

The Heterogeneous Effects of Workforce Diversity on Productivity, Wages and Profits

Andrea Garnero

Stephan Kampelmann

François Rycx

Septembre 2013

The Heterogeneous Effects of Workforce Diversity on Productivity, Wages and Profits

Abstract: We estimate the impact of workforce diversity on productivity, wages and productivity-wage gaps (i.e. profits) using detailed Belgian linked employer-employee panel data. Findings show that educational (age) diversity is beneficial (harmful) for firm productivity and wages. While gender diversity is found to generate significant gains in high-tech/knowledge intensive sectors, the opposite result is obtained in more traditional industries. Estimates neither vary substantially with firm size nor point to sizeable productivity-wage gaps except for age diversity.

Keywords: labour diversity, productivity, wages, linked panel data, GMM

JEL Classification: D24, J24, J31, M12

Les effets hétérogènes de la diversité de la main d'œuvre sur la productivité, les salaires et les profit

Résumé: Nous estimons l'impact de la diversité de la main-d'œuvre sur la productivité, les salaires et l'écart entre la productivité et les salaires (càd. les profits) à l'aide des données de panel appariées "employés-employeurs" pour la Belgique. Les résultats montrent que la diversité en termes de niveaux d'éducation (d'âges) est bénéfique (néfaste) pour la productivité des firmes et la rémunération des travailleurs. Quant à la diversité des genres, nous trouvons qu'elle génère des gains significatifs dans les secteurs à forte intensité technologique/de connaissance, tandis que le résultat inverse est obtenu pour les secteurs plus traditionnels. Les estimations ne varient pas de façon substantielle avec la taille des entreprises et ne montrent pas de décalage important entre la productivité et les salaires sauf pour la diversité des âges.

The Heterogeneous Effects of Workforce Diversity on Productivity, Wages and Profits*

Andrea Garnero[§]

ENS, Paris School of Economics, and SBS-EM (CEB, DULBEA)

Stephan Kampelmann

Université Libre de Bruxelles, SBS-EM (CEB, DULBEA)

François Rycx

Université Libre de Bruxelles, SBS-EM (CEB, DULBEA), and IZA

Abstract

We estimate the impact of workforce diversity on productivity, wages and productivity-wage gaps (i.e. profits) using detailed Belgian linked employer-employee panel data. Findings show that educational (age) diversity is beneficial (harmful) for firm productivity and wages. While gender diversity is found to generate significant gains in high-tech/knowledge intensive sectors, the opposite result is obtained in more traditional industries. Estimates neither vary substantially with firm size nor point to sizeable productivity-wage gaps except for age diversity.

Keywords: labour diversity, productivity, wages, linked panel data, GMM

JEL Classification: D24, J24, J31, M12

* We would like to thank Statistics Belgium for giving access to the data. We are grateful to Mahmood Arai, Philippe Askenazy, Andrew Clark, Jeremy Dawson, Patricia Garcia-Prieto, Pekka Ilmakunnas, Luca Marcolin, Sile O’Dorchai, Dario Pozzoli, Ilan Tojerow as well as audiences in Brussels (ULB), Paris (PSE and Paris I), Caserta (AIEL), Nuremberg (IAB), Leuven (Day for Labour Economists), Buch am Ammersee (IZA Summer School) and Turin (EALE) for helpful comments and discussions. All remaining errors are the authors’ responsibility. Funding for this research was provided by the Belgian Federal Government – SPP Politique scientifique, programme “Société et Avenir”, *Employment, wage discrimination and poverty*, research contract TA/00/046/EDIPO. Andrea Garnero gratefully acknowledges financial support from CEPREMAP.

[§] Corresponding author: ULB - Avenue F.D. Roosevelt 50, CP-140 - 1050 Brussels, Belgium, e-mail: agarnero@pse.ens.fr.

1. Introduction

Efficient management of human resources (HR) is a key issue for firms' economic success. It does not only consist in dealing appropriately with single workers' demands, bureaucratic procedures or institutional settings. Properly managing HR also (and perhaps mostly) implies finding the right workforce mix and to make the most of workers' skills. A diverse workforce, with respect to education, experience or physical stamina, is often needed due to the variety of tasks that have to be performed within firms. Labour diversity may also benefit firm productivity if it fosters complementarities (e.g. between high- and low-skilled workers), generates spillovers (e.g. knowledge transfers between more and less experienced workers), makes the workplace more enjoyable (e.g. educational/skills diversity could be appreciated by employees) or stimulates demand (e.g. customers may prefer companies that have a diverse workforce).¹ The downside of diversity, however, is that it may lead to misunderstandings, communication problems, personal conflicts or negative reactions from stakeholders that undermine performance (Akerlof and Kranton, 2000; Becker, 1957; Choi, 2007; Lazear, 1999).

Today's labour force is getting more and more heterogeneous: ageing, migration, women's increased labour participation and technological change are key drivers of this phenomenon (Ilmakunnas and Ilmakunnas, 2011; Kurtulus, 2012; Parrotta et al., 2012a). Moreover, in many countries companies are under legislative pressure to diversify their workforce either through quotas or affirmative action. Workforce diversity has thus become an essential business concern. Firms have to manage diversity both internally (i.e. among management and staff) and externally (i.e. by addressing the needs of diverse customers, suppliers or contractors). As a result, an increasing number of firms employ a 'diversity manager' whose task is to ensure that diversity does not hamper productivity but may contribute to attaining the firm's objectives. From the workers' point of view, labour diversity may also generate benefits or losses. The latter may be the result of a more (or less) enjoyable working environment, but they may also derive from a higher (or lower) wage. According to competitive labour market theory, workers are paid at their marginal revenue products. Hence, if labour diversity affects productivity, it may also influence workers' earnings.

The empirical evidence regarding the impact of labour diversity on productivity is very inconclusive and studies on wage effects are exceedingly rare (Ilmakunnas and Ilmakunnas, 2011). Moreover, findings must often be interpreted with caution because of methodological and/or data

¹ In the HR literature, "diversity management" refers to policies and practices that seek to include people within a workforce who are considered to be, in some way, different from those in the prevailing constituency. It usually refers to dimensions such as gender, age, sexual orientation, religion, ethnicity, social origin and physical appearance.

limitations. Only few papers examine how the diversity-productivity nexus is influenced by specific work environments. This is problematic since the optimal degree of diversity is likely to depend on the characteristics of the production unit, for instance the knowledge-intensity and technological content of production (Arun and Arun, 2012; Parrotta et al., 2012b; Pull et al., 2012) or the size of the firm (Fiegenbaum and Karnani, 1991; Konrad and Linneham, 1995; Levy and Powell, 1998; Rynes and Rosen, 1995; Stahl et al., 2010).

The aim of this paper is threefold. First, we put the relationship between labour diversity (measured through education, age and gender) and firm productivity to an updated test, using detailed Belgian linked employer-employee panel data for the years 1999-2006. These data offer several advantages. The panel covers a large part of the private sector, provides accurate information on average productivity (i.e. on the average value added per hour worked) and allows to control for a wide range of worker and firm characteristics. It also enables us to compute various diversity indicators and to address important methodological issues such as firm-level invariant heterogeneity and endogeneity (using both the generalized method of moments (GMM) and Levinsohn and Petrin (2003) estimators). A second aim is to examine how the benefits or losses of labour diversity are shared between workers and firms by estimating the impact of diversity on mean hourly wages and productivity-wage gaps (i.e. profits)² at the firm level. Finally, we investigate the link between diversity and productivity in different work environments defined by the technological and knowledge intensity (we use three complementary taxonomies developed by Eurostat (2012) and by O'Mahony and van Ark, 2003) and firm size.

The remainder of this paper is organized as follows. A review of the literature is presented in the next section. Sections 3 and 4 respectively describe our methodology and data set. The impact of workforce diversity on productivity, wages and productivity-wage gaps across work environments is analysed in Section 5. The last section discusses the results and concludes.

2. Review of the literature

2.1. Workforce diversity and firm productivity

There are different economic forces underlying the relationship between workforce diversity and productivity. As highlighted by Alesina and La Ferrara (2005), these forces may derive from: individual *preferences* (either people may attribute positive (negative) utility to the well-being of

² By definition, the gap between productivity and wages corresponds to the gross operation surplus (i.e. profits).

members of their own group (of other groups) or they may value diversity as a social good), individual *strategies* (even when individuals have no taste for or against diversity, it may be more efficient, notably in the presence of market imperfections, to interact preferably with members of one's own group), or the characteristics of the *production function* (i.e. the complementarity in individual inputs).

Theoretical predictions regarding the optimal workforce composition are mixed. Lazear (1999) follows the production function approach and develops a theoretical model in which a global (i.e. multinational) firm is presented as a diverse (i.e. multi-cultural) team. He argues that labour diversity is beneficial for firm performance if skills and information sets are group-specific. More precisely, he demonstrates theoretically that the gains from diversity are greatest when three conditions are fulfilled: a) individuals have completely disjoint skills and information sets, b) the latter are all relevant for the tasks that have to be performed within the firm, and c) individuals are able to communicate and understand each other.

The organizational demography literature stresses the importance of social similarity for interaction, communication and cohesion among the workforce (Pfeffer, 1985). For instance, diversity in terms of age, education or gender decreases social similarity and could hamper job satisfaction, communication and firm performance. In contrast, social comparison theory posits that people evaluate and compare their opinions and abilities with those of similar others, like individuals of the same age, education or gender (Festinger, 1954). More precisely, individuals may strive to outperform the members of their comparison group (Pelled et al., 1999), which in turn leads to rivalry and conflicts that could undermine organizational performance (Choi, 2007). But social similarity can also be beneficial: a decision may be of better quality when it is the outcome of a confrontation between competing views (Grund and Westergaard-Nielsen, 2008), and rivalry among similar workers may encourage workers to produce more effort in the context of intra-firm 'tournaments' (Lazear and Rosen, 1981).

Productivity effects of workforce diversity are likely to vary across work environments (Stahl et al., 2010). In particular, they may differ with respect to the knowledge intensity and high-tech content of the production. Firms which depend on the exploitation of new opportunities and the development of successful innovations may benefit more from diversity than traditional firms (Prat, 2002). The greater complexity of tasks within innovative sectors is also perceived as a feature likely to foster diversity-related benefits (Jehn, 1995; Stewart, 2006): provided that workforce diversity increases the set of ideas and potential solutions to a given problem, it may stimulate the innovative capacity of firms and hence their productivity (Parrotta et al., 2012b). In addition, the HR literature stresses that firms in innovative sectors may benefit from the promotion of diversity as it potentially

broadens the talent pool, widens perspectives and enlarges the customer base (Cox and Blake, 1991; Yang and Konrad, 2011).

Productivity effects of workforce diversity may also vary according to firm size. In general, workers are likely to be relatively more responsive to the dissimilarity of their close co-workers with whom they interact more frequently. As a consequence, the effects of diversity might be more pronounced in smaller firms in which all workers interact with each other more often (Stahl et al. 2010) and work organization is less rigid (Fiegenbaum and Karnani, 1991; Levy and Powell, 1998). In bigger firms, the diversity of the entire labour force is probably less likely to trigger productivity effects than diversity within teams or departments in which people interact more often. In addition to the frequency of interactions, another factor related to firm size is the capacity to manage diversity. Smaller firms may be less efficient regarding diversity management as their HR departments (if existent) may typically screen workers less systematically during the hiring process, allocate workers to less optimal positions, face more difficulties to recruit diverse workers (Carrington et al. 2000; Chay, 1998; Holzer, 1998) and devote less resources to diversity management (Konrad and Linnehan, 1995; Rynes and Rosen, 1995). The possibilities to relocate workers inside the company in case of disputes are also likely to be more limited in smaller organizations.

In what follows we examine more closely a range of specific arguments for why age, educational or gender diversity may impact firm productivity.

Age diversity

Younger workers are thought to learn faster (Skirbekk, 2003) and to have better cognitive and physical abilities (Hoyer and Lincourt, 1998). Conversely, age is positively correlated with job experience and knowledge about intra-firm structures, relevant markets and networks (Czaja and Sharit, 1998; Grund and Westergaard-Nielsen, 2008, Leonard and Levine, 2013). These two sets of skills are complementary in many production processes, so that age diversity may generate productivity gains for firms able to harness various age-specific skills. The impact of age diversity is also sensitive to firms' working environment (Göbel and Zwick, 2012). It has notably been argued that the complementarity between younger and older workers is higher in knowledge-intensive firms (Cataldi et al, 2011).

Whether the net effect on productivity is positive depends on whether the gains of age complementarities outweigh the costs that come with a more diverse workforce. Lazear (1999) and Jehn et al. (1999) argue that the net effect is likely to be negative for demographic diversity, defined as heterogeneity in terms of age, gender or ethnicity: age diversity may notably increase

communication costs within the labour force, reduce peer pressure (Lazear, 1999), decrease social ties and lead to more absenteeism and turnover (Ilmakunnas et al., 2010).

Educational diversity

Educational diversity enhances firm productivity if there is sufficient mutual learning and collaboration among workers with different educational backgrounds (Hamilton et al., 2004). The gains associated to educational diversity are typically greater if the skills of different educational groups are both relevant and complementary for the tasks performed within the firm (Lazear, 1999).

The effect of educational diversity may also depend on work environments. Jehn et al. (1999), for instance, argue that group performance is more likely to benefit from educational (i.e. task-related) diversity if: i) the tasks that have to be accomplished within a group are complex rather than routine, and ii) the output is of innovative character. These results suggest that the benefits of diversity are more likely to outweigh the costs in high-tech or knowledge-intensive sectors than in more traditional industries.

Gender diversity

Explanations for a potential relationship between gender diversity and firm productivity notably refer to concepts of group efficacy and identity. Social cognitive theory examines how the efficacy of a group (i.e. “a group’s belief in their conjoint capabilities to organize and execute the courses of action required to produce given levels of attainments” (Bandura, 1997: 477)) affects its performance. Mixed gender groups can foster the impact of group efficacy on performance (Lee and Farh, 2004). The argument is that gender diversity is likely to increase the heterogeneity in the values, beliefs and attitudes of the members of a group, which in turn may stimulate critical thinking and prevent the escalation of commitment, i.e. inflated perception of group efficacy resulting in poor decision making.

Akerlof and Kranton (2000) introduce the concept of identity (i.e. a person’s sense of self) into an economic model of behaviour to study how identity influences economic outcomes. Taking gender as an illustration of identity, the authors highlight that social categories such as ‘men’ and ‘women’ are associated to prescribed behaviours and ideal physical characteristics. In the context of work, they argue that a woman doing a “man’s” job (e.g. truck driver or carpenter) may deteriorate the self-image of her male co-workers. As a result, women in male-dominated occupations might

suffer from a strong hostility and be discriminated against by their male counterparts.³ They predict that increasing gender diversity may negatively affect firm performance, especially if men constitute a socially ‘dominant’ group (Haile, 2012). Under the hypothesis that the workforce is less gender-balanced and the environment more ‘macho’ in traditional companies than in high-tech/knowledge intensive firms, above arguments suggest that gender diversity will have a less favorable impact on performance in the former group of companies. This prediction could also be supported by the fact that high-tech/knowledge intensive sectors rely increasingly on inter-personal or ‘soft’ skills (that might be more effectively provided by women) and require generally less physical stamina than traditional (private sector) firms, e.g. construction companies (Arun and Arun, 2002; Webster, 2007).

2.2. Previous empirical studies

Harrison and Klein (2007: 1199) emphasized that the empirical evidence regarding the performance effects of workforce diversity is “weak, inconsistent or both”. This statement remains to a large extent valid. Indeed, findings are still quite inconclusive and often difficult to interpret due to methodological and/or data limitations.

Disaggregate versus aggregate approaches

The impact of workforce diversity on productivity can be analysed at the aggregate (e.g. firm, plant) or disaggregate (e.g. team, individual) level. Both approaches have their caveats and advantages.

Given that the effects of diversity are ultimately driven by individual behaviour and micro-level interactions, disaggregate approaches are in a better position to observe the mechanisms that link labour force diversity to labour productivity. Undoubtedly, micro-level information such as the composition of teams and the interaction patterns of individual employees constitute a precious source of empirical knowledge about diversity. Focusing on the individual or workplace level generally implies working with case studies based on personnel records from single companies or small samples of workers in specific occupations. Despite its high heuristic value, this type of empirical material has generally the disadvantage of limited external validity. Indeed, one of the main limitations of disaggregate studies is that they generally rely on “small samples of workers in narrow occupational fields that often lack a longitudinal component” (Kurtulus, 2011: 685) and that

³ The same reasoning can of course be applied to men employed in female-dominated occupations, e.g. nurses or primary school teachers. This case is, however, less relevant in our sample given that we work with empirical data covering the private sector and that female-dominated occupations are more frequent in the public sector (see Section 4).

“individual-level productivity measures are available only in very special cases” (Ilmakunnas and Ilmakunnas, 2011: 226). Moreover, the measurement of productivity poses additional conceptual and methodological challenges that can hardly be addressed in disaggregate studies. As noted by Vandenberghe (2011), “productivity is in essence a firm-level phenomenon” in which the contribution of each individual, or even of groups of individuals, can often not be disentangled from the overall process of joint value creation within a firm (Kampelmann and Rycx, 2012). As a consequence, “individual workers’ productivity is hardly ever observed” and “the alignment of productivity and pay at the individual level is hard to assess” (Vandenberghe, 2011). Moreover, in disaggregate studies it is methodologically difficult to account for issues such as the simultaneity of productivity and input variations that have attracted considerable attention in the literature on productivity measurement (Van Beveren, 2010). Despite these limitations, micro-level data have been used – mostly in the HRM literature – to measure the effects of diversity on outcomes such as ratings of group effectiveness, financial indicators, organizational commitment, turnover, creativity and frequency of communication. The implications of these studies are still widely debated (Jackson et al., 2003), but two main views emerge from the HRM literature. A first strand supports the “business case for diversity” (Cox, 1993) according to which diversity pays off and thus represents a compelling interest for employers. Other papers reject this hypothesis and depict diversity as “a process loss” and conclude that diversity leads to more conflicts, increased absenteeism and turnover (Stahl et al., 2010).

We argue that studies using aggregate data provide a complementary angle on the productivity effects of personnel diversity. The nature of aggregate empirical data tends to confound the underlying individual-level mechanisms, to the point that they can be criticised for hiding behavioural phenomena inside a black box. However, working with firm-level aggregates allows to apprehend diversity at the collective level while resolving numerous issues related to productivity measurements (Vandenberghe, 2011). Firm-level data notably enables: i) to estimate the relationships between firm-level labour diversity⁴ and statistically reliable measures of plant- or firm-level productivity (e.g. total factor productivity (TFP) or value-added) and ii) the use of large representative datasets. The use of more aggregate data tends to prevail in diversity studies carried out by labor economists, increasingly using linked employer-employee data covering large parts of

⁴ It can be argued that the performance of a firm-level diversity measure in terms of capturing actual interactions between diverse workers decreases with firm size. While workers in small firms are likely to be exposed to the entire diversity within the establishment, in big firms there may be a difference between overall firm-level diversity and the diversity within the groups in which most of the social interactions occur (e.g. departments or teams). This issue will be explored further in Section 5.3.

the economy. While the empirical literature on diversity in labor economics is substantially less voluminous compared to the HRM literature, it has so far also led to mixed conclusions.

In general, the comparison of aggregate and disaggregate approaches in the empirical literature reveals several trade-offs. From an econometric point of view, the measurement of productivity and the external validity of results tend to be more robust in aggregate studies based on firm-level diversity and performance measures. Conversely, disaggregate studies are likely to foster our understanding of the inter-personal mechanisms linking diversity and productivity. Especially in large entities the use of firm-level aggregates can mask intra-firm variations in the way that workers in different units or departments are exposed to diversity. While our study focuses on the advantages of firm-level aggregates, it is therefore important to make use of hypotheses and insights generated by studies conducted at the micro-level and to analyse potential biases inherent to aggregate data (see Section 5.3).

In the remainder of this section, we briefly review results from earlier empirical studies, concentrating on those that: a) use either detailed personnel records from individual companies or large representative linked employer-employee data and, b) address the productivity effects of at least one of the diversity dimensions investigated in this paper, i.e. education, age and gender.⁵

Evidence based on case studies

A series of papers investigates the diversity-performance nexus using case studies, often based on personnel records from single companies. Hamilton et al. (2004) use weekly data from a Californian garment manufacturing plant for the years 1995-1997. Their results indicate that teams with greater diversity in workers' abilities and composed of only one ethnicity (namely Hispanics) are more productive (i.e. sew more garments per day). In contrast, team heterogeneity in workers' age is found to decrease productivity. Yet, results for team demographics (age and ethnicity) should be considered with caution as they become insignificant when applying fixed effects (FE). Leonard and Levine (2006) rely on longitudinal data (collected in 1996-1998) from a low-wage service-sector employer with establishments (retail stores or restaurants) throughout the U.S. They study the influence of demographic (race, gender and age) diversity between a workgroup and its customers and within a workgroup on an indirect measure of productivity, namely individual turnover within workgroups. Results (controlling for individual FE) show that diversity does not consistently predict turnover. In contrast, belonging to a numerical minority with respect to co-workers and customers, especially

⁵ For extended reviews of the literature see e.g. Horwitz and Horwitz (2007), Ilmakunnas and Ilmakunnas (2011) and Roberge and van Dick (2010).

with respect to race, often leads to higher turnover. Kurtulus (2011) uses detailed personnel records of a large U.S. firm in the health service industry for the years 1989-1994. Her FE estimates highlight that diversity within organisational divisions with respect to age, firm tenure and performance is associated with lower labour productivity as measured through subjective evaluation by managers. In contrast, worker's performance appears to be influenced positively by intra-divisional wage inequality.

Evidence based on linked employer-employee data

Another strand of the literature relies on linked employer-employee data (LEED). Barrington and Troske (2001) examine the impact of plant-level diversity with respect to age and gender on plant-level productivity (i.e. value-added and sales per worker and TFP) in the manufacturing, retail trade and services industry. Based on cross-sectional LEED for 1999, their OLS estimates reject the hypothesis that workforce diversity would be detrimental for the productivity of U.S. plants. Grund and Westergaard-Nielsen (2008) use LEED for the Danish private sector over the period 1992-1997. Their FE estimator suggests that firms with a medium age dispersion obtain the highest value-added and profits per employee. Leonard and Levine (2013) examine workforce diversity in terms of age, gender, and race in 800 establishments of a US retail firm. Their between-establishment estimator suggests that age diversity is negatively related to sales, whereas the impact of race and gender diversity is insignificant or not substantial in their sample. Giuliano et al (2011) further explore the complex dynamics of ethnic diversity and find that employees usually have better outcomes when they are the same race as their hierarchical superiors.

The studies of Navon (2009), Ilmakunnas and Ilmakunnas (2011) and Parrotta et al. (2012a) are more directly comparable to our investigation as they not only control for firm time-invariant unobserved heterogeneity but also for endogeneity. Navon (2009) relies on LEED for the Israeli manufacturing industry over the period 2000-2003 and employs semi-parametric estimation techniques developed by Olley and Pakes (1996, hereafter OP) and Levinsohn and Petrin (2003, hereafter LP) to control for plant FE and endogeneity. He finds that within-plant educational diversity among higher educated workers, measured as the variability in academic disciplines in which the latter obtained their university degrees, is beneficial for plant-level added value. Ilmakunnas and Ilmakunnas (2011) investigate whether firms and employees benefit from diversity using Finnish LEED covering the industrial sector (i.e. mining, manufacturing, energy and construction) for the years 1990-2004. Plant-level regressions (estimated with FE, generalized methods of moments (GMM) and OP estimators) show that TFP depends positively (negatively) on

age (educational) diversity. In contrast, the latter variables turn out to be statistically insignificant when the authors estimate wage regressions for individuals. Parrotta et al. (2012a) use register-based LEED covering most of the Danish private sector between 1995 and 2005. Their results are based on an estimator developed by Akerberg et al. (2006) and show that diversity in education (ethnicity, age and gender) enhances (deteriorates) firm's value added.⁶ Moreover, dividing industries into two groups according to their aggregate level of R&D expenditures, they find no evidence that the impact of diversity is different for firms in high-tech industries (i.e. in industries with above-average R&D expenditures), although the latter are typically thought to require more creative thinking and problem-solving skills.⁷

Our paper contributes to this literature by investigating how diversity with respect to education, age and gender affects productivity, wages and productivity-wage gaps at the firm level. We also examine how the diversity-productivity-wage nexus varies according to the knowledge-intensity and high-tech status of firms and whether estimates are sensitive to firm size. Our empirical approach relies on longitudinal LEED from the Belgian private sector, uses various diversity indicators, controls for a large set of covariates and implements both GMM and LP estimation techniques. This approach differs from natural or quasi-natural experiments in which the intra-firm diversity would be the result of a truly exogenous shock. While natural experiments have of course very attractive features in that they allow to clearly identify the effects of changes in firm diversity in a specific context, it also has its downsides such as limited external validity. On any account, it should be noted that the approach presented in this paper controls for a range of measurement issues that improve considerably the reliability of estimation results compared to existing research. For instance, the consequential issue of simultaneity and of time-invariant unobserved firm heterogeneity is rarely addressed in previous studies.

⁶ A few recent papers (e.g. Vandenberghe, 2011) test for gender wage discrimination and investigate with LEED how the share of women within firms influences the latter's productivity and labour costs. Yet, results from these studies are not straightforward to interpret from a diversity perspective. Indeed, whether a growing share of women corresponds to more or less gender diversity depends on the initial intra-firm proportion of women.

⁷ In a companion paper, Parrotta et al. (2012b) merge the Danish LEED set with information on firms' innovation ability for the years 1995-2003. Using an instrumental variable approach, they find that ethnic diversity within firms is valuable for the latter's capacity to innovate. In contrast, educational, age and gender diversity turn out to be statistically insignificant. Based on similar data for the period 1980-2002 and controlling for endogeneity, Marino et al. (2012) show in addition that intra-firm diversity in terms of education and ethnicity (age and gender) increases (decreases) workers' transition probability from employment to self-employment, i.e. their propensity to become entrepreneurs.

3. Methodology

Our empirical approach is based on the separate estimation of an added value function and a wage equation at the firm level. The latter provide parameter estimates for the impact of labour diversity (with respect to education, age and gender) on average productivity and wages, respectively. Given that both equations are estimated on the same sample with identical control variables, the parameters for marginal products and wages can be compared and conclusions can be drawn on how the benefits or losses of diversity are shared between workers and firms. This technique was pioneered by Hellerstein and Neumark (1995) and refined by Hellerstein et al. (1999), Hellerstein and Neumark (2004), Aubert and Crépon (2009) and van Ours and Stoeldraijer (2011). It is now standard in the literature on the productivity and wage effects of labour heterogeneity (see e.g. Cataldi et al. 2012; Göbel and Zwick 2012; Vandenberghe 2013).

The estimated firm-level productivity and wage equations are the following:

$$\log\left(\frac{\text{Value Added}}{\text{Hours}}\right)_{j,t} = \alpha + \beta_1 A_{j,t}^\sigma + \beta_2 E_{j,t}^\sigma + \beta_3 G_{j,t}^\sigma + \beta_4 \bar{A}_{j,t} + \beta_5 \bar{E}_{j,t} + \lambda X_{j,t} + \varepsilon_{j,t} \quad (1)$$

$$\log\left(\frac{\text{Total Wages}}{\text{Hours}}\right)_{j,t} = \alpha^* + \beta_1^* A_{j,t}^\sigma + \beta_2^* E_{j,t}^\sigma + \beta_3^* G_{j,t}^\sigma + \beta_4^* \bar{A}_{j,t} + \beta_5^* \bar{E}_{j,t} + \lambda^* X_{j,t} + \varepsilon_{j,t}^* \quad (2)$$

The dependent variable in equation (1) is firm j 's hourly added value, obtained by dividing the total added value (at factor costs) of the firm j in period t by the total number of work hours (taking into account paid overtime hours) that have been declared for the same period. The dependent variable in equation (2) is firm j 's average hourly gross wage (including premia for overtime, weekend or night work, performance bonuses, commissions, and other premia). It is obtained by dividing the firm's total wage bill by the total number of work hours. Hence, the dependent variables in the estimated equations are firm averages of added value and wage on an hourly basis.

Labour diversity indicators with respect to education, age and gender (E^σ , A^σ and G^σ) are the main variables of interest. A theoretical model justifying the inclusion of diversity indicators, on top of mean values, in a firm-level productivity equation is provided by Iranzo et al. (2008). The three firm-level diversity indicators used in this paper (i.e. the standard deviation, the average dissimilarity index and the alternative gender diversity index) are conceptually and mathematically relatively

similar and can be regarded as robustness tests for our regression results.⁸ In particular, all three diversity indicators share the property that diversity is maximal in case of a symmetrical bi-modal distribution with the modes occurring at the extreme values of the attribute under study (i.e. when observations are equally split between the modes); conversely, the minimum of all three indicators is reached when all workers belong to the same group. A more detailed description of the three indicators and specific numerical examples for their minimum, maximum and range are provided in Appendix 1.

In addition to testing robustness, the inclusion of the three indicators for diversity allows for better comparability with other studies: the standard deviation is a widely used indicator with high familiarity (Harrison and Klein, 2007); the average Euclidean distance on which the dissimilarity index is based has been included given that this measure is frequently used in the literature on diversity (Ilmakunnas and Ilmakunnas, 2011); finally, the alternative diversity measure based on sex shares is widely used in the literature on gender diversity (Hoogendoorn et al., 2013).^{9,10}

In line with earlier empirical work, we also add workers' average age and education at the firm-level (\bar{E} and \bar{A}) among regressors in equations (1) and (2).¹¹ Control variables in the vector X include the share of part-time workers, the fraction of workers with a fixed-term employment contract, the proportion of employees with at least 10 years of tenure, the percentage of white-collar workers, firm size (i.e. the number of employees) and capital stock¹², 8 industry dummies, and 7 year dummies.¹³

Estimating equations (1) and (2) allows gauging the effect of labour diversity on firm productivity and wages, but it does not allow testing directly whether the difference between the added value and the wage coefficients for a given diversity indicator is statistically significant. A simple method to obtain a test for the significance of productivity-wage gaps has been proposed by

⁸ To avoid multicollinearity, the effect of the three diversity indicators is estimated in separate regressions.

⁹ Although many studies on diversity make use of Euclidean distances, the exact definitions and notations of the employed indicators can differ. For example, some authors prefer to compute only $N-1$ distances and exclude the trivial distance between the focal worker and itself – an issue that has little empirical incidence in our case given that our sample does not include very small firms. For a more detailed discussion see Appendix 1.

¹⁰ The three dimensions of diversity in the model may also interact. We have tested for interdependent effects of age, gender and educational diversity using interaction variables and the faultlines approach (as described in Van Knippenberg et al. (2011)) and found no substantial evidence for interdependencies (results available upon request).

¹¹ We do not find any evidence of multicollinearity between average values and diversity indicators for age and education (the corresponding pair-wise correlations are around 0.3 and 0.1, respectively). However, multicollinearity is more of a concern in the case of gender, a variable with only two modalities and pair-wise correlations around 0.6, which is why we have excluded the firm's share of female employees from the regression.

¹² It is estimated through the "perpetual inventory method" (or PIM, see OECD (2009) for more details). The PIM incorporates the idea that the capital stock results from investment flows and corrects for capital depreciation and efficiency losses. Following standard practice, we assume a 5 percent annual rate of capital depreciation.

¹³ All independent variables are measured in terms of shares in total work hours. For instance, the fraction of part-time workers is computed on the basis of the proportion of hours worked by employees working less than 30 hours per week over the total amount of hours worked with the firm.

van Ours and Stoeldraijer (2011). We apply a similar approach and estimate a model in which the *difference* between firm j 's hourly value added and hourly wage (i.e. the hourly gross operating surplus) is regressed on the same set of explanatory variables as in equations (1) and (2). This produces coefficients for the diversity indicators and directly measures the size and significance of their respective productivity-wage gaps.

Equations (1) and (2), as well as the productivity-wage gap, can be estimated with different methods: pooled ordinary least squares (OLS), a fixed-effect (FE) model, the generalized method of moments (GMM) estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998), or a more structural approach suggested by Levinsohn and Petrin (2003, hereafter LP). This being said, pooled OLS estimators of productivity models have been criticized for their potential “heterogeneity bias” (Aubert and Crépon 2003: 116) due to the fact that firm productivity depends to a large extent on firm-specific, time-invariant characteristics that are not measured in micro-level surveys (advantageous location, firm-specific assets like the ownership of a patent, or other firm idiosyncrasies).

One way to remove unobserved firm characteristics that remain unchanged during the observation period is by estimating a FE model. However, neither pooled OLS nor the FE estimator address the potential endogeneity of our explanatory variables.¹⁴ Yet, labour diversity is likely to be endogenous. Indeed, any shock in wages or in productivity levels might generate correlated changes in the firm's workforce and in labour productivity that are not due to changes in the firm's workforce composition *per se*. For instance, one might expect that a firm undergoing a negative productivity shock would prefer not to hire new staff, which would increase the age of the workforce and affect the age diversity index. Similarly, during economic downturns, firms may be more likely to reduce personnel among women and less educated workers as adjustments costs are often lower for these categories of workers due to their relatively lower wages and/or tenure. In order to control for this endogeneity and for the presence of firm fixed effects, we estimated our model using system GMM (GMM-SYS) and LP estimators, respectively.

The GMM-SYS approach boils down to simultaneously estimating a system of two equations (one in level and one in first differences) and to relying on internal instruments to control for endogeneity. More precisely, diversity variables¹⁵ in the differenced equation are instrumented by their lagged levels and diversity variables in the level equation are instrumented by their lagged differences (Göbel and Zwick, 2012). The implicit assumption is that changes (the level) in (of) the

¹⁴ Expected biases associated to OLS and the relatively poor performance and shortcomings of the FE estimator in the context of firm-level productivity regressions are reviewed in Van Beveren (2010).

¹⁵ By ‘diversity variables’, we mean diversity variables *stricto sensu* and other endogenous input factors.

dependent variable – productivity or wages – in one period, although possibly correlated with contemporaneous variations (levels) in (of) diversity variables, are uncorrelated with lagged levels (differences) of the latter. Moreover, changes (levels) in (of) diversity variables are assumed to be reasonably correlated to their past levels (changes). One advantage of GMM-SYS is that time-invariant explanatory variables can be included among the regressors, while the latter typically disappear in difference GMM. Asymptotically, the inclusion of these variables does not affect the estimates of the other regressors because instruments in the level equation (i.e. lagged differences of diversity variables) are expected to be orthogonal to all time-invariant variables (Roodman, 2009). In order to find the correctly specified model, we start with the moment conditions that require less assumptions and increase the number of instruments progressively (Göbel and Zwick, 2012). To examine the validity of additional instruments, we apply the Hansen (1982) test of over-identifying restrictions. In addition, Arellano-Bond (1991) test for serial correlation (i.e. for second-order autocorrelation in the first differenced errors) is used to assess whether estimates are reliable. Practically, we choose the model with the lowest number of lags that passes the Hansen and Arellano-Bond tests.

As an alternative to the GMM-SYS method, Olley and Pakes (1996) have developed a consistent semi-parametric estimator. This estimator, particularly well suited for panels with small t and big N , controls for endogeneity and firm fixed unobserved heterogeneity by using the employer's investment decision to proxy for unobserved productivity shocks. The intuition is that firms respond to time-varying productivity shocks observed by managers (and not by econometricians) through the adjustment of their investments. Put differently, profit-maximizing firms react to positive/negative productivity shocks by increasing/decreasing their output, which requires more/less investments (or intermediate inputs, see below). The OP estimation algorithm relies on the assumptions that there is only one unobserved state variable at the firm level (i.e. its productivity) and that investments increase strictly with productivity (conditional on the values of all state variables). This monotonicity condition implies that any observation with zero investment has to be dropped from the data, which generally leads to a sharp decrease in sample size. To avoid this drawback, Levinsohn and Petrin (2003) use intermediate inputs (i.e. inputs such as energy, raw materials, semi-finished goods, and services that are typically subtracted from gross output to obtain added value) rather than investments as a proxy for productivity shocks. Given that firms typically report positive values for intermediate inputs in each year, most observations can be kept with the LP approach. An additional argument for using intermediate inputs rather than investments is that the former may adjust more smoothly to the productivity term than the latter, especially if adjustment costs are an important issue. For instance, “if adjustment costs lead to kink points in the investment demand function, plants

may not respond fully to productivity shocks, and some correlation between the regressors and the error term can remain” (Petrin et al., 2004: 114). Intermediate inputs would thus provide a better proxy for unobserved productivity shocks. In the basic LP model, labour is a fully variable and capital a fixed input. Given our focus on diversity, the variable inputs in our setup include first and/or second moments of workforce characteristics. Assuming that intermediate inputs depend on capital and the unobservable productivity shocks, this relationship can be solved for the productivity term (Ilmakunnas and Ilmakunnas, 2011). When relying on the LP estimation algorithm, standard errors are computed using a bootstrap approach taking the panel structure of the data into account (Petrin et al., 2004).

4. Data and descriptive statistics

Our empirical analysis is based on a combination of two large data sets covering the period 1999-2006. The first, carried out by Statistics Belgium, is the ‘Structure of Earnings Survey’ (SES). It covers all firms operating in Belgium which employ at least 10 workers and with economic activities within sections C to K of the NACE Rev.1 nomenclature.¹⁶ The survey contains a wealth of information, provided by the management of firms, both on the characteristics of establishments (e.g. sector of activity, number of employees) and their workers (e.g. age, education, sex, tenure, gross earnings, paid hours, occupation).¹⁷ The SES provides no financial information. It has therefore been merged with a firm-level survey, the ‘Structure of Business Survey’ (SBS). The SBS, also conducted

¹⁶ It thus covers the following sectors: (i) mining and quarrying (C), (ii) manufacturing (D), (iii) electricity, gas and water supply (E), (iv) construction (F), (v) wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods (G), (vi) hotels and restaurants (H), (vii) transport, storage and communication (I), (viii) financial intermediation (J), and (ix) real estate, renting and business activities (K).

¹⁷ The SES is a stratified sample. The stratification criteria refer respectively to the region (NUTS-groups), the principal economic activity (NACE-groups) and the size of the firm. The sample size in each stratum depends on the size of the firm. Sampling percentages of firms are respectively equal to 10, 50 and 100 percent when the number of workers is lower than 50, between 50 and 99, and above 100. Within a firm, sampling percentages of employees also depend on size. Sampling percentages of employees reach respectively 100, 50, 25, 14.3 and 10 percent when the number of workers is lower than 20, between 20 and 50, between 50 and 99, between 100 and 199, and between 200 and 299. Firms employing 300 workers or more have to report information for an absolute number of employees. This number ranges between 30 (for firms with between 300 and 349 workers) and 200 (for firms with 12,000 workers or more). To guarantee that firms report information on a representative sample of their workers, they are asked to follow a specific procedure. First, they have to rank their employees in alphabetical order. Next, Statistics Belgium gives them a random letter (e.g. the letter O) from which they have to start when reporting information on their employees (following the alphabetical order of workers' names in their list). If they reach the letter Z and still have to provide information on some of their employees, they have to continue from the letter A in their list. Moreover, firms that employ different categories of workers, namely managers, blue- and/or white-collar workers, have to set up a separate alphabetical list for each of these categories and to report information on a number of workers in these different groups that is proportional to their share in the firm's total employment. For example, a firm with 300 employees (namely, 60 managers, 180 white-collar workers and 60 blue-collar workers) will have to report information on 30 workers (namely, 6 managers, 18 white-collar workers and 6 blue-collar workers). For more details see Demunter (2000).

by Statistics Belgium, provides information on financial variables such as firm-level material inputs, investments, added value and gross operating surplus. The coverage of the SBS differs from that of the SES in that it does not cover the whole financial sector (NACE J) but only Other Financial Intermediation (NACE 652) and Activities Auxiliary to Financial Intermediation (NACE 67). The merger of the SES and SBS datasets has been carried out by Statistics Belgium using firms' social security numbers.

Information in the SES refers to the month of October in each year, while data in the SBS are measured over entire calendar years from January until December. To avoid running a regression where information on the dependent variable (collected for the entire calendar year) precedes the recording of the explanatory variables (collected in October), all explanatory variables in Equations (1) and (2) have been lagged by one year. In this way, information on diversity indices is recorded in October in year t and used to explain firm-level productivity and wages during the calendar year $t+1$. The imperfect synchronization of SBS and SES data might introduce some fuzziness into our estimates since we cannot exclude the occurrence of external events influencing productivity or wages in the intermediate period. This concern could only be completely eliminated if we had firm-level information on the average diversity for the entire calendar year. This being said, even if this information was available, there is a compelling argument for using asynchronized information on diversity and productivity: it is difficult to conceive how changes in diversity could generate *immediate* effects, so that potential productivity effects are more likely to occur after a certain adjustment period. The slightly asynchronised use of SBS and SES data is therefore arguably the best option in light of data availability and productivity dynamics.

As a consequence, our sample contains firms that are observed in at least two consecutive years and thus over-represents medium-sized and large firms given that sampling percentages of firms in our data increase with the size of the latter.¹⁸ Next, we exclude workers and firms for which data are missing or inaccurate.¹⁹ Finally, we drop firms with less than 10 observations, the reason for this being our use of the first and second moments of workers' characteristics at the firm level.²⁰ Our final sample consists of an unbalanced panel of 7,463 firm-year-observations from 2,431 firms. It is representative of all medium-sized and large firms in the Belgian private sector, with the exception of large parts of the financial sector (NACE J) and the electricity, gas and water supply industry (NACE E).

¹⁸ See footnote 17.

¹⁹ For instance, we eliminate a (very small) number of firms for which the recorded value added was negative.

²⁰ This restriction is unlikely to affect our results as it leads to a very small drop in sample size. The average number of observations per firm in each year is equal to 35 in our final sample.

[INSERT TABLE 1]

Table 1 sets out the means and standard deviations of selected variables. We observe that firms have a mean value added per hour worked of 61.06 EUR and that workers' mean gross hourly wage stands at 17.14 EUR. As regards diversity indicators, we find that the intra-firm standard deviation (the dissimilarity index) reaches respectively 9.33 (12.61) for age, 1.90 (2.54) for education, and 0.35 (0.46) for gender. For comparison, Ilmakunnas and Ilmakunnas (2011) report similar standard deviations (and average dissimilarity indices) for Finland of 10.04 (13.67) for age and 1.93 (2.71) for education.

Employees in our sample have on average 11.44 years of education, are 38.42 years old, and are essentially concentrated in the manufacturing industry (57 percent), wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods (12 percent), construction (10 percent) and real estate, renting and business activities (11 percent). Moreover, firms employ on average 132 full-time equivalent workers, 27 per cent of women, 45 percent of white-collar workers, 61 percent of workers with less than ten years of tenure, 4 percent of workers with a fixed-term employment contract, and 2 percent of part-time workers.

5. Empirical results

5.1. Benchmark specification

Given the above mentioned econometric issues associated with pooled OLS and FE estimations, we report findings based on the GMM-SYS and LP estimators. Table 2 shows the impact of diversity indicators (the standard deviation and the dissimilarity index, respectively) on productivity, mean wages and productivity-wage gaps at the firm-level.

[INSERT TABLE 2]

GMM-SYS estimates are reported in columns (1) to (6). To examine their reliability, we first apply the Hansen and Arellano-Bond tests. For all specifications, they respectively do not reject the null hypothesis of valid instruments²¹ and of no second-order autocorrelation in the first differenced errors. Results in columns (1) and (2) suggest that age and gender diversity have a significant

²¹ First and second lags of explanatory variables (except time dummies) are used as instruments.

negative influence on productivity. More precisely, they indicate that if age diversity increases by one standard deviation, productivity on average decreases by 4 percent.²² Such a change in diversity is equivalent to an increase in the standard deviation of age of 1.82 years and an increase in the dissimilarity index of 2.52 years. To give a numerical example of a hypothetical firm with four employees, such a change roughly corresponds to a shift from workers aged 25, 40, 45, and 55 years to workers aged 25, 40, 45 and 60 years. The mean impact on productivity of a standard deviation increase in gender diversity (that is by respectively 0.15²³ and 0.22 for the standard deviation and dissimilarity index) is also estimated at about minus 4 percent.²⁴ Concerning education diversity, we find that the regression coefficient is positive but statistically insignificant in both specifications.

LP estimates, reported in columns (7) and (8), confirm that age and gender diversity appear to be harmful for productivity. Point estimates indeed suggest that an increase in these variables of one standard deviation hampers productivity on average by 1.3 and 1.7 percent, respectively. As regards the coefficient on educational diversity, it is still positive but now also significantly different from zero. More precisely, results suggest that when educational diversity increases by one standard deviation (that is by respectively 0.84²⁵ and 1.15 years for the standard deviation and dissimilarity index), productivity on average rises by approximately 2.7 percent.

Findings in columns (3) and (4) show that GMM-SYS regression coefficients associated to diversity indices are of the same sign and order of magnitude in the wage and productivity equations. While age and gender diversity are found to depress mean workers' wages, the reverse finding is found for educational diversity. Results in columns (5) and (6) further indicate that educational and gender diversity have a non-significant impact on the productivity-wage gap. Gains (losses) due to educational (gender) diversity thus appear to be shared 'competitively' between workers and firms so that profits remain unaffected. In contrast, age diversity is found to have a stronger negative impact on productivity than on wages. More precisely, results show that an increase of one standard deviation in the age diversity index decreases the productivity-wage gap (i.e. profits) on average by about 2.3 percent.²⁶

²² $-0.022 \cdot 1.82 = -0.04 = -4\%$ and $-0.016 \cdot 2.52 = -0.04$.

²³ A change of 0.15 in the standard deviation of gender corresponds roughly to the case of a firm in which the share of women increases from 15 to 50%.

²⁴ $-0.260 \cdot 0.15 = -0.039 = -3.9\%$ and $-0.176 \cdot 0.22 = -0.039$.

²⁵ An increase of 0.84 in the standard deviation of educational attainment is roughly similar to the change of a workforce of five individuals with 8, 9, 10, 10 and 12 years of education to a workforce with 6, 9, 10, 10 and 12 years of education.

²⁶ Results presented in Table 3 remain quite stable when replacing the gender standard deviation/dissimilarity index by an alternative indicator, namely the share of women times the share of men within firms (see Appendix 2). We also tested for a non linear relationship between the dependent variables (productivity, wages and productivity-wage gaps) and diversity indices. Therefore, we respectively include diversity indices in level, squared (and cubed) and used dummy variables to test for structural breaks notably at the 33rd and 66th percentiles of the distribution of the diversity indices. Results, reported in Appendix 3, show no evidence of nonlinearities.

5.2 Does the technological/knowledge environment matter?

HT/KIS nomenclature

The diversity-productivity-wage nexus is likely to vary across different work environments. Various theoretical arguments (reviewed in section 2.2) suggest in particular that the former may differ between knowledge intensive sectors and more traditional industries. Given the scarcity of empirical evidence on this issue, in this section we first present estimates of our model for two distinct types of firms: those belonging to high-medium tech/knowledge intensive sectors (HT/KIS) and those that do not. The subdivision of firms is based on a taxonomy developed by Eurostat (2012) that classifies manufacturing industries (at NACE 2- and/or 3-digit level) according to their degree of technological intensity (primarily assessed through the ratio of R&D expenditures to value added) and services (at NACE 2-digit level) according to their degree of knowledge intensity (i.e. the share of tertiary educated people in the activity).²⁷

Applied to our sample, this taxonomy classifies 679 firms as HT/KIS and 1,778 as non-HT/KIS firms.²⁸ As shown in Table 1, these two types of firms differ along several dimensions. Both the average hourly value added and wage are higher in HT/KIS compared to non-HT/KIS firms, confirming the intuition that HT/KIS firms are in general more productive. Moreover, HT/KIS firms are found to have a significantly larger capital stock and to invest more. Differences in age, educational and occupational composition also exist: the workforce of HT/KIS firms is on average much more concentrated in white collar occupations (62 vs. 39 percent), somewhat more educated and slightly younger compared to non-HT/KIS firms. Interestingly, HT/KIS firms are also characterised by a more feminine labour force (33 vs. 24 percent). Both HT/KIS and non-HT/KIS employment is predominantly concentrated in the manufacturing sector (respectively around 53 and 59 percent). Yet, while almost 40 percent of HT/KIS employment is found in real estate, renting and business activities and financial intermediation, about a third of non-HT/KIS workers is employed in the construction and wholesale and retail trade industry (including repair of motor vehicles, motorcycles and personal and household goods).

To formally test for differences between HT/KIS and non-HT/KIS firms, we add to our benchmark specification: i) a dummy variable that indicates if the firm is classified as HT/KIS, and

²⁷ A detailed presentation is provided in Appendix 4.

²⁸ The sum of HT/KIS and non-HT/KIS firms (2,457) is greater than the total number of firms in the baseline model (2,431). This is due to a small number of firms that changed NACE codes during the period 1999-2006. Suppression of these firms does not affect our conclusions.

ii) interactions between this HT/KIS dummy and first and second moments of age, education and gender variables.

[INSERT TABLE 3]

Results based on GMM-SYS and LP estimators are reported in Table 3. The reliability of GMM-SYS estimates is supported by the outcomes of the Hansen and Arellano-Bond tests. For all specifications, they respectively do not reject the null hypothesis of valid instruments²⁹ and of no second-order autocorrelation in first differenced errors.

Overall, GMM-SYS and LP estimates again suggest that age (educational) diversity is detrimental (beneficial) for firm productivity. Moreover, given that interaction effects with the HT/KIS dummy variable are systematically insignificant, it appears that the size of the elasticity between productivity and diversity in age and education does not depend on firms' technological environment and knowledge-intensity. Furthermore, results indicate that age and educational diversity have a similar impact on wages and productivity. On the whole, they thus suggest that profitability (i.e. the productivity-wage gap) does not depend on the diversity of the workforce in terms of education or age.

We find remarkable results regarding the consequences of gender diversity on productivity. Indeed, while gender diversity is still found to hamper firms' productivity in more traditional sectors, firms belonging to high-medium tech/knowledge intensive sectors appear to be significantly more productive when employing a more gender-balanced workforce. More precisely, estimates suggest that if gender diversity – measured respectively through the standard deviation and dissimilarity index – increases by one standard deviation, productivity increases (decreases) on average by between 2.5 and 6 percent (3 and 5 percent) in HT/KIS firms (non-HT-KIS firms). Besides, results show that gender diversity has no significant influence on the productivity-wage gap in both types of environments.

KIA and ICT nomenclatures

To examine the robustness of these results, we used two alternative taxonomies enabling to distinguish between knowledge-intensive industries and more traditional sectors. The first alternative

²⁹ Yet, it should be acknowledged that Hansen over-identification tests for the firm-level wage regressions are significant at the 5% level (p-values are respectively equal to 0.055 and 0.065). Therefore, results for wages should be interpreted with caution.

taxonomy is Eurostat's classification of Knowledge Intensive Activities (KIA). It differs from the HT/KIS classification in that it applies the same methodology to all sectors of industries and services. Moreover, it focuses solely on the level of education of the labour force. Both the HT/KIS and KIA nomenclatures classify service industries as knowledge intensive if the share of tertiary educated workers represents more than one third of total employment. In our specification, being HT/KIS or KIA is therefore potentially correlated with our explanatory variables 'average education' and 'educational diversity', which could dilute the measurement of the actual impact of education on the dependent variables. To check if this problem affects our conclusions, we also use an alternative classification of industries that is not directly based workers' educational level, namely the ICT classification developed by O'Mahony and van Ark (2003). The latter classifies industries according to their ICT capital intensity at the NACE 3-digit level. Industry groups are based on whether they produce ICT goods and services and whether they intensively use ICT or not.³⁰

Results based on KIA and ICT classifications are shown in Appendices 6 and 7.³¹ They are very similar to those obtained on the basis of the HT/KIS classification. This is quite remarkable, particularly given that correlation coefficients between HT/KIS, KIA and ICT taxonomies are not very high (see Appendix 5). Overall, results again highlight that productivity depends positively (negatively) on educational (age) diversity. Moreover, they show that gender diversity is detrimental (beneficial) for firm added value in traditional (knowledge/ICT intensive) industries. In line with our benchmark specification (see Table 3), results also indicate that age (educational) diversity has a negative (no significant) impact on firm profits. As regards the influence of gender diversity on the productivity-wage gap, results depend on whether we use the ICT or KIA nomenclatures. In the former case, profits do not depend on whether the labour force is gender-balanced or not. In the latter, gender diversity is found to increase (decrease) profits in firms belonging to knowledge intensive (traditional) sectors.

5.3 Does the diversity effect vary according to firm size?

The effect of diversity may vary according to firm size (see Section 2). To formally test for differences between small and big firms, we interact the first and second moments of age, education and gender variables with firm size. Secondly we also add to our benchmark specification: i) a

³⁰ For more details see Appendix 4.

³¹ For the sake of brevity, we only report results based on the standard deviation of age, education and gender. Estimates using the dissimilarity index (available upon request) lead to similar conclusions.

dummy variable indicating if the firm employs more than 500 workers, and ii) interaction effects between this dummy and the first and second moments of age, education and gender variables.³²

[INSERT TABLES 4 and 5]

Results based on GMM-SYS and LP estimators are reported in Tables 4 and 5. The reliability of GMM-SYS estimates relative to the productivity and gap equations is supported by the Hansen and Arellano-Bond tests. GMM-SYS results for the wage equation should be interpreted with caution as the p-value associated to the Hansen test is below 0.05. Our regression results suggest that interaction effects between firm size and diversity variables are either not significant or quantitatively very small. Taken at face value, they indicate that the effects of diversity do not substantially depend on firm size.

6. Discussion and conclusion

This paper estimates the impact of workforce diversity (in terms of education, age and gender) on productivity, wages and productivity-wage gaps (i.e. profits). It contributes significantly to the existing literature as it is one of the first: i) to use large representative data (i.e. Belgian linked employer-employee panel data covering most private sector firms over the period 1999-2006), ii) to address important methodological issues such as firm-level invariant heterogeneity and endogeneity, iii) to examine how the benefits or losses of labour diversity are shared between workers and firms (i.e. to extend the analysis to wages and productivity-wage gaps), iv) to investigate whether the diversity-productivity-wage nexus depends on the degree of technological/knowledge intensity of firms, v) to test whether results vary according to firm size.

Findings, based on the generalized method of moments (GMM) and Levinsohn and Petrin (2003) estimators, show that educational diversity is beneficial for firm productivity and wages. In contrast, age and gender diversity are found to hamper firm-level added value and average earnings. The magnitude of these effects is relatively big: estimates notably suggest that when age or gender diversity (educational diversity) increases by one standard deviation, productivity drops (rises) on average by around 4 percent (almost 3 percent). Yet, the consequences of gender diversity are found to depend on the technological/knowledge intensity of firms. Gender diversity generates gains in

³² We also examined alternative specifications fixing the threshold for firm size respectively at 100, 200, 250 and 300 workers. In addition, we tested for interaction effects with firm size in technological/knowledge intensive sectors and more traditional industries (using respectively the HT/KIS, KIA and ICT nomenclatures). Results (available on request) support our conclusion, namely that the diversity-productivity-wage nexus does not substantially vary with firm size.

high-tech/knowledge intensive sectors: productivity is found to rise on average by between 2.5 and 6 percent following a one standard deviation increase in gender diversity. The reverse result is obtained in more traditional industries. Overall, findings do not point to sizeable productivity-wage gaps associated with educational and gender diversity. Age diversity, on the opposite, is generally found to decrease firm's profitability.

Belgium is no exception regarding the labour market trends that affect diversity (ageing, increase in education levels and female labour market participation) in most OECD countries. Our estimations for Belgium suggest that the effects of these changes are also similar to those found in other economies. Results are notably in line with those obtained for Denmark by Parrotta et al. (2012a) showing a negative effect of demographic diversity (age, gender and ethnicity) and a positive one of educational diversity. Also Navon (2009) finds a positive effect of education diversity in Israel. Negative effects of age diversity are also in line with those found for the U.S. at company level by Hamilton et al. (2004), Kurtulus (2011) and Leonard and Levine (2013). The latter also find insignificant (or no substantial) evidence of the impact of gender diversity on sales, which is similar to our results for profits that do not account for the knowledge intensity of firms (gender diversity is significant in high-tech/knowledge intensive sectors). Our findings only contrast with those of Ilmakunnas and Ilmakunnas (2011) for Finland who show a positive effect of age diversity and a negative one of educational diversity.

How can these findings be interpreted? Results from our benchmark specification showing that educational (age and gender) diversity improves (hamper) firm productivity are consistent with the theoretical predictions of Lazear (1999) and Jehn et al. (1999) highlighting that diversity benefits productivity if the gains of a more diverse workforce in terms of complementary skills and information sets outweigh additional costs related to communication and conflicts. Moreover, they argue that this condition is unlikely to be satisfied for demographic diversity (heterogeneity in terms of e.g. age and gender) but may well be fulfilled for educational (i.e. task related) heterogeneity. In line with our results, they indeed suggest that mutual learning and collaboration among workers with different educational backgrounds may be sufficient to enhance efficiency. Results for gender and age diversity are more in line with the conclusions of the organizational literature (see e.g. Pfeffer, 1985), which emphasize the importance of social similarity (notably in terms of gender and age) to stimulate interaction, communication and cohesion among the workforce.

Interaction effects between gender diversity and the technological/knowledge environment of firms can be reconciled with the predictions of Prat (2002) and Jehn et al. (1999). The latter argue that the benefits of diversity are more likely to exceed the costs when the work environment is predominantly characterized by complex (rather than routine) tasks, negative complementarities (i.e.

workers' actions are substitutes in the firm's payoff function) and innovative (rather than functional) output. Given that these features are more likely to be encountered in high-tech/knowledge intensive sectors than in more traditional industries, they may contribute to the explanation of our results. Although our approach differs from Kurtulus (2011) in that we look at diversity effects in different sectors while Kurtulus assesses the impact of diversity in different occupational groups (finance, marketing, operations, etc) within the same establishment, our findings are analogue to Kurtulus' observation that "it is evident that the impact of worker dissimilarity on worker performance is quite different for workers in different occupations".

Akerlof and Kranton (2000)'s model, introducing the concept of identity into an economic model of behavior, may also explain why productivity effects of gender diversity differ across environments with varying technological/knowledge intensity. The authors argue that gender diversity may negatively affect firm performance, especially if men constitute a socially 'dominant' group (Haile, 2012). Given that the workforce is less gender-balanced (see Table 1) and the environment potentially more 'macho' in traditional companies (e.g. construction) than in high-tech/knowledge intensive firms, their arguments appear to be in line with our results. Empirical findings are also consistent with the observation that high-tech/knowledge intensive sectors increasingly rely on inter-personal or 'soft' skills (that may be more effectively provided by women) and generally require less physical stamina than traditional firms, e.g. construction companies (Arun and Arun, 2002; Webster, 2007).

Overall, our results regarding the impact of gender and educational diversity on the productivity-wage gap suggest that gains and losses associated with diversity are shared 'competitively' between workers and firms so that profits remain unaffected. In contrast, firm profitability is found to depend negatively on age diversity. According to Cataldi et al. (2012), older (younger) workers tend to be 'over-paid' ('under-paid') in Belgian private sector firms. Hence, the negative effect of age diversity on profitability is likely to derive from the fact that: i) increases in age diversity are essentially the consequence of an aging workforce, and ii) the 'over-payment' of older workers may outweigh the 'underpayment' of younger workers (as suggested by Cataldi et al., 2011).

Finally, our findings suggest that the effects of diversity do not substantially vary with firm size. This echoes the estimates by Kurtulus (2011), who measures similar diversity effects of gender and education across different firm sizes. By contrast, the study finds that the negative age effect is stronger in smaller units, a result that could be specific to the US retail firm analysed by Kurtulus (firms of different sizes also appear to differ with respect to diversity effects of race, occupational functions and performance dissimilarity, i.e. variables that we have not focused on in this paper).

One way to interpret our estimates according to firm size is that our indicators of firm-level diversity are relatively good proxies for the social interactions within smaller units in the firm (such as divisions, departments or teams). This could be the case if in many firms the diversity of smaller units reflects the overall diversity of firm's entire personnel. It should, however, be noted that firm size might be correlated with other features that could influence the relationship between diversity and productivity. For instance, bigger firms may have more efficient HR departments leading better job matches, more flexibility to transfer workers from one job to another in case of disputes and more resources to implement diversity management policies. While this paper provides evidence that diversity effects are relatively similar in small and big firms, future micro-level studies are needed to corroborate whether the size of a production unit affects the difference between team- and firm-level diversity. On any account, our results suggest that firm-level diversity is an important explanatory variable for firm productivity independent of firm size.

While diversity is thought to be beneficial in much of the literature in HRM, our findings suggest that in certain cases diversity may be detrimental for both companies and workers. Moreover, consequences of diversity are found to substantially depend on the firm's environment: production in high-tech/knowledge intensive sectors is more likely to benefit from gender diversity than those in more traditional industries. Accordingly, the latter could learn from best practices implemented in the former to make gender diversity work. More generally, personnel measures aimed at improving the impact of age diversity on economic outcomes deserve special attention. Our estimates indeed highlight that the size of the effects associated with diversity (in terms of age, but also gender and education) is substantial and effective diversity management thus remains crucial for a firm's success.

References

- Akerberg, Daniel, Kevin Caves and Garth Frazer. 2006. "Structural Identification of Production Functions." *MPRA Paper* No. 38349. Munich: University Library of Munich.
- Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715-53.
- Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3): 762-800.
- Arellano, Manuel, and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-components Models." *Journal of Econometrics* 68(1): 29-51.

- Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58(2): 277-97.
- Arun, Shoba, and Thankom Arun. 2002. "ICTs, Gender and Development: Women in Software Production in Kerala." *Journal of International Development* 14(1): 39-50.
- Aubert, Patrick, and Bruno Crépon. 2003. "Age, salaire et productivité: La productivité des salariés décline-t-elle en fin de carrière?" *Economie et statistique* 368: 95-119.
- Bandura, Albert. 1997. *Self-efficacy: The Exercise of Control*. New York: Freeman.
- Barrington, Linda, and Kenneth R. Troske. 2001. "Workforce Diversity and Productivity: An Analysis of Employer-Employee Match Data." Economics Program Working Paper No. 01-02. Ottawa: The Conference Board of Canada (Economics Program).
- Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Blundell, Richard, and Stephan Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87(1): 115-43.
- Carrington, William, Kristin McCue, and Brooks Pierce. 2000. "Using Establishment Size to Measure the Impact of Title VII and Affirmative Action." *Journal of Human Resources* 35(3): 503-23.
- Cataldi, Alessandra, Stephan Kampelmann, and François Rycx. 2011. "Productivity-Wage Gaps Among Age Groups: Does the ICT Environment Matter?" *De Economist* 159(2): 193-221.
- Cataldi, Alessandra, Stephan Kampelmann, and François Rycx. 2012. "Does It Pay To Be Productive? The Case of Age Groups." *International Journal of Manpower* 33(3): 264-83.
- Chay, Kenneth. 1998. "The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Opportunity Act of 1972." *Industrial and Labor Relations Review* 51(4): 608-32.
- Choi, Jin N. 2007. "Group Composition and Employee Creative Behaviour in a Korean Electronics Company: Distinct Effects of Relational Demography and Group Diversity." *Journal of Occupational and Organizational Psychology* 80(2): 213-34.
- Cox, Taylor. 1993. *Cultural Diversity in Organizations: Theory, Research, and Practice*. Berrett-Koehler.
- Cox, Taylor H., and Stacy Blake. 1991. "Managing Cultural Diversity: Implications for Organizational Competitiveness." *Academy of Management Executive* 5(3): 45-56.
- Czaja, Sara J., and Joseph Sharit. 1998. "Age Differences in Attitudes Toward Computers." *The Journals of Gerontology: Series B* 53B(5): 329-40.

- Dalton, Hugh. 1920. "The Measurement of the Inequality of Incomes." *Economic Journal* 30(119): 348-61.
- Demunter, Christophe. (2000), "Structure and Distribution of Earnings Survey: Analysis 1995." Working Paper. Brussels: Statistics Belgium.
- Eurostat. 2012. *High-tech statistics – Statistics explained*. Luxembourg: Eurostat.
- Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations* 7(2): 117-40.
- Fiegenbaum, Avi, and Adeel Karnani. 1991. "Output Flexibility - A Competitive Advantage for Small Frms." *Strategic Management Journal* 12(2): 101-14.
- Gini, Corrado. 1912. *Variabilità e mutabilità (Variability and Mutability)*. Bologna: Cuppini.
- Giuliano, Laura, David Levine, and Jonathan Leonard. 2011. "Racial Bias in the Manager-Employee Relationship. An Analysis of Quits, Dismissals, and Promotions at a Large Retail Firm." *Journal of Human Resources* 46(1): 26-52.
- Göbel, Christian, and Thomas Zwick. 2012. "Age and Productivity: Sector Differences." *De Economist* 160 (1): 35-57.
- Grund, Christian, and Niels Westergaard-Nielsen. 2008. "Age Structure of the Workforce and Firm Performance." *International Journal of Manpower* 29(5): 410-22.
- Haile, Getinet A. 2012. "Unhappy Working with Men? Workplace Gender Diversity and Employee Job-related Well-being in Britain: A WERS2004 Based Analysis." *Labour Economics* 19(3): 329-50.
- Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2004. "Diversity and Productivity in Production Teams." Working Paper. St. Louis: Olin Business School
- Hansen, Lars P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50(4): 1029-54.
- Harrison, David A., and Katherine J. Klein. 2007. "What's the Difference? Diversity, Constructs as Separation, Variety or Disparity in Organizations." *Academy of Management Review* 32(4): 1199-228.
- Hellerstein, Judith K., and David Neumark. 1995. "Are Earnings Profiles Steeper than Productivity Profiles? Evidence from Israeli Firm Data." *Journal of Human Resources* 30(1): 89-112.
- Hellerstein, Judith K., and David Neumark. 2004. "Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-employee Data Set." Working Paper No. 10365. Cambridge, MA: National Bureau of Economic Research.
- Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. "Wages, Productivity, and Worker Characteristics: Evidence from Plant-level Production Functions and Wage Equations." *Journal of Labor Economics* 17(3): 409-46.

- Holzer, Harry. 1998. "Why Do Small Establishments Hire Fewer Blacks Than Larger Ones?" *Journal of Human Resources* 33(4): 896-914.
- Hoogendoorn, Sander, Hessel Oosterbeek, and Mirjam Van Praag. 2013. "The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment." *Management Science*, forthcoming.
- Horwitz, Sujin K. and Irwin B. Horwitz. 2007. "The Effects of Team Diversity on Team Outcomes: A Meta-analytic Review of Team Demography." *Journal of Management* 33(6): 987-1015.
- Hoyer, William J., and Amy E. Lincourt. 1998. "Ageing and the Development of Learning." In *Handbook of Implicit Learning*, edited by Michael A. Stadler and Peter A. Frensch, pp. 445-70. California: Sage Publications.
- Ilmakunnas, Pekka, and Seija Ilmakunnas. 2011. "Diversity at the Workplace: Whom Does it Benefit?" *De Economist* 159(2): 223-55.
- 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(2): 247-85.
- Ilmakunnas, Pekka, Jan van Ours, Vegard Skirbekk and Matthias Weiss. 2010. "Age and Productivity". In Pietro Garibaldi, Joaquim Oliveira Martins, and Jan van Our (eds.). *Ageing, Health, and Productivity: The Economics of Increased Life Expectancy*. Oxford University Press.
- Jackson, Susan, Aparna Joshi and Niclas Erhardt. 2003. "Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications." *Journal of Management* 29(6): 801-30.
- Jehn, Karen. 1995. "A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict." *Administrative Science Quarterly* 40(4): 256-82.
- Jehn, Karen A., Gregory B. Northcraft and Margaret A. Neale. 1999. "Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups." *Administrative Science Quarterly* 44(4): 741-63.
- Kampelmann, Stephan, and François Rycx. 2012. "Are Occupations Paid What They Are Worth? An Econometric Study of Occupational Wage Inequality and Productivity". *De Economist* 160(3): 257-288.
- Konrad, Alison M., and Frank Linnehan. 1995. "Formalized HRM Structures: Coordinating Equal Employment Opportunity or Concealing Organizational Practices." *Academy of Management Journal* 3(38): 787-820.

- Kurtulus, Fidan A. 2011. "What Types of Diversity Benefit Workers? Empirical Evidence on the Effects of Co-Worker Dissimilarity on the Performance of Employees." *Industrial Relations* 50(4): 678-712.
- Kurtulus, Fidan A. 2012. "Affirmative Action and the Occupational Advancement of Minorities and Women During 1973-2003." *Industrial Relations* 51(2): 213-46.
- Lazear, Edward P. 1999. "Globalisation and the Market for Team-Mates." *Economic Journal* 109(454): C15-40.
- Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89(5): 841-64.
- Lee, Cynthia L., and Jiing-Lih Farh. 2004. "Joint Effects of Group Efficacy and Gender Diversity on Group Cohesion and Performance." *Applied Psychology* 53(1): 136-54.
- Leonard, Jonathan, and David Levine. 2006. "The Effect of Diversity on Turnover: A Very Large Case Study." *Industrial and Labor Relations Review* 39(4): 457-572.
- Leonard, Jonathan, and David Levine. 2013. "Diversity, Discrimination, and Performance." *Review of Economics and Statistics* forthcoming.
- Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70(2): 317-41.
- Levy, Margi, and Philip Powell. 1998. "SME Flexibility and the Role of Information Systems." *Journal of Small Business Economics* 11(2): 183-96.
- Marino, Marianna, Pierpaolo Parrotta, and Dario Pozzoli. 2012. "Does Labor Diversity Promote Entrepreneurship?" *Economics Letters* 116(1): 15-9.
- Navon, Guy. 2009. "Human Capital Spillovers in the Workplace: Labor Diversity and Productivity." MPRA Paper No. 7741. Munich: University Library of Munich.
- O'Mahony, Mary, and Bart van Ark. 2003. *EU Productivity and Competitiveness: An industry Perspective. Can Europe resume the catching-up process?* Luxembourg: European Commission.
- OECD. 2009. *Measuring Capital*. Paris: OECD Publishing.
- Olley, Steven G., and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64(6): 1263-97.
- Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2012a. "Does Labor Diversity Affect Firm Productivity?" Discussion Paper No. 6973. Bonn: Institute for the Study of Labor (IZA).
- Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2012b. "The Nexus between Labor Diversity and Firm's Innovation." Discussion Paper No. 6972. Bonn: Institute for the Study of Labor (IZA).

- Pelled, Lisa H., Kathleen M. Eisenhardt, and Katherine R. Xin. 1999. "Exploring the Black Box: An Analysis of Work Group Diversity, Conflict, and Performance." *Administrative Science Quarterly* 44(1): 1-28.
- Petrin, Amil, Brian P. Poi, and James Levinsohn. 2004. "Production Function estimation in Stata Using Inputs to Control for Unobservables." *Stata Journal* 4(2): 113-23.
- Pfeffer, Jeffrey. 1985. "Organizational Demography: Implications for Management." *California Management Review* 28(1): 67-81.
- Prat, Andrea. 2002. "Should a Team Be Homogeneous?" *European Economic Review* 46(7): 1187-207.
- Pull, Kerstin, Birgit Pferdmenges and Uschi Backes-Gellner. 2012. "Knowledge Production Process, Diversity Type and Group Interaction as Moderators of the Diversity-Performance-Link: An Analysis of University Research Groups." Working Papers No. 158. Zurich: Institute for Strategy and Business Economics (University of Zurich).
- Roberge, Marie-Elène, and Rolf van Dick. 2010. "Recognizing the Benefits of Diversity: When and How Diversity Increase Group Performance?" *Human Resource Management Review* 20(4): 295-308.
- Riordan, Christine, and Julie Wayne. 2008. "A Review and Explanation of Demographic Similarity Measures Used to Assess Relational Demography Within Groups." *Organizational Research Methods* 11(3): 562-92.
- Roodman, David. 2009. "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of Economics and Statistics* 71(1): 135-58.
- Rynes, Sara, and Benson Rosen. 1995. "A Field Survey of Factors Affecting the Adoption and Perceived Success of Diversity Training." *Personnel Psychology* 48(2): 247-70.
- Skirbekk, Vegard. 2003. "Age and Individual Productivity: A Literature Survey." Technical Report No. 2003-028. Rostock: Max Planck Institute for Demographic Research.
- Stahl, Günter, Martha Maznevski, Andreas Voigt and Karsten Jonsen. 2010. "Unraveling the Effects of Cultural Diversity in Teams: A Meta-analysis of Research on Multicultural Work Groups." *Journal of International Business Studies* 41(4): 690-709.
- Stewart, Greg. 2006. "A Meta-analytic Review of Relationships Between Team Design Features and Team Performance." *Journal of Management* 32(1): 29-55.
- van Beveren, Ilke. 2012. "Total Factor Productivity Estimation: A Practical Review." *Journal of Economic Surveys* 26(1): 98-128.

- van Knippenberg, Daan, Jeremy F. Dawson, Michael A. West, and Astrid C. Homan. 2011". Diversity Faultlines, Shared Objectives, and Top Management Team Performance." *Human Relations* 64(3): 307-36.
- van Ours, Jan C., and Lenny Stoeldraijer. 2011. "Age, Wage and Productivity in Dutch Manufacturing." *De Economist* 159(2): 113-37.
- Vandenberghe, Vincent. 2011. "Firm-level Evidence on Gender Wage Discrimination." *Labour* 25(3): 330-49.
- Vandenberghe, Vincent. 2013. "Are Firms Willing to Employ a Greying and Feminizing Workforce?" *Labour Economics* 22: 30-42.
- Webster, Juliet. 2007. "Diversity Management in the ICT Industry. Challenges and Issues for Social Dialogue." Report prepared for Union Network International. Brussels: European Commission.
- Yang, Yang, and Alison M. Konrad. 2011. "Understanding Diversity Management Practices: Implications of Institutional Theory and Resource-based Theory." *Group and Organization Management* 36(1): 6-38.

Table 1: Descriptive statistics at the firm level (1999-2006)

Variables	All firms		HT/KIS firms		Non-HT/KIS firms	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Hourly wage (€ ^a)	17.14	5.39	18.38	5.68	16.64	5.18
Value-added per hour (€ ^a)	61.06	458.61	64.49	239.10	59.71	520.20
Average age (years)	38.42	4.19	37.45	4.35	38.80	4.07
Standard deviation of age	9.33	1.82	9.01	2.01	9.45	1.73
Age dissimilarity index	12.61	2.52	12.16	2.77	12.79	2.39
Average education (years)	11.44	1.76	12.32	1.79	11.09	1.62
Standard deviation of education	1.90	0.84	1.79	0.77	1.94	0.86
Education dissimilarity index	2.54	1.15	2.40	1.05	2.60	1.18
Women (%)	0.27	0.24	0.33	0.25	0.24	0.23
Standard deviation of gender	0.35	0.15	0.38	0.14	0.34	0.16
Gender dissimilarity index	0.46	0.22	0.51	0.20	0.45	0.22
Workers with tenure >= 10 years (%)	0.39	0.24	0.33	0.25	0.42	0.24
White-collar workers (%)	0.45	0.34	0.62	0.36	0.39	0.31
Part-time (< 30h/week, %)	0.02	0.07	0.02	0.06	0.02	0.07
Fixed-term employment contacts (%)	0.04	0.10	0.05	0.12	0.04	0.09
Sector (%)						
Mining and quarrying (C)	0.01	0.09	0.00	0.00	0.01	0.11
Manufacturing (D)	0.57	0.49	0.53	0.50	0.59	0.49
Electricity, gas and water supply (E)	0.00	0.06	0.00	0.00	0.01	0.07
Construction (F)	0.10	0.29	0.00	0.00	0.13	0.34
Wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods (G)	0.12	0.33	0.00	0.00	0.17	0.37
Hotels and restaurant (H)	0.02	0.13	0.00	0.00	0.02	0.16
Transport, storage and communication (I)	0.06	0.24	0.05	0.21	0.07	0.25
Financial intermediation (J)	0.01	0.11	0.05	0.21	0.00	0.00
Real estate, renting and business activities (K)	0.11	0.31	0.38	0.49	0.00	0.01
Capital stock (€ ^a)	244,287	2,117,000	489,790	3,946,000	147,644	292,979
Investments (€ ^a)	18,543	254,447	40,205	476,648	10,019	24,221
Size of the firm (number of full-time equivalent workers)	131.85	336.37	203.76	551.76	116.63	267.12
Number of observations		7,463		2,108		5,355
Number of firms ^b		2,431		679		1,778

Notes: ^a At 2006 constant prices. ^b The sum of HT/KIS and non-HT/KIS firms exceeds the total number of firms due to a small number of them changing category during the observation period.

Table 2: Estimation results for the entire sample

	GMM-SYS						LP	
	Value added per hour worked (ln)		Mean wage per hour worked (ln)		Value added-wage gap (ln)		Value added per hour worked (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. dev. age	-0.022*** (0.008)		-0.010*** (0.004)		-0.013* (0.007)		-0.007** (0.003)	
Age dissimilarity		-0.016*** (0.006)		-0.007*** (0.003)		-0.009* (0.005)		-0.005* (0.003)
Std. dev. education	0.009 (0.015)		0.017** (0.007)		-0.008 (0.013)		0.032*** (0.008)	
Education dissimilarity		0.007 (0.011)		0.012** (0.005)		-0.005 (0.010)		0.024*** (0.006)
Std. dev. gender	-0.260** (0.102)		-0.140** (0.055)		-0.120 (0.094)		-0.113* (0.064)	
Gender dissimilarity		-0.176** (0.076)		-0.097** (0.041)		-0.079 (0.069)		-0.075* (0.039)
Average age	0.011*** (0.003)	0.011*** (0.003)	0.009*** (0.001)	0.009*** (0.001)	0.002 (0.003)	0.002 (0.003)	0.010*** (0.002)	0.010*** (0.002)
Average education	0.077*** (0.007)	0.077*** (0.007)	0.046*** (0.003)	0.046*** (0.003)	0.032*** (0.006)	0.032*** (0.006)	0.075*** (0.006)	0.075*** (0.005)
Hansen over-identification test, <i>p</i> -value	0.765	0.767	0.152	0.172	0.487	0.480		
Arellano-Bond test for AR(2), <i>p</i> -value	0.123	0.124	0.370	0.356	0.560	0.561		
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,461	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments.

Table 3: Estimation results for different technological/knowledge environments (HT/KIS nomenclature)

	GMM-SYS						LP	
	Value added per hour worked (ln)		Mean wage per hour worked (ln)		Value added-wage gap (ln)		Value added per hour worked (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. dev. age	-0.022** (0.010)		-0.011** (0.005)		-0.011 (0.009)		-0.001 (0.005)	
Age dissimilarity		-0.017** (0.007)		-0.007** (0.003)		-0.009 (0.007)		-0.001 (0.003)
Std. dev. education	0.011 (0.022)		0.001 (0.010)		0.010 (0.021)		0.025*** (0.009)	
Education dissimilarity		0.006 (0.016)		0.001 (0.007)		0.019*** (0.007)		0.019*** (0.007)
Std. dev. gender	-0.327** (0.136)		-0.172** (0.068)		-0.155 (0.123)		-0.194*** (0.069)	
Gender dissimilarity		-0.230** (0.100)		-0.119** (0.050)		-0.112 (0.089)		-0.142*** (0.039)
Std. dev. age*HT/KIS	0.011 (0.026)		0.006 (0.012)		0.005 (0.024)		-0.014 (0.009)	
Age dissimilarity*HT/KIS		0.011 (0.019)		0.004 (0.009)		0.007 (0.017)		-0.010 (0.007)
Std. dev. education*HT/KIS	-0.007 (0.056)		0.039* (0.022)		-0.047 (0.049)		0.033 (0.024)	
Education dissimilarity*HT/KIS		-0.001 (0.040)		0.026 (0.016)		-0.028 (0.034)		0.023 (0.017)
Std. dev. gender*HT/KIS	0.716* (0.398)		0.174 (0.139)		0.542 (0.361)		0.343** (0.147)	
Gender dissimilarity*HT/KIS		0.527* (0.283)		0.121 (0.102)		0.406 (0.255)		0.261*** (0.091)
Average age	-0.005 (0.016)	-0.003 (0.016)	0.003 (0.008)	0.003 (0.008)	-0.008 (0.014)	-0.006 (0.014)	0.008*** (0.003)	0.008*** (0.003)
Average education	0.055 (0.043)	0.048 (0.042)	0.002 (0.020)	0.002 (0.019)	0.053 (0.040)	0.046 (0.039)	0.063*** (0.005)	0.064*** (0.007)
Average age*HT/KIS	0.035* (0.021)	0.034 (0.021)	-0.001 (0.010)	-0.000 (0.010)	0.036** (0.018)	0.034* (0.018)	0.006 (0.004)	0.006 (0.004)
Average education*HT/KIS	0.066	0.073	0.064**	0.062**	0.002	0.011	0.037***	0.037***

HT/KIS	(0.064)	(0.064)	(0.029)	(0.029)	(0.053)	(0.052)	(0.010)	(0.013)
	-2.552***	-2.635***	-0.934**	-0.896**	-1.618*	-1.739**	-0.691***	-0.689***
	(0.981)	(0.972)	(0.453)	(0.452)	(0.868)	(0.860)	(0.213)	(0.212)
Hansen over-identification test, <i>p</i> -value	0.177	0.192	0.055	0.065	0.334	0.306		
Arellano-Bond test for AR(2), <i>p</i> -value	0.117	0.116	0.458	0.442	0.499	0.502		
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,461	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments. HT/KIS = 1 if the firm belongs to a high-medium tech/knowledge intensive sector, according to the taxonomy developed by Eurostat (2012).

Table 4: Estimation results including interaction effects with firm size

	GMM-SYS						LP	
	Value added per hour worked (ln)		Mean wage per hour worked (ln)		Value added-wage gap (ln)		Value added per hour worked (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. dev. age	-0.023** (0.011)		-0.014*** (0.005)		-0.010 (0.009)		-0.0040 (0.005)	
Age dissimilarity		-0.017** (0.008)		-0.001** (0.004)		-0.007 (0.006)		-0.002 (0.003)
Std. dev. education	-0.012 (0.022)		0.020** (0.010)		-0.032* (0.018)		0.027*** (0.010)	
Education dissimilarity		-0.007 (0.016)		0.015** (0.007)		-0.022* (0.013)		0.021*** (0.007)
Std. dev. gender	-0.261** (0.113)		-0.061 (0.063)		-0.200** (0.100)		-0.194*** (0.075)	
Gender dissimilarity		-0.176** (0.088)		-0.039 (0.048)		-0.137** (0.075)		-0.136** (0.058)
Std. dev. age * firm size	-0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		-0.000 (0.000)	
Age dissimilarity * firm size		-0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		-0.000 (0.000)
Std. dev. education * firm size	0.000 (0.000)		0.000 (0.000)		0.000* (0.000)		0.000** (0.000)	
Education dissimilarity * firm size		0.000 (0.000)		0.000 (0.000)		0.000* (0.000)		0.000* (0.000)
Std. dev. gender * firm size	0.000 (0.000)		0.000 (0.000)		0.000* (0.000)		0.000** (0.000)	
Gender dissimilarity * firm size		0.000 (0.000)		-0.000 (0.000)		0.000* (0.000)		0.000* (0.000)
Average age	0.018***	0.018***	0.013***	0.013***	0.005	0.005	0.014***	0.014***

	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Average education	0.058***	0.058***	0.045***	0.045***	0.013*	0.013*	0.068***	0.068***
	(0.009)	(0.009)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)
Average age * firm size	-0.000**	-0.000**	-0.000***	-0.000	-0.000	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Average education * firm size	0.000***	0.000***	0.000***	0.000	0.000***	0.000**	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hansen over-identification test, <i>p</i> -value	0.404	0.478	0.002	0.002	0.692	0.713		
Arellano-Bond test for AR(2), <i>p</i> -value	0.159	0.164	0.584	0.569	0.693	0.698		
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,461	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments.

Table 5: Estimation results including interaction effects for big firms, i.e. firms employing more than 500 workers

	GMM-SYS						LP	
	Value added per hour worked (ln)		Mean wage per hour worked (ln)		Value added-wage gap (ln)		Value added per hour worked (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. dev. age	-0.017*		-0.004		-0.013		-0.005	
	(0.009)		(0.004)		(0.008)		(0.004)	
Age dissimilarity		-0.012*		-0.000		-0.009		-0.003
		(0.007)		(0.001)		0.006)		(0.003)
Std. dev. education	0.006		0.021***		-0.015		0.035***	
	(0.013)		(0.007)		(0.012)		(0.008)	
Education dissimilarity		0.005		0.001		-0.011		0.026***
		(0.010)		(0.002)		(0.009)		(0.006)
Std. dev. gender	-0.225**		-0.101*		-0.124		-0.097	
	(0.112)		(0.060)		(0.099)		(0.066)	
Gender dissimilarity		-0.146*		0.027		-0.076		-0.066**
		(0.084)		(0.017)		(0.074)		(0.031)
Std. dev. age * size > 500	-0.025		-0.025		0.001		-0.008	
	(0.040)		(0.017)		(0.035)		(0.017)	
Age dissimilarity * size > 500		-0.022		-0.006*		-0.000		-0.005
		(0.030)		(0.003)		(0.026)		(0.011)
Std. dev. education * size > 500	0.047		0.005		0.042		0.036	
	(0.039)		(0.019)		(0.034)		(0.027)	
Education dissimilarity * size > 500		0.030		-0.004		0.029		0.026
		(0.028)		(0.005)		(0.024)		(0.017)
Std. dev. gender * size > 500	0.155		-0.108		0.263		0.200	
	(0.453)		(0.198)		(0.412)		(0.198)	
Gender dissimilarity * size > 500		0.062		-0.097***		0.175		0.137
		(0.325)		(0.037)		(0.286)		(0.140)
Average age	0.014***	0.014***	0.010***	0.001	0.004	0.004	0.011***	0.011***

	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)
Average education	0.074***	0.074***	0.049***	-0.001	0.026***	0.026***	0.075***	0.075***
	(0.006)	(0.006)	(0.003)	(0.002)	(0.004)	(0.004)	(0.006)	(0.007)
Average age * size > 500	-0.006	-0.005	0.004	-0.000	-0.009	-0.009	-0.003	-0.003
	(0.010)	(0.010)	(0.005)	(0.002)	(0.008)	(0.008)	(0.008)	(0.007)
Average education * size > 500	0.047**	0.047**	0.016**	0.003	0.030*	0.030*	0.046**	0.046**
	(0.020)	(0.020)	(0.008)	(0.003)	(0.016)	(0.016)	(0.022)	(0.019)
Size > 500	-0.108	-0.043	0.002	0.100	-0.110	-0.097	-0.414	-0.413
	(0.475)	(0.486)	(0.242)	(0.091)	(0.431)	(0.429)	(0.416)	(0.361)
Hansen over-identification test, <i>p</i> -value	0.736	0.748	0.003	0.004	0.681	0.658		
Arellano-Bond test for AR(2), <i>p</i> -value	0.114	0.115	0.692	0.684	0.655	0.650		
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,461	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments.

Appendix 1: Diversity indicators

1. The absolute standard deviation

This indicator refers to “the standard, or mean square, deviation from the arithmetic mean, i.e., the square root of the arithmetic average of the squares of deviations from the arithmetic mean” (Dalton, 1920). For the N employees i in firm j , this can be expressed as:

$$SD_j = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (\text{A.1})$$

This indicator is also referred to as ‘absolute standard deviation’ (the ‘relative standard deviation’ is obtained by dividing SD_j by the arithmetic mean). In the paper we follow conventional use and employ the absolute standard deviation, which means that standard deviations of variables with different units cannot be compared directly.

2. The dissimilarity index

The Euclidean distance between two points $X_1 = (x_{1,1}, x_{1,2}, \dots, x_{1,n})$ and $X_2 = (x_{2,1}, x_{2,2}; \dots; x_{2,n})$ in a Euclidean n -space is defined as:

$$\overline{X_1 X_2} = \sqrt{(x_{1,1} - x_{2,1})^2 + \dots + (x_{1,n} - x_{2,n})^2} = \sqrt{\sum_{d=1}^n (x_{1,d} - x_{2,d})^2} \quad (\text{A.2})$$

Considering the diversity between two employees in terms of a given characteristic is analogous to computing their Euclidean distance in a one-dimensional space. The Euclidean distance between the two employees is equal to the absolute difference between their coordinates:

$$\overline{X_1 X_2} = \sqrt{(x_1 - x_2)^2} = |x_1 - x_2| \quad (\text{A.3})$$

The dissimilarity index for employee i can be interpreted as the square root of the arithmetic mean of all Euclidean distances in a one-dimensional space. Formally, this can be written as follows:

$$DI_{i,j} = \sqrt{\frac{\sum_{k=1}^{N_j} (x_{i,j} - x_{k,j})^2}{N_j}} = \sqrt{(x_{i,j} - \bar{x}_j)^2 + \text{VAR}(x_j)} \quad (\text{A.4})$$

where N_j is the number of individuals in firm j . In other words, $DI_{i,j}$ is the square root of the average Euclidean distance of the worker i with respect to all co-workers in the firm j . Note that Equation A.4 includes the trivial distance of the focal worker to itself, which is by definition always zero. A convenient feature of $DI_{i,j}$ is that it can be rewritten in terms of the difference between the worker i and the arithmetic mean, on the one hand, and the variance within firm j on the other hand (see the development in Equation A.4). This manipulation reveals the relationship between the $DI_{i,j}$ and the standard deviation: if $x_{i,j}$ corresponds to age and if employee i has

exactly the average age in firm j , then $DI_{i,j}$ of this individual is equal to the standard deviation of age in firm j . If the age of the individual is below or above the arithmetic mean, then his or her $DI_{i,j}$ is higher than the standard deviation. The maximum $DI_{i,j}$ corresponds to either the oldest or youngest individual in firm j (depending on which one is further away from the mean). The dissimilarity index can be interpreted as capturing ‘relational demography’, i.e. how separate the individual i is from the other employees in firm j and is widely used in the literature on diversity (Harrison and Klein, 2007; Riordan and Wayne, 2008).

Computing the arithmetic mean of the individual dissimilarity indices yields the firm-level dissimilarity index used by Ilmakunnas and Ilmakunnas (2011). The main difference between the (absolute) standard deviation and the firm-level dissimilarity index is that the latter can be decomposed into individual-level dissimilarities because the firm-level index is simply the arithmetic mean over all N_j individuals in firm j :

$$DI_j = \frac{\sum_{i=1}^{N_j} \sqrt{\frac{\sum_{k=1}^{N_j} (x_{i,j} - x_{k,j})^2}{N_j}}}{N_j} = \frac{\sum_{i=1}^{N_j} \sum_{k=1}^{N_j} |x_{i,j} - x_{k,j}|}{N_j^2} \quad (\text{A.5})$$

The firm-level dissimilarity index DI_j in Equation A.5 is thus the average Euclidean distance between all possible dyads in a one-dimensional Euclidean space, including the trivial distances of the points with themselves. Rewriting the expression in absolute values shows that DI_j corresponds to the ‘absolute mean difference’, a measure of inequality defined by Corrado Gini in 1912.

3. Alternative gender diversity index

In addition to the standard deviation and dissimilarity index of gender, in line with standard practice in the literature (see e.g. Hoogendoorn et al., 2013), we also computed the following alternative dissimilarity index GDI_j based on the shares of women and men in firm j , $s_{w,j}$ and $s_{m,j}$, respectively. It can be written as follows:

$$GDI_j = s_{w,j}s_{m,j} = s_{w,j}(1 - s_{w,j}) \quad (\text{A.6})$$

4. Characteristics of the different indicators

Both the SD_j and the DI_j are not unit-free, i.e. results for attributes that are measured in different units cannot be directly compared. Consider the diversity attribute x ranging from a lower bound l to an upper bound u . For the case of age, this range could be equal to the interval [16;64]. The minimum of both the standard deviation and the dissimilarity index is equal to zero and corresponds to the absence of intra-firm variation on attribute x (e.g. all individuals have the same age). The maximum value of the standard deviation is equal to $(u-l)/2$, whereas the maximum of the dissimilarity index is $(u-l)/\sqrt{2}$. As a consequence, the range of the dissimilarity index is slightly bigger compared to the standard deviation. For both indicators the maxima are reached when firm j 's workforce is perfectly polarised into two equal-sized groups with the characteristics u and l , respectively (i.e. a bi-modal distribution with modes equal to the upper and lower bounds of the attribute x). For example, if firm j has 10 employees whose age

ranges from 16 to 64 years, then both SD_j and DI_j peak when half of the employees are 16 and the other half 64 years old: in this case, the SD_j would be equal to 24 $[=(64-16)/2]$ and the DI_j 33.94 $[=(64-16)/\sqrt{2}]$.

Given that both indicators attain their peak under the same conditions and in light of their general conceptual and mathematical similarities, Harrison and Klein (2007) conclude that “there is no particular advantage of either operationalization over the other, save for researcher familiarity with SD ”.

As for the GDI_j , its minimum value corresponds to a firm composed with either only male or only female workers, in which case the indicator equals zero. Its maximum is reached when all employees of firm j are distributed evenly among the two sexes. The indicator then equals 0.25 $(=0.5 \times 0.5)$. As a consequence, the range of the GDI_j is $[0;0.25]$. Given that each firm j is characterised by only one gender share, it is not possible to calculate a standard deviation or dissimilarity index on this variable. The gender standard deviation and dissimilarity index have been calculated with a dummy variable for sex that takes the value of one if the employee is a woman and zero otherwise. The range (and maximum value) of the GDI_j therefore differs with respect to the standard deviation and the dissimilarity index, which are 0.5 $[=(1-0)/2]$ and 0.71 $[=(1-0)/\sqrt{2}]$, respectively. The GDI_j is nevertheless conceptually similar to the other two indicators: its minimum corresponds to the absence of variability in the attribute x and the maximum to a bi-modal distribution.

Appendix 2: Estimates for the entire sample using ‘the share of women times the share of men’ as gender diversity index

	GMM-SYS						LP	
	Value added per hour worked (ln)		Mean wage per hour worked (ln)		Value added-wage gap (ln)		Value added per hour worked (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. dev. age	-0.022*** (0.008)		-0.009*** (0.004)		-0.013* (0.007)		-0.007 (0.005)	
Age dissimilarity		-0.016*** (0.006)		-0.007*** (0.003)		-0.009* (0.005)		-0.005* (0.003)
Std. dev. education	0.008 (0.015)		0.016** (0.007)		-0.008 (0.013)		0.031*** (0.010)	
Education dissimilarity		0.006 (0.011)		0.012** (0.005)		-0.005 (0.010)		0.023*** (0.006)
Share of women * share of men	-0.390* (0.229)	-0.389* (0.229)	-0.234** (0.115)	-0.230** (0.116)	-0.156 (0.197)	-0.159 (0.197)	-0.160 (0.109)	-0.162* (0.090)
Average age	0.011*** (0.004)	0.011*** (0.004)	0.009*** (0.001)	0.009*** (0.001)	0.002 (0.003)	0.002 (0.003)	0.010*** (0.002)	0.010*** (0.003)
Average education	0.077*** (0.007)	0.077*** (0.007)	0.046*** (0.003)	0.046*** (0.003)	0.031*** (0.006)	0.031*** (0.006)	0.075*** (0.007)	0.075*** (0.007)
Hansen over-identification test, <i>p</i> -value	0.866	0.840	0.117	0.140	0.468	0.461		
Arellano-Bond test for AR(2), <i>p</i> -value	0.131	0.130	0.349	0.343	0.564	0.564		
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,461	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments.

Appendix 3: Estimation results including nonlinearities

	Testing for nonlinearities at the 33rd percentile				Testing for nonlinearities at the 66th percentile				Testing for nonlinearities at 33rd and 66th percentiles			
	GMM-SYS			LP	GMM-SYS			LP	GMM-SYS			LP
	Value-added (1)	Wage (2)	Gap (3)	Value-added (4)	Value-added (5)	Wage (6)	Gap (7)	Value-added (8)	Value-added (9)	Wage (10)	Gap (11)	Value-added (12)
Std. dev. age	-0.022* (0.013)	-0.005 (0.006)	-0.017 (0.011)	-0.006 (0.004)	-0.023** (0.011)	-0.009* (0.005)	-0.014 (0.010)	-0.003 (0.006)	-0.046*** (0.017)	-0.009 (0.007)	-0.037** (0.015)	-0.006 (0.007)
Std. dev. education	0.032 (0.020)	0.003 (0.010)	0.028 (0.018)	0.035*** (0.008)	-0.010 (0.016)	0.013 (0.009)	-0.023 (0.015)	0.042*** (0.012)	-0.035 (0.031)	-0.010 (0.014)	-0.024 (0.030)	0.038* (0.020)
Std. dev. gender	-0.233** (0.109)	-0.145** (0.060)	-0.088 (0.101)	-0.082 (0.058)	-0.278*** (0.099)	-0.180*** (0.057)	-0.098 (0.091)	-0.178*** (0.062)	-0.387*** (0.123)	-0.203*** (0.068)	-0.185 (0.113)	-0.133 (0.084)
Std. dev. age*p33	0.001 (0.003)	-0.002 (0.001)	0.003 (0.002)	0.001 (0.001)					0.010*** (0.003)	-0.001 (0.002)	0.010*** (0.003)	0.001 (0.002)
Std. dev. education*p33	-0.012 (0.012)	0.011* (0.006)	-0.022** (0.010)	0.005 (0.007)					0.017 (0.021)	0.015* (0.009)	0.001 (0.020)	0.003 (0.014)
Std. dev. gender*p33	0.004 (0.060)	-0.012 (0.034)	0.016 (0.052)	-0.076* (0.039)					0.147* (0.075)	0.014 (0.041)	0.132** (0.066)	-0.040 (0.056)
Std. dev. age*p66					0.001 (0.002)	-0.001 (0.001)	0.001 (0.002)	-0.001 (0.001)	0.012** (0.005)	-0.001 (0.002)	0.013*** (0.005)	0.000 (0.003)
Std. dev. education*p66					0.015 (0.009)	0.006 (0.005)	0.009 (0.009)	-0.005 (0.008)	0.033 (0.025)	0.023** (0.011)	0.010 (0.023)	-0.002 (0.014)
Std. dev. gender*p66					0.018 (0.066)	0.034 (0.031)	-0.016 (0.055)	0.080* (0.045)	0.163 (0.105)	0.057 (0.052)	0.106 (0.088)	0.044 (0.069)
Average age	0.011*** (0.003)	0.009*** (0.001)	0.002 (0.003)	0.011*** (0.002)	0.011*** (0.003)	0.009*** (0.001)	0.001 (0.003)	0.011*** (0.002)	0.011*** (0.003)	0.010*** (0.001)	0.002 (0.003)	0.011*** (0.002)
Average education	0.078*** (0.007)	0.047*** (0.004)	0.031*** (0.006)	0.079*** (0.006)	0.079*** (0.008)	0.047*** (0.003)	0.032*** (0.006)	0.079*** (0.007)	0.080*** (0.007)	0.048*** (0.004)	0.032*** (0.006)	0.079*** (0.006)
Hansen test, <i>p</i> -value	0.735	0.069	0.809		0.669	0.339	0.413		0.711	0.176	0.685	
AR(2) test, <i>p</i> -value	0.125	0.339	0.588		0.123	0.355	0.565		0.107	0.374	0.595	
Number of observations	7,463	7,463	7,463	7,463	7,463	7,463	7,463	7,463	7,463	7,463	7,463	7,463
Number of firms	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments. HT/KIS = 1 if the firm belongs to a high-medium tech/knowledge intensive sector, according to the taxonomy developed by Eurostat (2012). p3 (p66) is a dummy variable that takes the value one if the variable is greater than the 33rd percentile (66th percentile). When p33 and p66 are included simultaneously, p33 takes the value one if the variable is greater than the 33rd percentile and smaller than the 66th percentile. The dependent variables are respectively: i) the value-added (i.e. the value added per hour worked (ln)), ii) the wage (i.e. the mean wage per hour worked (ln)), and iii) the gap (i.e. value added-wage gap (ln)).

Appendix 4: Description of HT/KIS, KIA and ICT taxonomies

a) High-medium tech/knowledge intensive sectors (HT/KIS - Eurostat, 2012)

Taxonomy that classifies manufacturing industries (at NACE 2- and/or 3-digit level) according to their degree of technological intensity (primarily assessed through the ratio of R&D expenditures to value added) and services (at NACE 2- digit level) according to their degree of knowledge intensity (i.e. the share of tertiary educated people in the activity)

HT/KIS firms are found in the following sectors: Aerospace (NACE 353); Computers, office machinery (NACE 30); Electronics-communications (NACE 32); Pharmaceuticals (NACE 244); Scientific instruments (NACE 33); Motor vehicles (NACE 34); Electrical machinery (NACE 31); Chemicals (NACE 24); Other transport equipment (NACE 352+354+355); Non-electrical machinery (NACE 29); Water transport (NACE 61); Air transport (NACE 62); Post and telecommunications (NACE 64); Financial intermediation, except insurance and pension funding (NACE 65); Insurance and pension funding, except compulsory social security (NACE 66); Activities auxiliary to financial intermediation (NACE 67); Real estate activities (NACE 70); Renting of machinery and equipment without operator and of personal and household goods (NACE 71); Computer and related activities (NACE 72); Research and development (NACE 73); Other business activities (NACE 74); Education (NACE 80); Health and social work (NACE 85); Recreational, cultural and sporting activities (NACE 92).

Non-HT/KIS firms are found in the following sectors: Rubber and plastic products (NACE 25); Shipbuilding (NACE 351); Other manufacturing (NACE 362 through 366); Non-ferrous metals (NACE 274+2753/54); Non-metallic mineral products (NACE 26); Fabricated metal products (NACE 28); Petroleum refining (NACE 23); Ferrous metals (NACE 271 through 273+2751/52); Paper printing (NACE 21+22); Textile and clothing (NACE 17 through 19); Food, beverages, and tobacco (NACE 15+16); Wood and furniture (NACE 20+361); Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel (NACE 50); Wholesale trade and commission trade, except of motor vehicles and motorcycles (NACE 51); Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods (NACE 52); Hotels and restaurants (NACE 55); Land transport; transport via pipelines (NACE 60); Supporting and auxiliary transport activities; activities of travel agencies (NACE 63); Public administration and defense; compulsory social security (NACE 75); Sewage and refuse disposal, sanitation and similar activities (NACE 90); Activities of membership organization n.e.c. (NACE 91); Other service activities (NACE 93); Private households with employed persons (NACE 95); Extra-territorial organizations and bodies (NACE 99).

b) Knowledge intensive activities (KIA - Eurostat, 2012)

Taxonomy that classifies industries (both manufacturing and services) according to their degree of knowledge intensity (assessed through the share of tertiary educated people at the NACE 2-digit level). An industry is classified as knowledge intensive if tertiary educated persons employed (according ISCED'97, levels 5+6) represent more than 33% of the total employment in that industry.

KIA firms are found in the following sectors: Manufacture of coke, refined petroleum products and nuclear fuel (NACE 23); Manufacture of chemicals and chemical products (NACE 24); Manufacture of office machinery and computers (NACE 30); Manufacture of radio, television and communication equipment and apparatus (NACE 32); Manufacture of medical, precision and optical instruments, watches and clocks (NACE 33); Air transport

(NACE 62); Financial intermediation, except insurance and pension funding (NACE 65); Insurance and pension funding, except compulsory social security (NACE 66); Activities auxiliary to financial intermediation (NACE 67); Computer and related activities (NACE 72); Research and development (NACE 73); Other business activities (NACE 74); Public administration and defence; compulsory social security (NACE75); Education (NACE 80); Health and social work (NACE 85); Activities of membership organizations n.e.c. (NACE 91), Recreational, cultural and sporting activities (NACE 92); Extra-territorial organizations and bodies (NACE 99).

c) Information and communication technology industries (ICT – O’Mahony and van Ark, 2003)

Taxonomy that classifies industries according to their ICT capital intensity at the NACE 3-digit level. It groups industries based on whether they produce ICT goods and services and whether they intensively use ICT or not.

ICT firms are found in the following sectors: Clothing (NACE 18); Printing and publishing (NACE 22); Mechanical engineering (NACE 29); Other electrical machinery and apparatus, except insulated wire (NACE 31); Other instruments, except scientific instruments (NACE 33); Building and repairing of ships and boats (NACE 351); Aircraft and spacecraft (NACE 353); Furniture, miscellaneous manufacturing; recycling (NACE 36-37); Wholesale trade and commission trade, except of motor vehicles and motorcycles (NACE 51); Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods (NACE 52); Financial activities, except insurance and pension funding (NACE 65); Activities auxiliary to financial intermediation (NACE 67); Renting of machinery and equipment (NACE 71); Legal, technical and advertising (NACE 741-743); Office machinery (NACE 30); Insulated wire (NACE 313); Electronic valves and tubes (NACE 321); Telecommunication equipment (NACE 322); Radio and television receivers (NACE 323); Scientific instruments (NACE 331); Communications (NACE 64); Computer and related activities (NACE 72).

Non-ICT firms are found in the following sectors: Quarrying (NACE 14); Food, drink and tobacco (NACE 15-16); Textiles (NACE 17); Leather and footwear (NACE 19); Wood and products of wood and cork (NACE 20); Pulp, paper and paper products (NACE 21); Mineral oil refining, coke and nuclear fuel (NACE 23); Chemicals (NACE 24); Rubbers and plastics (NACE 25); Non-metallic mineral products (NACE 26); Basic metals (NACE 27); Fabricated metal products (NACE 28); Motor vehicles (NACE 34); Construction (NACE 45); Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel (NACE 50); Hotels and restaurants (NACE 55); Inland transport (NACE 60); Water transport (NACE 61); Air transport (NACE 62); Supporting and auxiliary transport activities; activities of travel agencies (NACE 63); Real estate activities (NACE 70); Other business activities (NACE 749).

Appendix 5: Correlation coefficients between HT/KIS, KIA and ICT taxonomies

	HT/KIS	KIA	ICT
HT/KIS	1		
KIA	0.59	1	
ICT	0.49	0.22	1

Appendix 6: Estimates using the KIA taxonomy

	GMM-SYS			LP
	Value added per hour worked (ln)	Mean wage per hour worked (ln)	Value added-wage gap (ln)	Value added-wage gap (ln)
	(1)	(2)	(3)	(4)
Std. dev. age	-0.020** (0.008)	-0.004 (0.004)	-0.016** (0.007)	-0.005 (0.004)
Std. dev. education	0.017 (0.015)	0.013* (0.007)	0.005 (0.014)	0.024*** (0.009)
Std. dev. gender	-0.329*** (0.107)	-0.080 (0.055)	-0.249** (0.103)	-0.137** (0.069)
Std. dev. age *KIA	-0.000 (0.027)	-0.018* (0.010)	0.018 (0.023)	-0.005 (0.012)
Std. dev. education*KIA	-0.021 (0.042)	0.017 (0.020)	-0.038 (0.034)	0.039 (0.024)
Std. dev. gender*KIA	0.696** (0.330)	0.025 (0.140)	0.671** (0.288)	0.133 (0.148)
Average age	0.002 (0.003)	0.008*** (0.001)	-0.006* (0.003)	0.004* (0.002)
Average education	0.063*** (0.008)	0.031*** (0.004)	0.031*** (0.007)	0.059*** (0.006)
Average age *KIA	0.031*** (0.008)	0.007** (0.003)	0.024*** (0.007)	0.024*** (0.005)
Average education*KIA	0.037** (0.014)	0.038*** (0.007)	-0.001 (0.012)	0.051*** (0.010)
KIA	-1.605*** (0.340)	-0.448*** (0.159)	-1.156*** (0.301)	-1.459*** (0.211)
Hansen over-identification test, <i>p</i> -value	0.639	0.001	0.674	
Arellano-Bond test for AR(2), <i>p</i> -value	0.161	0.375	0.590	
Number of observations	7,463	7,463	7,463	7,463
Number of firms	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments. KIA = 1 if the firm belongs to a knowledge intensive industry, according to the taxonomy developed by Eurostat (2012).

Appendix 7: Estimates using the ICT taxonomy

	GMM-SYS			LP
	Value added per hour worked (ln)	Mean wage per hour worked (ln)	Value added-wage gap (ln)	Value added-wage gap (ln)
	(1)	(2)	(3)	(4)
Std. dev. age	-0.039*** (0.011)	-0.013*** (0.005)	-0.026*** (0.009)	-0.003 (0.006)
Std. dev. education	0.008 (0.016)	0.017** (0.008)	-0.008 (0.015)	0.033*** (0.010)
Std. dev. gender	-0.362*** (0.124)	-0.186*** (0.065)	-0.176 (0.114)	-0.208*** (0.069)
Std. dev. age *ICT	0.051** (0.024)	0.012 (0.011)	0.040** (0.020)	-0.005 (0.008)
Std. dev. education*ICT	0.017 (0.037)	0.008 (0.018)	0.009 (0.032)	0.000 (0.019)
Std. dev. gender*ICT	0.533** (0.265)	0.237* (0.138)	0.295 (0.233)	0.366*** (0.118)
Average age	0.019*** (0.004)	0.010*** (0.002)	0.009*** (0.003)	0.010*** (0.003)
Average education	0.067*** (0.008)	0.036*** (0.004)	0.031*** (0.007)	0.060*** (0.006)
Average age *ICT	-0.025*** (0.007)	-0.002 (0.003)	-0.023*** (0.005)	-0.002 (0.004)
Average education*ICT	0.036** (0.015)	0.032*** (0.007)	0.004 (0.012)	0.042*** (0.012)
ICT	-0.184 (0.313)	-0.481*** (0.149)	0.297 (0.269)	-0.482** (0.226)
Hansen over-identification test, <i>p</i> -value	0.553	0.088	0.183	
Arellano-Bond test for AR(2), <i>p</i> -value	0.063	0.336	0.509	
Number of observations	7,463	7,463	7,463	7,463
Number of firms	2,431	2,431	2,431	2,431

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Clustered standard errors are reported between brackets. Regressions also control for: % workers with 10 years of tenure or more, % white-collar workers, % employees with a fixed-term contract, % part-time workers, firm size and capital stock, industries (8 dummies), and years dummies (7). AR(2) refers to second-order autocorrelation in first-differenced errors. GMM-SYS specifications include first and second lags of explanatory variables (except time dummies) as instruments. ICT = 1 if the firm belongs to a sector using or producing intensively ICT (information and communication technology) goods and services, according to the taxonomy developed by O'Mahony and van Ark (2003)