

**NORFACE**  
MIGRATION



---

**NORFACE MIGRATION** Discussion Paper No. 2011-22

## Does Labor Diversity affect Firm Productivity?

Pierpaolo Parrotta, Dario Pozzoli and Mariola Pytlikova

# Does Labor Diversity affect Firm Productivity?\*

Pierpaolo Parrotta<sup>†</sup>, Dario Pozzoli<sup>‡</sup> and Mariola Pytlikova<sup>§</sup>

## Abstract

Using a matched employer-employee data-set, we analyze how workforce diversity in cultural background, education and demographic characteristics affects productivity of firms in Denmark. Implementing a structural estimation of the firms' production function (Akerberg et al., 2006) we find that labor diversity in education significantly enhances a firm's value added. Conversely, diversity in ethnicity and demographics induces negative effects on firm productivity. Hence, the negative effects, coming from communication and integration costs connected to a more culturally and demographically diverse workforce, seem to outweigh the positive effects coming from creativity and knowledge spillovers.

**JEL Classification:** C23, J24, L20.

**Keywords:** Labor diversity, skill complementarity, communication barriers, total factor productivity.

---

\*We thank Holger Bonin, Tor Eriksson, Michael Rosholm, Chad Syverson and Frederic Warzynski (alphabetical order) for helpful comments and suggestions. In addition, we appreciate comments from participants at seminars organized by the Copenhagen Business School, University of Bergamo, Aarhus School of Business, University of Lausanne, 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 London. Pierpaolo Parrotta acknowledges the financial support from the Graduate School for Integration, Production and Welfare. Mariola Pytlikova gratefully acknowledges funding from the NORFACE Migration Programme. The usual disclaimer applies.

<sup>†</sup>Business and Social Sciences, Aarhus University; Department of Economics, University of Lausanne. Email: pierpaolo.parrotta@unil.ch.

<sup>‡</sup>Business and Social Sciences, Aarhus University. Department of Economics and Centre for Research in Integration and Marginalization (CIM). Email: dpozzoli@asb.dk.

<sup>§</sup>Business and Social Sciences, Aarhus University. Department of Economics, Center for Corporate Performance (CCP) and Centre for Research in Integration and Marginalization (CIM). Email: marp@asb.dk.

# 1 Introduction

Diverse labor force is increasingly a reality in many developed countries. This results, among other things, from the following major factors: policy measures to counteract population aging, anti-discrimination measures, the growth in immigration from diverse origins experienced during the latest decades and educational and skill upgrading of workforces.<sup>1</sup> All that leads to an increasing diversity of labor force in terms of age, gender, ethnicity and skills.

From the demand side, we observe increasing diversity across many workplaces and we often hear about the importance of further internationalization and demographic diversification. In many countries firms' hiring decisions are affected by governmental affirmative action policies. Besides that, firms are often under pressure to be more diverse, because this is how they should socially look. At the same time, firms are challenged by constantly changing demand for goods and services, new customers and markets in today's globalized world. The diverse workforce may be a key factor in helping firms to understand and to meet the new needs.

Popular press usually emphasizes workforce diversity to be beneficial for firms, but is it really true? Do firms benefit from the labor diversity, so that it is translated into their competitive advantage? What is the relationship between workplace labor diversity and firm performance? Although the issue is very important, there is considerable ambiguity surrounding the topic.

So far the theory suggests that there are two forces pulling the effect of demographic and cultural diversity on firm performance in the opposite directions: (1) diversity can create negative effects due to worse communication, lower social ties and trust, and worse cooperation among workers (Becker, 1957; Lang, 1986; Lazear, 1998 and 1999), and conversely (2) diversity can be beneficial to the firm performance due to better decision making, improved problem solving, more creativity, and more information about global products markets (Alesina and La Ferrara, 2005; Hong and Page 2001 and 2004; Berliant and Fujita, 2008; Glaeser et al. 2000; Osborne, 2000; Rauch and Casella, 2003). Diversity in skills, education and tenure may generate knowledge spillovers and skill complementarity among the employees and thus it has a positive effect on firm

---

<sup>1</sup>Demographic projections by the United Nations suggest that during the next four decades populations in Europe might *ceteris paribus* decline by 12 per cent (United Nations, 2000). The main factor responsible for the population aging is a large decline in the total fertility rate over the last half century. As a consequence the governments have often adopted a number of measures to counteract the problem of population aging such as policies encouraging people to work longer, increasing female labor participation and attracting skilled immigrants. In many countries, governments increased the regular and early-retirement age, restricted access to early retirement by changing economic incentives, and promoted age anti-discrimination measures. Female labor participation has grown in most of the world during the last century (OECD, 2005). This is partly due to policies encouraging women to work, e.g. better childcare and parental leave provisions, and gender anti-discrimination measures. Furthermore, we observe an increase in the immigration towards the developed countries (and not only) and a broader diversity of immigrants with respect to migrant's origins (Adsera and Pytlikova, 2011; Pedersen et al., 2008). Subsequently, diversity of workforce with respect to gender, age and ethnicity has increased. Last but not least, as a consequence of the worldwide globalization process and skill biased technological change governments across many countries took a number of steps to increase the skill level of the workforce, by e.g. increasing the supply of university educated people and by enhancing the availability of lifelong learning.

performance (Lazear, 1998 and 1999). In the empirical literature, there seems to be some consensus on the positive contribution of skill diversity to firm productivity, whereas the evidence of diversity along ethnic and demographic lines on firm performance is rather mixed. Nevertheless, most of the previous studies were based on case studies within one firm (e.g. Hamilton et al., 2003, 2004; Kurtulus, 2009; Leonard and Levine, 2006), or on aggregate regional data (e.g. Ottaviano and Peri, 2006; Sparber, 2009, 2010; Suedekum et al., 2009). Evidence using more comprehensive micro-data is typically fairly scarce, and refers to single dimensions of diversity.

In this paper, we use a register-based linked employer-employee data-set (LEED) from Denmark, which provides us with a wide collection of information on individuals and firms' characteristics. Merged with a firm-level financial accounting data-set for the years 1995-2005, this LEED allows us to overcome many limitations of the previous studies and to shed some light on the rather unexplored research area of the effects of workforce diversity on firm productivity. Specifically, we investigate the effect of labor diversity on firm productivity by looking at three relevant dimensions: cultural background, education and demographics. It implies that we try to capture a multi-dimensionality of labor diversity and eventual different implications related to each of these dimensions in terms of productivity. In our analyses, we follow a recent structural estimation technique suggested by Akerberg, Caves and Frazen (2006), which allows us to address both collinearity and endogeneity issues, which previous literature has often ignored. In this model, firms are allowed to observe productivity shocks before deciding the optimal level of diversity, which ensures that the contribution of this input to productivity is estimated properly. In fact, the realization of a positive (unobservable) productivity shock may favor investments in labor diversity. Further, we dig deeper into the effects of workforce diversity on firm outcomes by testing different hypotheses derived from the existing theory. Specifically, we look at whether diversity plays a different role for distinct occupational groups as we expect that the beneficial effects of diverse problem-solving abilities and creativity materialize more in terms of higher productivity for white-collar occupations compared to blue-collar occupations. In addition, we test how important are the communication costs and costs of "cross-cultural dealing" by excluding certain groups of foreigners (i.e. individuals with tertiary education or speaking a language belonging to the Germanic group) from calculation of ethnic diversity measures. Finally, we evaluate the effects of the different dimensions of diversity on productivity for firms operating in more innovative or trade-open industries.

We find evidence that labor diversity in education significantly enhances firm productivity, whereas ethnic and demographic heterogeneity lowers it. However, diversity in demographics brings statistically insignificant results for the most complete specifications. These findings are in line with past relevant works by Lazear (1999), Glaeser et. al. (2000), and Alesina and La Ferrara (2002) and shows that the negative effects, coming from communication and integration costs connected to more demographically and culturally diverse work-

force, counteract the positive effects of diversity on firm productivity, coming from creativity and knowledge spillovers.

The structure of the paper is as follows: section 2 reviews related literature and derives hypotheses, section 3 briefly describes the data, section 4 provides details on the empirical strategy, sections 5 and 6 explain the results of our empirical analysis and section 7 offers some concluding remarks.

## 2 Previous literature and hypotheses development

Economic theory suggests that workforce diversity may affect firm performance differently and through various channels. Diversity in skills and education may generate knowledge spillovers and skill complementarity among the employees within a firm (as long as workers' information are relevant), which positively affects firm performance (Lazear, 1999). Similarly, diversity in age can be beneficial to firms because there are complementarity between the human capital of younger and older workers. Younger employees have knowledge of new technologies and IT and older employees have a better understanding and experience with the intra-firm structures and the operating process (Lazear, 1998). On the other hand, Becker's (1957) model of co-worker discrimination suggests that demographic heterogeneity among workers may create communication frictions if workers are prejudiced, and thus bring some costs connected to the frictions.

The theoretical contribution on the effect of ethnic and cultural diversity on firm performance brings mixed conclusions. Ethnic-cultural diversity may affect firm performance negatively as it may (i) hinder potential knowledge transfers among workers due to linguistic and cultural barriers, (ii) reduce peer pressure by weakening social ties and trust among them, and (iii) create non-pecuniary disutility of joining or remaining in a demographically diverse firm (Lazear, 1999). A similar point on trust is made by Glaeser et. al. (2000), and Alesina and La Ferrara (2002) showing that people often distrust members of other ethnic groups and tend to prefer interacting in culturally homogeneous communities. On the other hand, ethnic diversity can be beneficial to the firm performance through better decision making and improved problem solving (Hong and Page, 2001 and 2004). In their models, diverse groups of problem solvers consistently outperform the homogeneous groups of the individuals who are best at solving problems. The reason is that the diverse groups get stuck less often than homogeneous groups of high-ability solvers, who tend to think similarly. The authors argue that it is because more diverse groups have a broader spectrum of perspectives improving their decision-making (Hong and Page, 2001 and 2004). Berliant and Fujita (2008) also refer to the significance of cultural diversity for creation of new ideas and knowledge, and knowledge transfer. Further, Alesina and La Ferrara (2005) propose a simple theoretical framework, in which skills of ethnically heterogeneous groups of individuals are complementary in the production process for a private good, bringing more innovation and creativity, which translates diversity into increased productivity. However as individual utility also

depends on the consumption of a shared public good and as heterogeneous ethnic groups may have different public goods preferences, increased diversity lowers the utility from public good consumption (Alesina and La Ferrara, 2005). In addition, workforce diversity may provide useful information to the firm about the product's market, enhancing the firm's ability to compete in global markets (Osborne, 2000; Rauch and Casella, 2003).

To our knowledge, the empirical evidence concerning diversity and economic performance has been fairly scarce, and most of the previous studies were based on case studies within one firm (e.g. Hamilton et al., 2003, 2004; Kurtulus, 2011; Leonard and Levine, 2006), or on aggregate regional data (e.g. Ottaviano and Peri, 2006 and 2011; Suedekum et al., 2009), whereas evidence using more comprehensive data is fairly limited (Iranzo et al., 2008; Navon, 2009; Grund and Westergaard-Nielsen, 2008). Moreover, the majority of the previous studies has focused on only one dimension of diversity on firm performance, with the studies by Kurtulus (2011) and Leonard and Levine (2006) being the only exceptions.

Summarizing the key findings of the studies: (i) the former group of case studies find that diversity with respect to skills and knowledge has a positive effect on worker performance, whereas diversity in age and race lowers firm performance (Hamilton et al., 2003, 2004; Leonard and Levine, 2006; Kurtulus, 2011); (ii) studies using aggregated regional data find a positive effect of citizenship diversity on performance (e.g. Ottaviano and Peri, 2006 and 2011; Alesina and La Ferrara, 2005; Sparber, 2009; Suedekum et al., 2009; Peri, 2011); (iii) studies using the micro linked employer-employee data find a positive effect of skill diversity on firm performance (Iranzo et al., 2008; Navon, 2009), positive or no significant effect of ethnicity diversity on firm performance (Barrington and Troske, 2001) and inverse U-shaped relationship between age diversity and firm productivity (Grund and Westergaard-Nielsen, 2008). So there seems to be some consensus with respect to skill diversity being positively related to firm performance,<sup>2</sup> whereas the evidence of diversity along ethnic and demographic lines on performance is rather mixed.

Based on the different theoretical approaches and their predictions, we try to derive hypotheses for the effect of diversity on firm productivity. From the existing theoretical contributions it is clear that there are two forces driving the effect in opposite directions. On the one hand the demographic and ethnic diversity can benefit the firm with a more diverse spectrum of problem solving abilities, creativity and knowledge spillovers,

---

<sup>2</sup>There is quite a large amount of literature on the role of skill/educational distribution on firm performance and how it has changed over time, mostly due to skill biased technological change (SBTCH). Some argue that it is important to have few talented workers a la "superstar", which leads to more dispersed skill distribution of the workforce (Rosen, 1981). Others claim that tasks are performed at a certain level of competence leading to teams of workers with similar skills and more segregation (Kremer, 1993)). Some matching and sorting models argue that production has shifted from mode of hiring more diverse workers towards modes, where some firms hire only high-skilled (e.g. Microsoft) and other firms hire only low-skilled (e.g. McDonalds), resulting in segregation (Kremer and Maskin, 1996)). Some argue that SBTCH reduces communication costs and increases an optimal degree of skill dispersion (Garicano and Rossi-Hansberg, 2006). For some discussion and evidence of educational sorting see Eriksson et al. (2009). In our paper, we do not refer to educational diversity as overall educational distribution. By educational diversity we mean diversity in educational complementarity, i.e. we focus on different specializations, e.g. we distinguish between different sciences of the highest achieved education, see the educational diversity index described in the next section of the paper.

which in turn foster firm productivity (Lazear, 1998; Hong and Page, 2001 and 2004; Berliant and Fujita, 2004; Alesina and La Ferrara, 2005). We expect the inter-cultural learning and knowledge spillovers to materialize more easily for white-collar occupations compared to blue-collar occupations or for firms operating in more creative industries or in industries more open towards trade (Osborne, 2000; Rauch and Casella, 2003). On the other hand, the demographic and ethnic diversity may also reduce productivity because of higher costs connected to communication barriers and higher distrust levels, which arise if people of different cultural backgrounds, gender and ages have to interact and work together on projects (Lazear, 1999; Glaeser et. al., 2000; Alesina and La Ferrara, 2002). Specifically, we expect these costs of “cross-cultural dealing” to be more important when we exclude certain groups of foreigners who probably speak English (highly educated workers) or Danish from the calculation of ethnic diversity measure. Regarding skill diversity, there is a consensus across the existing theoretical contributions that because of the knowledge spillovers skill-related diversity shall bring a positive effect on firm productivity, as long as workers’ information sets are not overlapping but relevant to one another (Lazear, 1999).

## **3 Data**

### **3.1 Data description**

The data set for this empirical investigation is created by merging information from two different main sources. The first one is the "Integrated Database for Labor Market Research" (IDA henceforth) provided by Statistics Denmark. IDA is a longitudinal employer-employee register containing valuable information (age, demographic characteristics, education, labor market experience and earnings) on each individual employed in the recorded population of Danish firms during the period 1980-2005. Apart from deaths and permanent migration, there is no attrition in the data-set. The labor market status of each person is recorded at 30th November each year. The retrieved information is aggregated at the firm level to obtain variables such as firm size, workforce composition characteristics (average firm tenure, shares of managers, middle managers, males, highly skilled workers, technicians, shares of employees belonging to each age distribution quartiles), labor diversity (see the next section for more details) and partial/total foreign ownership or whether the firm is a multi-establishments firm.

The second data source refers to firms’ business accounts (REGNSKAB henceforth), which is also provided and compiled by Statistics Denmark. It covers the construction and the manufacturing industry from 1994, manufacturing from 1995, wholesale trade from 1998 and the remaining part of the service industry from 1999 onwards. From REGNSKAB, the following accounting items are retrieved for the estimation of the production

function: value added<sup>3</sup>, materials (intermediates), capital stock (fixed assets) and related industry.<sup>4</sup>

### 3.2 Firms' labor diversity

This section focuses on the measurement of employees' diversity at firm level. Labor's diversity is quantified by using information regarding workers' gender, age, work experience, highest fulfilled education and nationality. We use the Herfindahl index to measure the degree of diversity at firm level. Contrary to the traditional diversity measures, like the percentage of employees belonging to a specific group, the Herfindahl index combines two quantifiable measures: the "richness" (number of categories represented within the firm or the workplace) and "equitability" or evenness (how even are the numbers of the individual categories). Specifically, we calculate three separate indexes to measure diversity along the cultural, skill and demographic dimensions.

Cultural diversity is represented alternatively by the employees' nationality or language spoken. The nationality has been grouped in the following categories: North America and Oceania, Central and South America, Africa, West and South Europe, Formerly Communist Countries, Asia, East Asia, Muslim Countries.<sup>5</sup> It has been argued in previous literature<sup>5</sup> that linguistic distance serves as a good proxy for cultural distance (Guiso et al, 2009; Adsera and Pytlikova, 2011), therefore we have grouped the employees together by languages spoken in their country of origin. The linguistic classification is more detailed than the grouping by nationality. Specifically, we group countries (their major official language spoken by the majority) by 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 "Ethnologue: Languages of the World" (Lewis, 2009), see the Appendix section for more details about the list of countries and the linguistic groups included. The education-related diversity is represented by 6 categories based on information on employees' highest achieved educational level (tertiary education, secondary and vocational education, and below secondary education). We divide tertiary education into 4 categories making a distinction between Bachelor, Master and PhD degrees in social science, humanities, engineering and natural sciences. In a more disaggregated specification, we also distinguish secondary education into general high school, business high school, short and long vocational education. Finally, the demographic index is built on the intersection of gender and age quartiles or age quintiles (8 or 9 categories in total, depending on the level of aggregation).

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

---

<sup>3</sup>Computed as the difference between total sales and intermediates.

<sup>4</sup>The following sectors are excluded from the empirical analysis: i) agriculture, fishing and quarrying; ii) electricity, gas and water supply and iii) public services.

<sup>5</sup>Second generation immigrants are not treated as foreigners or as a distinct group from the natives in the main analysis. However, we show a specification where the second generation immigrants is included into the group of foreigners in the section on mechanisms driving the effect of workforce diversity on firm productivity.



as follows:

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

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  $i$ ,  $H$  is the total number of categories of the related diversity dimension,  $N_w$  and  $N_i$  are respectively 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  $p_{swt}$ .<sup>6</sup> The diversity index has a minimum value equal to 0 if there is only one category represented within the workplace, and a maximum value equal to  $\left(1 - \frac{1}{H}\right)$  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.

### 3.3 Descriptive statistics

Before discussing some descriptive statistics of the variables included in the main analysis, it is important to stress that (a) firms with imputed accounting variables and (b) firms with less than 10 employees are dropped from the main sample. The first choice is obviously to reinforce the reliability of our empirical analysis. The second one is to allow all investigated firms to potentially reach the highest degree of ethnic diversity at least when an aggregated specification is used.<sup>7</sup> All in all, we are able to analyze the total factor productivity of about 24,000 firms for years 1995 to 2005.

Table 1 provides some basic descriptive statistics on all variables used in our analysis for the main sample and by firm size. More specifically we split the sample into two main groups referring to firms above and below 50 employees. Consistently with the Danish industrial structure within the private sector, 78% of the observations is represented by small-sized firms.<sup>8</sup> Compared with larger firms, small companies are characterized by lower levels of value added, materials and capital stock.<sup>9</sup> Moreover, whereas higher shares of middle managers, relatively younger employees and personnel with vocational education characterizes

---

<sup>6</sup>For the ethnic diversity, the shares of foreign workers of different nationalities/linguistic groups in each workplace have been calculated as follows:

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

<sup>7</sup>When a linguistic classification is adopted, we adjust the ethnic diversity to take account of the firm size. 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 ( $H$ ).

<sup>8</sup>According to the OECD (2005), the structure of the Danish firm population is mainly composed of small and medium-sized companies as firms with less than 50 employees account for 97 per cent of the total number of firms and represent 42 per cent of the total employment in manufacturing and services.

<sup>9</sup>Values of accounting 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.

small firms, larger proportions of managers, women and foreigners distinguish companies with more than 50 employees. The two groups of firms are comparable in terms of average tenure of employees and firm ownership.

[Insert Table 1 and 2 around here]

Table 2 reports detailed descriptive statistics of all diversity indexes by industry, year and firm size. We observe higher values of diversity indexes for firms within the manufacturing and the financial and business services sectors, and lower diversity in all dimensions for small firms, no matter the level of aggregation used to measure workforce heterogeneity. Finally, diversity is slightly increasing over time, especially in terms of ethnicity. That is in line with the trend of growing immigration to Denmark during the latest decades.

## 4 Empirical strategy

### 4.1 Productivity estimation

As pointed out by the literature on the identification of firm production functions, the major issue in the estimation of parameters is the possibility that factors influencing production are unobserved by the econometrician but observed by the firm. In such case, asymmetrically observed shocks may be taken into account by firms to maximize their profits or minimize their costs. Specifically, it is expected that firms respond to positive (negative) productivity shocks by expanding (reducing) output, which requires higher quantity/quality production inputs. This implies that OLS estimates of coefficients on inputs observed by econometricians are biased and inconsistent, and error terms and regressors are correlated. Moreover, it is widely acknowledged that whereas fixed-effects (FE) estimation techniques (Mundlak, 1961) take account of firm heterogeneity, they do not solve the simultaneity problem when productivity shocks fluctuate over time.

Several remedies have been proposed to cope with simultaneity, including the recent structural approach advocated by Olley and Pakes (1996) (OP henceforth) and Levinsohn and Petrin (2003) (LP henceforth).<sup>10</sup> Both OP and LP suggest semi-parametric methods based on (i) the identification of a proxy variable, which is assumed to be a function of time-varying productivity shocks (total factor productivity) and (ii) the definition of conditions under which this function is invertible. The idea behind this, is to infer total factor productivity by using observed firms' input choices (Wooldridge, 2009).<sup>11</sup> Although, OP and LP are broadly

---

<sup>10</sup>See Akerberg et al. (2006) for a survey.

<sup>11</sup>Olley and Pakes (1996) approach present a two-step estimation method: in the first step, semiparametric methods are used to estimate the coefficients on the variable inputs along with the nonparametric function linking productivity to capital and investment; in the second step, parameters on capital inputs are identified under the assumptions of the dynamics of the

used methods for the structural identification of production function, they suffer from collinearity and even identification problems as pointed out by Akerberg, Caves and Frazen (2006) (ACF henceforth). Referring to the timing and dynamic implications of input choices, they cast doubts especially on the LP estimation techniques. Thus, ACF propose an estimation method built upon OP and LP approaches but not suffering from potential collinearity problems: the coefficient on labor is no longer estimated at the first stage (in a value added production function).

## 4.2 Methodology used

Referring to the literature on the identification of the production functions, we implement the structural techniques suggested by Akerberg et al. (2006). Specifically, in our analysis productivity is obtained from a Cobb-Douglas production function containing the real value added,  $Y$ , labor,  $L$ , capital,  $K$ , and a set of additional variable inputs, namely the workforce diversity indexes,  $index_{hit}$ , and a vector,  $X$ , of workforce composition characteristics (shares of foreigners, managers, middle managers, males, workers with either tertiary or secondary education, and differently aged workers belonging to the employees' age distribution quartiles, average firm tenure).<sup>12</sup> The log-linear production function is specified as follows:

$$\ln Y_{it} = \text{cons} + \alpha \ln L_{it} + \beta \ln K_{it} + \gamma(\text{index}_{hit}) + \delta(X_{it}) + u_{it}$$

The error term  $u_{it}$  consists of a time-varying firm specific effect  $v_{it}$  (unobserved by econometricians) and an idiosyncratic component  $\varepsilon_{it}$ . Following Akerberg et al. (2006), we assume that

$$E(\varepsilon_{it} \mid l_{it}, k_{it}, \text{index}_{hit}, X_{it}, m_{it}, l_{it-1}, k_{it-1}, \text{index}_{hit-1}, X_{it-1}, m_{it-1}, \dots, l_{i1}, k_{i1}, \text{index}_{hi1}, X_{i1}, m_{i1}) = 0,$$

with  $t = 1, 2, \dots, T$ , and where  $m$  refers to our proxy variable (materials) and lower-case letters to log-variables.

As past values of  $\varepsilon_{it}$  are not included in the conditioning set, it means that we allow for serial dependence in the pure shock term. However, we need to restrict the dynamics in the productivity process:

---

productivity process (where productivity is assumed to follow a first order Markov process), see Wooldridge, 2009. However, OP's estimation method presents two major drawbacks. First, because adjustment costs create lumpiness in investment levels, these levels may not respond smoothly to productivity shocks. Second, OP approach excludes firms reporting 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. Many benefits are associated with this choice. First, changes in intermediate inputs do not typically involve adjustment costs; intermediate inputs therefore respond better to productivity shocks than investments do. Second, intermediate inputs provide a simple link between estimation strategy and economic theory because they do not typically represent state variables. Third, because intermediate inputs are almost always used in production, LP's approach circumvents the above-mentioned data truncation problem. Moreover, LP suggest three specification tests for evaluating the proxy's performance (Petrin et al., 2004). However, the coefficient on the proxy is recovered at the second stage rather than in the first, as in the OP approach.

<sup>12</sup>We also specify other control variables identified by partial/total foreign ownership, multi-establishment dummy, year, 3-digit industry classification and regional dummies because they can potentially affect productivity.

$$E(v_{it} | v_{it-1}, v_{it-2}, \dots, v_{i1}) = E(v_{it} | v_{it-1}) = f(v_{it-1})$$

with  $t = 1, 2, \dots, T$ , and for given functions  $f(\cdot)$ . As in ACF's approach, we assume material input to be chosen after labor input. In addition, we assume that our diversity indexes and the other additional variable inputs,  $X$ , are set before or at the same time as material input is chosen. As a result, material demand will not only be a function of capital and productivity, but also of  $l$ ,  $index_h$  and  $X$ :

$$m_{it} = f(k_{it}, v_{it}, l_{it}, index_{hit}, X_{it})$$

and assuming that the material demand function is strictly increasing in productivity shock  $v_{it}$ , we get

$$v_{it} = f^{-1}(k_{it}, m_{it}, l_{it}, index_{hit}, X_{it}).$$

The key advantage of this approach is that it allows, for example, our key variables,  $index_{hit}$ , to have dynamic implications or to depend on unobserved input price shock that could potentially be serially correlated over time. In fact, it seems reasonable to assume that the hiring and firing costs for labor or the fixed costs of diversifying the workforce can last longer than a period (Parrotta et al., 2011). Plugging the inverse material demand into the production function, we get the first-stage equation, which here serves only to separate  $v_{it}$  from  $\varepsilon_{it}$ ,

$$y_{it} = cons + \alpha l_{it} + \beta k_{it} + \gamma index_{hit} + \delta X_{it} + f^{-1}(k_{it}, m_{it}, l_{it}, index_{hit}, X_{it}) + \varepsilon_{it}.$$

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

$$a_{it} = v_{it} - g(v_{it-1}),$$

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

$$E \left[ a_{it} \begin{matrix} k_{it} \\ l_{it-1} \\ index_{hit-1} \\ X_{it-1} \end{matrix} \right] = 0$$

to identify coefficients on  $k$ ,  $l$ ,  $index_{hit}$ , and  $X$ .<sup>13</sup>

## 5 Results

### 5.1 Effect of diversity on firm productivity

Our main results are shown in Table 3. As explained in section 3, we perform the analysis by using two different aggregation levels of the categories included in our diversity indexes. Results from a more aggregate level are shown in columns (1) to (6), whereas results using the disaggregated categories are presented in columns (7) to (12). Columns 1, 2, 7 and 8 do not include the additional variable inputs,  $X$ , beside our diversity indexes,  $index_h$ ; they are instead used in the remaining columns to check whether our parameters of interest change their sign, size or significance level.<sup>14</sup>

[Insert Table 3 around here]

Columns 1 and 7 in Table 3 show results from standard OLS while the other columns include the estimates from the ACF algorithm, which allows to properly address both endogeneity and simultaneity issues in the identification of all input coefficients. The first two rows report labor and capital elasticities, which are fairly stable over specifications. Specifically, labor elasticity is about 0.74, whereas the one referred to the capital stock fluctuates around 0.27. Consistently with previous studies (Akerberg et al. 2006; Konings and Vanormelingen 2009; Parrotta et al., 2011) a slightly lower (higher) labor (capital) contribution is estimated using OLS compared to the ACF algorithm.

Turning to our variables of interests, the coefficients to ethnic and demographic diversity are generally negative, whereas educational diversity positively affects productivity. Including additional input variables

<sup>13</sup>In a previous version of this study (Parrotta et al., 2010), we employ an instrumental variable (IV) approach to cope with potential simultaneity and endogeneity related to firm-level diversity indexes and we consider TFP as a measure of firm productivity. Specifically, we consider an index of labour diversity measured at the commuting area level, in which a given firm operates, as an instrument for firm level diversity index in the TFP equation. Results from this identification strategy are similar to those reported in this paper and they are available from the authors upon a request.

<sup>14</sup>However, all specifications include a firm foreign-ownership dummy, multi-establishment dummy and a full set of 3-digit industry, year and county dummies.

in the production function implies that only ethnic diversity remains statistically significant, independently of the level of aggregation. More specifically, a standard deviation increase in ethnic diversity reduces firm productivity by 1% (2%), when an aggregated (disaggregated) index is considered. Regarding the educational diversity, it significantly affects productivity only in the disaggregated specification. The results from this specification show that a standard deviation increase in educational diversity enhances firm productivity by 1%. Finally, the effects from demographic diversity turn out to be insignificant in the full model specification. Next, we run models with separate diversity dimensions one by one to check whether one dimension of diversity captures the effects associated with other indexes. For instance, ethnic diversity may pick up some of the skill diversity effects as individuals coming from different countries may represent different degrees of educational heterogeneity as well. The columns 4, 5 and 6 (9, 10 and 11) in Table 3 show results of aggregate (disaggregate) indexes of ethnic, educational and demographic diversity, respectively, entering the models separately. In those models, the coefficients to ethnic and educational diversity have lower magnitudes compared to the models with all three diversity dimensions included. This gives a support to the hypothesis that the ethnic diversity may also capture heterogeneity in a specific educational level.

## 5.2 Testing alternative hypotheses

In the next steps we test the different hypotheses as derived in section 2. In these analysis, we use disaggregated indexes only, as we think that the indexes based on a detailed categorization may be more adequate to represent workforce diversity.<sup>15</sup> Firstly, we calculate our diversity indexes for white- and blue-collar occupations separately. This is driven by the idea that diversity could play a different role for distinct occupational groups and consequently have diverse effects on firm productivity. In particular, we expect that the beneficial effects of diverse problem-solving abilities and creativity would materialize more in terms of higher productivity for white-collar occupations compared to blue-collar occupations. Second, we exclude/include certain groups of foreigners from/in the calculation of ethnic diversity measures to test how important are the communication costs and costs of “cross-cultural dealing”. The results of the effect of diversity indexes calculated separately for the two occupational groups on firm productivity are presented in the first two columns of Table 4. Our results show that the contribution of educational diversity on productivity is indeed much more important for white-collar than for blue-collar occupations. Moreover, the negative effect of ethnic diversity among the white-collar workers is lower than the effect associated with blue-collar occupations. Conversely, the effect of demographic diversity is insignificant for both occupational groups. Thus, our results support the creativity hypothesis proposed by the theoretical frameworks by Hong and Page (2001 and 2004) and Berliant

---

<sup>15</sup>The results using the aggregate indexes are qualitatively similar to the detailed categorization, and they are available from the authors upon request.

and Fujita (2008). To test the role of “cross-cultural dealing” we exclude from the calculation of ethnic diversity either foreigners with tertiary education or foreigners speaking a language belonging to the Germanic family of languages. Alternatively we include second generation immigrants in the calculation of ethnic diversity. All these groups of foreigners are likely to absorb Danish or English (which is the communication language in many businesses in Denmark) more quickly. Hence, it is plausible to expect that communication costs associated with ethnic diversity may increase (decrease) after subtracting out (including) foreigners who are likely to speak Danish or English. The results are shown in Table 4, columns 3, 4 and 5, for measures treating the second generation of immigrants as non-natives and foreigners with university education or with a language belonging to the Germanic group of languages as natives, respectively. Interestingly, the negative effect of ethnic heterogeneity on productivity strengthens (weakens) once we exclude (include) foreigners who probably speak Danish or English<sup>16</sup>, confirming the hypothesis that the communication costs and costs of “cross-cultural dealing” of ethnically heterogeneous workforces have an impact on firm outcomes.

[Insert Table 4 around here]

In the next step we examine different mechanisms by which diverse workforces affect firm productivity by looking at different industries and firm categories. Firstly, we look at whether there is any difference in the effect of diversity on productivity for firms in high-tech industries, which tend to have jobs characterized by higher levels of creative and problem-solving activities. Specifically, we divide industries into two groups defined on whether their aggregate level of R&D expenditure is above or below the average R&D level recorded for the overall economy.<sup>17</sup> As shown in Table 5, columns 1 and 2, the hypotheses on creativity is only partially supported, as the coefficient to education diversity is significantly positive only for industries with below mean expenditure on R&D, which is contrary of what we would expect. On the other hand, the coefficient to ethnic diversity is insignificant and of smaller magnitude for firms in industries with above mean expenditure on R&D, whereas it is significantly negative and of much larger magnitude for industries below mean expenditure on R&D. We also investigate whether the coefficients on diversity indexes differ for firms in more trade-open industries in line with the Osborne (2000) and Rauch and Casella (2003) hypothesis. Therefore, we divide industries according to their trade openness into above and below mean trade flows.<sup>18</sup> The results shown in Table 5, columns 3 and 4, reveal that the coefficient to ethnic diversity is significantly negative for both types of industries, however the coefficient is smaller for more trade open industries. Thus we cannot clearly reject nor accept the hypothesis that workforce diversity provides beneficial information to

<sup>16</sup>According to the existing literature individuals have easier time to acquire a foreign language if their mother language is linguistically closer to the foreign language to be learned than individuals speaking a language which is linguistically more distant (Isphording and Otten, 2011; Chiswick and Miller, 1995).

<sup>17</sup>Source: The Analytical Business Enterprise Research and Development Database ANBERD (OECD).

<sup>18</sup>Trade openness is measured as the sum of total exports and imports over value added. Data were retrieved from the Structural Analysis database (OECD).

firms about other countries and markets, and in this way brings positive effects on firm productivity. Finally, we check whether the effect of diversity varies between industries with raising or declining employment. It is in fact reasonable to expect that growing firms are more likely to benefit from diversity as they often hire younger and foreign talents compared to shrinking firms. Columns 5 and 6 in Table 5 support this hypothesis as the positive (negative) effects of educational (ethnic) diversity strengthen (weaken) in the subsample of industries with increasing employment.

[Insert Table 5 around here]

To sum up, we find evidence of positive effects of heterogeneity in education, which is consistent with the theory on knowledge spillovers, creativity and problem solving abilities (Lazear, 1999; Hong and Page, 1998 and 2001; Berliant and Fujita, 2004; Alesina and La Ferrara, 2005). However, in the case of ethnic and demographic diversity, the coefficients are either negative or insignificant. This is a mixture of two distinct forces pulling the effect of demographic and ethnic diversity in different directions: more culturally or demographically diverse workforce have better problem solving abilities, creativity and knowledge spillovers in line with propositions by Hong and Page (1998 and 2001), Berliant and Fujita (2004) and Alesina and La Ferrara (2005), but this positive effects are counteracted or even offset by the negative effects of diversity on firm productivity coming from communication and integration costs in line with proposition by Lazear (1998 and 1999).

## 6 Sensitivity analysis

In the next step, as a part of the sensitivity analysis we evaluate eventual variations in the effects of labor diversity when diversity is differently computed. In particular, we use two alternative diversity indexes: the Shannon-Weaver entropy and the richness indexes. The entropy index is considered as one of the most profound and useful diversity indexes in biology (Maignan et al., 2003), whereas the richness index is defined as a number of categories observed for each dimension of interest (it does not account for the “evenness” dimension). The results are shown in Table 6, columns 1 and 2, respectively, and both measures support the results from our main analysis.

Next, we include a Herfindhal index for the type of secondary and tertiary education (this index has now 8 categories: high school, business high school, short and long vocational education, engineering, natural sciences, social sciences and humanities) together with the standard deviation for the years of education and age. This allows us, on the one hand, to treat age as a cardinal variable and, on the other, to disentangle the effects associated with the amount of education from those related to the type of education. The results



from column 3 show that the effects of both educational and demographic diversity are never significant when measured as standard deviations. Conversely, the effects of the Herfindhal index for the type of education are positive and statistically significant.

We then divide firms by size and check whether any change in coefficients to workforce diversity occurs for small (less than 50 employees), middle (50-100 employees) and big firms (more than 100 employees). It might be that the effects of demographic and ethnic diversity could be more beneficial in larger firms as their organizational and management structures and practices are well established, and thus they are more likely to introduce policies, which can help to counteract the potential costs associated with the diversity. On the other hand, large-sized firms are likely to have many different types of jobs and occupations, in which diversity might affect firm productivity in different ways. As reported in Table 6, columns (4)-(6), the coefficients to the ethnic diversity index are significant negative for medium- and large-sized enterprises, with largest coefficient in case of large firms. Regarding educational diversity, the results show that it is more important for medium sized and compared to large firms. Interestingly, the diversity in all three dimensions is insignificant for firms with less than 50 employees, see column 4, Table 6.

Given that big cities have usually a lot of immigrants and high-skill workers and at the same time a high percentage of productive firms, we pursue an additional sensitivity check by taking firms from big city area out from our analyses. As the only real agglomeration area in Denmark is Copenhagen, we drop Copenhagen and environs from the analysis. Results from this robustness check are reported in column 7, Table 6, and do not qualitatively differ from the main results.

Lastly, as labor diversity has been computed at the firm level (weighting average of Herfindahl indexes computed at the workplace level), we evaluate how results change if multi-establishment firms are excluded from the sample. The last column of Table 6 reports information on such a check: the interpretation of these findings does not significantly differ from that related to the main results.

[Insert Table 6 around here]

Finally, we check whether there are any differences in the effect of diversity on productivity across different industries. The results are shown in Table 7, columns (1) to (5). We observe that for most industries the effects of workforce diversity are insignificantly different from zero. But few industries stand out above all - the effect of educational diversity is significantly positive for firms in manufacturing and in financial and business services. Ethnic diversity negatively affects firm productivity in transport, manufacturing and financial and business services industries.<sup>19</sup>

---

<sup>19</sup>Prior academic research suggests that diversity leads to economic gains or losses depending on the industrial characteristics (Sparber, 2009, 2010). More specifically, diversity seems to increase productivity in sectors that require creative decision-making,

[Insert Table 7 around here]

## 7 Conclusions

Using a comprehensive linked employer-employee data-set, this paper investigates the effect of labor diversity in ethnic-cultural, educational and demographic characteristics on firm productivity in Denmark. Contrary to the majority of previous empirical works, which focused on single aspects of labor diversity, we provide a number of findings that may concretely address as a whole the consequences of firm workforce heterogeneity on firm performance. For our analysis we use the well-known Herfindahl index to measure extensively the three above mentioned dimensions of workforce diversity. Regarding the methodology, we follow the approach suggested by Akerberg et al. (2006) to deal with simultaneity and endogeneity problems concerning the diversity indexes.

Controlling for a wide set of additional variable inputs and firm specific characteristics and performing different robustness checks, we find that diversity in education significantly enhances firm productivity. Specifically, we find that a standard deviation increase in educational diversity increases productivity by approximately 1%. The result gives support to the existing theory on knowledge spillovers. On the contrary, diversity in demographics and ethnicity brings either no or negative effects on firm productivity. Thus, the negative effects coming from communication and integration costs connected to a more demographically and culturally diverse workforce seems to counteract the positive effects of diversity coming from better problem solving abilities, creativity and knowledge spillovers. These findings are consistent in part with past relevant studies by Lazear (1999), Glaeser et. al. (2000), and Alesina and La Ferrara (2002).

Thus, if our empirical analysis clearly provides evidence of the positive contribution of educational diversity to firm productivity, it also does not support a general statement saying that diversity in ethnic and demographics is beneficial for businesses in terms of firm performance. Our findings may imply that firms strengthening their efforts to decrease the “obvious” costs associated with the workforce diversity, e.g. by implementing diversity management, modern techniques and integration practices, could turn the workforce heterogeneity into a substantial competitive advantage.

---

problem solving, and customer service, but it may decrease it in industries characterized by high levels of group or team work and efficiency. The current division by industries is however too rough to test the hypothesis as there are likely both types of jobs (jobs requiring creativity and efficiency) prevailing in those aggregate industries. Unfortunately our model did not converge for more detailed industries.

## References

- [1] Akerberg, Daniel A., Kevin Caves, and Garth Frazer. 2006. Structural Identification of Production Functions, Department of Economics, University of California.
- [2] Adsera, Alicia, and Mariola Pytlikova. 2011. The Role of Language in Shaping International Migration: Evidence from OECD Countries 1985-2006. Unpublished manuscript, Department of Economics, University of Aarhus.
- [3] Alesina, Alberto, and Eliana La Ferrara. 2002. Who Trusts Others?. *Journal of Public Economics*, 85: 207-234.
- [4] Alesina, Alberto, and Eliana La Ferrara. 2005. Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43: 762-800.
- [5] Becker, Gary. 1957. *The Economics of Discrimination*. Chicago, University of Chicago Press.
- [6] Barrington, Linda, and Kenneth R. Troske. 2001. Workforce Diversity and Productivity: An Analysis of Employer-Employee Matched Data. Economics Program Working Papers No. 01-02.
- [7] Berliant, Marcus, and Masahisa Fujita. 2008. Knowledge Creation as a Square Dance on the Hilbert Cube. *International Economic Review*, 49: 1251–1295.
- [15] Chiswick, Barry R., and Paul W. Miller. 1995. “The Endogeneity between Language and Earnings: International Analyses”. *Journal of Labor Economics*, 2: 246-88.
- [8] Eriksson, Tor, Mariola Pytlikova, and Frederic Warzynski. 2009. Increased Sorting and Wage Inequality in the Czech Republic: New Evidence Using Linked Employer-Employee Dataset. ASB Working Paper No. 09-14.
- [16] Lewis, M. Paul. 2009. *Ethnologue: Languages of the World, 16th edition*. ISBN-13 978-1-55671-216-6. <http://www.ethnologue.com/web.asp>
- [9] Garicano, Luis, and Esteban Rossi-Hansberg. 2006. Organization and Inequality in a Knowledge Economy. *The Quarterly Journal of Economics*, 121: 1383-1435.
- [17] Glaeser, Edward, David I. Laibson, José A. Scheinkman, and Christine L. Soutter. 2000. Measuring Trust. *The Quarterly Journal of Economics*, 115: 811-846.
- [18] Grund, Christian, and Niels Westergaard-Nielsen. 2009. Age Structure of the Workforce and Firm Performance. *International Journal of Manpower*, 29: 410-422.

- [19] Guiso, Luigi, Paolo Sapienza, and Luigi Zingales. 2009. Cultural Biases in Economic Exchange?. *Quarterly Journal of Economics*, 124: 1095-1131.
- [10] Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2003. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, 111: 465-497.
- [20] Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2004. Diversity and Productivity in Production Teams. Unpublished manuscript, Department of Economics, Washington University in St. Louis.
- [21] Hong, Lu, and Scott E. Page. 2001. Problem Solving by Heterogeneous Agents. *Journal of Economic Theory*, 97: 123-163.
- [11] Hong, Lu, and Scott E. Page. 2004. Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers. *Proceedings of the National Academy of Sciences*, 101: 123-139.
- [22] Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. 2008. Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data. *Journal of Labor Economics*, 26: 247-285.
- [12] Ispording, Ingo E., and Sebastian Otten. 2011. Linguistic Distance and the Language Fluency of Immigrants, Ruhr Economic Papers No. 274.
- [13] Konings, Jozef, and Vanormelingen, Stijn, 2010. The Impact of Training on Productivity and Wages: Firm Level Evidence. IZA Discussion Papers 4731, Institute for the Study of Labor (IZA).
- [23] Kremer, Michael. 1993. The O-ring Theory of Economic Development. *The Quarterly Journal of Economics*, 108: 551-575.
- [24] Kremer, Michael, and Eric Maskin. 1996. Wage Inequality and Segregation by Skill. NBER Working Paper No. 5718.
- [25] Kurtulus, Ana F. 2011. What Types of Diversity Benefit Workers? Empirical Evidence on the Effects of Co-Worker Dissimilarity on the Performance of Employees. *Industrial Relations: A Journal of Economy and Society*, 50: 678-712.
- [26] Lang, Kevin. 1986. A Language Theory of Discrimination. *Quarterly Journal of Economics*, 101: 363-382.
- [14] Lazear, Edward P. 1998. *Personnel Economics for Managers*. New York, John Wiley & Sons.

- [28] Lazear, Edward P. 1999. Globalisation and the Market for Team-Mates. *The Economic Journal*, 109: 15-40.
- [15] Leonard, Jonathan S., and David I. Levine. 2006. Diversity, Discrimination, and Performance. Institute for Research and Employment Working Paper, No 147.
- [29] Levinsohn, James, and Amil Petrin. 2003. Estimating Production Functions using Inputs to Control for Unobservables. *Review of Economic Studies*, 70: 317-341.
- [16] Maignan, Carole, Gianmarco Ottaviano, Dino Pinelli and Francesco Rullani. 2003. Bio-Ecological Diversity vs. Socio-Economic Diversity: A Comparison of Existing Measures. Nota di Lavoro, Fondazione Eni Enrico Mattei.
- [32] Mundlak, Yair. 1961. Empirical Production Function Free of Management Bias. *Journal of Farm Economics*, 43: 44-56.
- [34] Navon, Guy. 2009. Human Capital Spillovers in the Workplace: labor Diversity and Productivity. MPRA Paper No. 17741.
- [35] OECD. 2005. *Economic Survey of Denmark*. Paris, 2005.
- [36] Olley, Steven G., and Ariel Pakes. 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64: 1263-1297.
- [37] Osborne, Evan. 2000. The Deceptively Simple Economics of Workplace Diversity. *Journal of Labor Research*, 21: 463-475.
- [38] Ottaviano, Gianmarco I. P. and Giovanni Peri. 2006. The Economic Value of Cultural Diversity: Evidence from US Cities. *Journal of Economic Geography*, 6: 9-44.
- [39] Ottaviano, Gianmarco I. P., and Giovanni Peri. 2011. Rethinking the Effects of Immigration on Wages. *Journal of European Economic Association*. doi: 10.1111/j.1542-4774.2011.01052.x
- [17] Parrotta, Pierpaolo and Dario Pozzoli. Forthcoming. The Effects of Learning-by-Hiring on Productivity. *RAND Journal of Economics*.
- [40] Parotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2010. Does Labour Diversity affect Firm Productivity? ASB Working Paper No. 10-12.
- [18] Pedersen, Peder J., Mariola Pytlikova, and Nina Smith. 2008. Selection and Network Effects - Migration Flows into OECD Countries 1990-2000. *European Economic Review*, 52: 1160-1186.

- [19] Peri Giovanni. 2011. The Effect of Immigration on Productivity: Evidence from U.S. States. *The Review of Economics and Statistics* (forthcoming).
- [20] Petrin, Amil, James Levinsohn, and Brian Poi. 2003. Production Function Estimation in Stata Using Inputs to Control for Unobservables. *Stata Journal*, 4: 113-123.
- [42] Rauch, James E., and Alessandra Casella. 2003. Overcoming Informational Barriers to International Resource Allocation: Prices and Ties. *Economic Journal*, 113: 21-42.
- [21] Rosen, Sherwin. 1981. The economics of superstars. *American Economic Review*, 71: 845–858.
- [44] Sparber, Chad. 2009. Racial Diversity and Aggregate Productivity in US Industries: 1980-2000. *Southern Economic Journal*, 75: 829-856.
- [45] Sparber, Chad. 2010. Racial Diversity and Macroeconomic Productivity across US States and Cities. *Regional Studies*, 44: 71-85.
- [46] Suedekum, Jens, Katja Wolf and Uwe Blien. 2009. Cultural Diversity and Local Labor Markets. IZA Discussion Paper No. 4619.
- [22] Wooldridge, Jeffrey Marc. 2009. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104: 112–114.
- [23] 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, 2000.

## Appendix: Measurement of Ethnic Diversity

- 1) The citizens in the different nationality groups are: **Danish**, Danish native including second generation immigrants; **North America and Oceania**, United States, Canada, Australia, New Zealand; **Central and South America**, 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, Bulgaria, Croatia, Rep. 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, Kazakhstan, Kirgizstan, Kuwait, Lebanon, Libyan Arab Jamahiriya, Malaysia, Maldives, Mali, Mauritania, Morocco, Nigeria, Oman, Pakistan, Palestine, Qatar, Saudi Arabia, Senegal, Sierra Leone, Somalia, Sudan, Syria, Tadzhikistan, Tunisia, Turkey, Turkmenistan, Uzbekistan, Yemen; **East Asia**, China, Hong Kong, Japan, Korea, Korea Dem. People's Rep. Of, Macao, Mongolia, Taiwan; **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; **West and South Europe**, all the other European countries not included in the Formerly Communist Countries category.
- 2) Using linguistic grouping: **Germanic West** (Antigua Barbuda, Aruba, Australia, Austria, Bahamas, Barbados, Belgium, Belize, Bermuda, Botswana, Brunei, Cameroon, Canada, Cook Islands, Dominica, Eritrea, Gambia, Germany, Ghana, Grenada, Guyana, Haiti, Ireland, Jamaica, Liberia, Liechtenstein, Luxemburg, Mauritius, Namibia, 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, Columbia, Costa Rica, Cote D'Ivoire, Cuba, Djibouti, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, France, French Guina, 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, Mauritania, 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, 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.).

Table 1: Descriptive statistics, main sample and by size

Variables	Definition											
	Total			Small size			Middle and big size					
IDA Variables:	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd
Share of males	0.709	0.783	0.237	0.724	0.810	0.238	0.694	0.741	0.21	0.694	0.741	0.21
Share of foreigners	0.048	0.022	0.091	0.045	0	0.098	0.054	0.031	0.083	0.054	0.031	0.083
Age15-32	0.232	0.105	0.25	0.315	0.294	0.178	0.299	0.275	0.144	0.299	0.275	0.144
Age33-41	0.263	0.262	0.129	0.262	0.25	0.123	0.281	0.278	0.077	0.281	0.278	0.077
Age42-50	0.253	0.25	0.126	0.205	0.200	0.112	0.214	0.217	0.070	0.214	0.217	0.070
Age51-65	0.252	0.178	0.15	0.218	0.189	0.172	0.206	0.188	0.162	0.206	0.188	0.162
Basic education	0.272	0.164	0.128	0.272	0.178	0.324	0.298	0.169	0.333	0.324	0.298	0.169
Secondary education	0.685	0.75	0.326	0.690	0.727	0.195	0.650	0.650	0.151	0.650	0.650	0.151
Tertiary education	0.043	0	0.1	0.038	0	0.101	0.052	0.018	0.097	0.052	0.018	0.097
Tenure	4.531	4.403	1.831	4.382	4.205	1.913	4.545	4.491	1.623	4.545	4.491	1.623
Share of managers	0.022	0	0.043	0.049	0.025	0.067	0.040	0.029	0.042	0.040	0.029	0.042
Share of middle managers	0.226	0.263	0.147	0.178	0.254	0.112	0.233	0.271	0.132	0.233	0.271	0.132
Share of blue collars	0.762	0.841	0.242	0.773	0.857	0.244	0.727	0.797	0.261	0.727	0.797	0.261
Index Ethnic Aggr	0.072	0	0.116	0.067	0	0.176	0.285	0.301	0.267	0.285	0.301	0.267
Index Edu Aggr	0.544	0.564	0.165	0.363	0.396	0.148	0.425	0.460	0.112	0.425	0.460	0.112
Index Demo Aggr	0.734	0.757	0.098	0.745	0.759	0.079	0.791	0.799	0.055	0.791	0.799	0.055
Index Ethnic Disaggr	0.198	0	0.313	0.127	0	0.271	0.425	0.500	0.327	0.425	0.500	0.327
Index Edu Disaggr	0.543	0.561	0.161	0.522	0.538	0.166	0.610	0.616	0.121	0.610	0.616	0.121
Index Demo Disaggr	0.882	0.896	0.075	0.891	0.885	0.077	0.920	0.928	0.053	0.920	0.928	0.053
<b>Accounting Variables:</b>												
Value added	34801.84	10792.59	197860.4	9980.425	7695.536	22326.87	100039.1	39730.27	365373.2	100039.1	39730.27	365373.2
Materials	90729.84	18894.45	662175	27862.83	11496.38	137692	259226.9	75699.06	1220144	259226.9	75699.06	1220144
Capital	107911.9	16889.7	1301386	29074.99	10885.32	580798	324841.5	73445.74	2408409	324841.5	73445.74	2408409
Foreign ownership	0.004	0	0.0602	0	0.004	0.067	0.005	0	0.069	0.005	0	0.069
Multi-establishment	0.262	0	0.4397	0	0.033	0.179	0.412	0	0.492	0.412	0	0.492
N	104056											

*Notes:* 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 (4.05 %); textiles (2 %), wood products (6.19 %), chemicals (3.95 %), other non-metallic mineral products (1.94 %), basic metals (18.95 %), furniture (3.46 %), construction (15.07 %), sale and repair of motor vehicles (3.64 %), wholesale trade (14.67 %), retail trade (6.06 %), hotels and restaurants (2.08 %), transport (6.12 %), post and telecommunications (0.40 %), financial intermediation (1.17 %) and business activities (10.25 %). Small size: Employees  $\leq$  49; Middle and big size: Employees  $\geq$  50.



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

<i>Aggregate specification</i>							
	<b>Manufacturing</b>	<b>Construction</b>	<b>Wholesale and retail trade</b>	<b>Transport</b>	<b>Financial and business services</b>	<b>Others</b>	
Index Ethnic	0.175	0.193	0.035	0.067	0.083	0.156	
Index Edu	0.406	0.413	0.293	0.341	0.441	0.455	
Index Demo	0.774	0.735	0.719	0.760	0.734	0.766	
N	39039	4291	18470	25906	6274	10711	
	<b>Small size</b>	<b>Middle size</b>	<b>Big size</b>	<b>1995</b>	<b>1999</b>	<b>2005</b>	
Index Ethnic	0.037	0.093	0.282	0.093	0.108	0.128	
Index Edu	0.348	0.377	0.424	0.382	0.379	0.381	
Index Demo	0.729	0.760	0.791	0.743	0.758	0.735	
N	39207	40660	24824	6014	10924	12083	
<i>Disaggregate specification</i>							
	<b>Manufacturing</b>	<b>Construction</b>	<b>Wholesale and retail trade</b>	<b>Transport</b>	<b>Financial and business services</b>	<b>Others</b>	
Index Ethnic	0.258	0.319	0.085	0.142	0.168	0.278	
Index Edu	0.564	0.611	0.417	0.528	0.548	0.686	
Index Demo	0.901	0.854	0.849	0.885	0.862	0.888	
N	39039	4291	18470	25906	6274	10711	
	<b>Small size</b>	<b>Middle size</b>	<b>Big size</b>	<b>1995</b>	<b>1999</b>	<b>2005</b>	
Index Ethnic	0.081	0.172	0.425	0.158	0.188	0.219	
Index Edu	0.502	0.542	0.610	0.514	0.543	0.560	
Index Demo	0.854	0.888	0.920	0.872	0.884	0.878	
N	39207	40660	24824	6014	10924	12083	

*Notes:* Small size: Employees  $\leq 49$ ; Middle size:  $50 \leq$  Employees  $\leq 99$ ; Big size: Employees  $\geq 100$ .

Table 3: The effects of labor diversity on firm productivity, main results.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	ACF	ACF	ACF	ACF	ACF	OLS	ACF	ACF	ACF	ACF	ACF
Log(L)	0.694*** (0.010)	0.732*** (0.032)	0.752*** (0.026)	0.749*** (0.022)	0.755*** (0.027)	0.738*** (0.021)	0.694*** (0.010)	0.743*** (0.038)	0.745*** (0.023)	0.749*** (0.023)	0.741*** (0.023)	0.740*** (0.021)
Log(K)	0.298*** (0.009)	0.270*** (0.021)	0.261*** (0.021)	0.259*** (0.020)	0.255*** (0.023)	0.269*** (0.018)	0.293*** (0.009)	0.259*** (0.024)	0.266*** (0.018)	0.265*** (0.021)	0.264*** (0.021)	0.268*** (0.018)
Index Ethnic Aggr	-0.060*** (0.013)	-0.051*** (0.010)	-0.073*** (0.012)	-0.010*** (0.003)								
Index Edu Aggr	0.187*** (0.025)	0.166*** (0.021)	0.050 (0.069)		0.040 (0.085)							
Index Demo Aggr	-0.304* (0.163)	-0.189*** (0.038)	-0.247 (0.282)		-0.259 (0.265)							
Index Ethnic Disaggr	-	-	-				-0.054*** (0.008)	-0.077*** (0.014)	-0.075*** (0.012)	-0.044*** (0.008)		
Index Edu Disaggr	-	-	-				0.168*** (0.021)	0.164*** (0.033)	0.118*** (0.056)		0.074*** (0.024)	
Index Demo Disaggr	-	-	-				-0.215*** (0.038)	-0.359* (0.214)	-0.283 (0.330)			-0.351 (0.311)
Share of middle managers			-0.161** (0.053)	-0.165** (0.054)	-0.160** (0.055)	-0.160** (0.053)			-0.140** (0.062)	-0.168** (0.056)	-0.152** (0.065)	-0.157** (0.053)
Share of managers			0.058 (0.069)	0.048 (0.069)	0.069 (0.073)	0.049 (0.068)			0.043 (0.075)	0.050 (0.072)	0.050 (0.078)	0.046 (0.072)
Tenure			0.011*** (0.003)	0.009*** (0.002)	0.010** (0.003)	0.009*** (0.003)			0.012*** (0.003)	0.155* (0.079)	0.171*** (0.030)	0.133*** (0.032)
Secondary education			0.270** (0.083)	0.305*** (0.050)	0.270** (0.092)	0.305*** (0.053)			0.299*** (0.061)	0.296*** (0.057)	0.316*** (0.063)	0.303*** (0.054)
Tertiary education			0.417*** (0.078)	0.394*** (0.083)	0.449*** (0.074)	0.412*** (0.088)			0.398*** (0.081)	0.424*** (0.090)	0.407*** (0.077)	0.412*** (0.087)
Age15-32			-0.015 (0.065)	-0.022 (0.058)	-0.035 (0.063)	-0.020 (0.057)			-0.012 (0.063)	-0.014 (0.061)	-0.039 (0.060)	-0.016 (0.057)
Age33-41			0.201** (0.061)	0.206*** (0.059)	0.194** (0.065)	0.199*** (0.054)			0.205*** (0.050)	0.203** (0.062)	0.189** (0.060)	0.217*** (0.048)
Age42-50			0.146** (0.070)	0.131* (0.074)	0.131 (0.083)	0.139** (0.062)			0.157** (0.061)	0.130* (0.080)	0.123 (0.079)	0.157** (0.058)
Share of males			0.082 (0.060)	0.114*** (0.027)	0.103** (0.035)	0.082 (0.056)			0.095 (0.066)	0.100** (0.034)	0.112** (0.042)	0.076 (0.056)
Share of foreigners			0.197*** (0.042)	0.133*** (0.030)	0.170*** (0.033)	0.151*** (0.030)			0.211*** (0.042)	0.009*** (0.003)	0.011*** (0.003)	0.009*** (0.003)
N	105539	60405	49880	49928	49928	49928	105539	60405	49880	49928	49928	49928
R2	0.858	0.878	0.882	0.880	0.880	0.880	0.858	0.878	0.880	0.879	0.880	0.879

Notes: The dependent variable is the log of value added. All regressions include whether the firm is foreign-owned, a multi-establishment dummy, a full set of 3-digit industry, year and county dummies. Standard errors are computed using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-group correlation. Significance levels: \*\*\*1%, \*\*5%, \*10%.

Table 4: Testing alternative hypothesis: estimates by occupation and alternative index definitions and "subtract-out" groups.

	(1)	(2)	(3)	(4)	(5)
	Occupation specific diversity		2nd gen. Imm. as foreigners	University graduates as natives	Germanic group as natives
	White collar	Blue collar			
Index Ethnic Disaggr	-0.051*** (0.009)	-0.077*** (0.010)	-0.050*** (0.007)	-0.078*** (0.010)	-0.080*** (0.010)
Index Edu Disaggr	0.070*** (0.010)	0.005* (0.003)	0.120** (0.053)	0.126** (0.053)	0.126** (0.053)
Index Demo Disaggr	-0.030 (0.040)	-0.069 (0.035)	-0.306 (0.329)	-0.296 (0.325)	-0.296 (0.325)
Observations	49880	49880	49880	49880	49880
R2	0.881	0.882	0.882	0.882	0.882

*Notes:* All regressions include all the firm specific characteristics, year and three-digit industry dummies. Standard errors are computed using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-group correlation. Significance levels: \*\*\*1%, \*\*5%, \*10%.

Table 5: Testing alternative hypothesis: estimates by relevant industrial aggregations.

	(1)	(2)	(3)	(4)	(5)	(6)
	Estimates by Industry R&D expenses		Estimates by Industry Trade Openness		Estimates by Industrial employment development	
	Above mean	Below mean	Above mean	Below mean	Declining employment	Increasing employment
Index Ethnic Disaggr	-0.033 (0.032)	-0.074*** (0.009)	-0.063** (0.026)	-0.079*** (0.012)	-0.079*** (0.014)	-0.054*** (0.012)
Index Edu Disaggr	0.008 (0.492)	0.123** (0.050)	0.139 (0.255)	0.103* (0.054)	0.030 (0.086)	0.164** (0.068)
Index Demo Disaggr	0.326 (3.380)	-0.295 (0.315)	0.186 (1.812)	-0.268 (0.326)	-0.139 (0.594)	-0.414 (0.480)
Observations	4043	45481	6611	42835	18907	30055
R2	0.919	0.872	0.925	0.871	0.879	0.883

*Notes:* The dependent variable is the log of value added. All regressions include all the firm specific characteristics, year and three-digit industry dummies. Standard errors are computed using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-group correlation. Significance levels: \*\*\*1%, \*\*5%, \*10%.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	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		
Index Ethnic Disaggr	-0.237*** (0.035)	-0.033*** (0.005)	-0.063*** (0.008)	-0.036 (0.119)	-0.047* (0.023)	-0.065*** (0.012)	-0.068*** (0.009)	-0.040*** (0.007)
Index Edu Disaggr	0.140** (0.044)	0.021*** (0.004)	0.137*** (0.035)	0.110 (0.463)	0.088** (0.039)	0.037** (0.016)	0.159** (0.060)	0.130** (0.056)
Std(years of education)			-0.008 (0.005)					
Index Demo Disaggr	-0.137** (0.049)	-0.027*** (0.005)		0.007 (1.350)	-0.348 (0.448)	-0.746 (1.825)	-0.293 (0.351)	-0.182 (0.327)
Std(age)			-0.013 (0.022)					
Share of males			0.109** (0.040)					
Observations	49880	49880	49880	15380	14005	9654	46647	40482
R2	0.882	0.878	0.882	0.355	0.445	0.867	0.886	0.831

*Notes:* All regressions include all the firm specific characteristics, year and three-digit industry dummies. Standard errors are computed using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-group correlation. Significance levels: \*\*\*1%, \*\*5%, \*10%.

Table 7: Robustness checks: estimates by industry

	(1)	(2)	(3)	(4)	(5)
	<b>Estimates by Industry</b>				
	Manufacturing	Construction	Wholesale and retail trade	Transport	Financial and business services
Log(L)	0.709*** (0.038)	0.868*** (0.066)	0.701*** (0.029)	0.642*** (0.114)	0.856*** (0.042)
Log(K)	0.289*** (0.032)	0.195*** (0.060)	0.298*** (0.032)	0.360*** (0.078)	0.226*** (0.036)
Index Ethnic Disaggr	-0.040*** (0.009)	-0.055 (0.043)	-0.159*** (0.031)	-0.121 (0.110)	-0.225*** (0.050)
Index Edu Disaggr	0.087*** (0.006)	0.046 (0.175)	0.252 (0.162)	0.019 (1.085)	0.198*** (0.042)
Index Demo Disaggr	0.059 (0.546)	-0.435 (1.304)	-0.290 (0.752)	-0.241 (3.119)	-0.833 (1.564)
Observations	22550	7699	11861	1762	4010
R2	0.912	0.905	0.848	0.883	0.848

*Notes:* The dependent variable is the log of value added. All regressions include all the firm specific characteristics, year and three-digit industry dummies. Standard errors are computed using a block bootstrap procedure with 300 replications and are robust against heteroskedasticity and intra-group correlation. Significance levels: \*\*\*1%, \*\*5%, \*10%.