Volume 66
February 2014
ISSN 0014-2921

EUROPEAN ECONOMIC REVIEW

CONTENTS

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
North-Holland

European Economic Review 66 (2014) 144–179

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

Labor diversity and firm productivity

Pierpaolo Parrotta [a], Dario Pozzoli [b,*], Mariola Pytlikova [b,c]

[a] Department of Economics, School of Business and Economics, Maastricht University, 6200 MD Maastricht, Netherlands
[b] The Danish Institute for Regional Government Research (KORA), Denmark
[c] VSB-Technical University of Ostrava, Czech Republic

ARTICLE INFO

Article history:
Received 24 October 2012
Accepted 12 December 2013
Available online 2 January 2014

JEL classification:
J24
J61
J82
L20

Keywords:
Labor diversity
Substitutability
Complementarity
Communication barriers
Total factor productivity

ABSTRACT

Using a matched employer–employee data-set, we analyze how workforce diversity associates with the productivity of firms in Denmark, following two main econometric routes. In the first one, we estimate a standard Cobb–Douglas function, calculate the implied total factor productivity and relate the latter to diversity statistics in a second stage. This reduced-form approach allows us to identify which types of labor heterogeneity appear to descriptively matter. In the second approach, we move toward a richer production function specification, which takes different types of labor as inputs and that allows for flexible substitution patterns, and possible quality differences between types. Both methods show that workforce diversity in ethnicity is negatively associated with firm productivity. The evidence regarding diversity in education is mixed.

1. Introduction

Diversity in the labor force is an increasing reality in many developed countries. This diversity results from, among other things, the following major factors: policy measures that counteract population aging and anti-discrimination measures, the growth in immigration from diverse countries experienced in recent decades and the educational and skill upgrading of workforces.[1] All of these factors lead to increasing diversity within the labor force in terms of age, gender, ethnicity and skills.

* Corresponding author at: KORA, The Danish Institute for Regional Government Research, Købmagergade 22, Copenhagen 2220, Denmark.
E-mail addresses: p.parrotta@maastrichtuniversity.nl (P. Parrotta), dapo@kora.dk (D. Pozzoli), mapy@kora.dk (M. Pytlikova).

[1] Demographic projections by the United Nations suggest that during the next four decades, populations in Europe might *ceteris paribus* decline by 12% (United Nations, 2000). The main factor responsible for population aging is a large decline in the total fertility rate over the last half century. As a consequence of this trend, governments have adopted a number of measures to counteract the problem of population aging, including policies that encourage people to work longer, to increase female labor participation and to attract skilled immigrants. In many countries, governments have increased the regular and early retirement age, restricted access to early retirement by changing economic incentives and promoted anti-discrimination measures related to age. Female labor participation has grown in most of the world during the last century (OECD, 2005). This growth is partly due to policies encouraging women to work, e.g., better childcare and parental leave provisions and gender anti-discrimination measures. Furthermore, we can observe an increase in immigration, including to developed countries, and a broader diversity of immigrants with respect to their countries of origin (Adsera and Pytlikova, 2011; Pedersen et al., 2008). As a result of this change, the diversity of the workforce with respect to gender, age and ethnicity has increased. Finally, as a consequence of the worldwide globalization process and skill-biased technological changes, governments in many countries have taken steps to increase the skill level of the workforce (e.g., by increasing the supply of university-educated people and enhancing the availability of lifelong learning).

0014-2921/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.euroecorev.2013.12.002

We observe increasing diversity across many workplaces and often hear about the importance of further internationalization and demographic diversification for firms. In many countries, firms′ hiring decisions are affected by governmental affirmative action policies. Additionally, firms are under social pressure to increase diversity. At the same time, firms are challenged by the constantly changing demand for goods and services, as well as by new customers and markets, in today′s globalized world. A diverse workforce may be a key factor in helping firms to understand and meet these new needs.

The popular press usually emphasizes workforce diversity as beneficial for firms, but is this really true? Do firms benefit from labor diversity and does it generate competitive advantage? What is the relationship between workplace labor diversity and firm performance? Although the issue is very important, there is considerable ambiguity surrounding this topic.

Economic theory suggests that workforce diversity may affect firm performance differently and through various channels. Diversity in skills and education may generate knowledge spillover among the employees within a firm (as long as workers′ knowledge sets do not overlap and are relevant to one another), which positively affects firm performance (Lazear, 1999). However there are certain activities for which having workers with similar skills and education is preferable, as in the case of Kremer′s O-ring production function, where profit-maximizing firms should match workers of similar skills/education together. Similarly, diversity in age can be beneficial to firms because the human capital of younger and older workers can complement each other. Younger employees have knowledge of new technologies and IT, and older employees have a better understanding of (and more experience with) intra-firm structures and the operating process (Lazear, 1998). However, Becker′s model of co-worker discrimination suggests that demographic heterogeneity among workers may create communication friction if workers are prejudiced and may thus result in some productivity costs.

The expected contribution of ethnic and cultural diversity to firm performance is also unclear. Ethnic-cultural diversity may affect firm performance negatively because it may (i) hinder potential knowledge transfer among workers due to linguistic and cultural barriers, (ii) reduce peer pressure by weakening social ties and trust, and (iii) create non-pecuniary disutility associated with joining or remaining in a ethnically diverse firm (Lazear, 1999). A similar point regarding trust is made by Glaeser et al. (2000) and Alesina and La Ferrara (2002), who show that people often distrust members of other ethnic groups and tend to prefer interacting in culturally homogeneous communities. Conversely, ethnic diversity can be beneficial to firm performance, improving decision making and problem solving (Hong and Page, 2001, 2004), stimulating the creation of new ideas and favoring knowledge transfers (Berliant and Fujita, 2008). Further, workforce diversity may provide useful information to a firm about a product market, which can enhance the firm′s ability to compete in global markets (Osborne, 2000; Rauch and Casella, 2003).

To the best of our knowledge, the empirical evidence concerning diversity and economic performance is fairly scarce, and most of the previous work in this area has employed case studies of one firm (e.g., Hamilton et al., 2003, 2004; Kurtulus, 2011; Leonard and Levine, 2006) or has used aggregate regional data (e.g., Ottaviano and Peri, 2006, 2012; Suedekum et al., 2009). The use of more comprehensive data in this field is fairly rare (Barrington and Troske, 2001; Iranzo et al., 2008; Navon, 2009; Grund and Westergaard-Nielsen, 2009; Garnero and Rycz, 2013). Furthermore, most previous studies have focused on only one dimension of diversity, with the studies reported by Barrington and Troske (2001), Kurtulus (2011) and Leonard and Levine (2006) being the only exceptions, and none of these studies has determined the effect of diversity on firm performance. Within this largely “explorative” and “descriptive” literature, there seems to be some consensus with respect to skill diversity as a positive factor in firm performance (Hamilton et al., 2003, 2004; Leonard and Levine, 2006; Iranzo et al., 2008; Navon, 2009; Kurtulus, 2011; Garnero and Rycz, 2013), but the evidence regarding diversity along ethnic and demographic lines is rather mixed. Case studies, for example, find that diversity with regard to age and race is negatively associated with firm performance (Hamilton et al., 2003, 2004; Leonard and Levine, 2006; Kurtulus, 2011), whereas studies using aggregated regional data find a positive correlation between ethnic diversity and performance (e.g., Ottaviano and Peri, 2006, 2012; Alesina and La Ferrara, 2005; Sparber, 2009; Suedekum et al., 2009; Peri, 2012). As this study, Fox and Smeets (2011) make use of the Danish matched employer–employee data-set and consider different skills levels of workers. Their work is primarily focused on quality dispersion within labor rather than on the role that diversity of inputs plays in making firms more or less productive.

In this paper, we use a unique register-based linked employer–employee data-set (LEED) from Denmark, which allows us to overcome many of the limitations of previous studies and to contribute to the literature in several ways. We follow two main econometric routes to investigate the association between diversity and firm productivity. First, we estimate a standard Cobb–Douglas function that includes labor as a single undifferentiated input, calculate the implied total factor productivity and in a second stage relate the latter to three relevant dimensions of diversity, i.e., cultural background, education and demographics, and using two alternative specifications of diversity, i.e., an aggregate and a disaggregate one. Implementing this “reduced-form” approach, we also explore the possible mechanisms through which workforce diversity affects firm productivity by attempting to test a set of hypotheses derived from existing theories. Specifically, we look at whether the impact of diversity on productivity arises from diversity within distinct occupational groups rather than the establishment′s labor force in total, because we expect that diverse problem-solving abilities and creativity will be more strongly related to productivity in white-collar occupations than in blue-collar occupations. Additionally, we investigate the importance of communication costs and the costs of “cross-cultural dealing” by excluding certain groups of foreigners (i.e., individuals with tertiary education or those who speak a Germanic language) in calculating the ethnic diversity measures. The reduced-form approach allows us to identify which types of labor heterogeneity appear to descriptively matter but it does not formally take into account that the labor input is non-homogeneous in the production function, i.e., labor of different types is of different quality (Hellerstein and Neumark, 2004; Iranzo et al., 2008; Fox and Smeets, 2011; Irarrazabal et al., 2014). We therefore move toward a richer production function specification, which takes different types of labor as inputs and that allows for flexible substitution patterns, and possible quality differences between types. Specifically, we proceed by modeling a value-added production function that, as in the reduced-form approach, is Cobb–Douglas in capital and labor but in which the contribution of the labor aggregate also depends on different types of labor in a CES specification.

Our results generally show that labor diversity in ethnicity is negatively associated with firm productivity, while the demographic diversity seems not to matter. These findings are consistent with earlier research of Lazear (1999), Glaeser et al. (2000), and Alesina and La Ferrara (2002), and may provide evidence that the negative effects of the communication and integration costs that are associated with a more demographically and culturally diverse workforce counteract the positive effects of diversity on firm productivity (i.e., the effects of creativity and knowledge spillover). The evidence regarding labor heterogeneity in terms of education is mixed instead. On one hand, our reduced-form analysis reveals that labor diversity in education is significantly and positively associated with firm productivity. On the other hand, the estimated parameters of the structural production function, governing the substitutability between labor types, suggest that it is not generally optimal to have dispersion in labor types along the educational dimension. However estimating a modified specification of the structural model separately for each 2-digits industry suggests that, for about half of the sectors, skill diversity arising only among highly educated workers positively associates with firm productivity.

The structure of the paper is as follows: Section 2 briefly describes the data as well as the methods used to calculate labor diversity at the firm level; Section 3 describes the main econometric routes we follow to measure firm productivity and its association with labor diversity. Section 4 reports results on the relationship between diversity and productivity using the reduced-form approach. Section 5 includes results from the structural production function estimation which allows for labor heterogeneity. Section 6 offers concluding remarks.

2. Data

2.1. Data description

The data-set for this empirical investigation is created by merging information from three different main sources. The first source is the Integrated Database for Labor Market Research (henceforth IDA), provided by Statistics Denmark. The IDA is a longitudinal employer–employee register that contains valuable information (age, education, other demographic characteristics, labor market experience and earnings) about each individual employed in the recorded population of Danish firms for the period 1980–2005. Only attrition due to death and permanent migration is included in the data-set. The labor market status of each person is his or her status as of the 30th of November of each year. The retrieved information is aggregated at the firm level to obtain variables such as firm size, workforce composition (including average firm tenure; shares of managers, middle managers, men, highly skilled workers, and technicians; and the shares of employees belonging to each age distribution quartile), labor diversity (see the next section for more details) partial/total foreign ownership and whether the firm is multi-establishment.

The second data source (henceforth referred to as REGNSKAB), also compiled by Statistics Denmark, provides information on firms′ business accounts. These data cover the construction and manufacturing industries beginning in 1994, manufacturing beginning in 1995, wholesale trade beginning 1998 and the remaining portions of the service industry from 1999 onwards. From REGNSKAB, the following accounting items are used to estimate the production function: value added,[2] materials (intermediates), capital stock (fixed assets) and related industries.[3] Furthermore, linking these variables to a third data source, i.e., the Foreign Trade Statistics Register, we can retrieve information on whether the firm engages in export activities.

2.2. Firm level labor diversity

This section focuses on employee diversity at the firm level. Labor diversity is quantified using information regarding workers′ gender, age, work experience, highest level of education achieved and nationality. We use the Herfindahl index to measure the degree of diversity at the firm level. Unlike traditional diversity measures such as the percentage of employees belonging to a specific group, the Herfindahl index combines two quantifiable measures: the “richness” (the number of categories represented within the firm or the workplace) and “equitability” or evenness (how even the numbers are for the individual categories). We calculate three separate indices to measure the cultural, skill and demographic dimensions of diversity.

Cultural diversity is represented either by the employees′ nationalities or by the languages they speak. The various nationalities have been grouped into the following categories: North America and Oceania, Central and South America,

[2] Computed as the difference between the total sales and the intermediate costs.

[3] The following industries are excluded from the empirical analysis: (i) agriculture, fishing and quarrying; (ii) electricity, gas and water supply; and (iii) public services.

Africa, West and South Europe, former Communist countries, East Asia, Other Asia, and Muslim countries.[4] It has been argued in the previous literature that linguistic distance serves as a good proxy for cultural distance (Guiso et al., 2009; Adsera and Pytlikova, 2011). Therefore, we have grouped employees together by the languages spoken in their countries of origin. This linguistic classification is more detailed than the grouping by nationality. We group countries (using the major official language spoken by the majority) at the third linguistic tree level, e.g., Germanic West vs. Germanic North vs. Romance languages. The information on languages is drawn from the encyclopedia of languages entitled *Ethnologue: Languages of the World* (Lewis, 2009); see the Appendix for more details about the list of countries and the linguistic groups included. Education-related diversity is represented by 6 categories based on information concerning the employees′ highest educational level completed (tertiary education, secondary and vocational education, or pre-secondary education). We divide tertiary education into 4 categories, making a distinction between Bachelor′s, Master′s and postgraduate degrees in the social science, the humanities, engineering and the natural sciences. In a more disaggregated specification, we also decompose secondary education into general high school, business high school and short and long vocational education programs. Finally, the demographic index is built from the intersection of gender and age quartiles or quintiles (8 or 9 categories in total, depending on the level of aggregation).

To measure diversity at the firm level for each dimension, we sum the Herfindahl indices calculated for each workplace belonging to the same firm, which are weighted by the number of employees employed in each workplace:

$$index_{hit} = \sum_{w=1}^{W} \frac{N_w}{N_i}\left(1-\left(\sum_{s=1}^{S} p_{swt}^2\right)\right)$$

where $index_{hit}$ is the Herfindahl diversity index of firm i at time t calculated along the h-th dimension (education-related and demographic), W is the total number of workplaces belonging to firm i, S is the total number of categories of the related diversity dimension, and N_w and N_i are the total number of employees of workplace w and of firm i. The proportion of the workplace′s labor force that falls into each category s of the h-th dimension at time t is represented by the term p_{swt}.[5] The diversity index has a minimum value equal to 0 if only one category is represented within the workplace and a maximum value equal to $(1-1/S)$ if all categories are equally represented. The index is interpreted as the probability that two randomly drawn individuals in a workplace belong to different groups.

2.3. Descriptive statistics

Before discussing descriptive statistics for the variables included in the main analysis, we should stress that (a) firms with imputed accounting variables and (b) firms with fewer than 10 employees have been omitted from the main sample.[6] The former choice was made to reinforce the reliability of our empirical analysis. The latter was made to allow all of the investigated firms to potentially reach the highest degree of ethnic diversity when an aggregated specification is used.[7] Altogether, we are able to analyze the productivity of approximately 28,000 firms for the years 1995–2005.

Table 1 provides basic descriptive statistics for all of the variables used in our analysis for the main sample by firm size. We split the sample into two main groups: firms above and below 50 employees. Consistent with the overall character of the Danish private sector, 78% of the observations corresponds to firms with fewer than 50 employees.[8] Compared with larger firms, smaller companies are engaged in export activities to a lesser extent and are characterized by lower levels of value added, materials and capital stock.[9] Moreover, whereas small firms feature larger shares of managers, relatively younger employees and personnel with secondary education, firms with more than 50 employees present larger proportions of women, foreigners and employees with longer tenure. The two groups of firms are comparable in terms of share of employees with a tertiary education, a key element in our empirical analysis, and in terms of firm ownership.

Table 2 reports detailed descriptive statistics for all of the diversity indices by industry, by firm size and by year. As mentioned in the previous subsection, we calculate our diversity measures using two different aggregation levels for the categories included in the indices. Cultural diversity is represented by the employees′ nationalities in the aggregate specification and by the languages they speak in a more disaggregate one. Education-related diversity is based on the

[4] Second-generation immigrants are not treated as foreigners in the main analysis. However, we employ a specification in which second-generation immigrants are included in the group of foreigners in the section on the mechanisms driving the effect of workforce diversity on firm productivity.

[5] For ethnic diversity, the shares of foreign workers of different nationalities/linguistic groups in each workplace have been calculated as follows:

$$p_{swt} = \frac{foreigners_{swt}}{foreigners_{wt}}.$$

[6] Approximately 9000 observations corresponding to almost 2000 firms are lost by limiting the sample to firms with at least ten employees. Descriptive statistics for the excluded firms are reported in Table A1 of Appendix A.2. We have also performed the main analysis including firms with fewer than ten employees (Table A2 of Appendix A.2) and we found that the results obtained from the unrestricted sample are qualitatively similar to those reported in the paper.

[7] When a linguistic classification is adopted, we adjust the ethnic diversity to take firm size into account. Specifically, we standardize the index for a maximum value equal to $(1-1/N)$ when the total number of employees (N) is lower than the number of linguistic groups (S).

[8] According to the OECD (2005), the population of Danish firms is mainly composed of small and medium-sized companies; firms with fewer than 50 employees account for 97% of firms and provide 42% of the total employment in manufacturing and services.

[9] Accounting values are reported in thousands of real DKK. Monetary values are deflated by using the GDP deflator for the base year 2000 retrieved from the World Bank database.

Table 1
Descriptive statistics, main sample and by size.

Variables	Definition	Total			Size1			Size2		
		Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd
IDA variables										
foreigners1	Non-Danish employees from Muslim countries, as a proportion of foreign employees	0.079	0	0.228	0.065	0	0.221	0.127	0	0.243
foreigners2	Non-Danish employees from East Asia, as a proportion of foreign employees	0.037	0	0.156	0.031	0	0.153	0.06	0	0.166
foreigners3	Non-Danish employees from Other Asia, as a proportion of foreign employees	0.024	0	0.123	0.018	0	0.118	0.044	0	0.152
foreigners4	Non-Danish employees from formerly communist countries, as a proportion of foreign employees	0.037	0	0.156	0.031	0	0.154	0.059	0	0.164
foreigners5	Non-Danish employees from West/South Europe, as a proportion of foreign employees	0.257	0	0.395	0.212	0	0.384	0.41	0.344	0.391
foreigners6	Non-Danish employees from Africa, as a proportion of foreign employees	0.010	0	0.069	0.006	0	0.067	0.013	0	0.075
foreigners7	Non-Danish employees from North America/Oceania, as a proportion of foreign employees	0.014	0	0.098	0.012	0	0.096	0.013	0	0.077
foreigners8	Non-Danish employees from Central/South America, as a proportion of foreign employees	0.008	0	0.075	0.007	0	0.074	0.013	0	0.075
Males	Men, as a proportion of all employees	0.715	0.785	0.232	0.722	0.8	0.237	0.692	0.736	0.21
age1	Employees aged 15–32, as a proportion of all employees	0.243	0.2	0.191	0.254	0.214	0.195	0.204	0.157	0.169
age2	Employees aged 33–41, as a proportion of all employees	0.238	0.231	0.123	0.234	0.222	0.131	0.253	0.245	0.095
age3	Employees aged 42–50, as a proportion of all employees	0.192	0.187	0.107	0.186	0.176	0.113	0.21	0.214	0.077
age4	Employees aged 51–65, as a proportion of all employees	0.181	0.211	0.178	0.201	0.198	0.165	0.229	0.213	0.191
skill0	Employees with compulsory education, as a proportion of all employees	0.331	0.307	0.191	0.33	0.305	0.193	0.329	0.321	0.178
skill1	Employees with a secondary/post-secondary education, as a proportion of all employees	0.632	0.653	0.178	0.635	0.666	0.184	0.621	0.628	0.157
skill2_1	Employees with a humanities tertiary education, as a proportion of all employees	0.01	0	0.031	0.011	0	0.031	0.009	0	0.025
skill2_2	Employees with a scientific tertiary education, as a proportion of all employees	0.003	0	0.021	0.003	0	0.018	0.005	0	0.024
skill2_3	Employees with a social sciences tertiary education, as a proportion of all employees	0.016	0	0.052	0.015	0	0.051	0.021	0	0.055
skill2_4	Employees with a engineering tertiary education, as a proportion of all employees	0.01	0	0.047	0.01	0	0.047	0.012	0	0.046
Tenure	Average tenure	4.541	4.91	0.49	4.363	4.627	0.501	5.153	5.752	4.507
Managers	Managers, as a proportion of all employees	0.045	0.021	0.058	0.045	0	0.063	0.037	0.025	0.039
Middle managers	Middle managers, as a proportion of all employees	0.166	0.087	0.208	0.153	0.076	0.204	0.211	0.138	0.214
bluecoll	Blue-collars, as a proportion of all employees	0.741	0.829	0.276	0.739	0.811	0.265	0.75	0.832	0.289
Index_ethnic_aggr	Diversity index based on employees′ nationality (8 categories)	0.102	0	0.207	0.054	0	0.159	0.265	0.244	0.263
Index_edu_aggr	Diversity index based on employees′ education (6 categories)	0.415	0.444	0.111	0.407	0.436	0.115	0.442	0.468	0.090
Index_demo_aggr	Diversity index based on employees′ demographic characteristics (8 categories)	0.742	0.759	0.088	0.731	0.75	0.090	0.779	0.790	0.068

Index_ethnic_disaggr	Diversity index based on employees' language (40 categories)	0.121	0	0.407	0.058	0	0.374	0.336	0.432	0.440
Index_edu_disaggr	Diversity index based on employees' education (9 categories)	0.573	0.578	0.131	0.558	0.562	0.134	0.622	0.624	0.109
Index_demo_disaggr	Diversity index based on employees' demographic characteristics (10 categories)	0.868	0.886	0.086	0.857	0.876	0.088	0.907	0.919	0.066
Accounting variables										
Total sales	(1000 kr.)	106,455.6	24,424.69	773,081	33,501.14	17,806.29	102,538	357,089	119,981.1	1,591,768
Intermediates	(1000 kr.)	77,091.01	14,369.88	626,067.9	24,598.44	10,012.36	96,835.73	257,427.8	73,243.01	1,290,107
Capital	(1000 kr.)	45,632.38	3,661.496	978,689.2	11,263	2,516.854	543,658	163,708	21,980.64	1,793,078
Export	1, if the firm export	0.418	0	0.493	0.343	0	0.475	0.676	0.761	0.467
foreign_ownership	1, if the firm is foreign owned	0.004	0	0.059	0.004	0	0.06	0.003	0	0.058
Multi	1, if the firm is multi-establishment	0.118	0	0.323	0.032	0	0.178	0.413	0	0.492
Observations		126,788			98,203			28,585		

Note: All IDA and accounting variables are expressed as time averages from 1995 to 2005. The industrial sectors included in the empirical analysis are the following: food, beverages and tobacco (3.43%); textiles (1.81%), wood products (5.82%), chemicals (3.08%), other non-metallic mineral products (1.28%), basic metals (16.68%), furniture (3.11%), construction (19.69%), sale and repair of motor vehicles (4.16%), wholesale trade (14.11%), retail trade (8.92%), hotels and restaurants (3.10%), transport (4.59%), financial intermediation (1%) and business activities (9.22%). Small size (Size 1): employees ≤ 49; middle and big size (Size 2): employees ≥ 50.

Table 2
Descriptive statistics of diversity indices by industry, by size and by year.

	Aggregate specification				
	Manufacturing	**Construction**	**Wholesale and retail trade**	**Transport**	**Financial and business services**
Index ethnic	0.154	0.031	0.077	0.080	0.138
Index edu	0.430	0.387	0.402	0.437	0.449
Index demo	0.766	0.716	0.730	0.738	0.751
Observations	45,766	24,605	37,991	5,841	12,667
	Small size	**Middle size**	**Big size**	**1,995**	**2,005**
Index ethnic	0.030	0.080	0.259	0.081	0.121
Index edu	0.399	0.417	0.441	0.427	0.414
Index demo	0.717	0.747	0.777	0.735	0.739
Observations	50,564	46,630	29,676	7,461	13,819
	Disaggregate specification				
	Manufacturing	**Construction**	**Wholesale and retail trade**	**Transport**	**Financial and business services**
Index ethnic	0.191	0.046	0.080	0.083	0.156
Index edu	0.578	0.503	0.573	0.567	0.698
Index demo	0.891	0.847	0.854	0.865	0.877
Observations	45,766	24,605	37,991	5,841	12,667
	Small size	**Middle size**	**Big size**	**1,995**	**2,005**
Index ethnic	0.028	0.090	0.328	0.081	0.137
Index edu	0.545	0.573	0.622	0.546	0.588
Index demo	0.841	0.876	0.906	0.862	0.866
Observations	50,564	46,630	29,676	7,461	13,819

Note: Small size: employees ≤ 49; middle size: $50 \leq$ employees ≤ 99; big size: employees ≥ 100.

information concerning the employees′ highest educational level completed. In the aggregate index, we distinguish between different types of tertiary education while in a more disaggregate one, we also make a field-related distinction at the level of secondary education. Finally, the demographic index is based on gender and age quartiles or quintiles, depending on the level of aggregation. We observe greater diversity for firms within the manufacturing and the financial and business service industries, whereas small firms present lower diversity in all dimensions no matter the level of aggregation used. Finally, diversity is increasing slightly over time, especially in ethnicity. This result is consistent with the increasing migration to Denmark observed in recent decades.

3. Empirical strategy

3.1. Productivity estimation

As highlighted in the literature concerning the identification of the firm production function, the major issue in the estimation of input parameters is the possibility that there are factors influencing production that are unobserved by the econometrician but observed by the firm. In such a case, firms may use asymmetrically observed shocks to maximize their profits or to minimize their costs. More specifically, it is expected that firms respond to positive (negative) productivity shocks by expanding (reducing) their output, which requires a higher quantity and/or quality of production inputs. Thus, the OLS estimates of the coefficients of the inputs that are observed by econometricians may be biased and inconsistent, and error terms and regressors may be correlated. Moreover, it is widely acknowledged that whereas fixed-effect (FE) estimation techniques (Mundlak, 1961) consider firm heterogeneity, FE techniques do not solve the simultaneity problem when productivity shocks fluctuate over time.

Several methods to address simultaneity have been proposed, such as the structural approach advocated by Olley and Pakes (1996) (OP henceforth) and Levinsohn and Petrin (2003) (LP henceforth).[10] Both OP and LP suggest semiparametric methods based on (i) the identification of a proxy variable, which is assumed to be a function of time-varying productivity

[10] See Ackerberg et al. (2006) for a survey.

shocks (total factor productivity) and (ii) the definition of conditions under which this function is invertible. The aim is to infer the total factor productivity by using the observed firms′ input choices (Wooldridge, 2009).[11] Although OP and LP are broadly used approaches in the structural identification of the production function, they suffer from collinearity and even identification problems, as noted by Ackerberg et al. (2006) (henceforth ACF). Given the timing and dynamic implications of input choices, these researchers raise questions about the LP estimation techniques in particular. Therefore, ACF propose an estimation method built on OP and LP approaches that does not suffer from potential collinearity problems: the coefficient of labor is no longer estimated during the first stage (in a value added production function).

Gandhi et al. (2011) (henceforth GNR) have recently expressed some concerns about the fact that the ACF approach may lead to misleading inferences, as the latter is based on the value-added specification of the production function, which requires fairly specialized assumptions. They therefore propose an alternative method to ACF where the identification problem caused by flexible inputs is solved by a transformation of the firm′s short run first-order condition in a gross-output specification.

3.2. Methodological approaches

We follow two main methodological routes to investigate the link between workforce diversity and productivity. In the first one, we estimate a standard Cobb–Douglas function, calculate the implied total factor productivity and relate the latter to our labor diversity indices in a second stage. This "reduced-form" approach allows us to evaluate which dimensions of workforce diversity seem to descriptively matter but it does not formally account for the fact that labor of different types is of different quality (Hellerstein and Neumark, 2004; Iranzo et al., 2008; Fox and Smeets, 2011; Irarrazabal et al., 2014). We therefore move toward a richer production function specification that (i) accounts for different types of labor as inputs and (ii) adjusts for both quality differences and complementarity/substitutability between labor inputs. Specifically, we proceed by modeling a value-added production function that is Cobb–Douglas in capital and labor but in which the contribution of the labor aggregate also depends on different types of labor in a CES specification. One of the main advantages of the second approach compared to the reduced-form one is that the former is more structurally motivated and theoretically grounded, as it explicitly allows for flexible substitution patterns and possible quality differences between labor types. However, the structural estimation approach is computationally demanding and forces us to model labor dispersion with less flexibility and with a higher level of aggregation compared to the reduced form approach, which allows us to test a set of hypotheses and the robustness of our main results by attempting several alternative specifications as well as by looking at within occupation diversity. The reduced-form approach furthermore allows to simultaneously look at the three dimensions of diversity (ethnic, educational and demographic) within the same equation, while with the structural method we need to make the stronger assumption that each dimension of diversity enters separately into the production function. Given that both approaches have costs and benefits, we view their results as complementary.

3.2.1. Reduced form approach: the empirical association between diversity and productivity

Referring to the literature concerning the estimation of the production functions, we regard the method suggested by ACF as our main approach. Productivity is obtained from a Cobb–Douglas production function containing the value added, Y, the labor, L, and the capital, K. Because input characteristics differ across industries, the production function parameters are estimated separately for each 1-digit sector j. The log-linear firm i production function is specified as follows:

$$\ln Y_{ijt} = cons + \alpha \ln L_{ijt} + \beta \ln K_{ijt} + u_{ijt} \tag{1}$$

with $t = 1, 2, \ldots, T$. The error term u_{ijt} consists of a time-varying firm specific effect ν_{ijt} (unobserved by econometricians) and an idiosyncratic component ε_{ijt}. Following ACF, we assume that

$$E(\varepsilon_{ijt} | l_{ijt}, k_{ijt}, m_{ijt}, l_{ijt-1}, k_{ijt-1,}, m_{ijt-1}, \ldots, l_{ij1}, k_{ij1}, m_{ij1}) = 0 \tag{2}$$

where m refers to our proxy variable (materials) and lower-case letters to log-variables. Because past values of ε_{ijt} are not included in the conditioning set, we allow for serial dependence in the pure shock term. However, we need to restrict the dynamics of the productivity process:

$$E(\nu_{ijt} | \nu_{ijt-1}, \nu_{ijt-2}, \ldots, \nu_{ij1}) = E(\nu_{ijt} | \nu_{ijt-1}) = f(\nu_{ijt-1}) \tag{3}$$

[11] The approach advocated by Olley and Pakes (1996) is a two-step estimation method. In the first step, semiparametric methods are used to estimate the coefficients of the variable inputs along with the nonparametric function linking productivity to capital and investment. In the second step, parameters of capital inputs are identified based on the assumed dynamics of the productivity process (where productivity is assumed to follow a first-order Markov process, see Wooldridge, 2009). However, the OP estimation method presents two major drawbacks. First, because adjustment costs create lumpiness in the investment levels, these levels may not respond smoothly to productivity shocks. Second, the OP approach excludes firms that report zero investment levels: it induces a de facto truncation bias. To overcome these drawbacks, LP use a measure of intermediate inputs as a proxy for investment levels. This choice has many benefits. First, changes in the intermediate inputs do not typically involve adjustment costs, and the intermediate inputs therefore respond better to productivity shocks than investments. Second, the intermediate inputs provide a simple link between the estimation strategy and the economic theory because they do not typically represent state variables. Third, because intermediate inputs are almost always used in production, the LP approach circumvents the above-mentioned data truncation problem. Moreover, the LP approach suggests three specification tests for evaluating the proxy′s performance (Petrin et al., 2003). However, the coefficient of the proxy is recovered during the second stage rather than the first (as in the OP approach).

for given functions $f(\cdot)$. As in the ACF approach, we assume that the material input is selected after the labor input. As a result, material demand will be a function not only of capital and productivity but also of l:

$$m_{ijt} = f(k_{ijt}, v_{ijt}, l_{ijt}) \tag{4}$$

and assuming that the material demand function is strictly increasing in productivity shock v_{ijt}, we obtain

$$v_{ijt} = f^{-1}(k_{ijt}, m_{ijt}, l_{ijt}). \tag{5}$$

Plugging the inverse material demand into the production function, we obtain the first-stage equation, which serves here only to separate ν_{ijt} from ε_{ijt}:

$$y_{ijt} = cons + \alpha l_{ijt} + \beta k_{ijt} + f^{-1}(k_{ijt}, m_{ijt}, l_{ijt}) + \varepsilon_{ijt}. \tag{6}$$

The function $f^{-1}(\cdot)$ is proxied with a polynomial in materials, capital and labor. Therefore, the estimated net output of the idiosyncratic component is used to identify parameters for the inputs in the second stage. Recalling that ν_{ijt} is a first-order Markov process, we define a_{ijt} as an innovation that can be correlated with the current values of the proxy variable m_{ijt} and l_{ijt}

$$a_{ijt} = v_{ijt} - g(v_{ijt-1}), \tag{7}$$

where a_{ijt} is mean independent of all information known at $t-1$ and $g(\cdot,\cdot)$ is also proxied with a low-degree polynomial in the dependent variables.[12] Given our timing assumption, we proceed by using the moments

$$E[a_{ijt} | k_{ijt}, l_{ijt-1}] = 0 \tag{8}$$

to identify coefficients for k and l. Using the estimates of the production function parameters, the total factor productivity (henceforth TFP) for firm i at time t in industry j is defined as

$$TFP_{ijt} = y_{ijt} - \alpha l_{ijt} - \beta k_{ijt} \tag{9}$$

Following the computation of TFP values, the relationship between these and alternative measures of diversity can be estimated with OLS in the following equation separately for each 1-digit sector j

$$\begin{aligned} TFP_{ijt} = {} & \zeta_{0j} + \zeta_{1j}(index\_ethnic_{ijt}) + \zeta_{2j}(index\_edu_{ijt}) + \zeta_{3j}(index\_demo_{ijt}) \\ & + \zeta_{cj}(C_{ijt}) + \zeta_{tj} + \zeta_{rj} + \zeta_{nj} + \zeta_{nj} \star \zeta_{tj} + x_{ijt} \end{aligned} \tag{10}$$

where ζ_{1j}, ζ_{2j}, and ζ_{3j} measure the associations between TFP and employees′ diversity in terms of ethnicity, education and demographic characteristics, respectively; and C_{ijt} is a vector including workforce composition characteristics, such as the shares of employees belonging to each category included in our diversity indices.[13] We think that the inclusion of such shares partly control for the fact that different labor types may have different qualities. Failing to control for labor quality, the estimated contribution of diversity in the TFP equation (10) confounds the direct effects of diversity (ζ_{1j}, ζ_{2j}, and ζ_{3j}) with such quality differences. We further explicitly address the issue of labor quality in the structural estimation approach described in the next sub-section. The same vector C_{ijt} also includes the share of managers and middle managers and average firm tenure, whether the firm is foreign-owned, whether the firm exports and a multi-establishment dummy, whereas ζ_{tj}, ζ_{rj}, ζ_{nj} are time, regional and two-digit industry controls, respectively.

Independent estimations of Eq. (10) by 1-digit industry allow us to rule out the possibility that workplace diversity only reflects an industry technology choice. Factor intensities and the mix of capital and labor may vary substantially across industries. For example, some technologies might require a set of highly skilled employees working in concert with a set of mid-level employees and a set of low-skill workers. Other technologies might only require high-skill or low-skill labor. Considering industry-specific results will therefore ensure that variations in the observed diversity of education levels across firms within the same industry will also reflect cross-firm differences in the makeup of the workforce, rather than merely reflecting which type of technology the firm has chosen.[14]

[12] To keep the number of regressors manageable, we always use a fourth-degree polynomial (with interactions) in the first stage and a third-degree polynomial in ν_{ijt} to compute $g(\cdot)$.

[13] Specifically, in the aggregate specification of diversity we control for the shares of foreigners from North America and Oceania, Central and South America, Africa, Western and Southern Europe, former Communist countries, East Asia, Other Asia, and Muslim countries; the shares of employees with compulsory education, with secondary education, and with tertiary education split into 4 main categories (humanities, natural sciences, social sciences and engineering); and the shares of female and male employees belonging to various age distribution quartiles. In the disaggregate specification of diversity, we include the shares of foreigners belonging to each linguistic group, as described in Appendix A.1, and the shares of employees with different types of education and belonging to various gender-specific age distribution quintiles, as explained in Section 3.2.

[14] Prior academic research suggests that diversity leads to economic gains or losses depending on industry characteristics (Sparber, 2009, 2010). More specifically, diversity seems to increase productivity in sectors that require creative decision making, problem solving, and customer service, but ethnic diversity may decrease productivity in industries characterized by high levels of group work or teamwork and efficiency. Our current industry categorizations, however, are too rough for us to test the hypothesis, as jobs of both types (jobs that require creativity versus efficiency) are likely to be in each aggregate industry.

3.2.2. Structural estimation approach: production function with different types of labor

The reduced-form approach described above allows to identify which types of labor heterogeneity appear to descriptively matter but does not formally recognize that labor input is non-homogeneous, i.e., labor of different types is of different quality (Hellerstein and Neumark, 2004; Fox and Smeets, 2011; Irarrazabal et al., 2014). Therefore we move toward a richer value-added production function, which continues to be Cobb–Douglas in capital and a labor aggregate, but the contribution of the labor aggregate also depends on different types of labor in a CES specification. To account for differences in firms′ labor types, we use the observed shares of labor types within a given dimension, i.e. demographic or educational (for a similar specification see Fox and Smeets, 2011; Irarrazabal et al., 2014). Specifically, we use the following generalized Cobb–Douglas production function in capital and labor:

$$Y_{ijt} = A_{ijt} \cdot [L_{ijt} \cdot E(H_{1ijt} \cdots H_{Wijt})]^{\alpha} \cdot K_{ijt}^{\beta} \tag{11}$$

where W represents the maximum number of labor types in a given dimension and the term $E(H_{1jt} \cdots H_{Wjt})$ represents the overall efficiency of the labor force of firm i belonging to industry j depending on the shares of these labor types H_{wijt}.[15] Treating the overall efficiency as a CES function of the observed shares of labor types, we can express the term E as

$$E(H_{1ijt} \cdots H_{Wijt}) = [(H_{1ijt})^{\gamma} + (\eta_2 H_{2ijt})^{\gamma} + \cdots + (\eta_W H_{Wijt})^{\gamma}]^{1/\gamma} \tag{12}$$

where the parameters $\eta_2, \ldots, \eta_W$ model the relative difference in quality between worker type w and worker type 1, i.e. if workers of type w are more (less) productive than type 1, then η_w is greater (less) than one. The parameter γ governs instead the complementarity/substitutability between labor types, as the elasticity of substitution is given by $1/(1-\gamma)$. Assuming constant returns to each labor type, a parameter of γ larger than one would imply that the labor types are substitutable, the isoquants are concave, the technology is submodular and exhibits a taste for employing workers of different types (Grossman and Maggi, 2000; Bombardini et al., 2012). Put it differently, if γ is larger than one, then dispersion of labor types increases productivity and it is optimal to combine workers of different types along a specific dimension (Iranzo et al., 2008). By contrast, if γ is smaller than one, there is complementarity (or imperfect substitutability) between labor types (Iranzo et al., 2008). This means that the technology is supermodular, the isoquants are convex and dispersion in labor types has a negative effect on productivity.

As in the case of the standard production function (1), we estimate Eq. (11) with the ACF approach. Specifically we use the current value of capital, the lagged values of labor and the shares of workers for each labor type w to form a set of the following moment conditions:

$$E[\tilde{a}_{ijt} | k_{ijt}, l_{ijt-1}, H_{1ijt-1}, \ldots, H_{Wijt-1}] = 0 \tag{13}$$

to identify coefficients for k, l and the parameters γ and $\eta_2, \ldots, \eta_W$. As in Section 3.2.1, we assume that the innovation term $\tilde{a}_{ijt}$ to the productivity shock $\tilde{v}_{ijt} = f^{-1}(k_{ijt}, m_{ijt}, l_{ijt}, H_{1ijt}, \ldots, H_{Wijt})$ is mean independent of all information known at $t-1$.

4. Results from the reduced form approach

4.1. Main results

As mentioned in the previous section, in the reduced form approach measures of TFP are computed as residuals from the first step estimation in which the firms′ value added is regressed on their capital and labor stocks. The industry-specific elasticities of capital and labor obtained by implementing the ACF approach are reported in the first panel of Table 3. In the second and third panel of Table 3, we also report the same elasticities estimated using two alternative approaches to ACF. The first one (OP) allows for the control of sample selection issues and deals with firm exit.[16] The second method (GNR) employs a gross-output instead of a value-added production function. Comparing the estimated elasticities across these methods, we find that the OP and GNR estimates of the labor coefficients are more often smaller than their ACF counterparts while the capital elasticities obtained from OP (GNR) are larger (smaller). These results are generally in line with what has been found in Ackerberg et al. (2006) and Gandhi et al. (2011).

The main results for the second step using three alternative measures of TFP (ACF, OP and GNR) are shown separately in the three panels of Table 4. As explained in Section 3, we describe the empirical association between labor heterogeneity and firm productivity, using two different aggregation levels for the categories included in our diversity indices. The results obtained using the more aggregated level are shown in the first sub-panel of each TFP panel, whereas the results obtained using the disaggregated categories are presented in the second sub-panel.[17] All of the estimated coefficients of our diversity measures are reported in standard deviation units in order to compare the relative contributions of each dimension of diversity, thereby easing the comparison across magnitudes.

[15] We have to make sure that every $0 < H_{wijt} < 1$, as H_{wijt} taking the value 0 or 1 is not consistent with the Cobb–Douglas framework. We sort this problem out by adding or subtracting a small ε to any $H_{wijt} = 0$ or $H_{wijt} = 1$ (Irarrazabal et al., 2014).

[16] We have also investigated whether firm diversity plays a role in terms of firm survival. In most industries, our diversity indices are not significantly correlated with firm probability of exiting the market. The results are reported in Table A3 of Appendix A.2.

[17] The estimated coefficients on the share variables used as controls in the aggregate specification of diversity are reported in Tables A4.1–A4.3 of Appendix A.2.

Table 3
Production function estimates.

Dependent variable: log of value added	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
ACF					
Log(*L*)	0.913***	0.928***	0.984***	0.842***	0.906***
	(0.020)	(0.024)	(0.209)	(0.104)	(0.114)
Log(*K*)	0.102***	0.115***	0.089***	0.140***	0.156***
	(0.005)	(0.024)	(0.032)	(0.020)	(0.009)
Observations	36,079	18,095	26,604	4,041	8,097
R2	0.865	0.836	0.711	0.797	0.713
OP					
Log(*L*)	0.828***	0.955***	0.809***	0.697***	0.792***
	(0.014)	(0.006)	(0.009)	(0.020)	(0.014)
Log(*K*)	0.123***	0.157***	0.161***	0.243***	0.029
	(0.029)	(0.012)	(0.036)	(0.037)	(0.078)
Observations	10,684	6,589	9,225	1,528	2,985
R2	0.812	0.805	0.698	0.783	0.694
GNR					
Log(*L*)	0.487***	0.467***	0.246***	0.528***	0.538***
	(0.032)	(0.020)	(0.015)	(0.013)	(0.047)
Log(*K*)	0.096***	0.047***	0.055***	0.092***	0.069***
	(0.023)	(0.014)	(0.017)	(0.014)	(0.034)
Log(Intermediates)	0.441***	0.491***	0.546***	0.383***	0.182***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	35,387	17,823	26,016	3,928	7,626
R2	0.852	0.905	0.848	0.883	0.848

Note: The dependent variable is the log of value added in the first two sub-panels (ACF, OP). In the last sub-panel (GNR), the dependent variable is the log of total sales. Standard errors are computed by using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-firm correlation.
Significance levels: ***1%, **5%, *10%.

Ethnic diversity is generally negatively associated with firm TFP, whereas the coefficients of educational and demographic diversity are positive. The empirical associations between educational diversity are precisely estimated in all industries but transport, whereas the coefficient of ethnic diversity is statistically significant only in the case of manufacturing, construction and wholesale trade. Demographic diversity is never significantly correlated with firm productivity. All of these results are qualitatively robust across diversity specifications and are not substantially affected by the measure of TFP employed, although the estimated correlations between ethnic diversity and productivity are slightly smaller in both the OP and GNR specifications compared to the ACF counterpart. For the sake of brevity, we proceed by discussing the results for the manufacturing and service industries and those obtained using TFP calculated with the ACF approach only. In the manufacturing sector, a standard deviation increase in ethnic diversity is associated with a decrease in firm TFP by 1.3%, while a standard deviation increase in educational diversity is associated with an increase in firm TFP by 1%, when an aggregated index is considered. If we focus on the disaggregated index instead, a standard deviation increase in ethnic diversity is associated with a decrease in firm TFP by 1.6%, while a standard deviation increase in educational diversity is associated with an increase in firm TFP by 2.9%. The magnitudes involved in the wholesale and retail trade are qualitatively similar to those estimated in the manufacturing sector.

4.2. *Testing alternative hypotheses*

In the next steps, we attempt to assess which mechanisms drive our results on employee diversity and firm TFP by exploiting the variation in occupations, nationalities and industry characteristics. While these exercises provide useful information on the channels through which diversity is associated with productivity, it is important to underline that they are not conclusive evidence of a particular mechanism. To simplify the presentation of these exercises, we discuss the results for the manufacturing and the wholesale and retail trade industry only[18] and focus on the disaggregated indices.[19]

First, we separately calculate the diversity indices for white- and blue-collar occupations and include them all in the same specification. We use this strategy based on the supposition that diversity could play a different role for distinct

[18] The results for the other industries are reported in Table A5 of Appendix A.2.

[19] The results obtained using the aggregate indices are qualitatively similar to those obtained using the detailed categorization and are reported in Table A6 of Appendix A.2.

Table 4
Labor diversity and firm total factor productivity, main results.

Dependent variable: log of TFP	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
TFP (ACF)					
Index ethnic aggr	−0.013***	−0.012**	−0.033***	−0.009	−0.011
	(0.003)	(0.005)	(0.006)	(0.018)	(0.008)
Index edu aggr	0.014**	0.010*	0.010**	0.048	0.017**
	(0.006)	(0.006)	(0.004)	(0.027)	(0.008)
Index demo aggr	0.023	−0.026	−0.004	0.035	0.018
	(0.013)	(0.015)	(0.005)	(0.022)	(0.012)
Observations	35,887	18,024	26,418	4,007	7,931
R2	0.281	0.235	0.553	0.185	0.347
Index ethnic disaggr	−0.016***	−0.012**	−0.015***	−0.008	0.001
	(0.003)	(0.005)	(0.004)	(0.008)	(0.006)
Index edu disaggr	0.029***	0.012*	0.053***	0.007	0.054***
	(0.007)	(0.007)	(0.006)	(0.022)	(0.013)
Index demo disaggr	0.021	−0.027	−0.016	0.032	−0.010
	(0.011)	(0.015)	(0.009)	(0.019)	(0.012)
Observations	35,887	18,024	26,418	4,007	7,931
R2	0.290	0.247	0.558	0.203	0.361
TFP (OP)					
Index ethnic aggr	−0.008*	−0.007*	−0.015**	−0.038	−0.005
	(0.005)	(0.004)	(0.008)	(0.024)	(0.014)
Index edu aggr	0.012*	0.011*	0.031***	0.063*	0.032**
	(0.007)	(0.007)	(0.008)	(0.029)	(0.013)
Index demo aggr	0.030	−0.020	0.010	0.080	0.014
	(0.018)	(0.016)	(0.007)	(0.042)	(0.014)
Observations	10,662	6,572	9,199	1,526	2,985
R2	0.268	0.230	0.432	0.253	0.351
Index ethnic disaggr	−0.014***	−0.008**	−0.012***	−0.003	−0.001
	(0.005)	(0.003)	(0.005)	(0.011)	(0.014)
Index edu disaggr	0.029***	0.014*	0.053***	0.033	0.051**
	(0.008)	(0.008)	(0.008)	(0.025)	(0.016)
Index demo disaggr	0.028	−0.018	0.009	0.076	−0.002
	(0.018)	(0.011)	(0.007)	(0.046)	(0.016)
Observations	10,662	6,572	9,199	1,526	2,985
R2	0.271	0.245	0.444	0.271	0.364
TFP (GNR)					
Index ethnic aggr	−0.007***	−0.008**	−0.011**	−0.012	−0.004
	(0.002)	(0.003)	(0.004)	(0.014)	(0.009)
Index edu aggr	0.008**	0.009**	0.015***	0.027	0.028**
	(0.004)	(0.003)	(0.003)	(0.018)	(0.011)
Index demo aggr	0.007	−0.017	−0.006	0.040	0.003
	(0.004)	(0.013)	(0.004)	(0.029)	(0.012)
Observations	35,477	17,769	26,086	3,954	7,818
R2	0.262	0.262	0.513	0.331	0.265
Index ethnic disaggr	−0.009***	−0.008**	−0.010***	−0.008	−0.001
	(0.002)	(0.003)	(0.002)	(0.008)	(0.006)
Index edu disaggr	0.007*	0.010**	0.031***	0.017	0.067***
	(0.004)	(0.004)	(0.004)	(0.019)	(0.015)
Index demo disaggr	0.009	−0.016	0.008	0.039	−0.020
	(0.006)	(0.010)	(0.007)	(0.023)	(0.013)
Observations	35,477	17,769	26,086	3,954	7,818
R2	0.279	0.252	0.52	0.344	0.278

Note: The dependent variable is the log of total factor productivity estimated by using the ACF, OP and GNR approach, alternatively. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

Table 5
Testing alternative hypotheses: estimates by using occupation-specific diversity, alternative index definitions, and "subtract-out" groups.

Dependent variable: log of TFP	*Occupation specific diversity*		2nd gen. Imm. as foreigners	University graduates as natives	Germanic group as natives
	White collar	Blue collar			
Manufacturing					
Index ethnic disaggr (1)			−0.014**	−0.004	−0.006**
			(0.004)	(0.003)	(0.003)
Index ethnic disaggr (2)	−0.007**	−0.014***	−0.005*	−0.015***	−0.012***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Index edu disaggr	0.010**	0.008	0.019**	0.029***	0.029***
	(0.003)	(0.006)	(0.007)	(0.007)	(0.007)
Index demo disaggr	−0.002	−0.007	−0.008	−0.010	−0.011
	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)
Hypothesis tests (chi2, p-value)					
Index ethnic (white collar)=index ethnic (blue collar)	5.57; 0.047				
Index edu (white collar)=index edu (blue collar)	2.67; 0.087				
Index demo (white collar)=index demo (blue collar)	0.521; 0.322				
Index ethnic (1)=index ethnic (2)			6.78; 0.035	7.01; 0.021	5.54; 0.048
Observations	35,887		35,887	35,887	35,887
R2	0.29		0.291	0.291	0.291
Wholesale and retail trade					
Index ethnic disaggr (1)			−0.041***	−0.004	−0.002
			(0.006)	(0.003)	(0.003)
Index ethnic disaggr (2)	−0.014**	−0.038***	−0.005*	−0.039***	−0.039***
	(0.004)	(0.006)	(0.003)	(0.006)	(0.006)
Index edu disaggr	0.039***	0.018***	0.054***	0.054***	0.054***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Index demo disaggr	−0.015	−0.016	−0.015	−0.014	−0.014
	(0.014)	(0.015)	(0.009)	(0.009)	(0.008)∏
Hypothesis tests (chi2, p-value)					
Index ethnic (white collar)=index ethnic (blue collar)	9.77; 0.000				
Index edu (white collar)=index edu (blue collar)	6.07; 0.042				
Index demo (white collar)=index demo (blue collar)	0.06; 0.812				
Index ethnic (1)=index ethnic (2)			25.99; 0.000	22.29; 0.000	20.89; 0.000
Observations	26,418		26,418	26,418	26,418
R2	0.56		0.56	0.56	0.56

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Index ethnic disaggr (1) is calculated as in the main analysis whereas index ethnic disaggr (2) is separately calculated for white- and blue-collar workers in columns 1 and 2, or by excluding/including a specific group of employees, as described in the column head, in columns 3 and 4. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

occupational groups and could consequently have varying effects on firm productivity. In particular, we expect that diverse problem-solving abilities and creativity will generate higher productivity for white-collar occupations than for blue-collar occupations.[20] Second, we exclude or include certain groups of foreigners in calculating ethnic diversity to test the importance of communication costs and the costs of "cross-cultural dealing." The results regarding the association between diversity and firm productivity, calculated separately for the two occupational groups and included in the same regression are reported in the first two columns of Table 5. Our results show that the correlation of educational diversity with firm productivity is indeed much larger for white-collar occupations than for blue-collar ones. Moreover, the negative coefficient

[20] This is grounded on the fact that (i) white-collar workers typically (manage) interact with a larger number of employees than blue-collar ones; (ii) white-collar employees are, on an average, more educated than employees in other occupations and are therefore more likely to access and exploit their colleagues′ knowledge heterogeneity; and (iii) white-collar employees are typically more influential in firms business plans and strategies.

of ethnic diversity among white-collar workers is lower than the coefficient associated with blue-collar occupations. Conversely, the effect of demographic diversity is insignificant for both occupational groups.[21] Therefore, our results are consistent with the creativity and knowledge spillover hypotheses proposed in the theoretical frameworks developed by Hong and Page (2001, 2004) and Berliant and Fujita (2008).

To investigate the role of "cross-cultural dealing," we exclude either foreigners with tertiary education or foreigners who speak a Germanic language. Alternatively, we include second-generation immigrants in calculating ethnic diversity. All of these groups of foreigners are likely to absorb Danish or English (which is the communication language in many businesses in Denmark) more quickly. Therefore, it is plausible that the communication costs associated with ethnic diversity may increase (decrease) after we remove (include) these foreigners, who are likely to speak Danish or English.[22] The results presented in Table 5, columns 3–5, are obtained by including both the standard ethnic diversity, as calculated in the main analysis, and an alternative one in which the second generation of immigrants is treated as non-native and where foreigners with university education or those who speak a Germanic language are included as natives, respectively. Interestingly, the coefficient of ethnic heterogeneity is larger (smaller) in absolute terms, once we exclude (include) foreigners who most likely speak Danish or English, compared to the coefficient estimated on the standard ethnic diversity.[23] These findings are in line with the hypothesis that the communication costs and the costs of "cross-cultural dealing" within ethnically heterogeneous workforces play a role in terms of firm productivity. However, the results obtained by excluding second generation immigrants may also be explained by the fact that the latter generally have stronger labor market networks in addition to lower communication costs compared to other foreigners.

We next examine the different mechanisms by which diverse workforces are associated with firm productivity by looking at different industries and firm categories.[24] First we do not find that the coefficients of the diversity indices differ for firms that are more open to trade.[25] Therefore, our analysis does not support the hypothesis that workforce diversity provides beneficial information to firms from other countries and markets and, in this way, creates positive effects on firm productivity (Osborne, 2000; Rauch and Casella, 2003). Second, we look at whether the correlation of diversity and TFP is different for firms in high-tech industries, which tend to require higher levels of creative and problem-solving activities.[26] Our results reveal that the hypothesis on creativity is not supported by this empirical exercise because the coefficients of both the education and the ethnic diversity index are not statistically different across the two groups of industries, namely those with below and above average R&D intensity. Finally, we find that the positive (negative) effects of educational (ethnic) diversity are statistically stronger (weaker) in the subsample of industries with increasing employment compared to industries with declining employment. This is in line with the hypothesis that "growing" firms are more likely to benefit from diversity because they hire younger people and foreign individuals more often than do "shrinking" firms.

4.3. Sensitivity analysis

In the next step, as a part of our sensitivity analysis, we evaluate the variations in the coefficients estimated on labor diversity that result when diversity is computed in various ways. As in the previous sub-section, we discuss the results for the manufacturing and the wholesale and retail trade industry only and focus on the disaggregated indices.[27] In particular, we use two alternative diversity indices: the Shannon–Weaver entropy index and the richness index. The entropy index is considered to be one of the most profound and useful diversity indices in biology (Maignan et al., 2003). The richness index includes a number of categories observed for each dimension of interest; it does not include the "evenness" dimension. The results are shown in Table 6, columns 1 and 2, and both sets of results are consistent with our main findings. Next, we include a Herfindahl index for the type of tertiary education (this index now has only 4 categories: engineering, natural sciences, social sciences and humanities) and the standard deviation of the years of education and age. This allows us to treat age as a cardinal variable, and to disentangle the effects associated with the amount of education from those related to the type of tertiary education. Table 6, column 3, reports statistically insignificant coefficients on both standard deviations and an estimate on educational diversity (the Herfindahl index for the type of tertiary education) that is in line with the finding discussed in the main analysis. These findings, together with the results obtained from occupation-specific diversity (Table 5, columns 1 and 2), may suggest that the positive association between educational diversity and firm TFP is mainly driven by the white-collar occupations, who are very likely to have tertiary education, or by workers with different types of tertiary education.

---

[21] Hypothesis testing also reveals that the coefficients of diversity for white-collar occupations are statistically different from the coefficients of diversity for blue-collar occupations, especially for the ethnic dimension.

[22] According to the existing literature, individuals have an easier time acquiring a foreign language if their mother language is linguistically closer to the foreign language (Isphording and Otten, 2011; Chiswick and Miller, 1995).

[23] Hypothesis testing reveals that the two estimated coefficients are significantly different from one another.

[24] The results of these additional empirical exercises are all reported in Table A7 of Appendix A.2.

[25] Specifically we augment Eq. (2) by including interactions between the export dummy and diversity in ethnic, educational and demographic dimensions and we find that the interaction terms are not precisely estimated and are not jointly significantly different from zero.

[26] Specifically, we divide industries into two groups defined by whether their aggregate level of R&D expenditure as a share of employment is above or below the median recorded for the overall economy.

[27] The results for the aggregated indices are reported in Table A8 of Appendix A.2.

Table 6
Robustness checks: estimates by size, under alternative samples and index definitions.

Dependent variable: log of TFP	Shannon entropy index	Richness	Edu and demo diversity as sd	Estimates by firm size			"Copenhagen" county is excluded	Mono-establishment firms
				Small firms	Medium firms	Large firms		
Manufacturing								
Index ethnic disaggr	−0.019***	−0.014***	−0.015***	−0.009*	−0.009*	−0.030***	−0.017***	−0.012***
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)
Index edu disaggr	0.045***	0.017**	0.012***	0.013*	0.025**	0.093***	0.028***	0.019**
	(0.009)	(0.006)	(0.004)	(0.007)	(0.011)	(0.014)	(0.007)	(0.006)
Sd(years of education)			0.004					
			(0.003)					
Index demo disaggr	0.016	0.017		0.011	0.025	−0.006	0.013	0.007
	(0.010)	(0.012)		(0.007)	(0.015)	(0.014)	(0.009)	(0.007)
Sd(age)			−0.002					
			(0.005)					
Share of men			0.231***					
			(0.025)					
Hypothesis tests (chi2, p-value)								
Index edu = sd(years of education)			6.78; 0.031					
Sd (age) = share of men			82.31; 0.000					
Observations	35,887	35,887	35,887	9,727	13,446	12,714	34,763	31,279
R2	0.292	0.289	0.286	0.287	0.304	0.304	0.293	0.289
Wholesale and retail trade								
Index ethnic disaggr	−0.034***	−0.046***	−0.015***	−0.011*	−0.013***	−0.041***	−0.015***	−0.009**
	(0.006)	(0.008)	(0.004)	(0.006)	(0.004)	(0.010)	(0.004)	(0.004)
Index edu disaggr	0.051***	0.017**	0.033***	0.049***	0.057***	0.083***	0.051***	0.054***
	(0.010)	(0.009)	(0.006)	(0.009)	(0.010)	(0.019)	(0.007)	(0.007)
Sd(years of education)			0.007					
			(0.004)					
Index demo disaggr	−0.019	−0.004		−0.007	−0.009	−0.005	−0.013	−0.008
	(0.012)	(0.006)		(0.007)	(0.009)	(0.013)	(0.008)	(0.005)
Sd(age)			−0.010					
			(0.007)					
Share of men			0.142***					
			(0.025)					
Hypothesis tests (chi2, p-value)								
index edu = sd(years of education)			15.84; 0.000					
Sd (age) = share of men			25.42; 0.000					
Observations	26,418	26,418	26,418	10,042	10,492	5,884	24,329	21,720
R2	0.559	0.558	0.56	0.485	0.505	0.506	0.567	0.526

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

We then divide firms by size and evaluate whether there is any change in the coefficients of workforce diversity for small firms (those with fewer than 50 employees), medium-sized firms (those with 50–100 employees) and large firms (those with more than 100 employees).[28] The effects of diversity could be more beneficial to larger firms because the organizational and diversity management practices of such firms are well established and formalized, and thus, they are more likely to introduce policies that can help to sustain a diversified workforce and to counteract the potential costs associated with diversity. On one hand, the coefficients on the ethnic diversity index are negative for differently sized firms,

[28] It is important to clarify that the scope for diversity is not mechanically increasing in firm size. This can be explained with a simple example. Let us assume that there are 5 possible categories of employees, and let us compare two firms with 10 and 100 employees, respectively. The two firms would feature exactly the same level of diversity, if their workforces equally represented all possible categories, i.e. if there were 2 and 20 employees for each category in the first and second firm, respectively. For both firms the diversity index would equal $(1-((0.2)^2*5))$.

with the largest coefficient associated with large firms, as reported in columns 4–6 of Table 6. On the other hand, educational diversity is more important for large than for medium-sized firms. We can therefore conclude that larger firms are more likely to benefit from educational diversity but the latter are not necessarily more successful in counteracting the costs of ethnic diversity compared to smaller firms.

Given that large cities usually host many immigrants and highly skilled workers and also house a high percentage of productive firms, we conduct an additional sensitivity check by removing the only real agglomeration area in Denmark, namely Copenhagen and its environs. The results concerning this robustness check are reported in column 7 of Table 6 and do not qualitatively differ from the main results.

Furthermore, because labor diversity has been computed at the firm level (by weighting the average of the Herfindahl indices computed at the workplace level), we evaluate how the results change if multi-establishment firms are excluded from the sample. The last column of Table 6 reports the results. These findings do not significantly differ from the main results.

4.4. Endogeneity

Even if it is not clear whether firms always have control over their workforce diversity,[29] we cannot completely rule out that firms endogenously choose the level of labor diversity in order to improve their productivity. This would imply that the findings in the previous section are likely to be biased estimates. We therefore attempt a causal effect analysis by using an instrumental variable (IV) strategy to address these potential endogeneity issues. More specifically, we instrument our diversity variables with indices of workforce diversity in cultural background, education and demographic characteristics, computed at the commuting area,[30] where the firm is located.[31] Given that diversity in a given commuting area may be a function of the current firms′ demand for diversity, we predict the current composition of the labor supply at the commuting area level by using its historical composition and the current population stocks (for similarly computed instruments see Card and Di Nardo, 2000; Dustmann et al., 2005; Cortes, 2008; Foley and Kerr, 2011). Pre-existing workforce diversity at the commuting area level may be not correlated with a firm′s current labor demand and productivity, if measured with a sufficient time lag.[32] In particular we use workforce composition at the commuting areas from the year 1990.[33] In this approach, for example, the instrument for ethnic diversity is calculated using the predicted share of immigrants from country c and living in a commuting area l at time t, $\hat{m}_{clt}$. The latter is computed using the early 1990′s stock of immigrants from country c living in l and its current population of immigrants at time t:

$$\hat{m}_{clt} = \frac{stock_{cl1990}}{\sum_{c=1}^{C} stock_{clt}} \tag{14}$$

We believe that diversity at the commuting area level may represent a suitable supply driven instrument for workplace level diversity[34] because commuting areas in Denmark (except for the area around Copenhagen)[35] are typically relatively

[29] It seems reasonable to assume that the hiring and firing costs for labor or the fixed costs of changing the workforce characteristics can generally last longer than a period. This suggests that the workforce composition and diversity are very likely to be persistent over time. Therefore, firms may not promptly respond to TFP shocks with immediate changes to their diversity mix.

[30] The so-called functional economic regions or commuting areas are identified using a specific algorithm based on the following two criteria: firstly, a group of municipalities constitutes a commuting area if the interaction within the group of municipalities is high compared to the interaction with other areas; second, at least one municipality in the area must be a center, i.e., a certain share of the employees living in the municipality must also work in the municipality (Andersen, 2000). In total, 104 commuting areas are identified.

[31] Unfortunately, in our data-set, it is not possible to observe in which area each establishment of a multi-establishment firm is located. For multi-establishments firms, location information is only provided at the headquarter level. However, we do not think this represents a serious problem as multi-establishment firms constitute only 11% of our sample. This is also confirmed by the fact that estimating our IV models on the sub-sample of mono-establishment firms provides qualitatively similar results to the ones obtained from the full sample. These additional results are reported in Table A9 of Appendix A.2. Note that the estimation of the parameters on diversity cannot be carried out for transport sector because of the low number of observations.

[32] Reverse causality however may be still an issue for big firms employing a large fraction of the local labor force. However, running our IV models on the subsample of firms with fewer than 50 employees provides qualitatively similar results to those obtained from the full sample, as shown in Table A9 of Appendix A.2. Note that the estimation of the parameters on diversity cannot be carried out for transport sector because of the low number of observations.

[33] We choose the year 1990 as a historical base for our predictions because we believe that the lag of 5–13 years should be a sufficient lag for the purposes of our IV construction. In addition, the development in immigration to Denmark also supports the choice. The 1980s and 1990s were characterized by rather restrictive immigration policy with respect to economic migrants from countries outside the European Union (EU), which made it difficult for firms in Denmark to hire applicants from the international pool of applicants (due to the consequences of the oil crisis). Immigration to Denmark from those countries from the 1980s to the mid-1990s was characterized by immigration on the basis of humanitarian reasons and family reunion. However, since then, Denmark has further tightened its immigration policy (even laws concerning family reunification and asylum). In particular since the 2001 election, in which the right-wing Danish People′s Party (DF) with its anti-immigration agenda acquired significant political power, Denmark′s immigration policy became one of the strictest in the world. For firms, it meant almost zero possibilities to hire international workers from countries outside the EU, which has often been criticized by the Confederation of Danish Industry (DI). Given these historical developments, we decided to use shares of immigrants from 1990 as a base for our predictions.

[34] Summary statistics of our supply driven instruments reveal that they are persistent over time, i.e. most of the variation is between commuting areas and that the overtime within variation is modest. Specifically, the overall, between, and within variations of the commuting area ethnic diversity are 0.163, 0.161, and 0.026, respectively. The same statistics are 0.072, 0.071, and 0.012 for the commuting area educational diversity and 0.045, 0.043, and 0.01 for the commuting area demographic diversity. Similar descriptive statistics are obtained by excluding Copenhagen and its environs.

[35] Excluding firms located in Copenhagen and its environs from the IV estimations provides similar results to those obtained from the main sample, as shown in Table A9 of Appendix A.2.

Table 7
IV results: first stage regressions.

Dependent variable: firm level diversity	Index ethnic disaggr	Index edu disaggr	Index demo disaggr
Manufacturing			
Index ethnic com	1.601***	−0.120**	−0.056**
	(0.177)	(0.051)	(0.028)
Index edu com	−0.884***	0.273***	−0.190***
	(0.234)	(0.066)	(0.051)
Index demo com	−2.134***	−0.039	1.009***
	(0.342)	(0.196)	(0.097)
F test (excluded instruments); *p*-value	36.83; 0.000	18.14; 0.000	29.44; 0.000
Observations	48,238	48,238	48,238
R2	0.227	0.594	0.385
Construction			
Index ethnic com	0.351***	−0.243***	−0.038
	(0.082)	(0.062)	(0.030)
Index edu com	0.336**	0.314***	−0.035
	(0.113)	(0.077)	(0.060)
Index demo com	−0.494	0.542**	0.585**
	(0.327)	(0.173)	(0.173)
F test (excluded instruments); *p*-value	26.69; 0.000	38.90; 0.000	6.08; 0.011
Observations	26,969	26,969	26,969
R2	0.131	0.574	0.348
Wholesale and retail trade			
Index ethnic com	0.909***	−0.342***	0.166***
	(0.112)	(0.060)	(0.046)
Index edu com	0.431**	0.938***	−0.496***
	(0.135)	(0.102)	(0.057)
Index demo com	−0.192	−0.256**	1.238***
	(0.195)	(0.078)	(0.086)
F test (excluded instruments); *p*-value	19.77; 0.000	37.81; 0.000	90.43; 0.000
Observations	41,493	41,493	41,493
R2	0.187	0.570	0.481
Transport			
Index ethnic com	0.212***	0.022	−0.187**
	(0.023)	(0.236)	(0.062)
Index edu com	0.510	0.017**	0.158
	(0.486)	(0.008)	(0.126)
Index demo com	−0.389	0.202	0.576**
	(2.076)	(0.280)	(0.245)
F test (excluded instruments); *p*-value	16.70; 0.000	9.89; 0.000	25.78; 0.000
Observations	6,287	6,287	6,287
R2	0.261	0.635	0.285
Financial and business services			
Index ethnic com	1.585***	−0.407***	−0.076
	(0.230)	(0.081)	(0.058)
Index edu com	−0.021	0.903***	0.080
	(0.224)	(0.074)	(0.063)
Index demo com	−0.319	−0.225	1.027***
	(0.366)	(0.206)	(0.097)
F test (excluded instruments); *p*-value	16.81; 0.000	37.02; 0.000	25.87; 0.000
Observations	14,008	14,008	14,008
R2	0.396	0.468	0.353

Note: The dependent variable is diversity at the firm level. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the commuting area level. Significance levels: ***1%, **5%, *10%.

small, and are therefore very likely to recruit workers from a given local supply of labor, which is characterized by a certain degree of heterogeneity. This argument is further reinforced by the role of networks in the employment process (Montgomery, 1991; Munshi, 2003). Thus, firms placed in areas with high labor diversity are also more likely to employ a more diverse workforce. It is important to emphasize that although the commuting areas are not closed economies, in the sense that workers are free to move in and out, there is clear evidence of low residential mobility (Deding et al., 2009), which seems to support the appropriateness of our IV strategy.

Table 8
Labor diversity and productivity, IV results: second stage regressions.

Dependent variable: log of TFP	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
index ethnic disaggr	−0.026*	−0.038*	−0.028**	−0.031	0.009
	(0.014)	(0.019)	(0.014)	(0.084)	(0.012)
index edu disaggr	0.061**	0.037	0.095**	0.047	0.078*
	(0.028)	(0.019)	(0.040)	(0.149)	(0.038)
index demo disaggr	0.093	−0.048	−0.056	−0.085	−0.048
	(0.086)	(0.049)	(0.033)	(0.070)	(0.033)
Observations	35,887	18,024	26,418	4007	7931
R2	0.310	0.123	0.252	0.189	0.200

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Firm level diversity is instrumented by using the predicted level of diversity at the commuting area, where the firm is located. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the commuting area level.
Significance levels: ***1%, **5%, *10%.

The results of the first and the second stage of this IV exercise are shown in Tables 7 and 8, respectively. In addition to the economic motivation for the instruments presented above, their statistical validity is largely confirmed by the *F*-statistics reported in Table 7.

The estimation adopting the IV strategy yields qualitatively similar results to those reported in the main analysis (Table 4) and are in line with the conclusions drawn in the previous sub-sections. Table 8 reveals a statistically significant and positive (negative) relationship between educational (ethnic) diversity and firm productivity, especially within the manufacturing and the wholesale and retail trade sectors. However, the estimated parameters on diversity and their standard errors are generally larger compared to the results, in which diversity is treated as exogenous.

Although our instrument is based on the historical composition of the local labor supply, however, it may still be the case that our identification strategy is invalidated by the fact that firms choose locations that are historically richer in population diversity.[36] To indirectly assess the extent to which the endogeneity of location affects our results, we have also estimated the IV models on a sub-sample of firms, for which this issue may be less important, i.e. firms which enter the market before the reference year used to predict diversity at the commuting area level (1990). Unfortunately the information about the establishment year is available only for about 30% of the total sample. Findings obtained from this additional robustness check are reported in Table A9 of Appendix A.2 and they are qualitatively in line with those reported in Table 8, but they are generally less precisely estimated due to the fact that the sample size considerably shrinks. Note that the estimation of the parameters on diversity cannot be carried out for construction and transport sectors because of the low number of observations.

All in all, we think that, although our IV strategy presents some potential flaws and we should interpret our IV results with caution, this additional analysis provides us with a useful tool to assess the robustness of the directions of the associations between diversity and productivity in terms of the methodology used to estimate them. Moreover, the fact that the results are qualitatively similar whether or not instruments are used supports the initial assumption that the estimated TFP represents an exogenous first-order Markov process, i.e. these results support the assumption for the first stage of the reduce-form approach (see Eqs. (1)–(8)) to be valid.

5. Results from the structural estimation approach

The previous section has explored the empirical associations between labor diversity and TFP in a reduced-form fashion and have served to identify which types of labor heterogeneity appear to descriptively matter. On the basis of these results, it seems that labor diversity in ethnicity and education are strongly correlated with firm productivity with opposite signs, while demographic diversity is not significantly associated with firm TFP.

While our results show that diversity in both the ethnic and the educational dimensions plays an important role, we immediately recognize that labor input is non-homogeneous (Hellerstein and Neumark, 2004; Fox and Smeets, 2011; Irarrazabal et al., 2014) and that dispersion in these labor types may have an impact on firm output (Iranzo et al., 2008). Therefore in this section we discuss the results obtained by estimating a richer production function specification that takes

[36] Previous studies on firm localization (Krugman, 1991; Audretsch and Feldman, 1996; Adams and Adam, 1996; Alcacer and Chung, 2010; Delgado et al., 2010), however, have shown that the location choices are mainly driven by the access to local innovation potential and knowledge spillovers and also by the size of the local demand, the proximity to customers and suppliers and the quality of local physical infrastructure.

Table 9
Production function estimates with shares of labor types along the ethnic dimension.

Dependent variable: log of value added	Manufacturing	Wholesale and retail trade	Financial and business services
Log(L)	0.915***	0.981***	0.924***
	(0.020)	(0.209)	(0.114)
Log(K)	0.098***	0.091***	0.151***
	(0.004)	(0.032)	(0.009)
γ	0.601	0.702*	0.542**
	(0.438)	(0.305)	(0.223)
η_muslim	0.002	0.001	0.001
	(3.018)	(3.721)	(3.765)
η_East_Asia	0.993***	0.124	0.690***
	(0.138)	(0.078)	(0.134)
η_West/South_Europe	1.010***	1.431***	2.333***
	(0.188)	(0.378)	(0.630)
η_Africa	0.797***	0.675***	0.634***
	(0.270)	(0.260)	(0.280)
η_Central/South_America	0.966	0.648	0.464
	(1.037)	(0.690)	(0.763)
η_North_America/Oceania	1.016	2.776**	3.611***
	(0.937)	(1.119)	(1.424)
η_Formerly_communist_countries	0.998	0.944	0.343
	(0.818)	(0.718)	(0.600)
η_Asia	1.040**	0.856*	1.314***
	(0.501)	(0.444)	(0.548)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	6.57; 0.029	6.27; 0.031	6.78; 0.022
Observations	36,079	26,604	8097
R2	0.891	0.737	0.779

Note: The dependent variable is the log of value added. Input elasticities are estimated by using the ACF approach. Standard errors are computed by using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-firm correlation.
*** 1%.
** 5%.
* 10%.

different types of labor as inputs and directly models a set of parameters adjusting for quality differences and governing complementarity/substitutability between these different labor inputs. Specifically, we show the results obtained from two alternative specifications. In the first one, we include in the overall efficiency term E (see Eq. (12)) the shares of workers belonging to each ethnic group used in the aggregate specification of our ethnic diversity index, including natives. In this case we can re-write E as follows:

$$E(N_{ijt} \cdots F_{8ijt}) = [(N_{ijt})^{\gamma} + (\eta_1 F_{1ijt})^{\gamma} + \cdots + (\eta_8 F_{8ijt})^{\gamma}]^{1/\gamma} \tag{15}$$

where N_{ijt} is the share of natives, while $F_{1ijt}, \ldots, F_{8ijt}$ are the share of foreigners belonging to the eight categories used to construct the ethnic diversity index in the aggregate specification. Table 9 shows the estimated coefficients of the production function (11), where the overall efficiency of the labor force, E, is expressed by Eq. (15). As in the reduced form approach, production function (11) is estimated with ACF.[37] The coefficients on labor (capital) are slightly larger (smaller) but in line with those obtained from a standard specification of the production function, as reported in Table 3. These findings are consistent with Hellerstein and Neumark (2004), who state that firms′ labor heterogeneity and quality is not "a necessary component to the estimation of the rest of the production function.″ We also investigate how the introduction of the term E in the production function affects the measurement of TFP, by comparing TFP calculated from a standard production function (see Eq. (9)) with TFP calculated from the non-homogenous production function (11). Standard TFP estimates have a tendency to be larger than the non-standard TFP measures. This is clearly shown by plotting the kernel densities of standard and non-standard TFP measures (see Fig. 1). Applying Kolmogorov–Smirnov (KS) tests confirm that the distribution of standard TFP stochastically dominates the distribution of non-standard TFP.[38] The fact that standard TFP may constitute an up-ward biased measure of productivity can be explained by the fact that standard TFP measures do not take into account the term E, which allows for flexible substitution patterns and quality differences between labor types.

The estimates of $\eta_1, \ldots, \eta_W$ parameters governing relative quality difference between native and each type of foreign workers reveal that only a few groups of foreigner workers are relatively more productive than native workers, i.e. foreign workers from West/South Europe, Other Asia and North America/Oceania. The other groups are generally less productive

[37] Note that the estimation of production function (11) cannot be carried out for construction and transport sectors due to non-convergence of the estimation algorithm.
[38] The p-values associated with the null hypothesis for the equality of the distributions take value 0.000 in all cases.

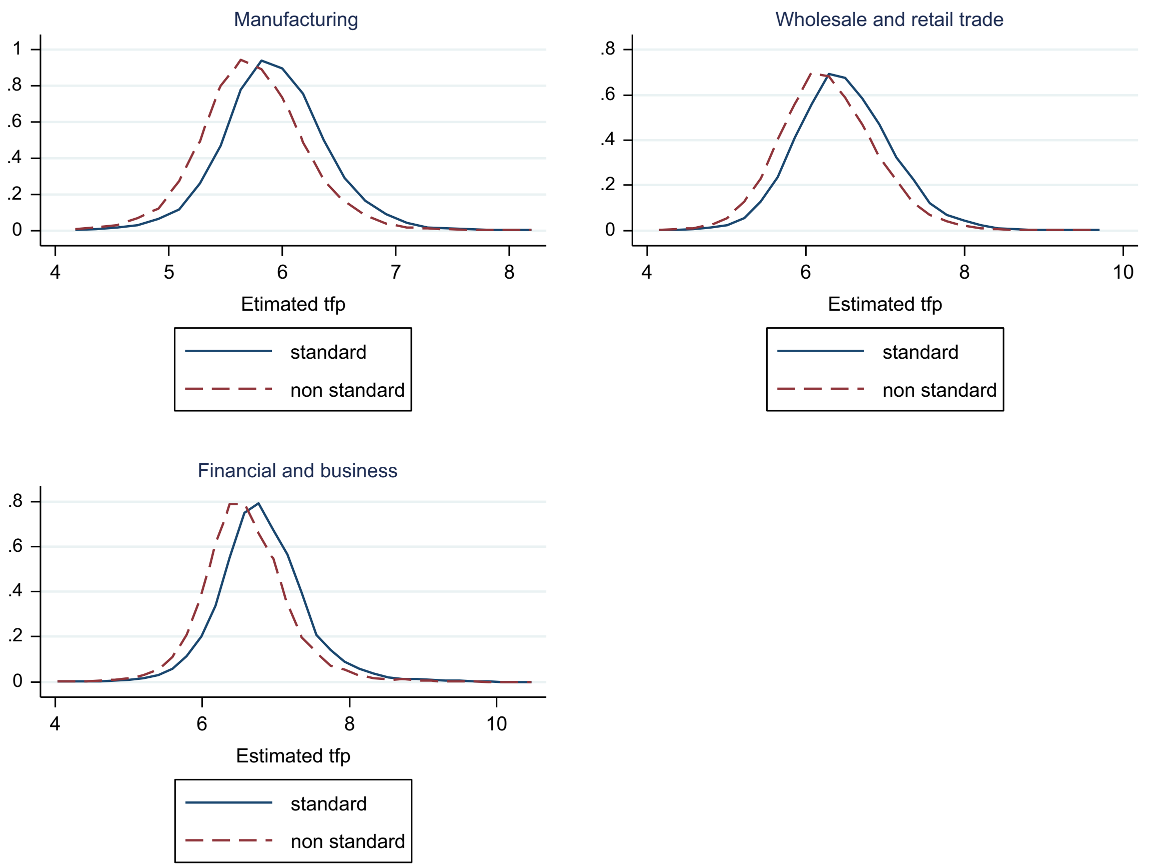

Fig. 1. Kernel density of firm TFP, calculated with or without worker heterogeneity in terms of nationality. *Note*: TFP is estimated by using the ACF approach.

Table 10
Production function estimates with shares of labor types along the educational dimension.

Dependent variable: log of value added	**Manufacturing**	**Wholesale and retail trade**	**Financial and business services**
Log(L)	0.921***	1.002***	0.976***
	(0.022)	(0.237)	(0.122)
Log(K)	0.098***	0.087***	0.155***
	(0.005)	(0.032)	(0.009)
γ	0.686***	0.764***	0.807***
	(0.073)	(0.078)	(0.064)
η_secondary/postsecondary	1.801**	1.927***	1.605***
	(0.972)	(0.337)	(0.876)
η_humanities_tertiary	0.988	1.252***	1.259
	(0.620)	(0.465)	(1.209)
η_scientific_tertiary	2.642***	2.286***	2.259***
	(0.976)	(0.479)	(0.709)
η_social_sciences_tertiary	1.346**	1.234***	1.365***
	(0.602)	(0.237)	(0.643)
η_engineering_tertiary	3.361***	3.686***	2.821**
	(0.435)	(0.681)	(1.047)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	6.49; 0.030	6.20; 0.032	6.14; 0.032
Observations	36,079	26,604	8,097
R2	0.891	0.735	0.756

Note: The dependent variable is the log of value added. Input elasticities are estimated by using the ACF approach. Standard errors are computed by using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-firm correlation.
Significance levels: ***1%, **5%, *10%.

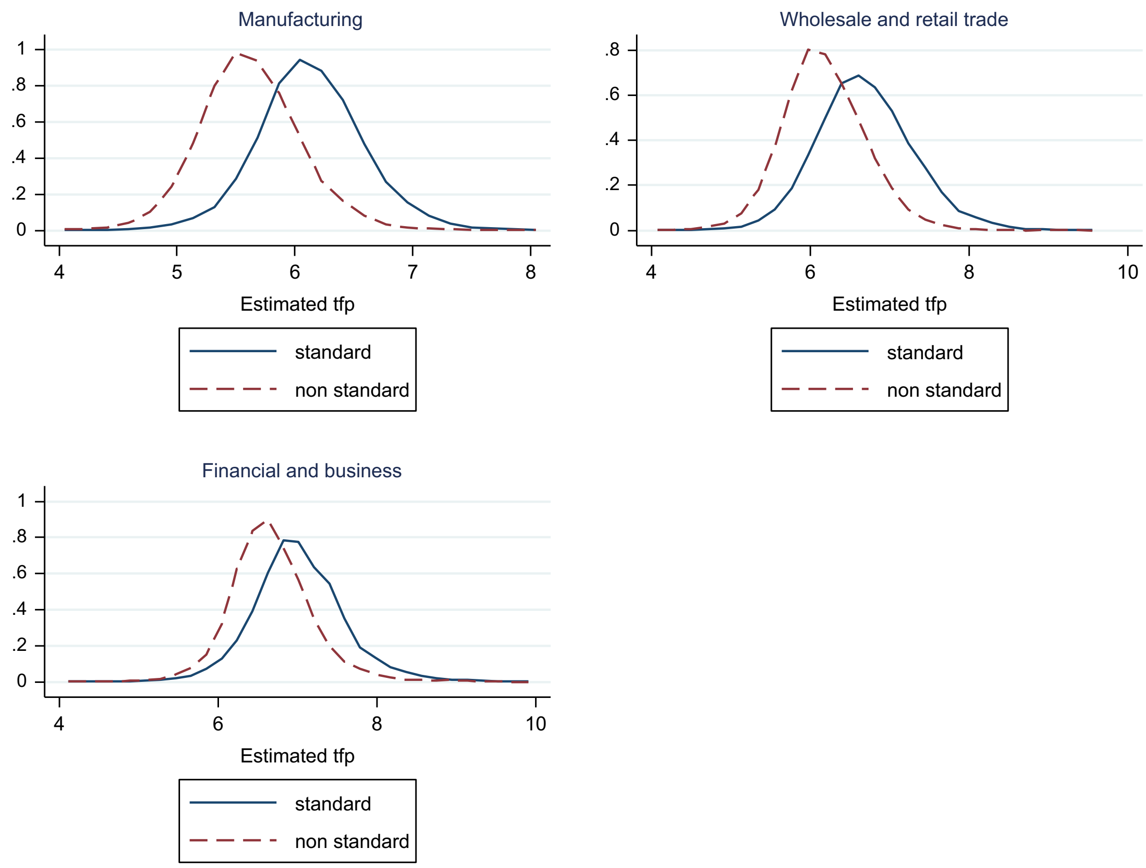

Fig. 2. Kernel density of firm TFP, calculated with or without worker heterogeneity in terms of education. *Note*: TFP is estimated by using the ACF approach.

instead. The industry-specific estimates of the parameter γ are precisely estimated in the wholesale and retail trade, and financial and business sector only, and are lower than one in all industries. Note that we always reject at reasonable significance levels the null hypothesis that the coefficients γ is greater than one, implying between labor types imperfect substitutability or complementarity. This is consistent with the hypothesis that dispersion in labor types along the ethnic dimension has a negative association with firm output, in line with the results found in the previous section with the reduced form approach.

In a second specification, we categorize workers according to their education as we did in the construction of our educational diversity index using the aggregate specification. In this case E can be re-written as

$$E(C_{ijt}, S_{ijt}, T_{1ijt}, \ldots, T_{4ijt}) = [(C_{ijt})^{\gamma} + (\eta_0 S_{1ijt})^{\gamma} + (\eta_1 T_{1ijt})^{\gamma} + \cdots + (\eta_4 T_{4ijt})^{\gamma}]^{1/\gamma} \tag{16}$$

where C_{ijt} is the share of workers with compulsory education, S_{ijt} the share of workers with secondary education, while $T_{1ijt}, \ldots, T_{4ijt}$ are the share of workers with different types of tertiary education, as described in the aggregate specification of education diversity. Table 10 shows the result obtained by estimating production function (11), in which the overall efficiency of the labor force is expressed by Eq. (16).[39] As in the previous table the coefficients on labor and capital are in line with those obtained from a standard specification of the production function, as reported in Table 3, and standard TFP estimates are larger than non standard TFP measures (see Fig. 2). The η parameters show that workers with either secondary or tertiary education are much more productive than workers with compulsory education, especially workers with an engineering or a scientific tertiary education. We find that the parameter γ governing substitutability among employees with different educational levels is below one in all industries, suggesting a supermodular technology with a distaste for dispersion in types along the educational dimension. We therefore interpret these results as evidence that highly educated workers and low-skilled workers are imperfect substitutes and that dispersion of labor types along the educational dimension has no positive impact on productivity in line with what has been found by Iranzo et al. (2008).

The latter result is starkly contradictory with the descriptive analysis reported above in Section 4, where we find that our education diversity statistic is positively related to productivity. This inconsistency can be partly explained by the fact that in the descriptive analysis we do not appropriately control for quality differences between labor types, which seem to play an

[39] Note that the estimation of the production function (11) cannot be carried out for construction and transport sectors due to non-convergence of the estimation algorithm.

Table 11
Production function estimates with shares of labor types along the educational dimension, labor aggregate splitted into two groups.

Dependent variable: log of value added	**Manufacturing**	**Wholesale and retail trade**	**Financial and business services**
Log (L_1)	0.266***	0.211***	0.193***
	(0.006)	(0.007)	(0.009)
Log (L_2)	0.339***	0.428***	0.493***
	(0.006)	(0.009)	(0.016)
Log(K)	0.124***	0.074***	0.050***
	(0.004)	(0.005)	(0.009)
γ	0.737***	0.768***	0.926***
	(0.225)	(0.103)	(0.073)
η_secondary/postsecondary	4.198	1.590*	2.266**
	(2.533)	(0.938)	(0.983)
η_scientific_tertiary	3.313	4.810**	3.366
	(2.097)	(1.912)	(3.015)
η_social_sciences_tertiary	1.165	1.056**	3.275***
	(1.097)	(0.431)	(1.820)
η_engineering_tertiary	3.773**	4.923***	4.657**
	(1.643)	(1.014)	(2.825)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	6.22; 0.032	6.21; 0.032	5.20; 0.025
Observations	36,079	26,604	8097
R2	0.895	0.779	0.777

Note: The dependent variable is the log of value added. Input elasticities are estimated by using the LP approach. Standard errors are computed by using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-firm correlation.
*** 1%.
** 5%.
* 10%.

important role in the production function, as indicated by the estimates on the η parameters. Furthermore in the reduced-form analysis we show that by calculating occupation-specific diversity (Table 5, columns 1 and 2) or the Herfindahl index based only on the type of tertiary education (Table 6, column 3), the positive association between educational diversity and firm TFP is largely driven by the white-collar workers, who are very likely to be highly educated, or by workers with different types of tertiary education. These results may therefore suggest that the contribution of educational diversity is most likely due to the combination of skills of highly educated employees rather than by combining the overall skills of workers with different levels of education. A possible test for the latter conjecture is to estimate a modified version of the structural production function (11) that splits labor aggregate L into two groups, i.e. employees with compulsory education L_1, and workers with more than compulsory education L_2, and that estimates the role of skill dispersion only among L_2 types. Treating L_1 as a standard Cobb–Douglas input with unit elasticity of substitution to the L_2 types, we can re-write production function (11) as follows:

$$Y_{ijt} = A_{ijt} L_{1ijt}^{\alpha_1} \cdot [L_{2ijt} \cdot [(\eta_0 S_{1ijt})^{\gamma} + (T_{1ijt})^{\gamma} + (\eta_1 T_{2ijt})^{\gamma} + (\eta_2 T_{3ijt})^{\gamma} + (\eta_4 T_{4ijt})^{\gamma}]^{1/\gamma}]^{\alpha 2} \cdot K_{ijt}^{\beta} \tag{17}$$

where S_{ijt} is the share of workers with secondary education, $T_{1ijt}, \ldots, T_{4ijt}$ are the share of workers with different types of tertiary education, as in Eq. (16), and γ governs the complementarity/substitutability only arising among workers with more than compulsory education, i.e. L_2. Table 11 shows the estimates of production function (17), using the LP approach.[40] The estimates of the parameter γ slightly increase compared to those reported in Table 10, but they never exceed the threshold of one, suggesting imperfect substitutability or complementarity also when we focus on dispersion in labor types within highly educated workers.

A possible explanation behind the latter result may be that the industrial classification used to separately estimate production function (17) is very aggregate and may group industries with a super-modular technology together with those with a sub-modular technology under the same 1-digit industry. Estimating Eq. (17) separately for each 2-digits sector in fact reveals that for about half of the industries the γ coefficient is estimated to be larger than one.[41] This is the case for the following industries: (i) food, beverages and tobacco, (ii) sales and repair of motor vehicles, (iii) retail trade, (iv) hotels and restaurants and (v) financial intermediation. However the γ coefficient is precisely estimated only for 3 industries (food, beverages and tobacco; sales and repair of motor vehicles; hotels and restaurants) (Table 12).

[40] Note that the estimation of production function (17) cannot be carried out with either ACF or OP due to non-convergence of the estimation algorithm.

[41] Note that the estimation of production function (17) cannot be carried out for textiles, other nonmetallic mineral products, construction and transport due to non-convergence of the estimation algorithm.

Table 12
Production function estimates with shares of labor types along the educational dimension, labor aggregate splitted into two groups and estimates by 2 digits industry.

Dependent variable: log of value added	Food, beverages and tobacco	Wood products	Chemicals
Log (L_1)	0.273***	0.232***	0.249***
	(0.021)	(0.010)	(0.023)
Log (L_2)	0.291***	0.369***	0.319***
	(0.019)	(0.013)	(0.022)
Log(K)	0.198***	0.085***	0.188***
	(0.015)	(0.009)	(0.015)
γ	1.153**	0.705	0.731
	(0.548)	(0.398)	(0.484)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	2.38; 0.112	5.23; 0.024	5.04; 0.026
Observations	3,067	5,269	2,897
R2	0.913	0.901	0.906
	Basic metals	**Furniture**	**Sales and repair of motor vehicles**
$Log(L_1)$	0.254***	0.278***	0.274***
	(0.027)	(0.018)	(0.018)
$Log(L_2)$	0.197***	0.314***	0.501***
	(0.024)	(0.021)	(0.018)
Log(K)	0.180***	0.121***	0.055***
	(0.021)	(0.134)	(0.011)
γ	0.938	0.755	1.082**
	(1.512)	(1.805)	(0.418)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	5.31; 0.024	5.02; 0.027	1.79; 0.153
Observations	1264	3001	3478
R2	0.909	0.897	0.812
	Wholesale trade	**Retail trade**	**Hotels and restaurants**
Log (L_1)	0.192***	0.137***	0.230***
	(0.009)	(0.012)	(0.029)
Log (L_2)	0.475***	0.365***	0.253***
	(0.013)	(0.015)	(0.033)
Log(K)	0.069***	0.089***	0.057***
	(0.007)	(0.008)	(0.011)
γ	0.764***	5.422	1.256*
	(0.105)	(3.885)	(0.732)
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	5.00; 0.025	0.17; 0.684	3.12; 0.076
Observations	11,994	6,720	1,730
R2	0.777	0.845	0.817
	Financial intermediation	**Business activities**	
Log (L_1)	0.118***	0.207***	
	(0.035)	(0.010)	
Log (L_2)	0.474***	0.509***	
	(0.037)	(0.017)	
Log(K)	0.013	0.045***	
	(0.042)	(0.009)	
γ	2.109	0.933***	
	(1.329)	(0.075)	
Hypothesis testing (chi2, p-value)			
$\gamma > 1$	0.70; 0.404	5.21; 0.024	
Observations	619	6,144	
R2	0.542	0.798	

Note: The dependent variable is the log of value added. Input elasticities are estimated by using the LP approach. The estimates of η parameters governing relative quality difference between labor types are not reported in the table and are available on request from the authors. Standard errors are computed by using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-firm correlation.

*** 1%.

** 5%.

* 10%.

By taking together all these results, we cannot thus conclude that dispersion in labor types in terms of education for the overall workforce is positively correlated with productivity. However estimating a modified specification of the structural model separately for each 2-digits industry suggests that, for about half of the sectors, skill diversity arising only among highly educated workers positively associates with firm productivity, in line with the descriptive evidence on the role of diversity among white-collar or tertiary education employees.

6. Conclusions

Using a comprehensive linked employer–employee data-set, this paper primarily investigates the empirical relationship of diversity in workers′ ethnic-cultural, educational and demographic characteristics with firm productivity in Denmark. Unlike the majority of previous empirical studies, which focused on single aspects of labor diversity, we provide a number of findings that extensively explore the overall consequences of firm workforce heterogeneity for firm performance. Specifically, we follow two main methodological routes to investigate the link between workforce diversity and productivity. In the first one, we estimate a standard Cobb–Douglas function, calculate the implied total factor productivity and relate the latter to our labor diversity indices in a second stage. The results from this reduced-form approach allow us to identify which types of labor heterogeneity appear to descriptively matter and suggest that labor diversity in ethnicity (education) is negatively (positively) associated with firm productivity, whereas the demographic diversity seems not to matter. Several robustness checks and the results obtained by implementing the IV method are in line with these descriptive findings. The

Table A1
Descriptive statistics, sample including firms with less than 10 employees.

Variables	Definition	Firms with less than 10 employees		
		Mean	Median	Sd
IDA variables				
foreigners1	Non-Danish employees from Muslim countries, as a proportion of foreign employees	0.037	0	0.181
foreigners2	Non-Danish employees from East Asia, as a proportion of foreign employees	0.014	0	0.111
foreigners3	Non-Danish employees from Other Asia, as a proportion of foreign employees	0.017	0	0.123
foreigners4	Non-Danish employees from formerly communist countries, as a proportion of foreign employees	0.01	0	0.081
foreigners5	Non-Danish employees from West/South Europe, as a proportion of foreign employees	0.122	0	0.319
foreigners6	Non-Danish employees from Africa, as a proportion of foreign employees	0.004	0	0.059
foreigners7	Non-Danish employees from North America/Oceania, as a proportion of foreign employees	0.007	0	0.081
foreigners8	Non-Danish employees from Central/South America, as a proportion of foreign employees	0.004	0	0.059
Males	Men, as a proportion of all employees	0.736	0.8	0.234
age1	Employees aged 15–32, as a proportion of all employees	0.206	0.2	0.183
age2	Employees aged 33–41, as a proportion of all employees	0.197	0.2	0.146
age3	Employees aged 42–50, as a proportion of all employees	0.325	0.3	0.21
age4	Employees aged 51–65 as a proportion of all employees	0.181	0.211	0.178
skill0	Employees with compulsory education, as a proportion of all employees	0.325	0.3	0.21
skill1	Employees with a secondary/ post-secondary education, as a proportion of all employees	0.643	0.7	0.204
skill2_1	Employees with a humanities tertiary education, as a proportion of all employees	0.007	0	0.038
skill2_2	Employees with a scientific tertiary education, as a proportion of all employees	0.003	0	0.025
skill2_3	Employees with a social sciences tertiary education, as a proportion of all employees	0.013	0	0.051
skill2_4	Employees with a engineering tertiary education, as a proportion of all employees	0.01	0	0.052
Tenure	Average tenure	4.93	4.78	0.024
Managers	Managers, as a proportion of all employees	0.034	0	0.07
Middle managers	Middle managers, as a proportion of all employees	0.137	0.1	0.19
bluecoll	Blue-collars, as a proportion of all employees	0.741	0.829	0.276
Index_ethnic_aggr	Diversity index based on employees′ nationality (8 categories)	0.045	0	0.09
Index_edu_aggr	Diversity index based on employees′ education (6 categories)	0.386	0.42	0.137
Index_demo_aggr	Diversity index based on employees′ demographic characteristics (8 categories)	0.694	0.72	0.268
Index_ethnic_disaggr	Diversity index based on employees′ language (40 categories)	0.074	0	0.102
Index_edu_disaggr	Diversity index based on employees′ education (9 categories)	0.379	0.406	0.124
Index_demo_disaggr	Diversity index based on employees′ demographic characteristics (10 categories)	0.701	0.739	0.129
Accounting variables				
Total sales	(1000 kr.)	14,114.76	8,048.458	49,583.6
Intermediates	(1000 kr.)	10,310.94	4,495	48,335.78
Capital	(1000 kr.)	3,329.596	1,142.857	28,779.17
Export	1, if the firm export	0.224	0	0.417
foreign_ownership	1, if the firm is foreign owned	0.001	0	0.049
Multi	1, if the firm is multi-establishment	0	0	0
N		9405		

Note: All IDA and accounting variables are expressed as time averages from 1995 to 2005.

main limit of the reduced-form approach is that it does not formally take into account that the labor input is non-homogeneous in the production function, i.e., labor of different types is of different quality (Hellerstein and Neumark, 2004; Iranzo et al., 2008; Fox and Smeets, 2011; Irarrazabal et al., 2014). We therefore move toward a richer production function specification, which takes different types of labor as inputs and that allows for flexible substitution patterns, and possible

Table A2
Labor diversity and firm total factor productivity, sample including firms with fewer than 10 employees.

Dependent variable: log of TFP	**TFP (ACF)**				
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
Index ethnic aggr	−0.021***	−0.014**	−0.016**	−0.024	−0.045***
	(0.004)	(0.005)	(0.006)	(0.020)	(0.010)
Index edu aggr	0.011**	−0.015**	−0.013**	0.057	0.019**
	(0.005)	(0.005)	(0.006)	(0.029)	(0.009)
Index demo aggr	0.034	0.021	0.005	0.066	0.005
	(0.018)	(0.016)	(0.007)	(0.041)	(0.010)
Observations	36,855	20,298	28,192	4,335	8,233
R2	0.261	0.222	0.479	0.190	0.344
Index ethnic disaggr	−0.022***	−0.014**	−0.010**	−0.006	−0.012*
	(0.004)	(0.005)	(0.003)	(0.006)	(0.007)
Index edu disaggr	0.036***	0.008	0.048***	0.007	0.062***
	(0.007)	(0.006)	(0.006)	(0.018)	(0.013)
Index demo disaggr	0.022	−0.021	−0.011	0.049	−0.019
	(0.013)	(0.015)	(0.010)	(0.027)	(0.012)
Observations	36,855	20,298	28,192	4,335	8,233
R2	0.271	0.230	0.490	0.210	0.357

Note: The dependent variable is the log of total factor productivity estimated using the ACF approach. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.
** 5%.
* 10%.

Table A3
Labor diversity and firm exit.

Dependent variable: firm exit	**Manufacturing**	**Construction**	**Wholesale and retail trade**	**Transport**	**Financial and business services**
Index ethnic aggr	0.001	−0.006	−0.007*	−0.010	−0.008
	(0.003)	(0.004)	(0.003)	(0.012)	(0.006)
Index edu aggr	−0.004	−0.014	−0.008*	−0.009	0.003
	(0.004)	(0.009)	(0.004)	(0.007)	(0.003)
Index demo aggr	−0.011*	−0.025*	−0.016	−0.009	−0.022
	(0.006)	(0.013)	(0.009)	(0.007)	(0.014)
Observations	35,887	18,024	26,418	4,007	7,931
R2	0.034	0.04	0.047	0.055	0.065
Index ethnic disaggr	−0.002	−0.006	−0.002	−0.012	−0.002
	(0.003)	(0.004)	(0.001)	(0.008)	(0.002)
Index edu disaggr	−0.003	−0.012	−0.004	−0.003	0.002
	(0.003)	(0.008)	(0.003)	(0.010)	(0.003)
Index demo disaggr	−0.010*	−0.024**	−0.016	−0.011	−0.023
	(0.006)	(0.012)	(0.008)	(0.008)	(0.017)
Observations	48,238	26,969	41,493	6287	14,008
R2	0.035	0.042	0.048	0.057	0.068

Note: The dependent variable is a dummy variable equal to 1 if the firm exits the market. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.
Significance levels: ***1%, **5%, *10%.

quality differences between types. The results obtained from this structural estimation approach suggest that labor heterogeneity in terms of ethnicity decreases firm output and that it is not optimal to have dispersion in terms of employees' education.

Thus, we find evidence that the negative effects of the communication and integration costs associated with a more culturally diverse workforce seem to counteract the positive effects of ethnic diversity (e.g., better problem-solving abilities, more creativity, and knowledge spillover). These findings are consistent with those of previous studies reported by Lazear

Table A4.1
Additional results of Table 4, workforce composition characteristics and firm total factor productivity (aggregate specification).

Dependent variable: log of TFP	***TFP (ACF)***				
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
age1_female	−0.686	0.031	0.163	2.204***	0.025
	(0.522)	(0.588)	(0.396)	(0.627)	(0.169)
age2_female	−0.472	0.370	0.698*	2.469***	0.651***
	(0.520)	(0.586)	(0.395)	(0.601)	(0.169)
age3_female	−0.801	0.210	0.381	2.443***	0.277*
	(0.520)	(0.587)	(0.395)	(0.618)	(0.161)
age4_female	−1.347***	−0.185	−0.141	1.002	−0.429***
	(0.520)	(0.588)	(0.399)	(0.626)	(0.166)
age1_male	1.199**	0.003	0.187***	1.683**	0.291*
	(0.525)	(0.589)	(0.398)	(0.671)	(0.180)
age2_male	−0.623	0.145	0.610	0.234	0.341**
	(0.519)	(0.593)	(0.393)	(0.657)	(0.168)
age3_male	0.883*	0.258	0.309	2.067***	0.096
	(0.524)	(0.596)	(0.396)	(0.658)	(0.172)
age4_male	−1.201**	−0.075	−0.012	0.606	−0.245
	(0.521)	(0.595)	(0.399)	(0.679)	(0.183)
skill0	−0.225	0.324	−1.487***	−2.205**	−0.042
	(0.431)	(0.501)	(0.264)	(1.017)	(0.195)
skill1	0.092	0.315	−1.039***	−1.861*	0.313
	(0.433)	(0.504)	(0.265)	(1.010)	(0.199)
skill2_1	0.513*	–	0.663*	2.339**	0.428*
	(0.278)	–	(0.362)	(1.107)	(0.236)
skill2_2	–	0.618**	–	1.866	–
	–	(0.207)	–	(2.163)	–
skill2_3	0.472*	0.707	−0.273	1.619	0.933***
	(0.235)	(0.625)	(0.314)	(1.264)	(0.190)
skill2_4	0.101	0.687	1.127***	–	0.472**
	(0.492)	(0.620)	(0.392)	–	(0.200)
foreigners1	−0.023*	−0.020	−0.025	−0.065	−0.033
	(0.012)	(0.014)	(0.016)	(0.045)	(0.034)
foreigners2	0.010	0.004	0.038*	0.095	0.032
	(0.017)	(0.029)	(0.023)	(0.119)	(0.057)
foreigners3	0.001	−0.002	0.018*	0.092***	0.059***
	(0.008)	(0.008)	(0.011)	(0.027)	(0.023)
foreigners4	−0.025	0.042	0.141***	0.201*	0.032
	(0.033)	(0.058)	(0.044)	(0.107)	(0.089)
foreigners5	0.021	−0.081*	0.059	0.138	−0.044
	(0.041)	(0.045)	(0.052)	(0.141)	(0.057)
foreigners6	0.024	0.037	0.116**	0.169*	0.058
	(0.032)	(0.035)	(0.050)	(0.100)	(0.041)
foreigners7	0.037*	−0.024	0.015	0.073	−0.007
	(0.022)	(0.026)	(0.022)	(0.053)	(0.041)
foreigners8	−0.009	−0.083**	−0.017	−0.230	−0.097
	(0.018)	(0.035)	(0.034)	(0.233)	(0.069)
manager	0.250***	0.468***	0.304***	−0.071	−0.000
	(0.067)	(0.082)	(0.068)	(0.258)	(0.106)
middle managers	0.195***	0.492***	0.229***	0.271***	0.229***
	(0.035)	(0.069)	(0.031)	(0.082)	(0.035)
Observations	35,887	18,024	26,418	4007	7931
R2	0.281	0.235	0.553	0.185	0.347

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. All regressions include average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

(1999), Glaeser et al. (2000), and Alesina and La Ferrara (2002). Instead, the results concerning labor heterogeneity in terms of education are mixed. On one hand, the reduced-form analysis provides descriptive evidence of a positive association between education diversity and productivity, consistently with Lazear (1999). On the other hand, the structural estimation approach reveals that the overall dispersion of labor types in terms of education decreases firm output, i.e. high-educated and low-educated workers′ skills are imperfect substitutes, in line with what has been found in Iranzo et al. (2008). However estimating a modified specification of the structural model separately for each 2-digits industry suggests that, for

Table A4.2
Additional results of Table 4, workforce composition characteristics and firm total factor productivity (aggregate specification).

Dependent variable: log of TFP	***TFP* (*OP*)**				
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
age1_female	−0.634	1.135	0.128	0.559	0.071
	(0.965)	(0.859)	(0.594)	(1.230)	(0.191)
age2_female	−0.466	1.374*	0.409	2.233*	0.591***
	(0.964)	(0.857)	(0.591)	(1.227)	(0.198)
age3_female	−0.713	1.277	0.019	1.300	0.500***
	(0.957)	(0.860)	(0.594)	(1.227)	(0.191)
age4_female	−1.231	0.846	−0.421	−1.791	−0.343*
	(0.960)	(0.860)	(0.592)	(1.218)	(0.199)
age1_male	1.078	1.056	1.124*	3.061**	1.291***
	(0.974)	(0.858)	(0.597)	(1.252)	(0.215)
age2_male	−0.668	1.257	0.349	1.131	0.406*
	(0.961)	(0.855)	(0.594)	(1.246)	(0.209)
age3_male	0.695	1.276	1.051*	2.936**	0.139
	(0.957)	(0.864)	(0.591)	(1.265)	(0.204)
age4_male	−1.159	1.036	−0.323	−1.903	−0.344
	(0.960)	(0.867)	(0.593)	(1.277)	(0.216)
skill0	0.536	0.668	−0.723*	−0.947	−0.248
	(0.461)	(0.965)	(0.389)	(2.121)	(0.215)
skill1	0.863*	0.686	−0.319	−0.533	0.109
	(0.465)	(0.966)	(0.392)	(2.118)	(0.203)
skill2_1	1.137*	–	1.274**	−2.225	–
	(0.609)	–	(0.481)	(2.189)	–
skill2_2	–	0.796	0.276	–	−0.074
	–	(1.371)	(0.528)	–	(0.366)
skill2_3	1.230**	1.880*	1.085**	1.574*	0.691***
	(0.578)	(0.998)	(0.456)	(0.792)	(0.244)
skill2_4	0.906*	0.330	–	1.659**	0.163
	(0.549)	(1.177)	–	(0.701)	(0.194)
foreigners1	−0.008	−0.011	−0.026	−0.156**	−0.024
	(0.018)	(0.028)	(0.026)	(0.071)	(0.053)
foreigners2	0.032	−0.011	0.053	−0.045	0.003
	(0.026)	(0.037)	(0.033)	(0.118)	(0.087)
foreigners3	0.018	−0.001	0.032**	0.035	0.064*
	(0.011)	(0.011)	(0.014)	(0.037)	(0.033)
foreigners4	0.012	0.094	0.139**	0.477**	−0.025
	(0.044)	(0.065)	(0.057)	(0.227)	(0.118)
foreigners5	−0.062	−0.128*	−0.001	0.241	0.030
	(0.070)	(0.067)	(0.085)	(0.251)	(0.088)
foreigners6	0.001	0.066*	0.059	0.051	0.167**
	(0.053)	(0.041)	(0.063)	(0.103)	(0.069)
foreigners7	0.019	−0.018	−0.016	−0.021	−0.035
	(0.034)	(0.043)	(0.030)	(0.065)	(0.063)
foreigners8	−0.015	−0.106*	−0.034	−0.011	−0.174
	(0.023)	(0.058)	(0.053)	(0.121)	(0.149)
manager	0.177*	0.589***	0.293***	−0.257	0.356**
	(0.094)	(0.096)	(0.084)	(0.358)	(0.158)
middle managers	0.290***	0.501***	0.293***	0.375***	0.273***
	(0.048)	(0.075)	(0.039)	(0.115)	(0.052)
Observations	10,662	6572	9199	1526	2985
R2	0.268	0.230	0.432	0.253	0.351

Note: The dependent variable is the log of total factor productivity estimated by using the OP approach. All regressions include average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

about half of the sectors, skill diversity arising only among highly educated workers positively associates with firm productivity.

Statement on data disclosure

The data used in this paper builds on anonymized micro-data-sets owned by Statistics Denmark (SD).

Table A4.3
Additional results of Table 4, workforce composition characteristics and firm total factor productivity (aggregate specification).

Dependent variable: log of TFP	***TFP* (*GNR*)**				
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
age1_female	−1.073***	−0.950**	0.152	0.268	−0.347
	(0.295)	(0.433)	(0.245)	(0.548)	(0.286)
age2_female	−0.920***	−0.698*	0.133	0.334	0.146
	(0.293)	(0.431)	(0.245)	(0.536)	(0.271)
age3_female	−1.092***	−0.718*	−0.052	0.366	−0.031
	(0.293)	(0.431)	(0.245)	(0.533)	(0.269)
age4_female	−1.354***	−0.872**	−0.258	0.135	−0.722**
	(0.295)	(0.430)	(0.245)	(0.540)	(0.282)
age1_male	1.460***	0.816*	1.048***	1.546**	0.660**
	(0.295)	(0.434)	(0.246)	(0.569)	(0.295)
age2_male	−0.898***	−0.703*	0.170	0.989*	−0.038
	(0.294)	(0.434)	(0.245)	(0.573)	(0.280)
age3_male	1.117***	0.759*	1.031***	0.685	0.550*
	(0.296)	(0.433)	(0.246)	(0.554)	(0.315)
age4_male	−1.382***	−1.005**	−0.303	−0.121	−0.752***
	(0.294)	(0.430)	(0.246)	(0.576)	(0.287)
skill0	0.350	0.121	−0.390**	−1.077	1.435***
	(0.263)	(0.306)	(0.164)	(0.983)	(0.371)
skill1	0.500*	0.171	−0.320*	−0.655	1.681***
	(0.265)	(0.307)	(0.164)	(0.977)	(0.379)
skill2_1	1.053***	–	1.136***	−1.200	1.773***
	(0.315)	–	(0.214)	(1.114)	(0.402)
skill2_2	–	1.830*	–	1.189*	–
	–	(0.934)	–	(0.613)	–
skill2_3	0.803***	0.819*	0.612***	−1.034	1.972***
	(0.281)	(0.375)	(0.201)	(1.120)	(0.366)
skill2_4	0.454	0.319	0.635***	–	1.535***
	(0.298)	(0.440)	(0.235)	–	(0.379)
foreigners1	0.002	−0.001	0.002	−0.065*	−0.077*
	(0.008)	(0.010)	(0.009)	(0.039)	(0.040)
foreigners2	0.006	0.005	0.024	0.104	−0.007
	(0.010)	(0.017)	(0.016)	(0.084)	(0.049)
foreigners3	0.007	0.000	0.024***	0.062***	0.078***
	(0.005)	(0.005)	(0.007)	(0.022)	(0.024)
foreigners4	0.001	0.010	0.035	0.007	0.048
	(0.025)	(0.037)	(0.033)	(0.117)	(0.093)
foreigners5	0.021	−0.041*	0.029	0.172	−0.116**
	(0.024)	(0.024)	(0.034)	(0.157)	(0.056)
foreigners6	0.008	0.005	0.027	0.143*	0.142**
	(0.018)	(0.018)	(0.030)	(0.084)	(0.061)
foreigners7	0.016	−0.028	0.004	0.050	0.042
	(0.013)	(0.019)	(0.016)	(0.041)	(0.049)
foreigners8	−0.013	−0.032*	−0.011	−0.082	−0.018
	(0.011)	(0.017)	(0.023)	(0.127)	(0.065)
manager	0.120***	0.295***	0.253***	−0.123	0.092
	(0.040)	(0.065)	(0.048)	(0.226)	(0.108)
middle managers	0.215***	0.353***	0.176***	0.192**	0.212***
	(0.022)	(0.047)	(0.022)	(0.078)	(0.038)
Observations	35,387	17,823	26,016	3928	7626
R2	0.852	0.905	0.848	0.883	0.848

Note: The dependent variable is the log of total factor productivity estimated by using the GNR approach. All regressions include average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

Table A5
Testing alternative hypotheses: estimates by using occupation-specific diversity, alternative index definitions, and "subtract-out" groups (results for all the other industries).

Dependent variable: log of TFP	***Occupation specific diversity***		**2nd gen. Imm. included in the ethnic diversity**	**University graduates excluded from the ethnic diversity**	**Germanic group excluded from the ethnic diversity**
	White-collar	Blue-collar			
Construction					
Index ethnic disaggr (1)			−0.015**	−0.007	−0.005
			(0.007)	(0.006)	(0.005)
Index ethnic disaggr (2)	−0.011**	−0.012**	−0.006	−0.016**	−0.015**
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Index edu disaggr	0.016**	0.004	0.012*	0.012*	0.012*
	(0.005)	(0.003)	(0.007)	(0.007)	(0.007)
Index demo disaggr	−0.021	−0.008*	−0.027	−0.027	−0.027
	(0.014)	(0.004)	(0.015)	(0.015)	(0.016)
Hypothesis tests (chi2, p-value)					
Index ethnic (white-collar)=index ethnic (blue-collar)	0.68; 0.041				
Index edu (white-collar)=index edu (blue-collar)	7.67; 0.005				
Index demo (white-collar)=index demo (blue-collar)	6.07; 0.013				
Index ethnic (1)=index ethnic (2)			5.99; 0.013	6.01; 0.013	6.1; 0.012
Observations	18,024		18,024	18,024	18,024
R2	0.247		0.247	0.247	0.247
Transport					
Index ethnic disaggr (1)			−0.007	−0.005	−0.005
			(0.008)	(0.018)	(0.007)
Index ethnic disaggr (2)	−0.008	−0.008	−0.003	−0.009	−0.013
	(0.014)	(0.020)	(0.019)	(0.008)	(0.017)
Index edu disaggr	0.008	−0.004	0.007	0.007	0.007
	(0.017)	(0.008)	(0.022)	(0.022)	(0.022)
Index demo disaggr	0.029	−0.003	0.032	0.033	0.032
	(0.019)	(0.006)	(0.019)	(0.019)	(0.019)
Hypothesis tests (chi2, p-value)					
Index ethnic (white-collar)=index ethnic (blue-collar)	0.00; 0.99				
Index edu (white-collar)=index edu (blue-collar)	0.33; 0.56				
Index demo (white-collar)=index demo (blue-collar)	0.50; 0.48				
Index ethnic (1)=index ethnic (2)			0.05; 0.83	0.01; 0.94	0.14; 0.70
Observations	4007		4007	4007	4007
R2	0.201		0.203	0.203	0.203
Financial and Business services					
Index ethnic disaggr (1)			−0.032**	0.010	0.007
			(0.010)	(0.006)	(0.006)
Index ethnic disaggr (2)	0.007	−0.014*	−0.013**	−0.025**	−0.017**
	(0.008)	(0.008)	(0.006)	(0.010)	(0.009)
Index edu disaggr	0.033**	0.017*	0.055***	0.055***	0.055***
	(0.010)	(0.009)	(0.013)	(0.013)	(0.013)
Index demo disaggr	−0.025	−0.008	−0.010	−0.009	−0.009
	(0.019)	(0.010)	(0.011)	(0.011)	(0.011)
Hypothesis tests (chi2, p-value)					
Index ethnic (white-collar)=index ethnic (blue-collar)	0.39; 0.53				
Index edu (white-collar)=index edu (blue-collar)	1.39; 0.23				
Index demo (white-collar)=index demo (blue-collar)	1.88; 0.17				
Index ethnic (1)=index ethnic (2)			9.51; 0.002	6.34; 0.012	3.90; 0.05
Observations	7931		7931	7931	7931
R2	0.361		0.361	0.361	0.361

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Index ethnic disaggr (1) is calculated as in the main analysis whereas index ethnic disaggr (2) is separately calculated for white- and blue-collar workers in columns 1 and 2, or by excluding/including a specific group of employees, as described in the column head, in columns 3 and 4. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

Table A6
Testing alternative hypotheses: estimates by using occupation-specific diversity, alternative index definitions, and "subtract-out" groups (aggregate specification of diversity).

Dependent variable: log of TFP	***Occupation specific diversity***		**2nd gen. Imm. included in the ethnic diversity**	**University graduates excluded from the ethnic diversity**
	White-collar	Blue-collar		
Manufacturing				
Index ethnic aggr (1)			−0.018**	−0.008
			(0.009)	(0.009)
Index ethnic aggr (2)	−0.005*	−0.013***	−0.006	−0.026***
	(0.003)	(0.003)	(0.009)	(0.007)
Index edu aggr	0.027***	0.009**	0.014**	0.014**
	(0.005)	(0.003)	(0.006)	(0.006)
Index demo aggr	−0.008	−0.002	0.024	0.024
	(0.005)	(0.003)	(0.016)	(0.016)
Hypothesis tests (chi2, p-value)				
Index ethnic (white-collar)=index ethnic (blue-collar)	4.57; 0.057			
Index edu (white-collar)=index edu (blue-collar)	8.53; 0.003			
Index demo (white-collar)=index demo (blue-collar)	2.52; 0.133			
Index ethnic (1)=index ethnic (2)			9.78; 0.000	8.71; 0.002
Observations	35,887		35,887	35,887
R2	0.21		0.291	0.281
Wholesale and retail trade				
Index ethnic aggr (1)			−0.024**	−0.007
			(0.009)	(0.010)
Index ethnic aggr (2)	−0.010**	−0.037***	−0.010	−0.027**
	(0.004)	(0.006)	(0.009)	(0.010)
Index edu aggr	0.036***	0.007**	0.012**	0.010*
	(0.005)	(0.003)	(0.006)	(0.006)
Index demo aggr	−0.015	−0.014	−0.004	−0.004
	(0.014)	(0.015)	(0.005)	(0.005)
Hypothesis tests (chi2, p-value)				
Index ethnic (white-collar)=index ethnic (blue-collar)	19.61; 0.000			
Index edu (white-collar)=index edu (blue-collar)	20.07; 0.000			
Index demo (white-collar)=index demo (blue-collar)	0.75; 0.381			
Index ethnic (1)=index ethnic (2)			17.91; 0.000	19.33; 0.000
Observations	26,418		26,418	26,418
R2	0.555		0.553	0.553

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Index ethnic disaggr (1) is calculated as in the main analysis whereas index ethnic disaggr (2) is separately calculated for white- and blue-collar workers in columns 1 and 2, or by excluding/including a specific group of employees, as described in the column head, in columns 3 and 4. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

Analyses of such data-sets must be done on servers hosted by Statistics Denmark and SD does not permit such data to be used elsewhere. The analysis data set will be archived for at least 5 years.

In the interest of scientific validation of analyses published using DS micro data the Department of Economics and Business, Aarhus University, will assist researchers in obtaining access to the data-set. The access is conditional on SD accepting a contract with the researcher in which he accepts the conditions of DS for using micro data, see http://www.dst.dk/en/TilSalg/Forskningsservice/Dataadgang.aspx.

Request for getting access can be emailed to datamanager_econ.au.dk@econ.au.dk.

Acknowledgments

We thank Sandra E. Black, Holger Bonin, Tor Eriksson, Luca Merlino, Andreas Moxnes, Jozef Konings, Jennifer Poole, Michael Rosholm, Fabiano Schivardi, Vanormelingen Stijn, Chad Syverson, Kenneth L. Sørensen, and Frederic Warzynski (alphabetical order) for their helpful comments and suggestions. Additionally, we appreciate the comments from participants at seminars organized by the Copenhagen Business School, University of Bergamo, Aarhus School of Business, University of Lausanne and ZEW (Mannheim) and from participants at the following conferences: ESPE 2010, the 5th Nordic Summer Institute in Labor Economics, the 2010 Ratio Young Scientist Colloquium, ESEM 2010 and CAED/COST 2010 in

Table A7
Testing alternative hypotheses: the role of export and estimates by relevant industrial aggregations.

Dependent variable: log of TFP	**Interactions with export dummy**					**Estimates by Industry R&D intensity**		**Estimates by Industrial employment development**	
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services	Above mean	Below mean	Declining employment	Increasing employment
Index ethnic disaggr	−0.014**	−0.010*	−0.014***	−0.024*	0.002	−0.010	−0.008***	−0.011***	−0.007**
	(0.005)	(0.005)	(0.004)	(0.014)	(0.007)	(0.007)	(0.002)	(0.003)	(0.003)
Index edu disaggr	0.018**	0.013*	0.042***	0.009	0.048***	0.035*	0.033***	0.018***	0.039***
	(0.008)	(0.007)	(0.007)	(0.023)	(0.013)	(0.018)	(0.004)	(0.004)	(0.006)
Index demo disaggr	0.029	−0.026	−0.016	0.041	−0.010	−0.006	−0.000	−0.006	0.007
	(0.016)	(0.019)	(0.009)	(0.025)	(0.012)	(0.018)	(0.003)	(0.004)	(0.005)
Export	0.042***	0.042**	0.061***	0.010	0.062*				
	(0.009)	(0.018)	(0.012)	(0.032)	(0.035)				
Export*index ethnic disaggr	−0.004	−0.033	−0.004	0.028*	0.007				
	(0.006)	(0.020)	(0.008)	(0.017)	(0.013)				
Export*index edu disaggr	0.015	−0.012	0.019	−0.016	0.017				
	(0.009)	(0.021)	(0.011)	(0.045)	(0.029)				
Export*index demo disaggr	−0.022	−0.008	−0.001	−0.057	−0.005				
	(0.014)	(0.024)	(0.011)	(0.053)	(0.027)				
Hypothesis tests (chi2, p-value)									
Interactions of diversity indices and export=0	1.70; 0.18	1.01; 0.45	1.47; 0.22	2.12; 0.14	1.65; 0.21				
Index ethnic (above)=index ethnic (below)						0.33; 0.56		2.56; 0.09	
Index edu (above)=index ethnic (below)						0.01; 0.92		6.08; 0.03	
Index demo (above)=index demo (below)						0.10; 0.75		0.58; 0.63	
Observations	35,887	18,024	26,418	4007	7931	6506	85,761	35,157	57,110
R2	0.292	0.249	0.56	0.204	0.363	0.190	0.673	0.665	0.654

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.

** 5%.

* 10%.

London. Pierpaolo Parrotta acknowledges financial support from the Graduate School for Integration, Production and Welfare, and the Swiss National Centres of Competence in Research LIVES. Mariola Pytlikova gratefully acknowledges funding from the NORFACE Migration Programme. The usual disclaimer applies.

Appendix A

A.1. Groups included in the measure of ethnic diversity

(1) *The citizens in the different nationality groups are as follows*: *Danish*: Danish native including second generation immigrants; *North America and Oceania*: United States, Canada, Australia, New Zealand; Central and South America,

Table A8
Robustness checks: estimates by size, under alternative samples, and index definitions (aggregate specification of diversity).

Dependent variable: log of TFP	Shannon entropy index	Richness	Edu and demo diversity as sd	Estimates by Firm Size			"Copenhagen" county is excluded	Mono-establishment firms
				Small firms	Medium firms	Large firms		
Manufacturing								
Index ethnic aggr	−0.015***	−0.014**	−0.014***	−0.019**	−0.014**	−0.020***	−0.014***	−0.009**
	(0.003)	(0.005)	(0.003)	(0.008)	(0.005)	(0.004)	(0.003)	(0.003)
Index edu aggr	0.023**	0.014**	0.012**	0.005	0.026**	0.048***	0.016***	0.013**
	(0.007)	(0.005)	(0.004)	(0.008)	(0.009)	(0.012)	(0.006)	(0.006)
Sd(years of education)			0.006					
			(0.004)					
Index demo aggr	0.024	−0.002		0.012	0.010	0.011	0.015	0.009
	(0.015)	(0.005)		(0.007)	(0.009)	(0.014)	(0.009)	(0.006)
Sd(age)			−0.001					
			(0.005)					
Share of men			0.203***					
			(0.025)					
Hypothesis tests (chi2, p-value)								
index edu = sd(years of education)			5.67; 0.036					
Sd (age) = share of men			64.92; 0.000					
Observations	35,887	35,887	35,887	9727	13,446	12,714	34,763	31,279
R2	0.282	0.279	0.286	0.287	0.293	0.290	0.283	0.277
Wholesale and retail trade								
Index ethnic aggr	−0.040***	−0.042***	−0.015***	−0.005*	−0.015*	−0.062***	−0.027***	−0.021***
	(0.007)	(0.009)	(0.004)	(0.003)	(0.008)	(0.010)	(0.006)	(0.006)
Index edu aggr	0.018**	0.023**	0.033***	0.004**	0.009*	0.017**	0.013**	0.019**
	(0.008)	(0.008)	(0.006)	(0.002)	(0.005)	(0.007)	(0.005)	(0.006)
Sd(years of education)			0.007					
			(0.004)					
Index demo aggr	−0.002	−0.001		−0.002	0.009	−0.026	−0.003	0.001
	(0.006)	(0.005)		(0.007)	(0.009)	(0.014)	(0.005)	(0.006)
Sd(age)			−0.010					
			(0.007)					
Share of men			0.142***					
			(0.025)					
Hypothesis tests (chi2, p-value)								
Index edu = sd(years of education)			8.88; 0.000					
Sd (age) = share of men			22.50; 0.000					
Observations	26,418	26,418	26,418	10,042	10,492	5884	24,329	21,720
R2	0.554	0.554	0.56	0.473	0.505	0.509	0.560	0.518

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the firm level.

*** 1%.
** 5%.
* 10%.

Guatemala, Belize, Costa Rica, Honduras, Panama, El Salvador, Nicaragua, Venezuela, Ecuador, Peru, Bolivia, Chile, Argentina, Brazil; *Formerly Communist Countries*: Armenia, Belarus, Estonia, Georgia, Latvia, Lithuania, Moldova, Russia, Tajikistan, Ukraine, Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia, Albania, Bosnia and Herzegovina, Croatia, Republic of Macedonia, Montenegro, Serbia, and Slovenia; *Muslim Countries*: Afghanistan, Algeria, Arab Emirates, Azerbaijan, Bahrain, Bangladesh, Brunei Darussalem, Burkina Faso, Camoros, Chad, Djibouti, Egypt, Eritrea, Gambia, Guinea, Indonesia, Iran, Iraq, Jordan, Kazakstan, Kirgizstan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Malaysia, Maldives, Mali, Mauritania, Morocco, Nigeria, Oman, Pakistan, Palestine, Qatar, Saudi Arabia, Senegal, Sierra Leone, Somalia, Sudan, Syria, Tadzhikstan, Tunisia, Turkey, Turkmenistan, Uzbekistan, Yemen; *East Asia*: China, Hong Kong, Japan, Korea, Korea Dem. People′s Rep. of Macao,

Table A9
Labor diversity and productivity, IV results: additional second stage regressions.

Dependent variable: log of TFP	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
All sample					
Index ethnic disaggr	−0.026*	−0.038*	−0.028**	−0.031	0.009
	(0.014)	(0.019)	(0.014)	(0.084)	(0.012)
Index edu disaggr	0.061**	0.037	0.095**	0.047	0.078*
	(0.028)	(0.019)	(0.040)	(0.149)	(0.038)
Index demo disaggr	0.093	−0.048	−0.056	−0.085	−0.048
	(0.086)	(0.049)	(0.033)	(0.070)	(0.033)
Observations	35,887	18,024	26,418	4007	7931
R2	0.310	0.123	0.252	0.189	0.200
Mono-establishment firms					
Index ethnic disaggr	−0.017	−0.062	−0.024	–	0.017
	(0.093)	(0.035)	(0.016)	–	(0.079)
Index edu disaggr	0.057*	0.028	0.082**	–	0.074
	(0.029)	(0.021)	(0.061)	–	(0.037)
Index demo disaggr	0.143	−0.046	−0.027	–	−0.066
	(0.088)	(0.085)	(0.033)	–	(0.157)
Observations	31,279	17,118	21,720	–	6479
R2	0.260	0.108	0.277	–	0.114
Small firms with fewer than 50 employees					
Index ethnic disaggr	−0.015*	−0.075	−0.033**	–	0.027
	(0.008)	(0.049)	(0.014)	–	(0.019)
Index edu disaggr	0.032*	0.055	0.123*	–	0.082
	(0.017)	(0.037)	(0.069)	–	(0.049)
Index demo disaggr	0.170	−0.084	−0.055	–	−0.041
	(0.161)	(0.079)	(0.046)	–	(0.035)
Observations	23,173	15,423	20,534	–	5,536
R2	0.435	0.096	0.188	–	0.109
Copenhagen county is excluded					
Index ethnic disaggr	−0.034	−0.047	−0.013	−0.027	0.025
	(0.023)	(0.030)	(0.016)	(0.068)	(0.012)
Index edu disaggr	0.053**	0.025*	0.060**	0.022	0.080*
	(0.022)	(0.013)	(0.021)	(0.123)	(0.042)
Index demo disaggr	0.144	−0.057*	−0.028	−0.127	−0.102
	(0.213)	(0.029)	(0.041)	(0.099)	(0.094)
Observations	34,763	17,302	24,329	3619	6149
R2	0.263	0.145	0.365	0.128	0.255
Firms established before 1990					
Index ethnic disaggr	−0.038	–	−0.045**	–	0.016
	(0.084)	–	(0.022)	–	(0.014)
Index edu disaggr	0.023*	–	0.186	–	0.142
	(0.014)	–	(0.269)	–	(0.089)
Index demo disaggr	0.105	–	−0.107	–	−0.029
	(0.180)	–	(0.139)	–	(0.142)
Observations	8962	–	5447	–	1717
R2	0.231	–	0.281	–	0.198

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Firm level diversity is instrumented by using the predicted level of diversity at the commuting area, where the firm is located. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies. Standard errors are clustered at the commuting area level.
Significance levels: ***1%, **5%, *10%.

Mongolia, Taiwan; *Other Asia*: all the other Asian countries non included in both East Asia and Muslim Countries categories and Africa, all the other African countries not included in the Muslim Country; *Western and Southern Europe*: all the other European countries not included in the Formerly Communist Countries category.

(2) *The linguistic groups are as follows*: *Germanic West* (Antigua Barbuda, Aruba, Australia, Austria, Bahamas, Barbados, Belgium, Belize, Bermuda, Botswana, Brunei, Cameroon, Canada, Cook Islands, Dominican Republic, Eritrea, Gambia, Germany, Ghana, Grenada, Guyana, Haiti, Ireland, Jamaica, Liberia, Liechtenstein, Luxemburg, Mauritius, Namibia,

Table A10
IV results: first stage regressions (aggregate specification of diversity).

Dependent variable: firm level diversity	Index ethnic aggr	Index edu aggr	Index demo aggr
Manufacturing			
Index ethnic com	0.949***	−0.064	−0.060**
	(0.080)	(0.045)	(0.027)
index edu com	−0.348**	0.072***	−0.172***
	(0.110)	(0.028)	(0.041)
index demo com	−1.181***	0.024	1.054***
	(0.246)	(0.182)	(0.093)
F test (excluded instruments); *p*-value	48.51; 0.000	11.67; 0.000	30.61; 0.000
Observations	48,238	48,238	48,238
R2	0.404	0.462	0.371
Construction			
Index ethnic com	0.247***	−0.301***	−0.041
	(0.063)	(0.062)	(0.032)
Index edu com	0.079	0.241**	−0.076
	(0.051)	(0.092)	(0.061)
Index demo com	−0.119	0.685***	0.590***
	(0.252)	(0.173)	(0.140)
F test (excluded instruments); *p*-value	9.19; 0.000	7.93; 0.003	7.12; 0.004
Observations	26,969	26,969	26,969
R2	0.223	0.513	0.397
Wholesale and retail trade			
Indexethnic com	0.463***	−0.411***	0.234***
	(0.099)	(0.050)	(0.049)
Index edu com	0.101	0.902***	−0.655***
	(0.190)	(0.092)	(0.058)
Index demo com	0.147*	−0.281***	1.392***
	(0.086)	(0.079)	(0.096)
F test (excluded instruments); *p*-value	30.82; 0.000	28.42; 0.000	81.58; 0.000
Observations	41,493	41,493	41,493
R2	0.331	0.487	0.47
Transport			
Index ethnic com	0.407***	0.157	−0.168**
	(0.106)	(0.239)	(0.067)
Index edu com	0.060	0.507**	0.155
	(0.166)	(0.231)	(0.145)
Index demo com	−1.030**	0.135	0.589**
	(0.326)	(0.249)	(0.188)
F test (excluded instruments); *p*-value	11.97; 0.000	6.94; 0.000	9.48; 0.000
Observations	6,285	6,285	6,285
R2	0.393	0.327	0.294
Financial and business services			
Index ethnic com	0.678***	−0.037	−0.139**
	(0.186)	(0.093)	(0.060)
Index edu com	−0.041	0.297***	0.070
	(0.133)	(0.086)	(0.072)
Index demo com	−0.018	−0.336*	1.004***
	(0.157)	(0.183)	(0.111)
F test (excluded instruments); *p*-value	10.35; 0.000	5.12; 0.002	51.19; 0.000
Observations	14,008	14,008	14,008
R2	0.43	0.593	0.441

Note: The dependent variable is diversity at the firm level. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the commuting area level.

*** 1%.

** 5%.

* 10%.

Table A11
Labor diversity and productivity, IV results: second stage regressions (aggregate specification of diversity).

Dependent variable: log of TFP	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
index ethnic aggr	−0.049**	−0.037*	−0.069**	−0.051	−0.027
	(0.016)	(0.018)	(0.025)	(0.058)	(0.036)
index edu aggr	0.029*	0.026*	0.048**	0.092	0.058*
	(0.016)	(0.014)	(0.013)	(0.069)	(0.029)
index demo aggr	0.027	0.070	−0.012	−0.068	−0.048
	(0.022)	(0.096)	(0.027)	(0.070)	(0.033)
Observations	35,887	18,024	26,418	4007	7931
R2	0.312	0.114	0.329	0.165	0.159

Note: The dependent variable is the log of total factor productivity estimated by using the ACF approach. Firm level diversity is instrumented by using the predicted level of diversity at the commuting area, where the firm is located. All regressions include the shares of employees belonging to each category considered in the diversity indices; shares of managers and middle-managers; average firm tenure; whether the firm is foreign owned, multi-establishment, and exports; a full set of 2-digit industry, year, size, and county dummies; and industry-year interactions. Standard errors are clustered at the commuting area level.
Significance levels: ***1%, **5%, *10%.

Netherlands, Netherlands Antilles, New Zealand, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and Grenadines, Seychelles, Sierra Leone, Solomon Islands, South Africa, St. Helena, Suriname, Switzerland, Trinidad and Tobago, Uganda, United Kingdom, United States, Zambia, Zimbabwe); *Slavic West* (Czech Republic, Poland, Slovakia); *Germanic Nord* (Denmark, Iceland, Norway, Sweden); *Finno-Permic* (Finland, Estonia); *Romance* (Andorra, Angola, Argentina, Benin, Bolivia, Brazil, Burkina Faso, Cape Verde, Chile, Colombia, Costa Rica, Cote D′Ivoire, Cuba, Djibouti, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, France, French Guyana, Gabon, Guadeloupe, Guatemala, Guinea, Guinea Bissau, Holy See, Honduras, Italy, Macau, Martinique, Mexico, Moldova, Mozambique, Nicaragua, Panama, Peru, Portugal, Puerto Rico, Reunion, Romania, San Marino, Sao Tome, Senegal, Spain, Uruguay, Venezuela); *Attic* (Cyprus, Greece); *Ugric* (Hungary); *Turkic South* (Azerbaijan, Turkey, Turkmenistan); *Gheg* (Albania, Kosovo, Republic of Macedonia, Montenegro); *Semitic Central* (Algeria, Bahrain, Comoros, Chad, Egypt, Irak, Israel, Jordan, Kuwait, Lebanon, Lybian Arab Jamahiria, Malta, Mauritiania, Morocco, Oman, Qatar, Saudi Arabia, Sudan, Syrian Arab Republic, Tunisia, Yemen, United Arabs Emirates); *Indo-Aryan* (Bangladesh, Fiji, India, Maldives, Nepal, Pakistan, Sri Lanka); *Slavic South* (Bosnia and Herzegovina, Croatia, Serbia, Slovenia); *Mon-Khmer East* (Cambodia); *Semitic South* (Ethiopia); *Slavic East* (Belarus, Bulgaria, Georgia, Mongolia, Russian Federation, Ukraine); *Malayo-Polynesian West* (Indonesia, Philippines); *Malayo-Polynesian Central East* (Kiribati, Marshall Islands, Nauru, Samoa, Tonga); *Iranian* (Afghanistan, Iran, Tajikistan); *Betai* (Laos, Thailand); *Malayic* (Malasya); *Cushitic East* (Somalia); *Turkic East* (Uzbekistan); *Viet-Muong* (Vietnam); *Volta-Congo* (Burundi, Congo, Kenya, Lesotho, Malawi, Nigeria, Rwanda, Swaziland, Tanzania, Togo); *Turkic West* (Kazakhstan, Kyrgystan); *Baltic East* (Latvia, Lithuania); *Barito* (Madagascar); *Mande West* (Mali); *Lolo-Burmese* (Burma); *Chadic West* (Niger); *Guarani* (Paraguay); *Himalayish* (Buthan); *Armenian* (Armenia); *Sino Tibetan* (China, Hong Kong, Singapore, Taiwan); *Japonic* (Japan, Republic of Korea, Korea D.P.R.O.).

A.2. Additional results

Aggregate specification for diversity has been included in Tables A10 and A11.

References

Ackerberg, Daniel A., Caves, Kevin, Frazer, Garth, 2006. Structural Identification of Production Functions. Department of Economics, University of California.

Alcacer, Juan, Chung, Wilburg, 2010. Location Strategies for Agglomeration Economies. HBS Working Paper 10-071. Harvard Business School, Boston.

Audretsch, David B., Feldman, Maryann P., 1996. R&D spillovers and the geography of innovation and production. Am. Econ. Rev. 86 (3), 630–640.

Adams, James D., Jaffe, Adam, 1996. Bounding the effects of R&D: an investigation using linked establishment and firm data. RAND J. Econ. 98, 673–702.

Adsera, Alicia, Pytlikova, Mariola, 2011. The Role of Language in Shaping International Migration: Evidence from OECD Countries 1985–2006. Unpublished Manuscript. Department of Economics, University of Aarhus.

Alesina, Alberto, La Ferrara, Eliana, 2002. Who trusts others? J. Publ. Econ. 85, 207–234.

Alesina, Alberto, La Ferrara, Eliana, 2005. Ethnic diversity and economic performance. J. Econ. Lit. 43, 762–800.

Andersen, Anne K., 2000. Commuting Areas in Denmark. AKF Working paper.

Becker, Gary, 1957. The Economics of Discrimination. University of Chicago Press, Chicago.

Barrington, Linda, Troske, Kenneth R., 2001. Workforce Diversity and Productivity: An Analysis of Employer-Employee Matched Data. Economics Program Working Papers No. 01-02.

Berliant, Marcus, Fujita, Masahisa, 2008. Knowledge creation as a square dance on the Hilbert cube. Int. Econ. Rev. 49, 1251–1295.

Bombardini, Matilde, Gallipoli, Giovanni, Pupato, German, 2012. Unobservable Skill Dispersion and Comparative Advantage. Mimeo. University of British Columbia.

Card, David, Di Nardo, John E., 2000. Do immigrant inflows lead to native outflows? Am. Econ. Rev.: Pap. Proced. 90, 360–367.
Cortes, Patricia, 2008. The effect of low-skilled immigration on U.S. prices: evidence from CPI data. J. Polit. Econ. 116 (3), 322–381.
Chiswick, Barry., Miller, Paul W., 1995. The endogeneity between language and earnings: international analyses. J. Labor Econ. 2, 246–288.
Deding, Mette, Filges, Trine, Ommeren, Jos Van, 2009. Spatial mobility and commuting: the case of two-earner households. J. Reg. Sci. 49, 113–147.
Delgado, Mercedes, Porter, Michael E., Stern, Scott, 2010. Clusters and entrepreneurship. J. Econ. Geogr. 10, 418–495.
Dustmann, Christian, Francesca, Fabbri, Ian, Preston, 2005. The impact of immigration on the British labor market. Econ. J. 115 (507), F324–F341.
Foley, C. Fritz, William, R. Kerr, 2011. Ethnic Innovation and US Multinational Firm Activity. National Bureau of Economic Research Working Paper 17336.
Fox, Jeremy T., Smeets, Valerie, 2011. Does input quality drive measured differences in firm productivity? Int. Econ. Rev. 52 (4), 961–989.
Gandhi, Amit, Navarro, Salvador, Rivers, David, 2011. On the Identification of Production Functions: How Heterogeneous is Productivity?. University of Wisconsin-Madison and University of Western Ontario.
Garnero, Andrea, Rycz, Francois, 2013. The Heterogeneous Effects of Workforce Diversity on Productivity, Wages and Profit. IZA Discussion Paper No. 7350.
Glaeser, Edward, Laibson, David I., Scheinkman, Jos A., Soutter, Christine L., 2000. Measuring trust. Q. J. Econ. 115, 811–846.
Grossman, Gene Michael, Maggi, Giovanni, 2000. Diversity and trade. Am. Econ. Rev. 20, 1255–1275.
Grund, Christian, Westergaard-Nielsen, Niels, 2009. Age structure of the workforce and firm performance. Int. J. Manpower 29, 410–422.
Guiso, Luigi, Sapienza, Paolo, Zingales, Luigi, 2009. Cultural biases in economic exchange? Q. J. Econo. 124, 1095–1131.
Hamilton, Barton H., Nickerson, Jack A., Owan, Hideo, 2003. Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation. J. Polit. Econ. 111, 465–497.
Hamilton, Barton H., Nickerson, Jack A., Owan, Hideo, 2004. Diversity and Productivity in Production Teams. Unpublished Manuscript. Department of Economics, Washington University in St. Louis.
Hellerstein, Judith, Neumark, David, 2004. Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer–Employee Data Set. NBER Working Paper 10325.
Hong, Lu, Page, Scott E., 2001. Problem solving by heterogeneous agents. J. Econ. Theory 97, 123–163.
Hong, Lu, Page, Scott E., 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proc. Natl. Acad. Sci. 101, 123–139.
Iranzo, Susana, Schivardi, Fabiano, Tosetti, Elisa, 2008. Skill dispersion and firm productivity: an analysis with employer–employee matched data. J. Labor Econ. 26, 247–285.
Irarrazabal, Alfonso, Moxnes, Andreas, Ulltveit-Moe, Karen H, 2014. Heterogeneous firms or heterogeneous workers? Implications for export premia and the gains from trade. Rev. Econ. Stud., 95 (3), 839–849, http://dx.doi.org/10.1162/REST_a_00285.
Isphording, Ingo E., Otten, Sebastian, 2011. Linguistic Distance and the Language Fluency of Immigrants. Ruhr Economic Papers No. 274.
Kurtulus, Ana F., 2011. What types of diversity benefit workers? Empirical evidence on the effects of co-worker dissimilarity on the performance of employees. Ind. Relat.: J. Econ. Soc. 50, 678–712.
Kremer, Michael, 1993. The O-ring theory of economic development. Q. J. Econ. 108, 551–575.
Krugman, Paul, 1991. Geography and Trade. MIT Press, Cambridge, MA.
Lazear, Edward P., 1998. Personnel Economics for Managers. John Wiley & Sons, New York.
Lazear, Edward P., 1999. Globalisation and the market for team-mates. Econ. J. 109, 15–40.
Leonard, Jonathan S., Levine, David I., 2006. Diversity, Discrimination, and Performance. Institute for Research and Employment Working Paper, No. 147.
Levinsohn, James, Petrin, Amil, 2003. Estimating production functions using inputs to control for unobservables. Rev. Econ. Stud. 70, 317–341.
Lewis, M. Paul, 2009. Ethnologue: Languages of the World, 16th edition. ISBN-13 978-1-55671-216-6. ⟨http://www.ethnologue.com/web.asp⟩.
Maignan, Carole, Ottaviano, Gianmarco, Pinelli, Dino, Rullani, Francesco, 2003. Bio-Ecological Diversity vs. Socio-Economic Diversity: A Comparison of Existing Measures. Nota di Lavoro, Fondazione Eni Enrico Mattei.
Montgomery, James D., 1991. Social networks and labor market outcomes: toward an economic analysis. Am. Econ. Rev. 81, 1408–1418.
Mundlak, Yair, 1961. Empirical production function free of management bias. J. Farm Econ. 43, 44–56.
Munshi, Kaivan, 2003. Networks in the modern economy: Mexican migrants in the US labor market. Q. J. Econ. 118, 549–599.
Navon, Guy. 2009. Human Capital Spillovers in the Workplace: labor Diversity and Productivity. MPRA Paper No. 17741.
OECD, 2005. Economic Survey of Denmark. Paris, 2005. Labor Research 21, 463–475.
Olley, G. Steven, Pakes, Ariel, 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. Econometrica 64, 1263–1297.
Osborne, E., 2000. The deceptively simple economics of workplace diversity. J. Labor Res. 21, 463–475.
Ottaviano, Gianmarco I.P., Peri, Giovanni, 2006. The economic value of cultural diversity: evidence from US cities. J. Econ. Geogr. 6, 9–44.
Ottaviano, Gianmarco I.P., Peri, Giovanni, 2012. Rethinking the effects of immigration on wages. J. Eur. Econ. Assoc. 10 (1), 152–197, http://dx.doi.org/10.1111/j.1542-4774.2011.01052.
Pedersen, Peder J., Pytlikova, Mariola, Smith, Nina, 2008. Selection and Network Effects—migration flows into OECD countries 1990–2000. Eur. Econ. Rev. 52, 1160–1186.
Peri, Giovanni, 2012. The Effect of immigration on productivity: evidence from US states. Rev. Econ. Stat. 94 (1), 348–358.
Petrin, Amil, Levinsohn, James, Poi, Brian, 2003. Production function estimation in Stata using inputs to control for unobservables. Stata J. 4, 113–123.
Rauch, James E., Casella, Alessandra, 2003. Overcoming informational barriers to international resource allocation: prices and ties. Econ. J. 113, 21–42.
Sparber, Chad, 2009. Racial diversity and aggregate productivity in US industries: 1980–2000. South. Econ. J. 75, 829–856.
Sparber, Chad., 2010. Racial diversity and macroeconomic productivity across US states and cities. Reg. Stud. 44, 71–85.
Suedekum, Jens, Wolf, Katja, Blien, Uwe, 2009. Cultural Diversity and Local Labor Markets. IZA Discussion Paper No. 4619.
Wooldridge, Jeffrey Marc, 2009. On estimating firm-level production functions using proxy variables to control for unobservables. Econ. Lett. 104, 112–114.
United Nations, 2000. Replacement Migration: Is it a Solution to Declining and Aging Populations? New York, UN Population Division. Department of Economic and Social Affairs.