

DISCUSSION PAPER SERIES

No. 5539

**SKILL DISPERSION AND FIRM
PRODUCTIVITY: AN ANALYSIS
WITH EMPLOYER-EMPLOYEE
MATCHED DATA**

Susana Iranzo, Fabiano Schivardi
and Elisa Tosetti

INDUSTRIAL ORGANIZATION



Centre for Economic Policy Research

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP5539.asp

SKILL DISPERSION AND FIRM PRODUCTIVITY: AN ANALYSIS WITH EMPLOYER-EMPLOYEE MATCHED DATA

Susana Iranzo, University of Sydney
Fabiano Schivardi, Banca d'Italia and CEPR
Elisa Tosetti, Cambridge University

Discussion Paper No. 5539
March 2006

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Susana Iranzo, Fabiano Schivardi and Elisa Tosetti

CEPR Discussion Paper No. 5539

March 2006

ABSTRACT

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

We study the relation between workers' skill dispersion and firm productivity using a unique dataset of Italian manufacturing firms from the early eighties to the late nineties with individual records on all their workers. Our measure of skill is the individual worker's effect obtained as a latent variable from a wage equation. Estimates of a generalized CES production function that depends on the skill composition show that a firm's productivity is positively related to skill dispersion within occupational status groups (production and non-production workers) and negatively related to skill dispersion between these groups. Consistently, the variance decomposition shows that most of the overall skill dispersion is within and not between firms. We find no change over time in the share of each component, in contrast with some evidence from other countries, based on less comprehensive data.

JEL Classification: D24, J24 and L23

Keywords: matched data, productivity, segregation and skills

Susana Iranzo
Assistant Professor
Department of Economics
University of Sydney
Sydney, 2006
AUSTRALIA

Fabiano Schivardi
Research Department
Banca d'Italia
Via Nazionale, 91
00814 Roma
ITALY

Tel: (61 2) 9036 9104
Fax: (61 2) 9351 4341
Email: s.iranzo@econ.usyd.edu.au

Tel: (39 06) 479 22168
Fax: (39 06) 4792 820
Email: Fabiano.Schivardi@bancaditalia.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157875

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145224

Elisa Tosetti
Faculty of Economics
University of Cambridge
Sidgwick Avenue
Cambridge
CB3 9DD

Email: et268@cam.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=164101

* We are indebted to Andrea Brandolini, Francesco Caselli, Francis Kramarz, Michael Kremer, Marco Magnani, Antoinette Schoar and Till Von Wachter for very useful comments and suggestions. We also thank seminar participants at the Bank of Italy, University of Barcelona, Luiss University, LaTrobe University and the 2005 Australian Conference of Economists, Melbourne. Many thanks to Giuseppe Bruno for helping us with the estimation routine and to the Italian Social Security Institute (INPS) for providing us their worker-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.

Submitted 01 February 2006

1 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, etc. However, less is known about the way firms' outcomes are related to the characteristics of the workers that firms employ. In this paper we focus on one aspect of workforce composition: the skill mix. Using a newly created matched Italian employer-employee dataset, 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 with workers of similar skill levels (Kremer 1993). Some recent matching and sorting models have shown that certain changes in the economy may alter the optimal production mode and thus the skill structure of firms. Following changes in the supply of skills (Kremer and Maskin 1996) and/or in technology (Acemoglu 1999, Caselli 1999), 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. As this literature suggests, the role of the skill mix in firms' performance also carries important implications for various fields, such as innovation, technological change, wage and income distribution and personnel economics.

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. We address this question using a new matched employer-employee dataset that is representative of Italian manufacturing firms with at least 50 employees, that covers almost 20 years (1981-1997) and, most

importantly, that includes individual information based on social security records on *all* the workers of each firm in the sample.¹ In addition to demographic characteristics and compensation of workers, we have detailed information on the characteristics of firms. This dataset, 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. However, wages also depend on 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 person-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, it is not only based on observable characteristics but also includes innate ability and informal skills not reflected in these.

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 tendency towards segregation. We then move on to study the relation between productivity and skill dispersion at the level of the firm directly. We

¹Previous 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).

estimate a generalized 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 sub-periods of the sample. In all our estimates we use the procedure of Olley and Pakes (1996) 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 towards skill segregation would imply. These results, robust to a number of checks, are in contrast with the evidence from other countries, such as the US, Britain and France, where some tendency towards segregation has been documented, although based on less comprehensive data (Dunne, Haltiwanger and Troske 1997, Dunne, Foster, Haltiwanger and Troske 2000, Kramarz et al. 1996).

The production function estimates show that 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. This is because P and NP workers are imperfect substitutes in production. We obtain an elasticity of substitution slightly below 1, a little lower but in line with the prevailing estimates in the literature, which unlike ours are obtained from the relative labor demand equation (Katz and Murphy 1992, Johnson 1997). By contrast, within each group of workers, the dispersion of skills, particularly that of NP workers, is beneficial for productivity: given an average skill level, it is preferable to have some highly skilled workers together with low-skilled ones than a uniform group. The results for the entire manufacturing sample are confirmed by sectoral estimates and are robust to a number of checks. Nor do we find any 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, 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 as "Ferrari and Fiat"

model instead of Kremer and Maskin (1996) “Microsoft and McDonald’s”. Ferrari and Fiat are both vertically integrated 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).

The rest of the paper is organized as follows. In section 2 we review the theoretical and empirical literature. Section 3 describes the data and presents the estimation of workers’ individual effects. Section 4 decomposes the variance of workers’ skills between and within firms. Section 5 examines the relationship between within-firm skill dispersion and productivity by estimating a generalized production function that allows for heterogeneity of workers skills. Section 6 concludes.

2 The literature

There is a good deal of empirical research on the connection between productivity and human capital at national and local level, but not, until recently, at the micro level of the firm. Using matched employer-employee datasets, Abowd et al. (1999) for France, Haltiwanger, Lane and Spletzer (1999) for the US, and Haskel, Hawkes and Pereira (2005) for the UK investigate the relation between productivity and workers’ skills. All of them find that the more productive firms have more highly skilled workers. Focusing on average skill levels within firms, these papers implicitly assume that workers’ skills are perfect substitutes. In reality, however, they may also be substitutes or complements, in which case not only the average level but also the particular combination of skills is important. We improve on these papers by explicitly considering firms’ skill mix in the production function, including higher moments of the within-firm skill distribution that identify the degree of complementarity/substitutability.

From a theoretical perspective, the impact of the skill distribution on firms' performance is related to the nature of the production function and depends on the degree of complementarity or substitutability between skills (Milgrom and Roberts 1990). 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 where workers' skills are substitutes and the performance of one subset of workers might be very important, as in the case of coordination and supervision.³ In this case, it is preferable to have teams with a few 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 to extend the production function to include higher moments of the skill distribution. To our knowledge, this is the first paper to analyze the role of skill dispersion on firm productivity for a representative sample of firms.

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 the hypothesis of Kremer and Maskin (1996) and Acemoglu (1999) of segregation of workers by skills between firms. Theoretically this phenomenon is explained as a move from a pooling equilibrium, in which in a labor market with search cost firms hire all types of workers and offer "middling" jobs with lax job descriptions, to a separating equilibrium, in which some firms offer high-quality jobs to high-skill workers and others hire mainly low-skilled workers (the Microsoft and McDonald models respectively). Two forces have been suggested as potentially responsible. One is changes in the economy-wide distribution of skills (Kremer and Maskin 1996, Acemoglu 1999). If the spread of workers skills or the productivity differential between different workers increase sufficiently, the skill mix might result in important differences in firm productivity. Consequently, firms find it profitable to move away from "middling" jobs to the Microsoft-McDonald model, where more attention is paid

²An 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.

³Grossman and Maggi (2000) argue that many of the goods and services exported by the U.S. 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.

on getting the right mix of workers skills. The second is skill-biased technological change (Acemoglu 1999, Caselli 1999), which can also induce skill segregation. Much of the technology adopted in recent decades requires highly-skilled operators. To the extent that firms differ in the rate of adoption of the new technology, we will find some firms making intensive use of it, with high productivity and mostly high-skilled workers, and others that use the technology less intensively, are less productive and target less skilled workers (Caselli 1999).

Theoretically appealing though the hypothesis is, the empirical evidence on segregation of workers by skill is scant and based on coarse and questionable measures of skills. Dunne et al. (1997) and Dunne et al. (2000) use the share of NP workers as a proxy for (high) skills and document secular increases in this share for all US manufacturing sectors from 1972 to 1988. Though it is a good proxy for 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 are determined by productivity and thus reflect workers' ability, Davis and Haltiwanger (1991) and Dunne et al. (2000) use wages as an alternative proxy for skills and analyze the dispersion of wages across and within US 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 due partly to differences in firms' compensation policies unrelated to differences in workers' ability and common to all workers in a firm. So 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 UK 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, which can be problematic, as we have argued. And their work, unlike ours, is not based on individual records for all workers within each firm.

Our work also contributes to the empirical literature on the degree of substitutability between skilled and unskilled workers (see for example Katz and Murphy 1992, Krusell and Violante 2000, Caselli and Coleman 2000, Ciccone and Peri 2005), to which our estimated elasticity of substitution between P and NP average skills is directly comparable.⁴ Mostly based on US data, these papers tend to find values of the elasticity of substitution between skilled and unskilled labor well above 1, at around 1.5. Manasse and Stanca (2003) get lower values, between 0.49 and 0.67, for P and NP workers in Italy. 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 method.

3 Sample construction

3.1 Data description

The data used in this paper 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 that ever transited in an INVIND firm for the period 1981-1997, including spells of employment in which they were employed at 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 year. The rest are employed in 100,000 other firms of which we only know the fiscal identifier.

The data on workers include age, gender, area where the employee works, occupational status (production, clerical, manager), annual gross earnings, number of weeks worked and

⁴Berman, 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.

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.⁵ 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. After this cleaning procedure, we are left with a total of 17,684,407 records, 1,683,854 individuals and 3,676,508 distinct employer-employee pairs, including non-INVIND firms. As the only information needed to estimate firm-fixed effects is the firm identifier, we use this dataset to estimate the wage equation that identifies the worker and firm fixed effects.

Table 1 gives the statistics on workers' characteristics for the total sample and for INVIND firms only. For the total sample average gross weekly earnings at 1995 constant prices are 350 euros, 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.⁶

Attrition in INVIND firms is substantial: on average 10% of workers enter and 12% exit the sample from one year to the next. Overall, approximately 80% of workers in an INVIND firm in 1981 had dropped out of the sample by 1997, and 72% of the workers in the 1997 sample had not been 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

⁵Extreme values of the earning distribution could be due to exceptional events (illness and the like) or to measurement error. Given that measures of dispersions can be very sensitive to such values, we decided to drop them from the analysis altogether.

⁶Guiso, Pistaferri and Schivardi (2005) report descriptive statistics for a different sample of workers, representative of the entire population of workers. The characteristics are highly similar to those of our sample of manufacturing firms with at least 50 employees.

shipments, value of exports and investment. In some years additional questions were asked: in 1995 one on organizational changes, in 1992-1995 one on number of establishments. We completed the data with the 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.⁷ For consistency with the capital data, in the estimation of the production function we take the 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 euros; 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 (see appendix, Table 8). 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 over-samples large firms. All the other characteristics are fairly similar to the unweighted data.

As we do not have plant level data, all our analysis is at firm level. From a theoretical point of view, it is unclear which unit would be 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,⁸ suggesting that this is not likely to be a major issue in our data. In any case, we will check all our results restricting the analysis to just single plant firms.

⁷See Cingano and Schivardi (2004) for a detailed account of the procedure.

⁸The 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 1992 and 1995, only 59 report more than one in other years, and 40 of these report only two.

3.2 Estimation of person-fixed effects

According to Abowd et al. (1999), wages can be decomposed into a component due to time-variant observable individual characteristics, a pure person 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, $J(i, t)$ is the firm where worker i works at time t . The person-fixed effect, θ , captures 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 effect, ψ , is interpreted as the component of wages specific to the firm where the employee works, and responds to efficiency wages or other particular compensation policies, rent-sharing or the bargaining power of workers in the firm.

Panel data allow us to identify firm and person effects as long as there is enough mobility of workers across firms. Following Abowd et al. (1999) and Abowd, Creedy and Kramarz (2002), we maintain the assumption of exogenous mobility conditional on the observables.⁹ 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 methods initially used in the literature were based on approximative methods, consisting of a two-stage procedure to estimate person 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), hereafter ACK, which simultaneously estimates person and firm effects. The ACK procedure estimates the full model in (1) by fixed-effect methods using the standard conjugate gradient (CG) algorithm with preconditioning as described in Dongarra, Duff, Sorenson and Van der Vorst (1991). The identification strategy consists of first determining the groups of

⁹This 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 solve a selection model, a computationally unfeasible problem.

connected workers and firms. A connected group comprises all the workers that any firm in the group has ever employed 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 person- 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 person or one firm-fixed effect equal to zero, so the estimated effects can only be interpreted in relative terms.

The first step of the estimation procedure was the identification of connected groups.¹⁰ Due to the sample design, based on the totality of workers for medium-sized and large firms, our dataset turns out to be one huge connected group: only 0.5% of the observations are disconnected. For computational simplicity, we only use the largest connected group, which contains 421,019 firms, 1,674,684 workers and 3,651,000 distinct firm-worker pairs. The relatively great mobility of workers (about 70% have more than one employer during the period) allows the identification of firm and person effects.

We estimated the wage decomposition of log weekly earnings into the three components of equation 1. The matrix of time-variant individual characteristics, X , includes age, age squared and occupational category, which changes for a substantial number of workers.¹¹ We also included seniority and calculated a unique coefficient on seniority for all firms.¹² The person effect is fixed over time. It captures unchanging personal attributes, such as the worker's innate ability and formal education (under the reasonable assumption, for Italy, that workers do not go back to school). As noted, the firm effect reflects a firm's compensation above the average for workers of comparable characteristics and can be explained as efficiency wages or other firm-specific compensation policies. Finally, the regression also controls for any trend or common time effect in wages by means of a full set

¹⁰We adapted the code implementing the CG and grouping algorithms kindly provided by Francis Kramarz.

¹¹The 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 re-classification of occupational status in the course of the employee's working life.

¹²Our data on seniority is left-censored as we do not have information on workers previous to 1981. To deal with this problem, we took the workers for whom we had information on their complete durations in jobs, that is, the workers who initiated 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 at any given year.

of year dummies.

The estimated coefficients of the covariates are reported in Table 3. We 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 $-.06\%$. This might be due to the measurement error embedded in our variable of seniority (see footnote 12) and to the correlation with the age.

We use person-fixed effects to proxy 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 person-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. It is worth noting that this proxy for skills does not include the returns to seniority in a firm, which supposedly reflect learning on-the-job. We chose to exclude seniority because, as is shown by Flabbi and Ichino (2001), wage increases related to seniority are not likely to reflect higher productivity but automatic upgrades due to typical Italian contractual arrangements.

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 person effects – the correlation between log earnings and person effects is the highest, 0.8. Firm effects play less of a role, with a 0.43 correlation with earnings. The correlation between the person and the firm effects is positive, but very small (0.044), which is also 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 2. On average, NP workers skills are 50% higher than P workers, with a fairly low dispersion across firms (standard errors are around .09 for both type of workers). The within-firm variance of skills is substantially higher for NP workers (.054 vs. .017).

4 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 US manufacturing (Davis and Haltiwanger (1991) and Dunne et al. (2000)) and for other countries (Kremer and Maskin 1996).

Our variance decomposition improves previous works in two directions. First, we use the worker fixed effects, which is 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 INVIND 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:

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

and

$$V_W = \sum_{f=1}^N l_f \cdot \sigma_f^2 \quad (3)$$

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 skills.

Kremer and Maskin (1996) index of segregation is the between-firm component of the variance of skills relative to the total variance, $\frac{V_B}{V_B+V_W}$. 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.¹³ 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-

¹³For comparability, we include in NP workers both clerks and managers. Our results do not change if we exclude the latter.

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 1981-1997, 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) and those of Dunne et al. (1997) and Dunne et al. (2000), who present some evidence of increasing segregation in the UK, France and the US. Moreover, the level of our segregation indexes is much lower than those for American manufacturing. The difference could be due to the fact that those studies are based on the dispersion of wages across plants, while we use the estimated person effects as our measure of skills. In order to make our results comparable, we re-calculated 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 person effects. Yet, as with the indexes based on worker effects, 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 towards the end of the decade and a flat pattern thereafter.

In terms of comparison with other countries, the index computed for the US is substantially higher, ranging from 0.48 in the late 1970s to 0.56 in 1992 (Dunne et al. 2000). The same holds for P and NP workers separately.¹⁴ The Italian values are closer, although still slightly lower, to those of Kramarz et al. (1996) for France for all workers – 0.36 in 1986 and 0.44 in 1992. Unfortunately, comparability is not complete, as the French sample also includes service workers and is based on firms with as few as 10 employees. Moreover, the

¹⁴Using wages, the segregation index for P and NP workers in Italian manufacturing is never higher than 0.45 and 0.15 respectively, while those for American manufacturing range between 0.76-0.84 and 0.47-0.69 respectively.

French indexes are based on an average of just 30 workers per firm in 1986 and 11 in 1992.

Considering manufacturing as a whole could mask important differences across sectors due, say, to technological differences. We re-calculated the segregation indexes for 4 sectoral groups according to the OECD technological classification (see appendix Table 8 for the list of sectors included in each category). Figure 3 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 10% to 13%). In terms of segregation levels, the less technologically intensive sectors display the highest segregation indexes. One explanation for this is that it is easier for low tech activities to separate the different phases of production physically 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.¹⁵

A further difference with respect to the US studies is that we consider segregation at the firm, rather than the plant or establishment, level.¹⁶ Unfortunately, we do not have plant-level data. Yet it is possible with our dataset to identify the single-plant firms and Figure 4 reports the segregation indexes for these alone. 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 multi-plant firms the distribution of skills is slightly smoothed out across establishments.

All in all, we find that, unlike other countries, Italy shows no tendency towards increasing segregation of workers by skill.

¹⁵This observation is consistent with the common view that the low tech sectors have undergone a process of delocalization of the production phase to countries with cheaper labor, while keeping in-house the activities with a higher value-added content.

¹⁶The French study by Kramarz et al. (1996) is also conducted at the firm level.

5 Within-firm skill dispersion and productivity

5.1 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 dataset and for P and NP workers separately. In all three 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 was explained earlier, there are certain activities for which having workers with similar skills is preferable. This is the case of Kremer (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 skills are dispersed.¹⁷ Although this is

¹⁷These ideas are formalized in Milgrom and Roberts (1990) with the concepts of supermodularity and

suggestive of 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.¹⁸ 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 should expect similar firms in different locations to employ different skill mixes. Second, labor regulations and other adjustment costs might prevent firms from fine-tuning their skill composition, again resulting in variation in the skill distribution across firms. Moreover, 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.

Based on the production function analysis, we address 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.¹⁹ 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, presumably reducing the optimal level of skill dispersion.²⁰

To formally investigate the relation between skill distribution and productivity, we use the following generalized Cobb-Douglas production function in capital and labor:

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

submodularity.

¹⁸In 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 (2000) 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, towards some "optimal" worker mix.

¹⁹Iranzo (2003) investigates this relation at the city level, finding that a more dispersed skill structure is beneficial for productivity.

²⁰For example, Caroli and Van Reenen (2001) show that ICT leads to a more decentralized organization of decision making, probably favoring a more homogeneous skill distribution within the firm.

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 $E(s_1, \dots, s_L)$ represents the overall efficiency of the labor force and depends on workers' skill levels s_i , $i = 1 \dots L_{ft}$ and the way they are combined in different firms. As P and NP workers differ not only in average skill level but also in the type of tasks they perform, we treat these two types of labor as distinct inputs. Consequently, the overall efficiency of the firm's labor force is a CES function of the efficiency of P and NP workers:

$$E(\mathbf{s}^P, \mathbf{s}^{NP}) = [l^P * (E^P(\mathbf{s}^P))^\gamma + l^{NP} * (E^{NP}(\mathbf{s}^{NP}))^\gamma]^\frac{1}{\gamma} \quad (5)$$

where \mathbf{s}^j is the vector of skills of workers in occupational status $j = P, NP$ and $l^j = \frac{L^j}{L}$ 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 $\frac{1}{1-\gamma}$. $E^j(\mathbf{s}^j)$ is in turn a CES function of workers skills within status j :

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

with the elasticity of substitution of skills for workers of status j given by $\frac{1}{1-\rho^j}$. 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 latter case is highly improbable. At least in the short run, the possibilities of substitution between 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 ($\rho^j > 1$). However, as the above discussion on production processes

has illustrated, within occupational status, substitutability of skills is less implausible.²¹

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 vector where all workers have the mean skill level as follows:

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

The first term is the overall skill mean, \bar{s} , the second term contains the *between-occupational-status* component, $\frac{(\bar{s}^P - \bar{s}^{NP})^2}{\bar{s}}$, 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:

$$\begin{aligned} \ln y_{ft} = & a_{ft} + \alpha \ln K_{ft} + \beta \ln L_{ft} + \beta \ln[\bar{s}_{ft} + \frac{1}{2}(\gamma - 1)l_{ft}^P l_{ft}^{NP} \frac{(\bar{s}_{ft}^P - \bar{s}_{ft}^{NP})^2}{\bar{s}_{ft}} \\ & + \frac{1}{2}(\rho^P - 1)l_{ft}^P \frac{\sigma_{ft}^{P2}}{\bar{s}_{ft}} + \frac{1}{2}(\rho^{NP} - 1)l_{ft}^{NP} \frac{\sigma_{ft}^{NP2}}{\bar{s}_{ft}}] \end{aligned} \quad (8)$$

which is a production-function-estimating equation augmented with the first and second moments of the skill distribution. Equation (8) makes the relation between skills complementarity/substitutability and productivity clear: controlling for the firm's average skill level \bar{s} , if γ , ρ^{NP} or $\rho^P > 1$ dispersion of skills increases productivity and vice-versa. Note that if we disregard the distinction between P and NP workers and assume that workers only differ in skill level, equation (8) simplifies to

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

²¹If $\rho^i > 1$, the optimal skill mix would be a corner solution, meaning that firms would want to hire only one worker with all the skills they require. As it is practically impossible for one worker to meet all the skills requirements of most firms, these are necessarily forced to hire several workers and match highly talented workers with less talented ones.

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.

5.2 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 (Figure 6). Labor productivity is constructed as output per worker, net of time and sectoral effects.²² 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, all indexes are increasing.²³ For example, firms in the last decile of the productivity distribution have an average person effect 25% greater than those in the first decile. The relation is highly similar for P workers alone and less strong for NP workers.²⁴

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 have 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 sub-periods: 1982-1990 and post-1990. The two graphs show very similar trends for the total and the single components, indicating that there was no significant change,

²²We have regressed output per worker on a set of year and two-digit sectoral dummies and used the residuals as a measure of productivity.

²³Abowd et al. (1999) and Haskel et al. (2005) also find that more productive firms have workers with higher average skill levels.

²⁴Note 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 increase and so does the share of NP workers, then the overall skill level will increase by more than the other two.

between the eighties and the nineties, in the relation between the skill mix and productivity.

We now turn to the econometric analysis. The main econometric problem in estimating equation (8) is that inputs are a choice variable and 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).²⁵ Using a standard dynamic programming approach, Olley and Pakes show that the unobservable productivity shock can be approximated by a non-parametric function of the investment and the capital stock, $a_{ft} = h(i_{ft}, k_{ft})$.²⁶ 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.²⁷ All the regressions include year, 13 sectoral and 4 area dummies, and observations are weighted by sampling weights. In order to account for the problem of generated 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 statuses. Results of the nonlinear least squares estimation are reported in Table 5. For comparability, in the first column we do not use the Olley and Pakes procedure. 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 (column 2) and the null hypothesis of ρ being larger than 1 is not rejected. In the 1,000 bootstrapped replications, the estimate was always above 1. This implies that overall skills are substitutes and that within-firm dispersion is positively correlated with productivity.

²⁵Ideally, 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.

²⁶Olley and Pakes show that the investment function takes the form $i_{ft} = i(k_{ft}, a_{ft})$ and that it is monotonous in both k and a , so that it can be inverted to express the productivity shock in terms of an unknown function of capital and investment, $h(i_{ft}, k_{ft})$.

²⁷Note 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. However, given that the coefficient on capital is of no particular interest to us, this is inconsequential for our purposes.

To check for any structural break in the coefficients, we split the sample into two time periods, pre- and post-1990, and run separate regressions. We obtain very similar estimates for ρ (1.87 and 1.82 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 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 1 in almost 90% of the bootstrap replications, which indicates that P and NP workers' skills are complements or imperfect substitutes. According to the point estimates, the elasticity of substitution between P and NP workers is .89. This value is lower than the preferred estimates for skilled and unskilled labor in the literature (between 1.3 and 1.6)²⁸ and more closely in line with the estimates of the elasticity of substitution between P and NP workers around 0.49-0.67 obtained by Manasse and Stanca (2003) for Italian manufacturing in the nineties.

In terms of within-status dispersion, we find that both ρ^{NP} and ρ^P are greater than one, implying within-status skill substitutability. The value is substantially greater for NP workers (5.5 vs. 1.4). In both cases we accept at reasonable significance levels the null hypothesis that the coefficient is greater than 1. 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. The last two columns of the table repeat the same exercise splitting 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 structural change in the way the skill distribution enters the production function.

An important critique to the previous regressions is that they assume the same underlying technology for firms in very different sectors. To tackle this potential problem, we run

²⁸See for example Katz and Murphy (1992), Johnson (1997), Krusell and Violante (2000), Caselli and Coleman (2000), Ciccone and Peri (2005).

the nonlinear estimation procedure separately for each of the three main Italian manufacturing sectors: metal and engineering products (metal products, machinery and transport equipment), textiles and leather, and chemicals, which account for 38%, 21% and 15% of total manufacturing production respectively. We thus have estimates for a traditional, low-tech sector (textiles and leather) and two sectors with a higher technological content.

The results are reported in Table 7. The positive impact of skill dispersion on firm productivity is confirmed by the sectoral regressions, with only a few caveats. First, when occupational status is not specified, we obtain a value of ρ that is positive and greater than 1, indicating overall skill substitutability, for the metal and engineering products and textile sectors (1.76 and 2.21 respectively), although this is significantly larger than 1 only in the case of textiles. For chemicals, the coefficient on overall dispersion is not significantly different from zero and we obtain estimates below 1 in 75% of the bootstrap replications. This suggests that this sector is different in the way skills are mixed, although the large standard errors do not allow for more precise inference. When we distinguish between the two types of labor, we get coefficients significantly greater than 1 for the dispersion of NP workers in all sectors. This indicates that NP dispersion enters the production function in a similar way in all three sectors. The dispersion of P workers appears to have a significantly positive impact on productivity only in the case of textiles, while results for metal and engineering products are mixed and complementarity seems to emerge in chemicals, possibly because here the production process is more sensitive to the contribution of all workers, as in the O-ring production function. Finally, the estimates on between-status skill dispersion appear imprecise, although in no case we can reject the null hypothesis that they are equal to zero and thus that P and NP workers are complementary or imperfect substitutes. Taken together, the sectoral estimates confirm that the skill dispersion of NP workers has a positive impact on productivity and that P and NP are imperfect substitutes in all sectors, while the results on dispersion of skills of P workers are mixed.

We performed a number of other robustness checks. The specification of our overall labor efficiency term does not allow for exogenous changes in the relative efficiency of P and NP workers that could arise, say, as a result of skill-biased technological change. Although for brevity we omit it here, we derived and estimated such a specification. Qualitatively

the results are similar to those of Table 6 in that ρ^P and ρ^N are greater than 1 and γ is less than 1. The only significant difference is in the magnitudes; as a consequence of having an extra term that measures the relative efficiency of NP workers with respect to that of P workers, γ is slightly higher while ρ^N is lower. We also re-ran the regressions for single-plant firms alone and 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. Finally, we ran some reduced-form regressions, in which dispersion is entered linearly, controlling for average skills and using the ratio of the 90th to the 10th percentile of the skill distribution as an alternative measure of dispersion. 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 of -1%. All in all, the results of a positive impact of within-status dispersion on productivity and a negative impact of between-status dispersion proved to be fairly robust.

5.3 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. This is true within each occupational status, particularly for NP workers, while the reverse holds between status groups. That is, the more productive firms tend to have P and NP workers of similar average skill levels. 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 attribute, such as the complexity or the technological content of the products, determines the optimal average skills of P and NP workers. On the other hand, within occupational status groups, workers' skills display a higher degree of substitutability, which is also sensible given that workers of the same occupational status are more likely to perform similar tasks and contribute to the same production objective. Our results go further: they imply that, controlling for the average skill level, it is optimal to have a dispersed skill distribution, particularly for NP workers. This could be due to the asymmetry in importance of the tasks performed by

workers. Some workers are managers or supervisors, and their duties have a powerful impact on the firm's performance, while others might be devoted to more straightforward and standardized tasks, for which skills are less important. This type of organization accords with a hierarchical view of the firm, where production depends on the skills of workers with decisional power.

Overall, our findings do not square with the "Microsoft vs. McDonald's" organizational mode but are a better fit with the "Fiat vs. 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.²⁹ 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 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 benefit from a dispersed skills mix.³⁰ This contrasts with a series of studies for France and the US (Caroli and Van Reenen 2001, Bresnahan, Brynjolfsson and Hitt 2002), showing that the ICT revolution decentralized decision-making power and reduced the number of hierarchical levels. During our sample period, ICT technologies were far from pervasive in Italian firms (Fabiani, Schivardi and Trento 2005), and consistent with that, at least up to the late nineties, there is no evidence of these more modern organizational modes in Italian manufacturing.

²⁹This 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 sub-contracting.

³⁰Tronti (1997) reports interviews with Fiat executives who say that it was not until the late nineties that the company pursued "(...) a new organizational mode based on participation with respect to the previous one, which had been based on hierarchical power.." (pg. 37, translation is our own).

6 Conclusion

We have assembled a matched employer-employee dataset for Italy to analyze the distribution of workers' skills within and between firms and its relation to firms' productivity. We first conducted 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 US, France and the UK, we find no evidence of a tendency towards 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 have a negative impact. 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.

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

References

- Abowd, J. M., Creecy, R. and Kramarz, F. (2002), Computing person and firm effects using linked longitudinal employer-employee data, Working paper, United States Census Bureau.
- Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999), ‘High wage workers and high wage firms’, *Econometrica* **67**, 251–333.
- Acemoglu, D. (1999), ‘Changes in unemployment and wage inequality: An alternative theory and some evidence’, *American Economic Review* **89**, 1259–1278.
- Baker, G., Gibbs, M. and Holmstrom, B. (1994), ‘The internal economics of the firm: Evidence from personnel data’, *Quarterly Journal of Economics* **109**(4), 921–955.
- Berman, E., Bound, J. and Griliches, Z. (1994), ‘Changes in the demand for skilled labor within u.s. manufacturing: Evidence from the annual survey of manufacturers’, *Quarterly Journal of Economics* **109**, 367–397.
- Bresnahan, T., Brynjolfsson, E. and Hitt, L. (2002), ‘Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence’, *Quarterly Journal of Economics* **117**, 339–376.
- Caroli, E. and Van Reenen, J. (2001), ‘Skill-biased organizational change? Evidence from a panel of British and French establishments’, *Quarterly Journal of Economics* **116**, 1449–1492.
- Caselli, F. (1999), ‘Technological revolutions’, *American Economic Review* **89**, 78–102.
- Caselli, F. and Coleman, W. (2000), The world technology frontier, NBER Working Paper 7904.
- Ciccone, A. and Peri, G. (2005), ‘Long-run substitutability between more and less educated workers: Evidence from u.s. states 1950-1990’, *Review of Economics and Statistics*.
- Cingano, F. and Schivardi, F. (2004), ‘Identifying the sources of local productivity growth’, *Journal of the European Economic Association* **2**, 720–742.
- Davis, S. and Haltiwanger, J. (1991), ‘Wage dispersion between and within u.s. manufacturing plants, 1963-86’, *Brookings Papers on Economic Activity. Microeconomics* **1991**, 115–180.
- Dongarra, J., Duff, I., Sorenson, D. and Van der Vorst, H. (1991), Solving linear systems on vector and shared memory computers, SIAM, Philadelphia.
- Dunne, T., Foster, L., Haltiwanger, J. and Troske, K. (2000), Wage and productivity dispersion in u.s. manufacturing: The role of computer investment, NBER Working Paper no. 7465.

- Dunne, T., Haltiwanger, J. and Troske, K. (1997), ‘Technology and jobs: secular changes and cyclical dynamics’, *Carnegie-Rochester Conference Series on Public Policy* **46**, 107–178.
- Fabiani, S., Schivardi, F. and Trento, S. (2005), ‘ICT adoption in italian manufacturing: Firm-level evidence’, *Industrial and Corporate Change* **14**, 225–249.
- Flabbi, L. and Ichino, A. (2001), ‘Productivity, seniority, and wages: New evidence from personnel data’, *Labour Economics* **8**, 359–387.
- Grossman, G. and Maggi, G. (2000), ‘Diversity and trade’, *American Economic Review* **90**, 1255–1275.
- Guiso, L., Pistaferri, L. and Schivardi, F. (2005), ‘Insurance within the firm’, *Journal of Political Economy* **113**, 1054–1087.
- Haltiwanger, J., Lane, J. and Spletzer, J. (1999), ‘Productivity differences across employers: the role of employer size, age, and human capital’, *American Economic Review* **89**, 94–98.
- Haltiwanger, J., Lane, J. and Spletzer, J. (2000), Wages, productivity and the dynamic interaction of businesses and workers, NBER Working Paper 7994.
- Haskel, J., Hawkes, D. and Pereira, S. (2005), Skills, human capital and the plant productivity gap: UK evidence from matched plant, worker and workforce data, CEPR Discussion Papers no.5334.
- Iranzo, S. (2003), City wages and dispersion of workers skills, mimeo, U.C. Davis.
- Johnson, G. (1997), ‘Changes in earning inequality: The role of demand shifts’, *Journal of Economic Perspectives* **11**, 41–54.
- Katz, L. and Murphy, K. (1992), ‘Change in relative wages 1963-1987: Supply and demand factors’, *Quarterly Journal of Economics* **107**, 35–78.
- Kramarz, F., Lollivier, S. and Pele, L. (1996), ‘Wage inequalities and firm-specific compensation policies in France’, *Annales d’Economie et de Statistique* **41/42**, 369–386.
- Kremer, M. (1993), ‘The O-ring theory of economic development’, *Quarterly Journal of Economics* **108**, 551–575.
- Kremer, M. and Maskin, E. (1996), Wage inequality and segregation by skill, NBER Working Paper 5718.
- Krusell, P., O. L. R.-R. J. and Violante, G. (2000), ‘Capital-skill complementarity and inequality: A macroeconomic analysis’, *Econometrica* **68**, 1029–1053.
- Manasse, P. and Stanca, L. (2003), Working on the train? The role of technical progress and trade in explaining wage differentials in Italian firms, mimeo, Università di Bologna.

- Milgrom, P. and Roberts, J. (1990), 'The economics of modern manufacturing: Technology, strategy, and organization', *American Economic Review* **80**, 511–528.
- Olley, S. and Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**, 1263–1297.
- Rosen, S. (1981), 'The economics of superstars', *American Economic Review* **71**, 845–858.
- Tronti, L. (1997), *Ristrutturazione industriale e struttura verticale dell'impresa*, Milano, Franco Angeli.

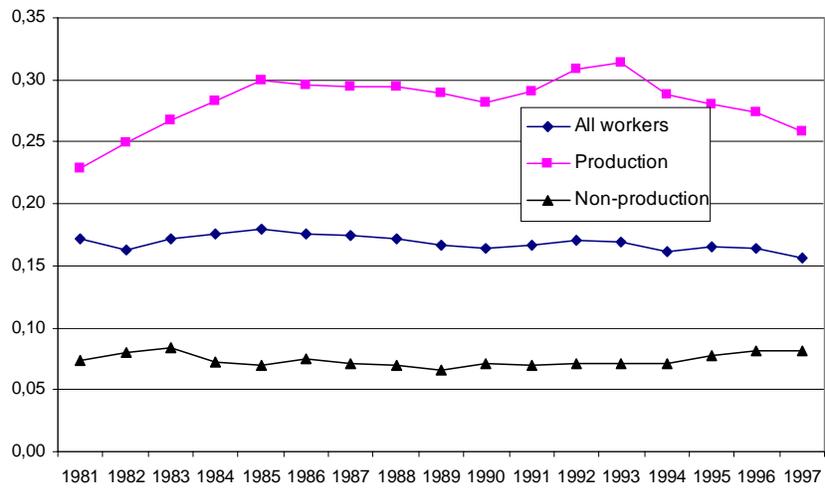


Figure 1: Segregation indexes, person effects

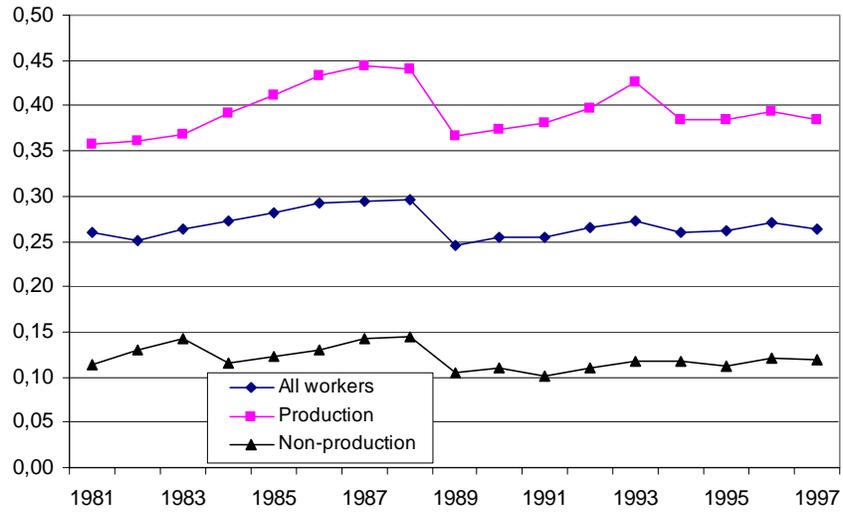


Figure 2: Segregation indexes, wages

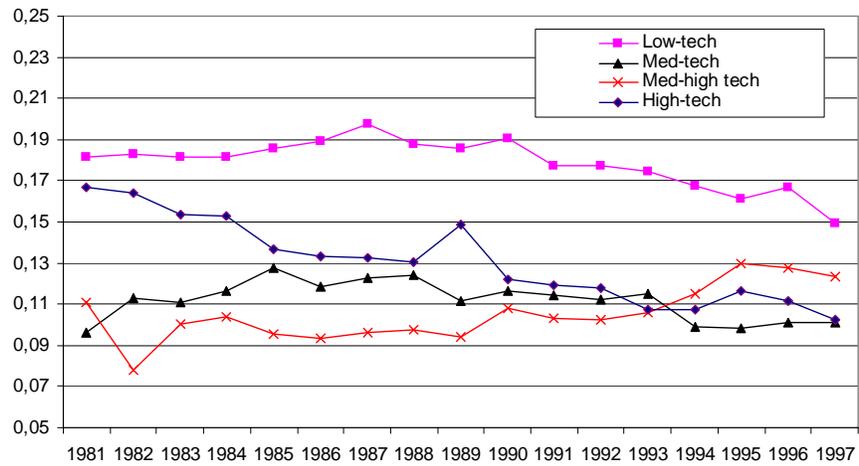


Figure 3: Segregation indexes by technological intensity, person effects

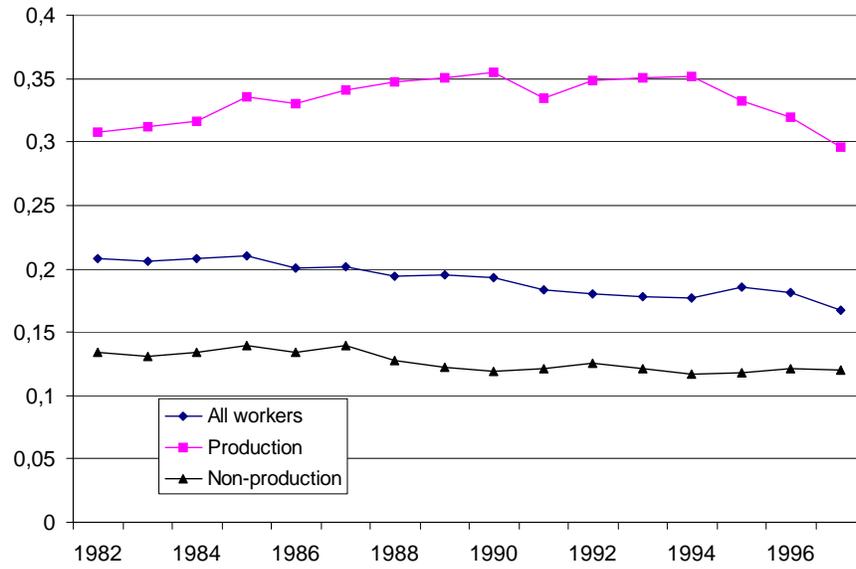


Figure 4: Segregation index, single plant firms

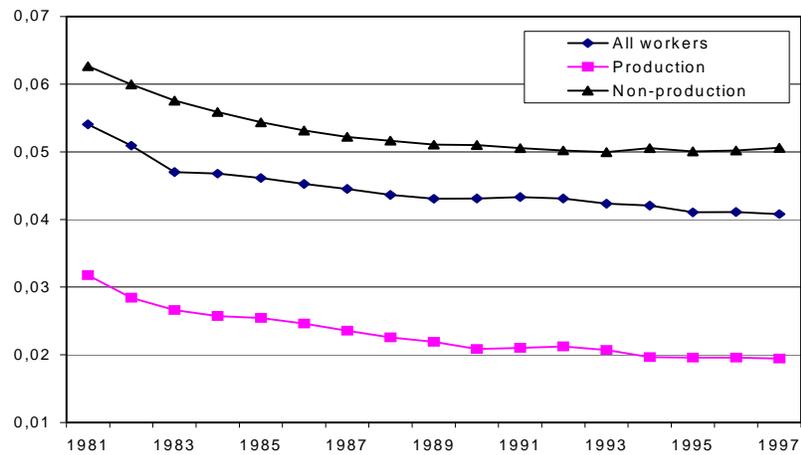


Figure 5: Variance of person effects

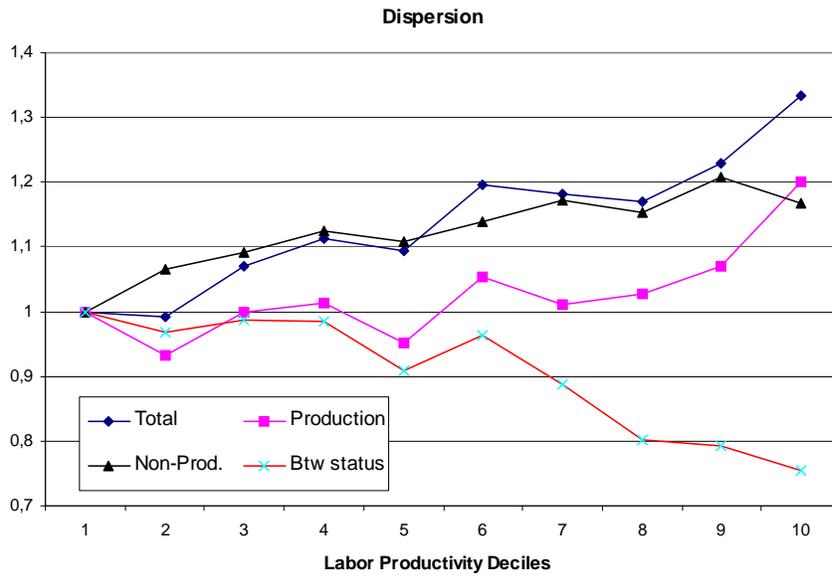
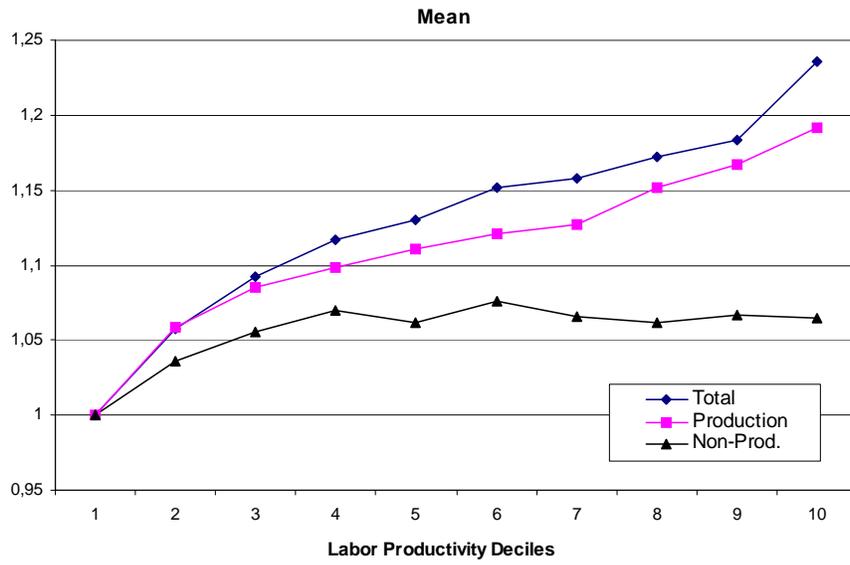


Figure 6: Mean and variance of firm-level skill distribution by deciles of productivity

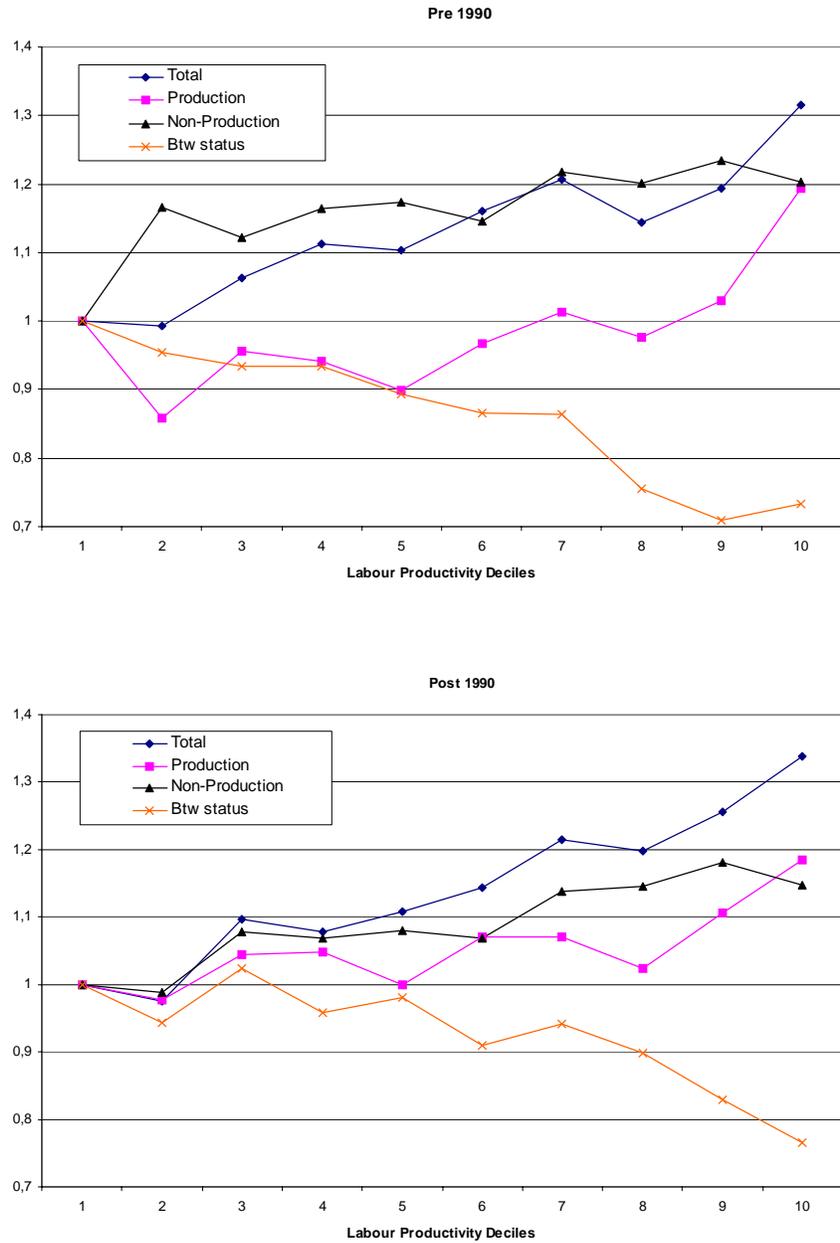


Figure 7: Variance of firm level skill distribution by deciles of productivity: pre and post 1990

Table 1: Workers' characteristics

	Whole sample		INVIND sample	
	Mean	Std. Dev.	Mean	Std. Dev.
Weakly Wage (1995 Euros)	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 prod. workers	66.0		69.17	
Share of non prod. workers	32.7		29.93	
No. of observations	17,593,816		9,559,271	

Table 2: Firms' characteristics

	Unweighted		Sampling weights	
	Mean	Std. Dev.	Mean	Std. Dev.
Value added	26.6	118.1	8.1	38.0
Capital stock	46.4	242.0	13.8	77.0
No. of workers	625	3,024	198	949
Ave. person effect \bar{s}	.46	.09	.45	.08
Ave. P workers effect \bar{s}^P	.40	.08	.40	.08
Ave. NP workers effect \bar{s}^{NP}	.61	.09	.60	.10
Variance of person effect σ^2	.036	.018	.034	.018
Variance of P effect σ^{P^2}	.017	.012	.017	.013
Variance of NP effect σ^{N^2}	.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				
North-West		.44		.47
North-East		.25		.27
Center		.20		.16
South		.11		.10
Share of single-plant firms		.63		.78
No. of observations		9,790		9,790

Value added and capital stock are in millions of 1995 euros. P stands for production and NP for non-production workers. See the appendix for the sectoral classification in terms of technological content.

Table 3: Estimated coefficients of the time-variant personal characteristics

	Coeff.	Std. Dev. (*)
Age	0.02349	0.0000261
Age squared	-0.00024	0.0000003
Dummy for prod. workers	-0.52908	0.0003147
Dummy for non prod.workers	-0.45683	0.0003149
Seniority	-0.00063	0.0000120

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

Table 4: Correlation between predicted value, firm effect, person effect adjusted for sex, person effect, y, error

	std. Dev.	y	predxb	effpers	effpers*	effirm	err
y	0.332	1					
predxb	0.091	0.605	1				
effpers	0.220	0.798	0.428	1			
effpers*	0.252	0.475	0.311	0.588	1		
effirm	0.120	0.426	0.131	0.044	-0.004	1	
err	0.129	0.388	0.001	0.000	0.000	0.000	1

Note: y: log(weekly earning) in 1995 Euros; predxb: predicted value of time-variant personal characteristics; effpers: person effect; effpers*: gender-adjusted person effect; effirm: firm effect; err: regression residual.

Table 5: Non-linear least squares: overall skill dispersion

		[1]	[2]	[3]	[4]
Variable	Param.	Whole sample	Whole sample	Pre-1990 sample	Post-1990 sample
Overall dispers.	ρ	1.750 (.391) [1.00]	1.800 (.414) [1.00]	1.869 (.544) [.999]	1.816 (.398) [1.00]
Labor	β	.766 (.019)	.722 (.019)	.711 (.0240)	.731 (.021)
Capital	α	.226 (.015)			
Olley-Pakes		NO	YES	YES	YES
R^2		.87	.88	.89	.87
# observations		9,790	9,790	4,180	5,610

Dependent variable: log value added. Results from estimating eq. (9) with non-linear least squares. Bootstrapped standard errors based on 1,000 replications in round brackets, probability that the parameter is greater than 1 (computed as the frequency of bootstrapped estimates greater than 1) in square brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include year, 2-digit sector dummies and 4 macro-region dummies.

Table 6: Non linear least squares: skill dispersion within and between occupational status groups

		[1]	[2]	[3]	[4]
Variable	Param.	Whole sample	Whole sample	Pre-1990 sample	Post-1990 sample
Btw disp.	γ	-.590 (.791) [.031]	-.118 (.900) [.112]	-.002 (1.178) [.231]	-.186 (.886) [.106]
Prod. dispers.	ρ^P	1.545 (.268) [.988]	1.416 (.276) [.931]	1.489 (.388) [.961]	1.357 (.380) [.753]
Non-prod dispers.	ρ^{NP}	5.099 (.691) [1.00]	5.551 (.786) [1.00]	5.838 (1.376) [1.00]	5.511 (.842) [1.00]
Labor	β	.766 (.019)	.718 (.018)	.710 (.023)	.723 (.020)
Capital	α	.225 (.015)			
Olley-Pakes		NO	YES	YES	YES
R^2		.88	.88	.90	.87
# observations		9,790	9,790	4,180	5,610

Dependent variable: log value added. Results from estimating eq. (8) with non-linear least squares. Bootstrapped standard errors based on 1,000 replications in round brackets, probability that the parameter is greater than 1 (computed as the frequency of bootstrapped estimates greater than 1) in square brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include year, 2-digit sector dummies and 4 macro-region dummies.

Table 7: Non-linear least squares: Sectoral estimates

		[1]	[2]	[3]	[4]	[5]	[6]
Variable	Param.	Metal-Engin.		Text-Footw.		Chemicals	
Overall dispers.	ρ	1.761 (1.091) [.791]		2.210 (.562) [.999]		.118 (1.155) [.264]	
Btw dispers.	γ		.364 (1.654) [.335]		.129 (1.093) [.199]		-.575 (2.765) [.330]
Prod. dispers.	ρ^P		.592 (1.322) [.413]		1.977 (.462) [.997]		-.764 (1.595) [.156]
Non prod dispers.	ρ^{NP}		5.145 (1.592) [.998]		5.923 (1.388) [1.00]		4.287 (1.734) [.978]
Labor	β	.755 (.030)	.739 (.029)	.730 (.041)	.727 (.034)	.761 (.044)	.721 (.046)
Olley-Pakes		YES	YES	YES	YES	YES	YES
R^2		.88	.88	.86	.87	.92	.93
# observations		3,698	3,698	2,059	2,059	1,474	1,474

Dependent variable: log value added. Results from estimating eq. (9) and (8) with non-linear least squares. Bootstrapped standard errors based on 1,000 replications in round brackets, probability that the parameter is greater than 1 (computed as the frequency of bootstrapped estimates greater than 1) in square brackets. Observations are weighted according to the sampling weights (see the main text for details). All regressions include years and 4 macro-region dummies.

Table 8: Appendix: OECD technology classification system

Classification	ISIC Rev.3 . Sectoral No.
HIGH-TECHNOLOGY MANUFACTURES	
PHARMACEUTICALS	2423
OFFICE, ACCOUNTING AND COMPUTING MACHINERY	30
RADIO, TELEVISION AND COMMUNICATION EQUIPMENT	32
AIRCRAFT AND SPACECRAFT	353
MEDIUM-HIGH TECHNOLOGY MANUFACTURES	
CHEMICALS EXCLUDING PHARMACEUTICALS	24ex2423
MACHINERY AND EQUIPMENT, N.E.C.	29
ELECTRICAL MACHINERY AND APPARATUS, NEC	31
MEDICAL, PRECISION AND OPTICAL INSTRUMENTS	33
MOTOR VEHICLES, TRAILERS AND SEMI-TRAILERS	34
RAILROAD EQUIPMENT AND TRANSPORT EQUIPMENT N.E.C.	352+359
MEDIUM-LOW TECHNOLOGY MANUFACTURES	
COKE, REFINED PETROLEUM PRODUCTS AND NUCLEAR FUEL	23
RUBBER AND PLASTICS PRODUCTS	25
OTHER NON-METALLIC MINERAL PRODUCTS	26
BASIC METALS AND FABRICATED METAL PRODUCTS	27-28
BUILDING AND REPAIRING OF SHIPS AND BOATS	351
MANUFACTURING N.E.C.	369
LOW-TECHNOLOGY MANUFACTURES	
FOOD PRODUCTS AND BEVERAGES	15
TOBACCO PRODUCTS	16
TEXTILES	17
WEARING APPAREL, DRESSING AND DYEING OF FUR	18
LEATHER, LEATHER PRODUCTS AND FOOTWEAR	19
WOOD AND PRODUCTS OF WOOD AND CORK	20
PULP, PAPER, PAPER PRODUCTS, PRINTING AND PUBLISHING	21-22
FURNITURE	361