Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data

Susana Iranzo, *Universitat Rovira Virgili*

Fabiano Schivardi, *University of Cagliari, EIEF, CRENoS, and CEPR*

Elisa Tosetti, *Cambridge University*

We study the relation between workers' skill dispersion and firm productivity using a unique data set of Italian manufacturing firms with individual records on all their workers. Our measure of skill is the individual worker's effect from a wage equation. We find that a firm's productivity is positively related to skill dispersion within occupational status groups (production and nonproduction workers) and negatively related to skill dispersion between these groups. Con-

We are indebted to Andrea Brandolini, Francesco Caselli, Hajime Katayama, Francis Kramarz, Michael Kremer, Marco Magnani, Antoinette Schoar, Till Von Wachter, and, especially, Thierry Lallemand for very useful comments and suggestions. We also thank seminar participants at the Bank of Italy; University of Barcelona; University of Cagliari; Luiss University; LaTrobe University; Universitat Rovira Virgili; the 2005 Australian conference of economists, Melbourne; and the "Labor Market Flows, Productivity and Wage Dynamics" conference, Turin. Many thanks to Giuseppe Bruno for helping us with the estimation routine, to the Italian Social Security Institute (INPS) for providing us their worker-level data, and to the Bank of Italy for the firm-level data. We are responsible for any mistakes. The views expressed here are our own and do not necessarily reflect those of the Bank of Italy or the INPS. Correspondence: Fabiano Schivardi, schiva@stanfordalumni.org; Susana Iranzo, susana.iranzo@urv.cat; Elisa Tosetti, et268@cam.ac.uk.

[*Journal of Labor Economics,* 2008, vol. 26, no. 2]

0734-306X/2008/2602-0002$10.00

sistently, most of the overall skill dispersion is within and not between firms. These findings are consistent with some recent hierarchical models of the firms' organizational structure.

I. Introduction

The factors driving firm productivity have been the subject of a good deal of research over the years. Persistent substantial differences in productivity across firms have been documented, and many empirical papers have provided a deeper understanding of the connection between productivity and observable characteristics of firms, such as size, technology, innovative activity, and so forth. However, less is known about the way firms' outcomes are related to the characteristics of the workers they employ. In this study, we focus on one aspect of workforce composition: the skill mix. Using a newly created matched Italian employer-employee data set, we examine the way in which firms' productivity is associated with the dispersion of skills within the firm.

The role of the skill distribution in determining firms' performance is intrinsically related to the nature of the production function and depends on the degree of complementarity or substitutability between skills (Milgrom and Roberts 1990). Some activities depend heavily on the performance of a few workers (Rosen's "superstars" [1981]), leading to a dispersed skill distribution of the workforce; others require that all tasks be performed at a certain level of competence, fostering the formation of teams of workers with similar skill levels (Kremer's "O-ring" theory [1993]). There is a lively theoretical debate on whether and how technological innovation has modified the optimal skill mix over the recent past. On the one hand, some recent matching and sorting models argue that production may have shifted from a mode in which firms hire workers with different skill levels to one in which some firms use mainly high-skill workers (Microsoft) and others only low-skill workers (McDonald's), resulting in low skill dispersion within firms and segregation between them (Kremer and Maskin 1996). On the other hand, hierarchical models of the firm organizational structure suggest that, to the extent that improvements in information and communication technologies reduce communication costs, they might also increase the optimal degree of skill dispersion (Garicano and Rossi-Hansberg 2006).

While the theory behind the role of skill dispersion in firm performance is fairly well developed, the evidence is scant, due to the heavy data requirements. Our goal in this study is to empirically assess the impact on a firm's productivity of its workers' skill mix. More precisely, we ask whether there are productivity effects derived from the particular combination of workers' skills and, if so, whether workers' skills are complementary or substitutable. We address this question using a new matched employer-employee data set that is representative of Italian manufacturing firms with at least 50 employees, that covers almost 20 years (1981–97), and, most importantly, that includes individual information based on social security records on all the workers of each firm in the sample.[1] We have detailed information on the characteristics of the firms as well as on workers' demographics and compensation. This data set, comprising 10 million worker-year and 10,000 firm-year records, offers a unique opportunity to study the skill distribution within and between firms and its role in production for a fairly long period and a representative sample of firms.

The right measure of skills is quite controversial. The most common proxies have been the educational attainment and experience, by themselves or as the basis for the construction of more sophisticated measures of human capital. However, these are mostly measures of formal skills that only imperfectly reflect innate differences in ability and informal skills, such as accuracy on the job or communication ability. Alternatively, some studies have used earnings as the proxy for skills, assuming that workers are paid the value of their marginal product (see, e.g., Davis and Haltiwanger 1991; Dunne et al. 2004). However, wages also entail an important firm component that reflects such things as the firm's compensation policies, rent sharing, and workers' bargaining power within the firm. To overcome these problems, we use the worker fixed effects obtained as a latent variable from a wage equation, as proposed by Abowd, Kramarz, and Margolis (1999). This is a better measure of workers' skills because, by including the firm fixed effect in the wage equation, we control for firm (and sector) idiosyncrasies; moreover, not only is it based on observable characteristics but it also includes innate ability and informal skills not reflected in these. In fact, this indicator is increasingly used in the literature to construct human capital measures (Abowd, Lengermann, and McKinney 2003; Haskel, Hawkes, and Pereira 2005).

With this measure, we first examine the distribution of workers' skills between and within firms. We compute the share of overall skill dispersion accounted for by the between-firm component (the segregation index) and its evolution from the early eighties to the late nineties, a period in which important changes in the firms' organization may have taken place. This gives us an idea of the relative importance of between- and within-firm skill dispersion and of any pattern over time. We then move on to study the relation between productivity and skill dispersion at the level of the firm directly. We estimate a generalized constant elasticity of sub-

[1] Previous studies using individual worker information at the level of the firm used either a small subsample of the total workforce of each firm (Kramarz, Lollivier, and Pele 1996) or the total workforce of just one firm (Baker, Gibbs, and Holmstrom 1994; Flabbi and Ichino 2001).

stitution (CES) production function to recover the parameters governing skill complementarity-substitutability which, as we show, are directly related to the second moments of the skill distribution. We also test for changes in these parameters by performing the estimation for different subperiods of the sample. The estimation is performed using the procedure of Olley and Pakes (1996) as well as the extension of Ackerberg, Caves, and Frazer (2006) to control for the endogeneity of inputs.

Our results are easily summarized. First, a variance decomposition exercise shows that most of the dispersion in workers' skills is within and not between firms: the between-firm component accounts for less than 20% of overall dispersion. Nor is there any evidence of an increase in this share over time, as a tendency toward skill segregation would imply. These results, robust to a number of checks, are in contrast with the evidence from other countries, such as the United States, Britain, and France, where some tendency toward segregation has been documented, although based on less comprehensive data (Kramarz et al. 1996; Kremer and Maskin 1996; Dunne, Haltiwanger, and Troske 1997; Dunne et al. 2004).

The production function estimates show that overall within-firm skill dispersion has a positive impact on productivity. Distinguishing between production workers (*P*) and nonproduction workers (*NP*), we find that differences in their average skill levels tend to have a negative impact on a firm's productivity. While the relative imprecision of the estimates leaves the ground open to further investigation, we interpret this as evidence that *P* and *NP* workers are imperfect substitutes in production, in line with the results of the vast literature that uses relative labor demand equations to estimate such parameters (Katz and Murphy 1992; Ciccone and Peri 2005). By contrast, within each group of workers, the dispersion of skills, particularly that of *NP* workers, is clearly beneficial for productivity: given an average skill level, it is preferable to have a mix of highly skilled and low-skilled workers than a uniform group. The results are robust to a number of checks. Furthermore, we find no evidence of significant changes over time in the parameters governing skill substitutability and, consequently, in the optimal way to combine skills within the firm. This finding, in agreement with the flat segregation index obtained, constitutes indirect evidence that in our sample there was no substantial change in the production mode during the period.

We have termed the production mode implied by our estimates the "Ferrari and Fiat" model. Ferrari and Fiat are both vertically integrated firms and are therefore likely to have quite a highly dispersed skill distribution. At the same time, reflecting the different technological content of the cars produced, Ferrari has both *P* and *NP* workers with higher average skill than Fiat. Finally, our findings on the connection between skill dispersion and productivity are consistent with a hierarchical organization of production, where it is optimal to concentrate skills in individuals with decision and supervisory power, on whom the firm performance is heavily dependent. According to case studies in the managerial literature, this is the organizational mode adopted by Fiat at least up to the mid-nineties (Tronti 1997).

Some recent international evidence based on matched employer-employee data supports our conclusions. In particular, Lazear and Shaw (forthcoming) in summarizing the findings of nine country studies on wage dispersion within firms, conclude (among other things) that (i) in all countries there is substantial within-firm wage dispersion, which is the primary source of the overall dispersion, and (ii) this dispersion seems to reflect workers' skill heterogeneity rather than different compensation policies across firms. These findings suggest that our results on the relation between skill distribution and productivity might generalize beyond the case of Italy.

The rest of the paper is organized as follows. In Section II we review the theoretical and empirical literature. Section III describes the data and presents the estimation of workers' individual effects. Section IV decomposes the variance of workers' skills between and within firms. Section V examines the relationship between within-firm skill dispersion and productivity by estimating a generalized production function that allows for heterogeneity of workers' skills. Section VI concludes.

II. The Literature

There is a good deal of empirical research that examines the connection between productivity and human capital at the national and local levels but not, until recently, at the micro level of the firm. Using matched employer-employee data sets, Abowd et al. (1999) for France, Haltiwanger, Lane, and Spletzer (1999) for the United States, and Haskel et al. (2005) for the United Kingdom investigate the relation between productivity and workers' skills. All of them find that the most productive firms have more skilled workers. Focusing on average skill levels within firms, these papers implicitly assume that workers' skills are perfect substitutes. In reality, though, they may be substitutes or complements, in which case not only the average level but also the particular combination of skills is important. For example, Kremer and Maskin (1996) use a production function where skills are complementary and where it is therefore optimal to combine workers of similar skills.[2] By contrast, there are activities in which workers' skills are substitutes and the performance of one subset

[2] An extreme case of complementarity is given by Kremer (1993): an O-ring production function where the value of the final product depends crucially on the way every task is performed, so that failure at any stage jeopardizes the entire project.

of workers might be very important, as in the case of coordination and supervision tasks.[3] In this case, it is preferable to have teams with some very talented workers, what Rosen (1981) calls "superstars." These different modes carry precise implications for the relation between skill dispersion and productivity and consequently constitute the basis for extending the production function to include higher moments of the skill distribution. To our knowledge, this is the first study to analyze the role of skill dispersion on firm productivity for a representative sample of firms.[4]

The production mode, and thus the optimal skill mix, might be altered in response to certain changes in the economy. This is the idea underlying Kremer and Maskin (1996) and Acemoglu's (1999) hypothesis of segregation of workers by skills between firms. Following changes in the supply of skills (Kremer and Maskin 1996) and/or in technology (Acemoglu 1999; Caselli 1999), production may shift from a pooling equilibrium in which firms hire workers with different skill levels to a separating one in which some firms use mainly high-skill workers (Microsoft) and others only low-skill workers (McDonald's), resulting in low skill dispersion within firms and segregation across them.

The optimal combination of workers' skills has also been analyzed by models of firm hierarchies such as Rosen (1982) and, more recently, Garicano and Rossi-Hansberg (2004, 2006). These models explain the existence of different occupational categories and hierarchies based on the asymmetry of tasks, so that less skilled workers are assigned more standardized and routine tasks and high-ability workers perform more complex ones. Garicano and Rossi-Hansberg (2004) obtain positive sorting between occupational categories: firms with high-skill managers also hire high-skill (production) workers, while Garicano and Rossi-Hansberg (2006) also obtain "skill stratification" within occupational categories: different hierarchically ranked layers coexist within the same firm.

On the empirical front, there is some evidence on segregation of workers by skill, even if scant and based on coarse and questionable indicators of skills. Dunne et al. (1997, 2004) use the share of *NP* workers as a proxy for (high) skills and document secular increases in this share for all U.S. manufacturing sectors from 1972 to 1988. Though it is a good proxy for

[3] Grossman and Maggi (2000, 1255) argue that many of the goods and services exported by the United States fall into this category, as they "reflect disproportionately the input of a few very talented individuals," and they cite the software industry and the financial services emanating from Wall Street as examples.

[4] Abowd and Kramarz (forthcoming) present some reduced-form regressions that include second moments of the ability distribution, finding inconclusive evidence. Their results are not directly comparable with ours, as they estimate a linear specification without controlling for the endogeneity of inputs and use a different decomposition of workers' skills.

pure skill groups, the classification of workers into P and NP is too coarse and does not reflect only differences in skill. These two types of workers also perform fundamentally different tasks and often work in separate units or departments. For example, in an automobile firm, a mechanic (P) will be working at the assembly line while an engineer (NP) will be working in the design department. At least in the short term, the possibilities of substitution between them are quite limited. We treat P and NP workers as different types of labor, but we also consider heterogeneous workers within each group.

Assuming that wages reflect workers' productivity, Davis and Haltiwanger (1991) and Dunne et al. (2004) use wages as an alternative proxy for skills and analyze the dispersion of wages across and within U.S. plants. But this proxy too is problematic, particularly when computing measures of segregation such as Kremer's index, which is the ratio of the between-firm component of the variance of skills to the total variance. As Abowd et al. (1999) argue, the between-firm variation in wages is partly due to differences in firms' compensation policies unrelated to differences in workers' ability and common to all workers in a firm. Thus, ignoring this results in an upward bias in the between-firm component of skill dispersion and, therefore, in the segregation index. Kremer and Maskin (1996) also reproduce some evidence of skill segregation across firms during the 1980s from studies in the United Kingdom and France. The measure of segregation used is the within-firm correlation among workers of different indicators of skill, such as occupational classification, experience, and wages. This, as we have argued, can be problematic. Moreover, unlike ours, those studies are not based on individual records for all workers within each firm.

This study also contributes to the empirical literature on the degree of substitutability between skilled and unskilled workers (see, e.g., Katz and Murphy 1992; Krusell et al. 2000; Ciccone and Peri 2005; Caselli and Coleman 2006), to which our estimated elasticity of substitution between P and NP average skills can be compared.[5] Unlike all these papers, we estimate the elasticity of substitution between P and NP workers directly from the production function rather than from relative labor demand functions, thus offering an important check to the robustness of these results to the estimation procedure.

Finally, this study is also related to the literature that considers the effect of wage inequality on firm performance, such as "tournament mod-

[5] Berman, Bound, and Griliches (1994) argue that the classification of workers into NP and P, which usually proxies for occupational status (white and blue collar), is also a good proxy for the skill level of workers based on educational attainment, as this classification shows trends similar to those found using education.

els" (Lazear and Rosen 1981), according to which wage dispersion is beneficial for effort extraction, and to models that stress fairness and cooperation (Akerlof and Yellen 1990), which predict the opposite. The evidence from this literature is mixed.[6] We are interested in skill rather than wage dispersion. However, given that most papers do not control for workers' unobserved characteristics but only observable ones, part of the effects attributed to residual wage dispersion might be due to unobserved workers' skills. Addressing this issue could be an interesting direction for future work.

III. Sample Construction

Our empirical strategy is based on two sequential steps. In the first step, we run a wage equation to compute an estimate of each worker's skills. In the second step, we use these estimates to compute the distribution of skills within each firm, and run production function regressions augmented with such measures of the firm-specific skill distribution in order to assess the degree to which workers' skills are complementary within the firm. In this section we discuss the data and the proxy we develop to measure individual-specific skills. In Section V.B, where we carry out the estimation of the production function, we explain why skills are imputed from a wage regression estimated independently of production.

A. Data Description

The data used in this study were constructed from the Bank of Italy's annual INVIND survey of manufacturing firms. INVIND is an open panel of around 1,200 firms per year representative of manufacturing firms with at least 50 employees. It contains detailed information on firms' characteristics (see below). The Social Security Institute (Inps) was asked to provide the complete work histories of all workers who ever transited in an INVIND firm for the period 1981–97, including spells of employment in which they were employed in firms not listed in the INVIND survey. We have information on about a million workers per year, more than half of whom are employed in INVIND firms in any given year. The rest are employed in 100,000 other firms of which we only know the unit identifier.

The data on workers include age, gender, area where the employee works, occupational status (production, clerical, manager), annual gross earnings (including irregular payments such as overtime, shift work, and

[6] See, e.g., Winter-Ebmer and Zweimüller (1999) for evidence on Austria and Lallemand, Plasman, and Rycx (2004) on Belgium; the latter paper contains a comprehensive review of the recent literature. See also the contributions in the book edited by Lazear and Shaw (forthcoming).

Table 1
Workers' Characteristics

	Whole Sample		INVIND Sample	
	Mean	SD	Mean	SD
Weekly wage (€1995)	347.86	138.72	348.54	127.66
Age	37.14	10.00	38.44	9.95
Seniority	4.04	4.02	3.028	2.65
Share of males	79.2		77.6	
Share of production workers	66.0		69.17	
Share of nonproduction workers	32.7		29.93	
No. of observations	17,593,816		9,559,271	

NOTE.—Whole sample refers to all workers in the data set, independently from the firm they are employed in; INVIND sample refers to workers that are currently employed by a firm that belongs to the INVIND survey.

bonuses), number of weeks worked, and the firm identifier. As is always the case with social security data, there is no information on education. We cleaned the data by eliminating the records with missing entries on either the firm or the worker identifier, those corresponding to workers younger than 15 and older than 65, those who had worked less than 4 weeks in a year, and those in the first and last percentiles of the earnings distribution.[7] We also avoided duplication of workers within the same year; when a worker changed employer, we considered only the job at which he had worked the longest and computed weekly earnings accordingly. We use this data set to estimate the wage equation that identifies the worker and firm fixed effects.

Table 1 shows the statistics on workers' characteristics for the total sample and for INVIND firms on which we base the analysis of Section V.[8] For the total sample, average gross weekly earnings at 1995 constant prices are €350, and the average age is 37 years. Almost 80% of the observations pertain to males, 66% to *P* workers, and 32.7% to *NP* workers. The INVIND sample consists of almost 10 million observations. The descriptive statistics are quite similar to those of the total sample, as they contain the same workers but observed only when employed by an INVIND firm.[9]

Attrition in INVIND firms is substantial: on average 10% of workers

[7] Extreme values of the earning distribution could be due to exceptional events (illness and the like) or to measurement error. Given that measures of dispersion are very sensitive to such values, we decided to drop them from the analysis altogether.

[8] We base the analysis of dispersion and productivity on INVIND firms only because these are the firms for which we have detailed information.

[9] Guiso, Pistaferri, and Schivardi (2005) report descriptive statistics for a different sample of workers, representative of the entire population of workers. The characteristics are very similar to those of our sample of manufacturing firms with at least 50 employees.

enter and 12% exit the sample from one year to the next. Overall, approximately 80% of the workers in an INVIND firm in 1981 had dropped out of the sample by 1997, and 72% of the workers in the 1997 sample were not present in 1981. This implies that even if our measure of skills is fixed over time, in principle the skill distribution could have changed significantly due to turnover.

The INVIND survey gives an extensive list of firm characteristics, including industrial sector, nationality, year of creation, average number of employees during the year, value of shipments, value of exports, and investment. In some years, additional questions were asked, for example, one on organizational changes in 1995 and one on number of establishments in 1992–95. We completed the data set with balance-sheet data collected by the Company Accounts Data Service (CADS) since 1982, from which it was possible to reconstruct capital series, using the perpetual inventory method.[10] Investment is at book value, adjusted using the appropriate two-digit deflators, derived from National Accounts published by the National Institute for Statistics. For consistency with the capital data, in the estimation of the production function we take value added and labor from the CADS database. Both the INVIND and the CADS samples are unbalanced, so that not all firms are present in all years.

Table 2 reports summary statistics for the firm data used in the regression analysis. The first two columns are unweighted. On average, firms employ 600 workers and hold a capital stock of €45 million; most are located in the north of Italy. By sector, our data confirm the specialization of Italian manufacturing in industries with low technological content. Only 7% are classified as high tech according to the OECD system.[11] The last two columns give sample-weighted statistics, which makes the sample representative of the population of firms with 50 or more employees. The average size is substantially smaller, as the survey oversamples large firms. All the other characteristics are fairly similar to the unweighted data.

Since we do not have plant level data, all our analysis is at the firm level. From a theoretical point of view, it is unclear which unit would be the most appropriate; arguments can be found for both the firm and the plant level. However, as table 2 shows, between 2/3 (unweighted) and 4/5 of the firms are single plants,[12] suggesting that this is not likely to be

[10] See Cingano and Schivardi (2004) for a detailed account of the procedure.

[11] See OECD (2003) for the details on how the classification system is constructed.

[12] The information on the number of plants was collected in the INVIND survey only between 1992 and 1995. We completed the series for this variable by extending backward the oldest and forward the latest number of plants of each firm. This procedure is not likely to introduce substantial bias for single-plant firms. In fact, out of the 842 firms that report single-plant in at least one year between

Table 2
Firms' Characteristics

	Unweighted		Sampling Weights	
	Mean	SD	Mean	SD
Value added	26.6	118.1	8.1	38.0
Investment	5.2	30.3	1.5	9.7
Capital stock	46.4	242.0	13.8	77.0
No. of workers	625	3,024	198	949
Average workers' skills, $\bar{s}$	.46	.09	.45	.08
Average *P* workers' skills, $\bar{s}^P$	.40	.08	.40	.08
Average *NP* workers' skills, $\bar{s}^{NP}$	.61	.09	.60	.10
Variance of workers' skills, σ^2	.036	.018	.034	.018
Variance of *P* workers' skills, σ^{P2}	.017	.012	.017	.013
Variance of *NP* workers' skills, σ^{N2}	.054	.025	.055	.028
Between-status dispersion, $(\bar{s}^N - \bar{s}^P)^2$	.052	.044	.048	.045
Sectoral shares:				
Low tech	.38		.41	
Medium low	.25		.26	
Medium high	.30		.29	
High	.07		.04	
Geographical shares:				
Northwest	.44		.47	
Northeast	.25		.27	
Center	.20		.16	
South	.11		.10	
Share of single-plant firms	.63		.78	
No. of observations	9,790		9,790	

Note.—Value added, investment, and capital stock are in millions of 1995 euros. *P* stands for production and *NP* for nonproduction workers. Skills are measured by the individual worker effects obtained estimating eq. (1). See OECD (2003) for the classification of sectors according to technological content.

a major issue in our data. In any case, we also check our results restricting the analysis to just single plant firms.

B. Estimation of Worker Fixed Effects

According to Abowd et al. (1999) wages can be decomposed into a component due to time-variant observable individual characteristics, a pure worker effect, a pure firm effect, and a statistical residual, as follows:

$$w_{it} = X_{it}\beta + \theta_i + \psi_{J(i,t)} + \varepsilon_{it}, \quad (1)$$

where the subscript i denotes the worker, t denotes time, and $J(i, t)$ is the firm where worker i works at time t. The worker fixed effect, θ, is the component of wages due to the worker's pure ability, irrespective of the characteristics of the particular firm and net of the personal time-variant characteristics included in the matrix of controls X. Likewise the firm

1992 and 1995, only 59 report more than one in other years, and 40 of these report only two.

effect, ψ, is interpreted as the component of wages specific to the firm where the employee works and might respond to particular compensation policies such as efficiency wages or rent sharing.

Panel data allow us to identify firm and worker effects as long as there is enough mobility of workers across firms. Following Abowd et al. (1999) and Abowd, Creecy, and Kramarz (2002), we maintain the assumption of exogenous mobility conditional on the observables.[13] Ordinary least squares (OLS) estimation of the fixed effects requires the computation of the inverse of the matrix in (1), which has dimensionality equal to the number of workers plus the number of firms plus that of the other covariates: in our case, 2,100,000 by 2,100,000. The methodology initially used in the literature was based on approximative methods, consisting of a two-stage procedure to estimate worker effects first and then, from the resulting residuals, firm effects or vice versa (Abowd et al. 1999). We use the direct method proposed in Abowd et al. (2002), which simultaneously estimates worker and firm effects. The Abowd et al. procedure estimates the full model in (1) by fixed-effect methods using the standard conjugate gradient (CG) algorithm with preconditioning as described in Dongarra et al. (1991). The identification strategy consists of first determining the groups of connected workers and firms. A connected group comprises all the workers ever employed by any firm in the group and all the firms that any worker in the group has ever worked for. The connected groups set the restrictions that allow for the identification of worker and firm fixed effects. Once the groups are formed, we apply the algorithm to each group. Uniqueness of the solution further requires setting either one worker or one firm fixed effect equal to zero, so that the estimated effects can only be interpreted in relative terms.

The first step of the estimation procedure was the identification of connected groups.[14] Due to the sample design, based on all workers for medium-sized and large firms, our data set turns out to be one big connected group: only 0.5% of the observations are disconnected. For computational simplicity, we use only the largest connected group, which contains 421,019 firms, 1,674,684 workers, and 3,651,000 distinct firm-worker pairs (see descriptive statistics in col. 1 of table 1). The relatively great mobility of workers (about 70% have more than one employer during the period) allows the identification of firm and worker effects.

We estimated the wage decomposition of log weekly earnings into the three components of equation (1). The matrix of time-variant individual

[13] This assumption can be defended on the grounds that the conditioning set controls for both worker and firm fixed effects, in addition to other time-varying observables. Dismissing the exogeneity assumption would require setting up and solving a selection model, a computationally unfeasible problem.

[14] We adapted the code implementing the CG and grouping algorithms kindly provided by Francis Kramarz.

Table 3
Estimated Coefficients of the Time-Variant Personal Characteristics

	Coefficient	SD
Age	.02349	.0000261
Age squared	−.00024	.0000003
Dummy for production workers	−.52908	.0003147
Dummy for nonproduction workers	−.45683	.0003149
Seniority	−.00063	.0000120

Note.—The omitted category is "managers." Standard deviations calculated according to the approximative method described in Abowd et al. (2002).

characteristics, X, includes age, age squared, seniority,[15] and occupational category, which changes for a substantial number of workers.[16] The worker effect is fixed over time. It captures unchanged personal attributes, such as the unobservable worker's innate ability, and observable characteristics, such as formal education, under the reasonable assumption, for Italy, that workers do not go back to school once they enter the labor market. Unlike Abowd et al. (2003), we do not decompose the worker fixed effect into the potentially observable (such as education) and unobservable (such as innate ability or propensity to exert effort) components, because we do not observe education in our data set.[17] As noted, the firm effect reflects a firm's compensation above the average for workers of comparable characteristics and can be explained based on firm-specific compensation policies or firm-specific skills. Finally, the regression also controls for any trend or common time effect in wages by means of a full set of year dummies.

The estimated coefficients of the covariates are reported in table 3. We

[15] Our data on seniority is left censored as we do not have information on workers prior to 1981. To deal with this problem, we took the workers for whom we had information on their complete job durations, that is, the workers who started and left jobs within the sample period. We estimated a job-duration model based on all the available workers' characteristics—geographical area, age, and occupational status—and ran separated regressions for men and women. We then used the estimated coefficients to compute predicted job durations for all the workers in our first sample year, 1981. From their predicted job durations we could impute the seniority in any given year.

[16] The inclusion of this dummy variable does not remove the wage premia due to occupational status but only the changes in wages due strictly to the reclassification of occupational status in the course of the employee's working life.

[17] It would be interesting to distinguish between education and other unobserved workers' skills in order to determine the contribution of workers' education to skill dispersion within and across firms. However, such distinction is not feasible with the present data set.

Table 4
Correlation between Earnings, Predicted Values, Firm Effects, and Worker Effects

	SD	y	predxb	effpers	effirm
y	.332	1			
predxb	.091	.605	1		
		(.00)			
effpers	.220	.798	.428	1	
		(.00)	(.00)		
effirm	.120	.426	.131	.044	1
		(.00)	(.00)	(.00)	

NOTE.—y = log of weekly earnings in 1995 euros; predxb = predicted value of time-variant personal characteristics; effpers = worker effect; effirm = firm effect. Standard errors are in parentheses.

find the usual concave profile of earnings in age, and lower wages for clerks and production workers than managers. Contrary to expectations, seniority is negatively related to earnings, but the coefficient is extremely low, with an elasticity of −0.06%. This might be due to the measurement error embedded in our variable of seniority (see n. 15) and to the correlation with age. Abowd, Kramarz, and Roux (2006), who estimate firm-level seniority coefficients, also find small and often negative effects of seniority in French firms.

We use worker fixed effects as the proxy for workers' portable skills. This improves on other proxies in the literature in a number of respects. First, it is clean of firm and sector idiosyncrasies, such as the particular compensation policies of the firm or union dominance. Second, it is a comprehensive measure of skills that includes innate ability and informal skills. Finally, given that the worker fixed effects are calculated on the basis of workers' wages over time and across firms, they are orthogonal to time-specific and firm-productivity shocks, and they are suitable for comparison throughout the period analyzed. Since in the estimation of (1) we also control for seniority, our proxy of skills is net of seniority effects. One could argue that returns to seniority in a firm supposedly reflect learning on the job and should therefore be reflected in the measure of skills. However, we chose to exclude seniority because, as shown by Flabbi and Ichino (2001), wage changes related to seniority in the Italian system are more likely to reflect automatic upgrades due to typical contractual arrangements than to changes in skills.

Table 4 presents summary statistics and correlations between the different components of wages. As in Abowd et al. (1999), a significant part of the variation in earnings is due to heterogeneity in worker effects: the correlation between log earnings and worker effects is the highest, 0.8. Firm effects play less of a role, with a 0.43 correlation with earnings. The correlation between the worker and the firm effects is positive but very small (0.044), which is similar to Abowd et al. (1999) for the case of French manufacturing.

Descriptive statistics of the skill distribution at the level of the firm are included in table 2. On average, *NP* workers' skills are 50% higher than those of *P* workers, with a fairly low dispersion across firms (standard deviations are around .09 for both types of workers). The within-firm variance of skills is substantially higher for *NP* workers (.054 vs. .017).

One potential problem with our measure of skills is that they are affected by estimation error. This can induce some bias when constructing the firm-level indicators, particularly for dispersion, as they will reflect both true heterogeneity across workers and also estimation errors. Estimation error, though, is likely to be limited in our analysis. First, contrary to most previous papers, we observe all workers in a firm, so that our measures are not affected by sampling error. Second, we observe workers for a fairly long period of time (the mean number of years in the sample is 13.3 and the median is 15), which also helps mitigate the estimation error.

IV. Dispersion of Workers' Skills: A Decomposition Exercise

As discussed above, a body of theoretical work predicts the segregation of workers by skill across production units following certain technological changes and/or changes in the supply of skills. In this section, we undertake a decomposition exercise of skill dispersion similar to that performed for U.S. manufacturing (Davis and Haltiwanger 1991; Dunne et al. 2004) and for other countries (Kremer and Maskin 1996).

Our variance decomposition improves previous works in two directions. First, we use worker fixed effects, which are a better measure of skills than raw wages. Second, previous measures of within-firm dispersion have generally been based on a subsample of firms' workers. Instead, we observe the entire labor force of our firms and thus we can obtain the actual measure of firm skill dispersion.

The total dispersion of skills in the labor force can be decomposed into two components, the between-firm and the within-firm components:

$$V_B = \sum_{f=1}^{N} l_f \cdot (\bar{s}_f - \bar{s})^2 \tag{2}$$

and

$$V_W = \sum_{f=1}^{N} l_f \cdot \sigma_f^2, \tag{3}$$

where l_f denotes the weight in total employment of firm f, $\bar{s}$ is the overall average skill, while $\bar{s}_f$ and σ_f^2 are firm f's mean and variance of workers' skills.

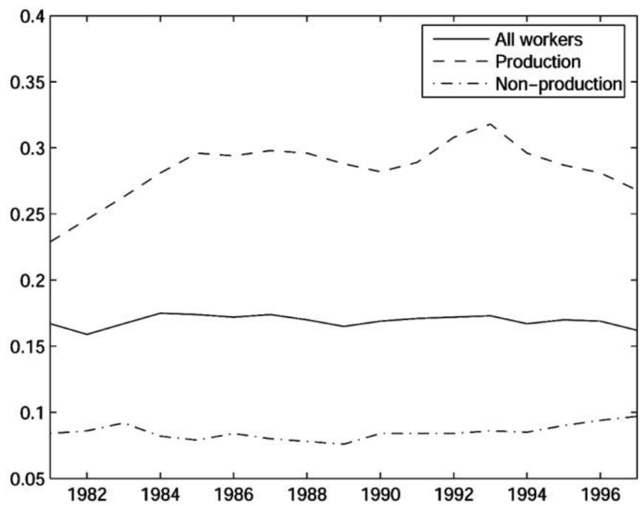

FIG. 1.—Segregation indexes—share of between-firm skill dispersion on total dispersion

Kremer and Maskin's (1996) index of segregation is the between-firm component of the variance of skills relative to the total variance, $V_B/(V_B + V_W)$. An increase in the relative importance of the between-firm component would constitute evidence of increased segregation of workers by skill at the firm level. Figure 1 reports the index for all workers and for *P* and *NP* workers separately.[18] As can be seen, most of the dispersion of skills takes place within and not between firms: less than 20% of the dispersion is accounted for by the between-firm component. There is an even more marked pattern for *NP* workers, for which the between-firm share is always below 10%.

In terms of time patterns, we find no evidence of an increase in segregation. Over the period 1981–97, the segregation index for all workers and for *NP* workers is basically flat. The index for *P* workers increases from less than 25% to around 30% between the early 1980s and the early 1990s, before declining to the values that had prevailed at the beginning of the period.

These findings are in contrast with those reproduced in Kremer and Maskin (1996), Dunne et al. (1997), and Dunne et al. (2004), who present some evidence of increasing segregation in the United Kingdom, France, and the United States. Moreover, the level of our segregation indexes is much lower than those for American manufacturing. Such indexes are based on the dispersion of wages across plants, while we use the estimated

[18] For comparability, we include both clerks and managers among *NP* workers. Our results do not change if we exclude managers.

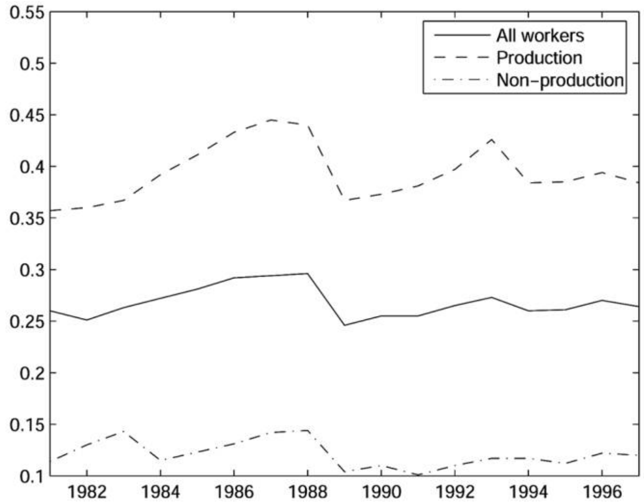

FIG. 2.—Segregation indexes—share of between-firm wage dispersion on total dispersion

worker effects as the measure of skills. Thus, in order to make our results comparable, we recalculated the segregation indexes using wages (the log of weekly earnings). Results are reported in figure 2. As expected, since wages also include the firm effect common to all workers, the between-firm component is now larger: it accounts for between 25% and 30% of the total dispersion, approximately 10 percentage points more than in the case of the dispersion of worker effects. Also with this index we find no pattern of increasing segregation over time. A moderate increase in overall segregation during the 1980s is followed by an equally moderate decline toward the end of the decade and a flat pattern thereafter, particularly after 1993.

One possible explanation of the lower degree of segregation that we find in Italy compared to the United States is labor market regulation. As argued by Dell'Aringa and Pagani (2007), the Italian nationally centralized collective bargaining system might limit cross-firm wage dispersion, as firms would tend to stick to the centrally bargained wage levels. This would be consistent with the stability of the segregation index apparent in figure 2 after 1993, when a wage protocol was signed, that induced wage moderation and strengthened the degree of centralization of the wage-setting process. The Italian segregation index based on earnings is closer in value to those reported in Kramarz et al. (1996) for France—0.36 in 1986 and 0.44 in 1992—where wage-setting mechanisms are similar to those in Italy. Unfortunately, comparability with the French indexes is not complete, as the French sample also includes service workers and is based on firms with as few as 10 employees. Moreover, the French

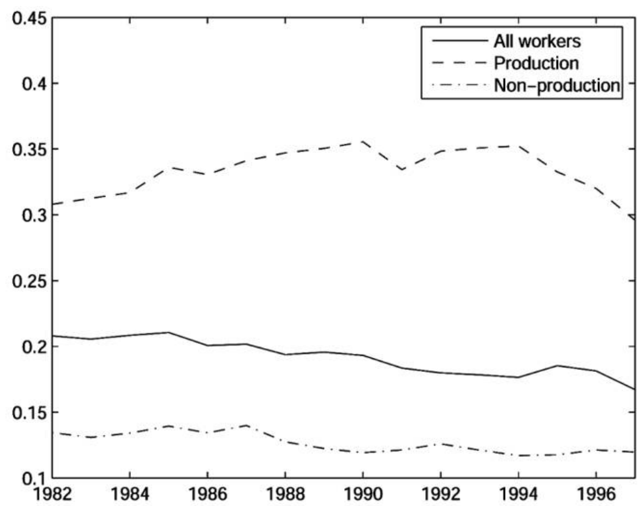

FIG. 3.—Segregation index—share of between-firm skill dispersion on total dispersion, single plant firms.

indexes are based on an average of just 30 workers per firm in 1986 and 11 in 1992. By contrast, our analysis relies heavily on larger firms, with 50 employees or more, among which the within-firm dispersion of skills tends to be more important than the between-firm component.

A further difference with respect to the U.S. studies is that we consider segregation at the firm level rather than the plant or establishment.[19] Although we do not have plant-level data, it is possible with our data set to identify the single-plant firms. Figure 3 reports the segregation indexes based on single-plant firms only. The time patterns are basically the same as those for the whole sample, indicating that the evolution of skill dispersion is similar across firms and plants. The only difference is that the level of segregation is on average 5 percentage points higher for single-plant firms, which implies that in multiplant firms the distribution of skills is slightly smoothed out across establishments.

Considering manufacturing as a whole could mask important differences across sectors due, say, to technological differences. We recalculated the segregation indexes for four sectoral groups according to the OECD technological classification (OECD 2003). Figure 4 shows that even at this lower level of aggregation there is no evidence of an increase in segregation. If anything, it has decreased substantially in high-tech industries and somewhat in low-tech industries as well; only in medium-high-technological industries has it increased moderately (from around

[19] The French study by Kramarz et al. (1996) is also conducted at the firm level.

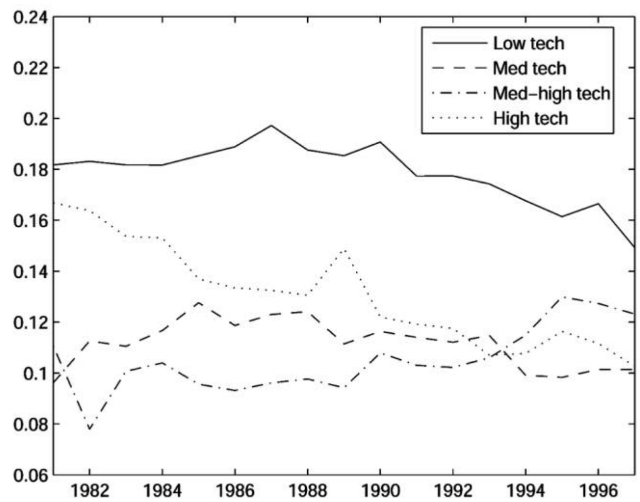

FIG. 4.—Segregation indexes—share of between-firm skill dispersion on total dispersion, by technological intensity.

10% to 13%). In terms of levels, the less technologically intensive sectors display the highest segregation indexes. This is not surprising if we take into account that it is easier for low-tech activities to physically separate the different phases of production and outsource the simplest tasks while more sophisticated industries require greater integration between the design (and other headquarter activities) and the production phase, making skill segregation less viable.[20] All in all, we find that, unlike other countries, Italy shows no tendency toward increasing segregation of workers by skill.

V. Within-Firm Skill Dispersion and Productivity

A. Workers' Skills in the Production Function

Why do we not observe a phenomenon of segregation by skill between firms in the Italian case? One possible explanation is that the structural changes that should lead to skill segregation did not take place. The first element we consider is overall skill dispersion. Kremer and Maskin (1996) show that an increase in dispersion might lead to greater segregation between production units. Figure 5 plots total skill dispersion for all workers in our data set and for *P* and *NP* workers separately. In all three

[20] The process of delocalization of the production phase to countries with cheaper labor, while keeping in house the activities with a higher value-added content, has indeed been prominent in Italian manufacturing in sectors such as textiles and apparel.

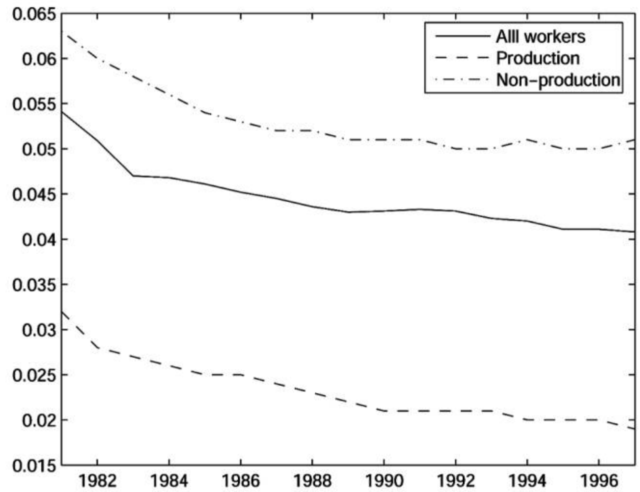

FIG. 5.—Overall skill dispersion (variance of worker effects)

cases we observe a moderate decline in overall dispersion, arguably due to the increase in educational attainments, mainly the steady increase in the share of college graduates. This finding indicates that the first potential change that might increase segregation is absent over the period of analysis.

We now consider the role of the skill mix in determining firms' productivity. In particular, we investigate whether, given a certain average skill level, skill dispersion within the firm increases or decreases productivity. From a theoretical point of view, the answer is fairly straightforward and rests on the parameters of the production function that govern the substitutability/complementarity of skills. As explained above, there are certain activities for which having workers with similar skills is preferable. This is the case of Kremer's (1993) O-ring production function, a process consisting of different tasks in which each task must be performed at a given level of competence for the project to attain full value. By contrast, there are activities where workers' skills are substitutable and output disproportionately reflects the contribution of a few very talented people. Activities such as research and innovation or design, where the achievement of a certain common goal is more important than the partial contributions of every individual in the team, are examples of this type of production processes. Another example is a production process involving tasks of different importance, such as complex tasks of coordination and supervision, together with more straightforward ones requiring less skill. In such activities the marginal product of a talented worker is greater when matched with less talented ones and thus, for a given average skill level, productivity is higher the more dispersed skills are.[21]

Although this points to an optimal skill mix for all firms using the same technology, in practice we observe significant variation in within-firm skill dispersion, in much the same way as there is significant variation between production units in firm characteristics such as capital intensity, size, and innovative activity, and in firm outcomes.[22] There are various reasons for such a variation in skill composition across firms. First, the labor force in Italy is hardly mobile geographically, so the local skill composition affects the availability and the relative price of skills. As a consequence, we expect similar firms in different locations to employ different skill mixes. Second, labor regulations and other adjustment costs, particularly in the firm organizational structure,[23] might prevent firms from fine tuning their skill composition, again resulting in variation in the skill distribution across firms. Finally, different managers might have different opinions on the optimal way to organize production and choose their skill mixes accordingly. It is precisely the between-firm variation in skill mix and productivity that allows us to identify any relation between the two.

The production function analysis addresses two issues. First, we study the effects of skill dispersion on productivity. This question has not been dealt with in previous empirical work on firm productivity, due to the lack of data.[24] Second, we investigate whether the role of skill dispersion has changed over time: according to the theories surveyed above, ICT and other innovations in the organization of production may have changed the way workers are mixed in the production process.

To formally investigate the relation between skill distribution and pro-

[21] These ideas are formalized in Milgrom and Roberts (1990) with the concepts of supermodularity and submodularity.

[22] In fact, in unreported regressions, we found that a significant fraction of skill dispersion remains unexplained when conditioning on sector and observable firm characteristics. Haltiwanger, Lane, and Spletzer (2007) also observe significant firm heterogeneity along other dimensions of the workforce composition, which tends to decrease as firms age, suggesting an adjustment process based on learning and exit of mistaken firms, toward some "optimal" worker mix.

[23] Ichniowski, Shaw, and Gant (2003) study human resource management policies in the steel-finishing lines. They find that organizational modes with different efficiency levels coexist in different plants even within a highly homogeneous production process. According to their evidence, based on field visits, the main reason for which less efficient modes remain in production is that the "work culture" of the plant is very persistent and can be a major obstacle to the adoption of the best practices.

[24] Iranzo (2003) investigates this relation at the city level, finding that a more dispersed skill structure is beneficial for productivity.

ductivity, we use the following generalized Cobb-Douglas production function in capital and labor:

$$y_{ft} = A_{ft} \cdot K_{ft}^{\alpha} \cdot [L_{ft} \cdot E(s_1, \ldots s_{L_{ft}})]^{\beta}, \tag{4}$$

where subscripts f and t denote firm and time, respectively; A is a Hicks-neutral technological factor; and K and L are capital and number of workers, respectively. The term $E(s_1, \ldots, s_L)$ represents the overall efficiency of the labor force and depends on workers' skill levels s_i, $i = 1 \ldots L_{ft}$, and the way they are combined in different firms. As P and NP workers differ not only in their average skill level but also in the type of tasks they perform, we consider these two types of labor as distinct inputs. More precisely, we treat the overall efficiency of the firm's labor force as a CES function of the efficiency of P and NP workers:

$$E(\mathbf{s}^P, \mathbf{s}^{NP}) = [l^P \times (E^P(\mathbf{s}^P))^{\gamma} + l^{NP} \times (E^{NP}(\mathbf{s}^{NP}))^{\gamma}]^{1/\gamma}, \tag{5}$$

where $\mathbf{s}^j$ is the vector of skills of workers in occupational status $j = P, NP$ and $l^j = L^j/L$ is the share of workers of status j in the firm's total labor force. The elasticity of substitution between P and NP workers is given by $1/(1-\gamma)$. The term $E^j(\mathbf{s}^j)$ is, in turn, a CES function of workers' skills within status j:

$$E^j(\mathbf{s}^j) = \left(\frac{1}{L^j}\sum_{i=1}^{L^j} s_i^{\rho^j}\right)^{1/\rho^j}, \tag{6}$$

with the elasticity of substitution of skills for workers of status j given by $1/(1-\rho^j)$. In other words, the parameters γ, ρ^P, and ρ^{NP} govern the substitutability of skills. If $\gamma < 1$, the elasticity of substitution between P and NP workers is positive, implying complementarity (or imperfect substitutability) between the two types of labor. A parameter of $\gamma > 1$ would imply that P and NP workers are substitutable, in which case the isoquants are concave and then only one type of worker would be employed with the relative wages determining which one is to be used. This case is highly improbable. At least in the short run, the possibilities of substitution between the two types of worker are rather limited, because P basically cannot do NP workers' jobs and vice versa. In effect, all the available estimates on this elasticity of substitution suggest that P and NP workers are imperfect substitutes. Similarly, ρ^j indicates whether skills within each occupational status are complementary ($\rho^j < 1$) or substitutes ($\rho^j > 1$). As the above discussion on production processes has illustrated, within each occupational status substitutability of skills is less implausible.

The importance of the dispersion of workers' skills for total output can be seen more clearly by rewriting expression (5) as a function of the first and second moments of the skill distribution. Using a second-order

Taylor series expansion, the expression for the overall efficiency of labor can be approximated around the mean skill level as follows:

$$E(s) = \bar{s} + \frac{1}{2}(\gamma - 1) l^P l^{NP} \frac{(\bar{s}^P - \bar{s}^{NP})^2}{\bar{s}} + \frac{1}{2}(\rho^P - 1) l^P \frac{\sigma^{P2}}{\bar{s}} + \frac{1}{2}(\rho^{NP} - 1) l^{NP} \frac{\sigma^{NP2}}{\bar{s}}. \tag{7}$$

The first term in (7) is the overall skill mean, $\bar{s}$; the second term contains the between-occupational-status component, $(\bar{s}^P - \bar{s}^{NP})^2/\bar{s}$, weighted by the product of the shares of *P* and *NP* workers; while the third and fourth terms are the within-firm dispersion of *P* and *NP* workers' skills, respectively, divided by the overall skill mean and weighted by their shares in the total labor force. Using (7) and taking logs in the production function in (4), we obtain:

$$\ln y_{ft} = a_{ft} + \alpha \ln K_{ft} + \beta \ln L_{ft} + \beta \ln \left[\bar{s}_{ft} + \frac{1}{2}(\gamma - 1) l^P_{ft} l^{NP}_{ft} \frac{(\bar{s}^P_{ft} - \bar{s}^{NP}_{ft})^2}{\bar{s}_{ft}} + \frac{1}{2}(\rho^P - 1) l^P_{ft} \frac{\sigma^{P2}_{ft}}{\bar{s}_{ft}} + \frac{1}{2}(\rho^{NP} - 1) l^{NP}_{ft} \frac{\sigma^{NP2}_{ft}}{\bar{s}_{ft}} \right]. \tag{8}$$

Equation (8) differs from a standard production function because of the term in brackets, which includes the first and second moments of the skill distribution. If there are no interaction effects derived from the combination of workers' skills, then the dispersion of skills does not matter for productivity: γ, ρ^{NP}, and ρ^P are equal to one, and only the average skill has an impact on productivity. In terms of the elasticity of substitution discussed above, this means that workers' skills are perfect substitutes for each other. By contrast, if γ, ρ^{NP}, and ρ^P differ from one, the way skills are combined also matters for productivity. In particular, if a parameter is larger than one, the dispersion of the related skills increases productivity, as it is optimal to combine workers of different skill levels, whereas if it is smaller than one, dispersion has a negative effect on productivity.

Note that if we disregard the distinction between *P* and *NP* workers and assume that workers differ only in their skill level, equation (8) simplifies to

$$\ln y_{ft} = a_{ft} + \alpha \ln K_{ft} + \beta \ln L_{ft} + \beta \ln \left[\bar{s}_{ft} + \frac{1}{2}(\rho - 1) \frac{\sigma^2_{ft}}{\bar{s}_{ft}} \right], \tag{9}$$

where the parameter ρ now governs the degree of substitutability among all workers' skills. We will also estimate equation (9) as a benchmark for our main results.

B. Estimation Results

Before going into the estimation of (8) and (9), we report some preliminary evidence on the relation between labor productivity and the distribution of skills (fig. 6). Labor productivity is constructed as output per worker, net of time and sectoral effects.[25] We compute the average of the variables for each decile of the productivity distribution and plot the indexes of each variable with respect to the values of the first decile. The first panel reports the relation between productivity and average skills. As expected, more productive firms have workers with higher skills on average. For example, firms in the last decile of the productivity distribution have an average worker effect 25% greater than those in the first decile. When the two groups are considered separately, the relation is very similar for *P* workers and less strong for *NP* workers.[26]

The second panel of figure 6 plots firm dispersion of skills and its decomposition in the within- and between-occupational status components. More productive firms tend to have a more heterogeneous labor force: firms in the last decile show a skill dispersion almost 35% greater than those in the first decile. In terms of occupational status, there is a clear contrast between the within-status components—positively related to productivity—and the between-status component, which decreases with firm productivity. This preliminary evidence thus suggests that skills are complements between status groups and substitutes within them.

For a preliminary gauge of any change over time in the relation between skill mix and productivity, figure 7 replicates the lower panel of the previous figure, splitting the sample into two subperiods: 1982–90 and post-1990. The two graphs show very similar trends for the total and the single components, indicating that there was no significant change between the 1980s and the 1990s.

Before turning to the econometric results, there is an issue regarding our estimation exercise that deserves some discussion. As equation (8) makes clear, individual wages are not independent from the particular combination (match) of workers' skills within the firm: certain combinations might result in higher productivity, generating a rent that the firm and the workers share. More specifically, skill complementarity implies

[25] We have regressed output per worker on a set of year and two-digit sectoral dummies and used the residuals as a measure of productivity.

[26] Note that the index of overall average skills does not need to lie between the two components. In fact, if both the average skill of *P* and *NP* workers increases and so does the share of *NP* workers, then the overall skill level will increase by more than the other two.

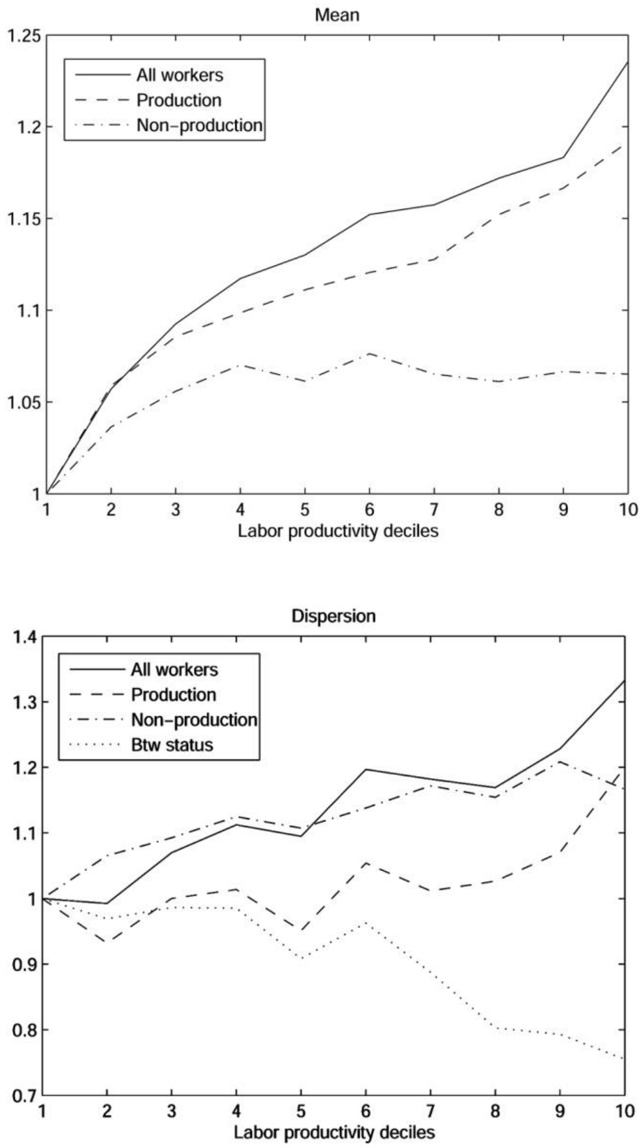

Fig. 6.—Mean and variance of firm-level skill distribution by deciles of productivity

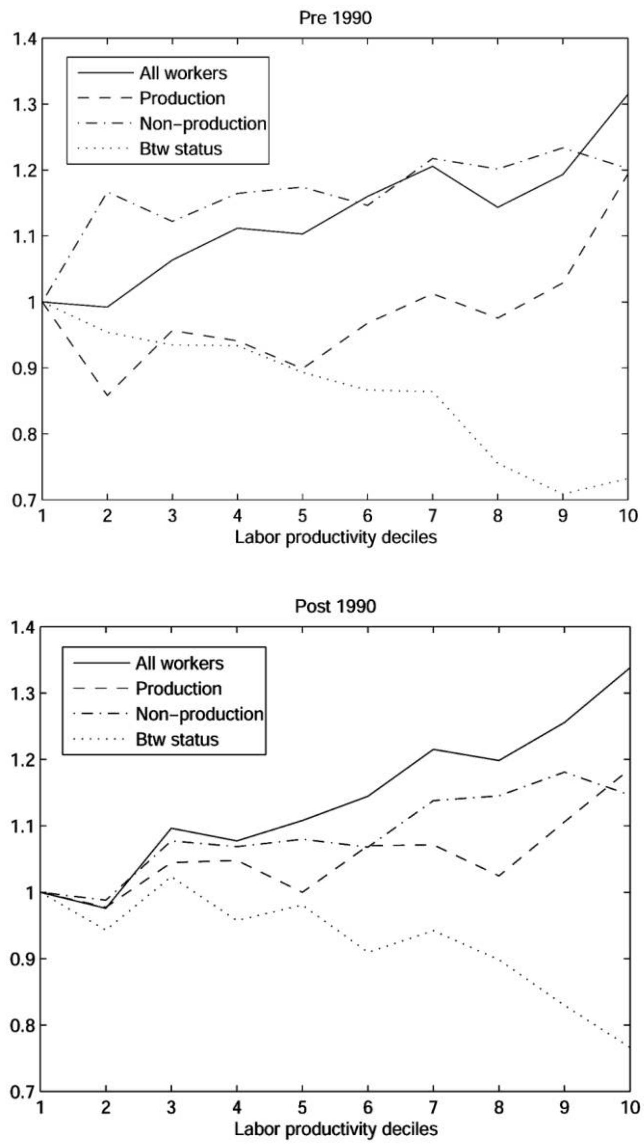

FIG. 7.—Variance of firm-level skill distribution by deciles of productivity: pre- and post-1990.

that the productivity of each worker (and thus her wages) depends on the skills of other workers in the firm as well. Therefore, in principle, one should consider estimating the production function and the wage equation jointly. However, that would require a theoretical model of job matching between workers and firms and also a model of rent sharing and thus negotiation of wages assuming different degrees of bargaining power between the worker and the firm. Such a theoretical model is beyond the scope of this study. Instead, we estimate the individual worker fixed effects, our proxy for skills, from the simple additive equation in (1).[27]

The main econometric problem in estimating equation (8) is that inputs are a choice variable and thus are likely to be correlated with unobservables, particularly the productivity shock a_{ft}. This is the classical problem of endogeneity in the estimation of production functions. To deal with it we follow the procedure proposed by Olley and Pakes (1996).[28] Using a standard dynamic programming approach, Olley and Pakes show that the unobservable productivity shock can be approximated by a nonparametric function of the investment and the capital stock, $a_{ft} = h(i_{ft}, k_{ft})$. We therefore include in the regression a third-degree polynomial series in i and k and their interactions, which should approximate the unobserved productivity shock a_{ft} and take care of the endogeneity issue.[29] All the regressions include year dummies as well as 13 sectoral and four geographical area dummies, and observations are weighted using the sampling weights. In order to account for the problem of generated

[27] The exclusion from eq. (1) of possible rents deriving from the quality of the skill mix might bias the estimates of the workers' fixed effect. This potential problem, though, is lessened by the fact that the equation controls for firms' fixed effects. In fact, as long as firms' skill composition is stable over time, any interaction effects of workers' skills on wages will be captured by the firm effect. In some unreported regressions we have verified that most of the variation of the skill composition is cross sectional and that the firm skill composition tends to be persistent over time. Firm fixed effects alone explain almost 90% of the variation in the within-firm skill dispersion. Similarly, the first-order autocorrelation coefficient of within-firm dispersion is .95, which implies a substantial degree of persistence.

[28] Ideally, one would like to have instruments for the skill distribution. Unfortunately, it is very hard to come up with variables that are correlated with skills at the firm level while orthogonal to productivity shocks. For example, educational attainment in the local labor force will shift the skill distribution but is also likely to be correlated with the unobserved component of productivity via human capital externalities.

[29] Note that when the nonparametric term in capital and investment is included, the capital coefficient can no longer be interpreted as the parameter of the production function in the first stage of the procedure. However, given that the coefficient on capital is of no particular interest to us, this is inconsequential for our purposes.

Table 5
Nonlinear Least Squares: Overall Skill Dispersion

Variable	Parameter	Sample: Whole (1)	Whole (2)	Whole (3)	Pre-1990 (4)	Post-1990 (5)
Overall dispersion	ρ	1.75	1.80	1.49	1.87	1.81
		(.391)	(.414)	(.64)	(.544)	(.398)
		[1.00]	[1.00]	[1.00]	[.999]	[1.00]
Labor	β	.766	.722	.742	.711	.731
		(.019)	(.019)	(.007)	(.0240)	(.021)
Capital	α	.226		.257		
		(.015)		(.006)		
Estimation method		NLS	OP	ACF	OP	OP
R^2		.87	.88		.89	.87
No. of observations		9,790	9,790	8,575	4,180	5,610

Note.—Dependent variable: log value added. Results from estimating eq. (9) with nonlinear least squares. Estimation method: NLS is nonlinear least squares, OP is Olley and Pakes (1996), ACF is Ackerberg et al. (2006). Bootstrapped standard errors based on 1,000 replications in parentheses; probability that the parameter is greater than one (computed as the frequency of bootstrapped estimates greater than one) in brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include year, two-digit sector dummies, and four macro-region dummies.

regressors in the nonlinear estimation procedure, which makes the computation of standard errors problematic, we base our inference on blocked bootstrap (i.e., sampling full firm histories rather than single observations) with 1,000 replications.

We start by estimating equation (9) where we consider the overall impact of dispersion on productivity, without distinguishing between occupational status. Results of the nonlinear least squares estimation are reported in table 5. For comparability, in the first column we report the results of simple nonlinear least squares (NLS).[30] Given that the parameters of interest are scarcely affected, throughout we only comment on those with the Olley-Pakes procedure. The results are in line with the evidence of figure 6. We obtain estimates for ρ of 1.8 (col. 2); in no bootstrap repetition did we obtain a value below one, which is the threshold that determines substitutability (if $\rho > 1$) or complementarity of skills (if $\rho < 1$). This implies that overall skills are substitutes and that within-firm dispersion is positively correlated with productivity.

The Olley-Pakes procedure assumes that labor is a perfectly variable input. This might be a strong assumption; due to adjustment costs and other frictions, labor might not be perfectly variable. Ackerberg et al. (2006) propose an extension of the Olley-Pakes methodology that allows

[30] The Olley-Pakes procedure requires excluding observations with zero investment. In our data, these represent less than 6% of the total, because we only have medium and large firms. To maximize comparability between Olley-Pakes and NLS, we exclude such observations from the latter, so that the estimates are computed on exactly the same sample.

for the estimation of the production function coefficients when all inputs are not perfectly variable. The procedure entails obtaining an estimate of the innovation to productivity, ξ_{ft}, where $\xi_{ft} \equiv a_{ft} - E[a_{ft}|I_{t-1}]$ and I_t is the information set at t, and then exploit the fact that such innovation is by construction uncorrelated with all variables decided up to $t-1$. As before, we maintain the assumption that capital is decided one period in advance, so that k_{ft} is unrelated to ξ_{ft}. As far as labor is concerned, its lagged value belongs to the information set at $t-1$ and therefore is uncorrelated to ξ_{ft}. We can apply the same reasoning to the skill distribution terms and use their lagged values to form a set of four moment conditions:

$$E(\xi_{ft} \cdot \mathbf{Z}'_{\mathbf{t}}) = 0, \tag{10}$$

where $\mathbf{Z}_{\mathbf{t}} = [\ln K_{ft}\ \ \ln L_{ft-1}\ \ \bar{s}_{ft-1}\ \ (\sigma^2_{ft-1}/\bar{s}_{ft-1})]$. To keep the number of regressors manageable, we use a second-degree polynomial (with interactions) in the first stage and a third-degree polynomial in a_{ft-1} to compute $E[a_{ft}|I_{t-1}]$. We construct our GMM estimator using the identity matrix as weighting matrix. As before, we block-bootstrap standard errors.[31] The results, reported in column 3 of table 5, confirm those obtained by the other estimation method. The coefficient on overall dispersion drops from 1.8 to 1.5 but is still well above 1; the capital and labor coefficients are hardly affected. This makes us confident that our estimation results are robust with respect to the estimation method.

To check for any structural break in the coefficients, we split the sample into two time periods, pre-1990 and post-1990, and run separate regressions with the Olley-Pakes procedure. We obtain very similar estimates for ρ (1.87 and 1.81, respectively), which indicates that there was no structural break in the relation between overall skill dispersion and productivity. In particular, there is no evidence of a decrease in ρ, which would make dispersion more detrimental to productivity and would therefore foster segregation.

Table 6 reports the estimation results of equation (8), where we distinguish between the two occupational statuses. The estimate of the parameter γ, governing the elasticity of substitution between P and NP, is -0.12 with a large standard error (.9). Despite the imprecise estimation, the estimates are below one in almost 90% of the bootstrap replications. This suggests that P and NP workers' skills are imperfect substitutes. According to this point estimate, the elasticity of substitution between P and NP workers is .89. This value is lower than the preferred estimates for skilled

[31] We have experimented with the optimal weighting matrix, obtaining similar point estimates but rather larger standard errors.

Table 6
Nonlinear Least Squares: Skill Dispersion Within and Between Occupational Status Groups

		Sample				
Variable	Parameter	Whole (1)	Whole (2)	Whole (3)	Pre-1990 (4)	Post-1990 (5)
Between dispersion	γ	−.590	−.118	−.242	−.002	−.186
		(.791)	(.900)	(.585)	(1.178)	(.886)
		[.031]	[.112]	[.037]	[.231]	[.106]
Production dispersion	ρ^P	1.55	1.42	1.56	1.49	1.36
		(.268)	(.276)	(.441)	(.388)	(.380)
		[.988]	[.931]	[.995]	[.961]	[.753]
Nonproduction dispersion	ρ^{NP}	5.10	5.55	5.11	5.84	5.51
		(.691)	(.786)	(.615)	(1.376)	(842)
		[1.00]	[1.00]	[.997]	[1.00]	[1.00]
Labor	β	.766	.718	.741	.710	.723
		(.019)	(.018)	(.008)	(.023)	(.020)
Capital	α	.225		.260		
		(.015)		(.007)		
Estimation method		NLS	OP	ACF	OP	OP
R^2		.88	.88		.90	.87
No. of observations		9,790	9,790	8,575	4,180	5,610

Note.—Dependent variable: log value added. Results from estimating eq. (8) with nonlinear least squares. Estimation method: NLS is nonlinear least squares, OP is Olley and Pakes (1996), ACF is Ackerberg et al. (2006). Bootstrapped standard errors based on 1,000 replications in parentheses; probability that the parameter is greater than one (computed as the frequency of bootstrapped estimates greater than one) in brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include year, two-digit sector dummies, and four macro-region dummies.

and unskilled labor in the literature (between 1.3 and 1.6)[32] and closer to the elasticity of substitution between *P* and *NP* workers estimated by Manasse and Stanca (2003) to be around 0.49–0.67 for Italian manufacturing in the 1990s.

In terms of within-status dispersion, we find that both ρ^{NP} and ρ^P are greater than one, implying within-status skill substitutability, and in both cases we accept at reasonable significance levels the null hypothesis that the coefficient is greater than one. The value of this coefficient is substantially larger for *NP* than for *P* workers (5.5 vs. 1.4). This might be due to the wider range of complexity of the tasks that *NP* workers perform, which increases the returns from having some high ability individuals in the team and leads to "skill-stratification" (Garicano and Rossi-Hansberg 2006) among *NP* workers. Taken together, these results indicate that it is optimal to have a dispersed skill composition within each occupational status group, particularly for *NP* workers, while matching the average skill levels of *P* and *NP* workers.

As before, we check for the robustness of the results with respect to

[32] See, e.g., Katz and Murphy (1992), Johnson (1997), Krusell et al. (2000), Ciccone and Peri (2005), and Caselli and Coleman (2006).

the estimation procedure applying the Ackerberg et al. methodology. We construct the six moment conditions:[33]

$$E(\xi_{ft} \cdot \mathbf{Z}'_{\mathbf{t}}) = 0, \tag{11}$$

where $\mathbf{Z}_{\mathbf{t}} = [\ln K_{ft} \quad \ln L_{ft-1} \quad \bar{s}_{ft-1} \quad [(\bar{s}^{NP}_{ft-1} - \bar{s}^{P}_{ft-1})^2/\bar{s}_{ft-1}] \quad (\sigma^{P^2}_{ft-1}/\bar{s}_{ft-1})$ $(\sigma^{NP^2}_{ft-1}/\bar{s}_{ft-1})]$. Again we find that the results are very similar to those obtained applying the Olley-Pakes methodology (table 6, col. 3). In particular, the estimate of γ becomes more precise and, although the standard error is still rather large, we can accept with a good degree of confidence the hypothesis that P and NP workers are imperfect substitutes.[34]

Finally, in the last two columns of the table we split the sample before and after 1990. Again, we find that the estimates of all coefficients are extremely similar over the two periods, a further confirmation of the absence of a structural change in the way the skill distribution enters the production function.

We performed a number of other robustness checks. First, we ran some reduced-form regressions in which dispersion entered linearly, and we used alternative measures for skill dispersion such as the ratio of the 90th to the 10th percentile of the skill distribution. The results broadly confirm those of the structural regressions. In particular, we obtain elasticities of value added to dispersion within each occupational status in the range of 2%–5% and of the between-status component around −1%. Second, we reran the regressions for single-plant firms alone and also excluding managers from the NP group as their inclusion could overplay the role of dispersion. In both cases we find no qualitative change in the results. All in all, the results of a positive impact of within-status dispersion on productivity, especially for NP workers, and a negative impact of between-status dispersion proved to be fairly robust.

C. Extensions

An important critique to the previous regressions is that they impose the same coefficients for all firms. One possibility would be to run our estimates by sectors. Unfortunately, sample size is greatly reduced when considering single two-digit sectors, which makes the estimation problematic. An alternative is to group firms according to some other char-

[33] An additional moment condition could be obtained using the lagged value of the share of P workers. We have experimented by adding this additional condition, but the results become less precise.

[34] Woolridge (2005) proposes a further refinement of the Olley-Pakes procedure, consisting in estimating contemporaneously the first and second step in a GMM framework. In our case, this turned out to be unfeasible, as it implies running a GMM estimation with a very large set of parameters (the basic parameters and those of the first step procedure, which includes the interaction between all the regressors to the second degree and the year, sector, and area dummies).

Table 7
Nonlinear Least Squares: Sample Splits

Variable	Parameter	Size: Medium (1)	Size: Large (2)	ICT Intensity: ICT (3)	ICT Intensity: NO ICT (4)	Tech Level: Medium Low (5)	Tech Level: Medium High (6)
Between dispersion	γ	.587	−2.16	.406	−.77	−.223	.180
		(.887)	(1.29)	(1.192)	(1.190)	(1.13)	(1.63)
		[.323]	[.039]	[.295]	[.108]	[.184]	[.277]
Production dispersion	ρ^P	1.39	1.59	1.53	1.52	1.46	.854
		(.307)	(1.009)	(.162)	(.314)	(.280)	(1.22)
		[.859]	[.940]	[.759]	[.897]	[.943]	[.491]
Nonproduction dispersion	ρ^{NP}	5.30	7.91	5.25	5.52	6.01	3.68
		(.928)	(1.295)	(1.25)	(1.042)	(1.00)	(1.353)
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[.989]
Labor	β	.702	.715	.771	.693	.706	.746
		(.026)	(.025)	(.031)	(.021)	(.022)	(.029)
R^2		.74	.86	.89	.88	.87	.90
No. of observations		5,424	4,366	3,345	6,199	6,158	3,632

Note.—Dependent variable: log value added. Results from estimating eq. (8) using the Olley-Pakes procedure. "Medium" are firms with 50–250 employees, "large" with more than 250 employees; ICT intensity follows the classification of O'Mahony and van Ark (2003) and technological level that of OECD (2003). Bootstrapped standard errors based on 1,000 replications in parentheses; probability that the parameter is greater than one (computed as the frequency of bootstrapped estimates greater than one) in brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include year, two-digit sector dummies, and four macro-region dummies.

acteristics. We examine the relationship between skill dispersion and productivity along three dimensions: firm size, information and communication technologies (ICT) intensity, and technological level.

The role of within-firm skill dispersion on productivity might be different depending on firm size. There is ample evidence that within-firm wage dispersion rises with firm size (Davis and Haltiwanger 1996; Lallemand and Rycx 2006) due to higher heterogeneity in the workforce of the large firms, particularly for *NP* workers. Table 7 reports the results of estimating equation (8) distinguishing between medium-sized firms—less than 250 employees, which is approximately the median size of our sample—and large firms—with more than 250 employees (we do not observe firms with fewer than 50 employees). We find that *P* and *NP* workers are imperfect substitutes in large firms: in 94% of the bootstrapped replications the coefficient on the between-occupational-status dispersion was less than one. That is, in such organizations productivity increases when the average skills of *P* and *NP* workers are similar. By contrast, for smaller firms, although less than one, the coefficient is very imprecisely estimated with only a little less than 68% of the bootstrapped estimates being less than one. At the same time, dispersion of skills within the group of *NP* workers is associated with higher productivity in both groups, more so in large firms. This is in line with the existence of what Garicano and Rossi-Hansberg (2006) call "knowledge hierarchies" among *NP* workers, which seems to be more prominent among large firms.

Another interesting dimension of firms to be considered is ICT intensity. As argued above, there is some controversy in the literature on the relationship between ICT use and the hierarchical structure of the firm. According to the literature on ICT adoption and workplace transformation (Caroli and Van Reenen 2001; Bresnahan, Brynjolfsson, and Hitt 2002), ICT intensive firms should have a flatter, more decentralized structure, together with a more uniform skill distribution. On the contrary, Garicano and Rossi-Hansberg (2006) suggest that to the extent that ICT reduces the costs of communication within the firm, it might also lead to more centralized structures. We use the taxonomy proposed by O'Mahony and van Ark (2003) that classifies sectors according to the intensity of usage of ICT. More precisely we break the sample into two groups: ICT-using and non–ICT-using sectors.[35] The estimation results, shown in columns 3 and 4 of table 7, do not allow us to discriminate between the alternative theories. In fact, we find that between-status dispersion is more detrimental to productivity for non-ICT firms, possibly also because such firms are on average larger in our sample (648 vs. 548 employees), even if the large standard errors prevent clear inference. The within-status coefficients are very similar. In general, therefore, the two groups of firms look fairly similar in terms of the relation between skill distribution and productivity. One possible explanation is that the process of ICT adoption was only beginning toward the end of our study period (Fabiani, Schivardi, and Trento 2005), so that its effects (if any) cannot be detected in our data. This is also in line with the stability of the production function coefficient found for the pre- and post-1990 period.

Finally, we have also checked for differences according to the technological level (OECD 2003), distinguishing between low- and medium-low tech on one side and medium-high and high tech on the other.[36] In this case, we find that between status has a similar impact for the two groups, while the within-status coefficients are smaller for the high-tech firms, particularly for *NP* workers. Consistently with the O-ring theory of production, this might be due to the fact that in such firms more *NP* workers perform nonroutine tasks, such as R&D, product development, and so forth, so that a less dispersed skill composition is preferable.

[35] We leave out the category ICT-producing sectors, as it amounts to less than 300 observations, too few to obtain stable estimates.

[36] Unfortunately, the relatively low specialization of Italian firms in high-tech sectors is reflected in the small number of observations in the latter group, so that such results should be interpreted with care.

D. Discussion

The regression results are fairly clear cut, and they agree with the graphical evidence in figure 6. By and large, they suggest that for a given average skill level, a firm's productivity is higher the more dispersed the skill distribution of its labor force. When distinguishing between occupational status groups, productivity is associated with a disperse distribution within status. The effect is very strong for *NP* workers ($\rho^{NP} = 5.5$ in our preferred estimate) and less so for *P* workers, for which deviation from perfect substitutability ($\rho = 1$) is not too large: in fact, we estimate $\rho^{P} = 1.4$. The reverse holds between status groups: the more productive firms tend to have *P* and *NP* workers of similar average skill levels. This conclusion will have to be corroborated by further evidence, as the estimates of the between-status coefficients are imprecise. However, to the extent that *P* and *NP* workers perform tasks that are fundamentally different but correlated in terms of complexity, skill self-matching across occupational status groups appears sensible. This suggests that some other firm attributes, such as the complexity or the technological content of the products, determines the optimal average skills of *P* and *NP* workers. These results are in line with the theoretical models on firm hierarchies of Garicano and Rossi-Hansberg (2004, 2006), who predict that high-ability managers work with high-ability *P* workers, but also with stratification of skills within occupational categories, particularly managers and more generally *NP* workers. According to these theories, given the wider range of tasks that *NP* workers perform, and their asymmetric importance, it is optimal to have "knowledge hierarchies" in which the knowledge of senior personnel is leveraged with less skilled staff. That is, the most straightforward and standardized tasks are assigned to less skilled workers, while the high-ability ones are reserved for problem solving and decision making. This type of organization accords with a hierarchical and more centralized view of the firm, where production depends on the skills of workers with decisional power.

Overall, our findings do not square with the "Microsoft versus McDonald's" organizational mode but are a better fit with the "Fiat versus Ferrari" dichotomy. Both Fiat and Ferrari are vertically integrated firms and do all phases of the production process in house, from R&D to design to assembly line production.[37] Therefore, both are likely to have a dispersed skill distribution, consistent with our finding that the within-firm component explains most of the overall skill dispersion. Second, the different technological content of the products (luxury and racing cars in

[37] This is particularly remarkable in the case of Ferrari, whose racing department, unlike other companies, has always produced both engines and chassis of racing cars. This does not mean that the two companies do not subcontract single components; on the contrary, both make extensive use of subcontracting.

the case of Ferrari and family cars in the case of Fiat) explains the fact that Ferrari has better engineers and better mechanics than Fiat. This is also in line with the optimality of matching the average skill level of *P* and *NP* workers. Third, according to case studies in the managerial literature, at least for Fiat during our sample period, the organizational mode was hierarchical, which might have benefited from a dispersed skills mix.[38] This contrasts with a series of studies for France and the United States (Caroli and Van Reenen 2001; Bresnahan et al. 2002), showing that the ICT revolution decentralized decision-making power and reduced the number of hierarchical levels. At least up to the late 1990s, there is no evidence of these more decentralized organizational modes in Italian manufacturing.

VI. Conclusion

We have assembled a matched employer-employee data set for Italy to analyze the distribution of workers' skills within and between firms and its relation to firms' productivity. We first conduct a variance decomposition exercise, which reveals that most of the dispersion of skills takes place within firms and not between them. We find no significant change in this pattern between 1981 and 1997. Thus, unlike other studies for the United States, France, and the United Kingdom, we find no evidence of a tendency toward skill segregation. Second, we find that the dispersion within occupational groups (*P* and *NP* workers) is positively correlated with firm productivity, while differences in the average skill levels of *P* and *NP* workers tend to have a negative impact, even if the evidence is less conclusive on this aspect. This suggests a production process in which it is optimal to match *P* and *NP* workers of similar average skills while having a dispersed distribution of skills within each occupational status group. We argue that this evidence is consistent with a hierarchical organization of the firm in which it is optimal to concentrate skills in individuals with decision-making and supervisory power, on which the firm's performance is heavily dependent.

An important question is whether our results extend outside the manufacturing sector. To the extent that *NP* workers are more prominent in services, one could argue that our conclusion on the importance of within-status dispersion might be even stronger in that sector. At the same time, the great diversity of activities performed in services implies that it is very hard to draw general inference on this question without direct evidence:

[38] Tronti (1997, 37) reports interviews with Fiat executives who say that it was not until the late 1990s that the company pursued "a new organizational mode based on participation with respect to the previous one, which had been based on hierarchical power" (translation is our own).

only further research that extends the analysis outside manufacturing will provide us with insights on how general these patterns may be.

In terms of policy, the results can be taken to give both positive and negative news on the evolution of Italian manufacturing over the last 20 years. The lack of an increase in segregation implies that productivity, and therefore income, of the low-skilled workers continues to benefit from workplaces with a dispersed skill distribution. On the other hand, if segregation occurs within a process that increases productivity by reorganizing production, then up to the late 1990s Italian manufacturing firms show no sign of participating in this transformation. This interpretation is supported by more recent evidence on the slow-paced ICT diffusion and new organizational modes in the last 10 years in the Italian economy, as well as the disappointing productivity trend since the mid-1990s.

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. Unpublished manuscript, Department of Labor Economics, School of Industrial and Labor Relations, Cornell University.

Abowd, John M., and Francis Kramarz. Forthcoming. Human capital and worker productivity: Direct evidence from linked employer-employee data. *Annales d'Economie et de Statistique*.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67 (March): 251–333.

Abowd, John M., Francis Kramarz, and Sebastien Roux. 2006. Wages, mobility and firm performance: An analysis using matched employee and employer data from France. *Economic Journal* 116 (June): 245–85.

Abowd, John M., Paul Lengermann, and Kevin L. McKinney. 2003. The measurement of human capital in the U.S. economy. Unpublished manuscript, Department of Labor Economics, School of Industrial and Labor Relations, Cornell University.

Acemoglu, Daron. 1999. Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review* 89 (December): 1259–78.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2006. Structural identification of production functions. Unpublished manuscript, Department of Economics, UCLA.

Akerlof, George A., and Janet L. Yellen. 1990. The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics* 105 (May): 255–83.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994. The internal economics of the firm: Evidence from personnel data. *Quarterly Journal of Economics* 109 (November): 881–919.

Berman, Eli, John Bound, and Zvi Griliches. 1994. Changes in the demand for skilled labor within U.S. manufacturing: Evidence from the annual survey of manufacturers. *Quarterly Journal of Economics* 109 (May): 367–97.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics* 117 (February): 339–76.

Caroli, Eve, and John van Reenen. 2001. Skill-biased organizational change? Evidence from a panel of British and French establishments. *Quarterly Journal of Economics* 116 (November): 1449–92.

Caselli, Francesco. 1999. Technological revolutions. *American Economic Review* 89 (March): 78–102.

Caselli, Francesco, and Wilbur J. Coleman. 2006. The world technology frontier. *American Economic Review* 96 (June): 499–522.

Ciccone, Antonio, and Giovanni Peri. 2005. Long-run substitutability between more and less educated workers: Evidence from U.S. states 1950–1990. *Review of Economics and Statistics* 87 (November): 652–63.

Cingano, Federico, and Fabiano Schivardi. 2004. Identifying the sources of local productivity growth. *Journal of the European Economic Association* 2 (June): 720–42.

Davis, Steven J., and John Haltiwanger. 1991. Wage dispersion between and within U.S. manufacturing plants, 1963–86. *Brookings Papers on Economic Activity: Microeconomics* 1991:115–80.

———. 1996. Employer size and the wage structure in U.S. manufacturing. *Annales d'Economie et de Statistique* 41/42:311–23.

Dell'Aringa, Carlo, and Laura Pagani. 2007. Collective bargaining and wage dispersion. *British Journal of Industrial Relations* 45 (March): 29–54.

Dongarra, Jack J., Iain S. Duff, Danny C. Sorensen, and Henk Van Der Vorst. 1991. *Solving linear systems on vector and shared memory computers*. Philadelphia: Society for Industrial and Applied Mathematics.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. 2004. Wage and productivity dispersion in United States manufacturing: The role of computer investment. *Journal of Labor Economics* 22 (April): 397–429.

Dunne, Timothy, John C. Haltiwanger, and Kenneth R. Troske. 1997. Technology and jobs: Secular changes and cyclical dynamics. *Carnegie-Rochester Conference Series on Public Policy* 46 (June): 107–78.

Fabiani, Silvia, Fabiano Schivardi, and Sandro Trento. 2005. ICT adoption in Italian manufacturing: Firm-level evidence. *Industrial and Corporate Change* 14 (April): 225–49.

Flabbi, Luca, and Andrea Ichino. 2001. Productivity, seniority, and wages: New evidence from personnel data. *Labour Economics* 8 (June): 359–87.

Garicano, Luis, and Esteban Rossi-Hansberg. 2004. Inequality and the organization of knowledge. *American Economic Review P&P* 94 (May): 197–202.

———. 2006. Organization and inequality in a knowledge economy. *Quarterly Journal of Economics* 121 (November): 1383–1435.

Grossman, Gene M., and Giovanni Maggi. 2000. Diversity and trade. *American Economic Review* 90 (December): 1255–75.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi. 2005. Insurance within the firm. *Journal of Political Economy* 113 (October): 1054–87.

Haltiwanger, John C., Julia I. Lane, and James R. Spletzer. 1999. Productivity differences across employers: The role of empoyer size, age, and human capital. *American Economic Review P&P* 89 (May): 94–98.

———. 2007. Wages, productivity and the dynamic interaction of businesses and workers. *Labour Economics* 14 (June): 575–602.

Haskel, Jonathan, Denise Hawkes, and Sonia Pereira. 2005. Skills, human capital and the plant productivity gap: UK evidence from matched plant, worker and workforce data. CEPR Discussion Paper no. 5334, Center for Economic Policy Research, London.

Ichniowski, Casey, Kathryn L. Shaw, and Jon Gant. 2003. Working smarter by working together: Connective capital in the workplace. Unpublished manuscript, Stanford Graduate School of Business.

Iranzo, Susana. 2003. City wages and dispersion of workers skills. Unpublished manuscript, Department of Economics, University of California, Davis.

Johnson, George E. 1997. Changes in earning inequality: The role of demand shifts. *Journal of Economic Perspectives* 11 (Spring): 41–54.

Katz, Lawrence F., and Kevin M. Murphy. 1992. Change in relative wages 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107 (February): 35–78.

Kramarz, Francis, Stefan Lollivier, and Louis-Paul Pele. 1996. Wage inequalities and firm-specific compensation policies in France. *Annales d'Economie et de Statistique* 41/42:369–86.

Kremer, Michael. 1993. The O-ring theory of economic development. *Quarterly Journal of Economics* 108 (August): 551–75.

Kremer, Michael, and Eric Maskin. 1996. Wage inequality and segregation by skill. Working Paper no. 5718, National Bureau of Economic Research, Cambridge, MA.

Krusell, Per, Lee E. Ohanian, José-Victor Ríos-Rull, and Giovanni L. Violante. 2000. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68 (September): 1029–53.

Lallemand, Thierry, Robert Plasman, and Francois Rycx. 2004. Intra-firm

wage dispersion and firm performance: Evidence from linked employer-employee data. *Kyklos* 57, no. 4:533–58.
Lallemand, Thierry, and Francois Rycx. 2006. Establishment size and the dispersion of wages: Evidence from European countries. *Applied Economics Quarterly* 52, no. 4:309–36.
Lazear, Edward, and Sherwin Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89 (October): 841–64.
Lazear, Edward, and Kathryn Shaw. Forthcoming. Wage structure, raises and mobility, In *The structure of wages*, ed. Edward Lazear and Kathryn Shaw. Chicago: University of Chicago Press for the National Bureau of Economic Research.
Manasse, Paolo, and Luca Stanca. 2003. Working on the train? The role of technical progress and trade in explaining wage differentials in Italian firms. Unpublished manuscript, Department of Economics, University of Bologna.
Milgrom, Paul, and John Roberts. 1990. The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review* 80 (June): 511–28.
OECD (Organization for Economic Cooperation and Development). 2003. *OECD Science, Technology and Industry Scoreboard, Annex 1*. Paris: OECD.
Olley, Steven G., and Ariel Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64 (November): 1263–97.
O'Mahony, Mary, and Bart van Ark, eds. 2003. *EU productivity and competitiveness: An industry perspective; Can Europe resume the catching-up process?* Luxembourg: Office for Official Publications of the European Comission.
Rosen, Sherwin. 1981. The economics of superstars. *American Economic Review* 71 (September): 845–58.
———. 1982. Authority, control and the distribution of earnings. *Bell Journal of Economics* 13 (Autumn): 311–23.
Tronti, Leonello. 1997. *Ristrutturazione industriale e struttura verticale dell'impresa*. Milan: Franco Angeli.
Winter-Ebmer, Rudolf, and Josef Zweimüller. 1999. Intra-firm wage dispersion and firm performance. *Kyklos* 52, no. 4:555–72.
Woolridge, Jeffrey M. 2005. On estimating firm-level production functions using proxy variables to control for unobservables. Unpublished manuscript, Department of Economics, Michigan State University.