0.03	0.12
Contribution to productivity differentials (%):						
- Regional factors			161.49	64.23	29.00	34.60
- Skills			-2.22	2.16	0.83	0.77
- Technologies			-28.29	32.79	72.70	67.94
- Allocation			-30.98	0.82	-2.52	-3.31
ΔP under reallocation (%), all sectors	9.63	9.12	10.46	8.96	11.58	7.89
ΔP under reallocation (%), all sectors	7.55	7.71	7.39	7.09	11.54	1.72
ΔP under no migration (%)	-0.02	-0.04	0.06	0.04	-0.02	-0.18
ΔP under no firm relocation (%)	-0.43	-0.59	0.48	-0.11	-0.62	-1.55

Notes. Estimates of rank correlation and results of counterfactual exercises under the assumption of log-normal distribution of skills (panel A), of technologies (panel B), and of both skills and technologies (panel C).

Appendix

A Wage variance decomposition

Since the aim of this work is to separate the contribution of skills, technologies, and region-specific factors to heterogeneity in productivity, a decomposition of the wage variance due to fixed effects of workers, firms, and regions can be a useful reference. Consider a standard Mincerian wage equation with a fixed effect for the worker, for the firm and for the region where the job is located,

$$w_{ifrt} = Z_{ifrt}\beta + \theta_i + \phi_f + \psi_r + \epsilon_{ifrt}, \tag{7}$$

where w_{ifrt} is the logarithm of the weekly wage, $i = 1, \ldots, N$ is an index for workers, $f = 1, \ldots, F$ for firms, $r = 1, \ldots, R$ for regions and $t = 1, \ldots, T$ for years. The vector Z_{ifrt} includes a quadratic specification for experience and tenure, and year fixed effects. This regression is an extended version of the linear wage models with additive person and establishment fixed effects, introduced by AKM.[23] It estimates a time-invariant component θ_i for each worker i with at least two spells of employment, a time-invariant component ϕ_f for each firm f, and a time-invariant component ψ_r for each region r. The term ϵ_{ifrt} captures the variation that is not explained by the heterogeneity of the covariates and the time-invariant effects.

I fit the model on the INPS-INVIND sample by using the iterative algorithm for models with high-dimensional fixed effects proposed in Carneiro et al. (2012) and Correia (2016). The coefficients are similar to those of the standard AKM model reported in column 3 of Table 3, but adding the worker and region fixed effects increases the explanatory power (the R-squared is higher than 0.9). In the first column of panel A in Table A1, I compute the variance of each group of estimated fixed effects. To assess which fixed effect contributes the most to the dispersion of the log wage, in the second column I divide the variance of each group by the variance of the wage.[24] The variance of worker fixed effects accounts for 92% of the variance of wages, while the variance of firm fixed effects explains around 11% of the variance of wages, consistently with the low between-firm dispersion of the wages. In line with the almost null between region dispersion of wages, the variance of region fixed effects is almost null and has no power in explaining the variance of wages. As an alternative test, I run the wage regression excluding one variable at a time and then I compute the percentage decline in the R-squared with respect to the full regression. Results in the third column show that the explanatory power of fixed effects follows closely the share of variance explained. Finally, the remaining columns report the correlation between the considered variables. Notice that the

[23] Unfortunately, since these data are recorded for tax purposes information on workers' schooling is not collected. Whenever available, the usual set of observable characteristics includes year dummies as well as quadratic and cubic terms in age fully interacted with educational attainment as in Card et al. (2013).

[24] The sum of the variances of the model components is greater than the variance of the wages given the negative correlations between the model components.

correlation between worker and firm fixed effects is negative and strong. This finding can result from the negative bias characterizing AKM estimates, which is discussed in Gautier and Teulings (2006); Eeckhout and Kircher (2011); Lopes de Melo (2018) among others, while Andrews et al. (2008) and Kline et al. (2019) propose bias corrections. Panel B shows that the results remain unchanged by restricting the sample to workers with spells in more than one firm. They also hold when considering workers by their occupational status.

Table A1: Wage variance decomposition using fixed effects

Panel A: All workers

Variable	(1) $\sigma^2_\cdot$	(2) $\sigma^2_\cdot/\sigma^2_w$	(3) Expl.adj-R^2	(4) $Corr(\cdot,w)$	(5) $Corr(\cdot,Z'\hat{\beta})$	(6) $Corr(\cdot,\theta)$	(7) $Corr(\cdot,\phi)$	(8) $Corr(\cdot,\psi)$
Dep. Var.: Wage (w)	0.1765	1	.	1				
Observables ($Z'\hat{\beta}$)	0.0044	0.0248	0.0259	0.2569	1			
Worker FE (θ)	0.1624	0.9201	0.9169	0.8925	0.0984	1		
Firm FE (ϕ)	0.0204	0.1153	0.1158	0.1104	0.0241	-0.2418	1	
Region FE (ψ)	0.0002	0.0009	0.0009	-0.0034	-0.1055	-0.0031	-0.0383	1

Panel B: Movers

Variable	(1) $\sigma^2_\cdot$	(2) $\sigma^2_\cdot/\sigma^2_w$	(3) Expl.adj-R^2	(4) $Corr(\cdot,w)$	(5) $Corr(\cdot,Z'\hat{\beta})$	(6) $Corr(\cdot,\theta)$	(7) $Corr(\cdot,\phi)$	(8) $Corr(\cdot,\psi)$
Dep. Var.: Wage (w)	0.2595	1	-	1				
Observables ($Z'\hat{\beta}$)	0.0011	0.0043	0.0045	0.1673	1			
Worker FE (θ)	0.2674	1.0304	0.9639	0.9060	0.0901	1		
Firm FE (ϕ)	0.0370	0.1425	0.1347	0.0748	0.0211	-0.2979	1	
Region FE (ψ)	0.0005	0.0018	0.0018	-0.0333	0.0479	-0.0623	-0.0408	1

Notes. Statistics computed using the fixed effects estimated in a AKM regression augmented with region fixed effects (FE). Variables are indicated on the left. The first column reports the variance of the variable, the second column the share of wage variance explained by the variance of the variable, the third column the share of R-squared explained by a regression of wage on the considered variable, the remaining columns report the correlation between the considered variables.