PRELIMINARY DRAFT

DO NOT CITE WITHOUT AUTHORS' AUTHORIZATION

DIVERSITY TO FOSTER INNOVATION: USING THE LENS OF BRAZILIAN MICRODATA

Authors[1]

Filipe Lage de Sousa (BNDES and UFF – Universidade Federal Fluminense)[2]

Glaucia Estefânia de Sousa Ferreira (IPEA – Instituto de Pesquisa Econômica Aplicada)[3]

Leandro Justino Pereira Veloso (IBRD – The World Bank Group [4])

Synthia Kariny Silva de Santana (IBGE – Instituto Brasileiro de Geografia e Estatística)[5]

Abstract

Different factors enable firms to innovate and eventually to improve their performance. The contribution of workforce diversity to innovation is a venue under-explored in earlier empirical work. The main objective is to highlight the relationship between innovation in the Brazilian private sector and worker's diversity within the firm, measured by different angles, including gender, age and race. Using detailed firm-level data, including employer-employee dataset, our results suggest that even though there are some costs, the benefits of workforce diversity can offset them in most innovation outcomes.

Keywords: workforce diversity; innovation; gender; age; race.

JEL Codes: D24; J16, J21, O31.

[1] **Disclaimer:** The BNDES, UFF, IPEA, IADB and IBGE are exempt from any responsibility for the views, information, data and concepts presented in this article, which are solely the responsibility of the authors.

[2]Professor at Department of Economics, Universidade Federal Fluminense – corresponding author: fl_sousa@id.uff.br

[3] External Consultant at CEPAL and IPEA

[4] Research Assitant at World Bank

[5] Economist at the Brazilian Institute of Geography and Statistics (IBGE)

1. Introduction

Equal opportunity for men and women, as widely acknowledged a key element for decent work, is a challenge for many countries achieve sustainable growth, but it is even most critical in developing countries such as Brazil. Using detailed firm-level data for the manufacturing sector, this paper aims to assess the impact of workforce diversity on innovation and productivity of Brazilian companies.

Value-in-diversity supporters argued that "diversity enriches the workplace by broadening employee perspectives, strengthening their teams, and offering greater resources for problem resolution" (Cox Jr., 2001). Similarly, Huber (1991) also argued that diversity provides an organization with a broader set of cognitive maps, which is associated with an original and creative problem-solving background. These set of attributes could ultimately provide a better scenario for innovative activities.

However, critics of the diversity model are skeptical about its implementation because of the difficulties surrounding the measurement of the multidimensional nature of workforce diversity. According to this view, deep-level diversity is crucial because it focuses on the individual rather than merely perpetuate stereotypical characterizations, as demographic variables would allow us to do. In fact, personality, values, creativity, and abilities would enable us to provide a better link to innovation. Unfortunately, those attributes are rarely available on datasets.

Although it is expanding, the literature that investigates the impact of workforce diversity on firm performance has not led to clear conclusions. In fact, understanding the underlying connections between composition, matching, and sorting of workers (and businesses) is at the heart of the overlap between industrial organization and labor economics, where this paper aims to fit. However, little is known about these kinds of linkages, which is partially due to the rarity and confidentiality on these types of data. Moreover, while the empirical literature on the impact of the firm's characteristics on the worker's outcomes is settled (Abowd et al, 1999; Montgomery, 1991), the converse is still an avenue underexplored.

On the one hand, scholars advocate workforce diversity as a catalyzer for improvements in firms. For example, according to McMahon (2011), "Diversity in business helps in pooling the best talent, reduces the gap between increasingly diverse customer bases, unleashes creativity,

promotes innovation and thereby enhances the competitiveness of the organization". Additionally, diversity may influence absorptive capacity at the organizational level (Cohen and Levinthal, 1990) through the experience and variety of knowledge. On the other hand, other researchers argue that diversity does not help but rather disturbs. These authors list personal conflicts and communication problems as shortcomings of having a diverse workforce (Becker 1971; Akerlof and Kranton 2000; Choi 2007). However, studies linking workforce diversity and firm innovation are scarce, despite the seminal innovation–performance relationship provided by Griliches (1979).

Following these recent findings, we aim at investigating how workers-diversity at firm level affects innovation capacity in Brazilian firms. Low productivity and innovation is a critical barrier to developing-country growth (Cirera and Maloney, 2017). Therefore, understanding the determinants of the firm's innovative capacities is critical to design the appropriate policymaker's interventions.

Our paper contributes to the literature on different aspects. First, we are not aware of any paper investigating the impact of workforce diversity using the most popular model in the innovation literature, which is that of Crepon, Duguet and Mairesse (1998), the CDM model. Second, we make use a longitudinal database for a longer period compared to any previous study, which enables us to investigate the dynamics of innovation. Last, this paper appears to be the first to address the effect of workforce diversity on innovation using data from a developing country, which its inhabitants are completely diverse (Brazil).

Our outcomes suggest the importance of workforce diversity in promoting innovation directly, especially in terms of gender and age, and productivity gains indirectly. Results in terms of race diversity show that costs of diversity might be neutralizing or even surprasing its benefits. Our outcomes are similar to those found in the empirical works in developed countries in terms of gender and race, yet different in terms of age. Therefore, promotion of government policies aiming towards higher integration between different backgrounds, especially in terms of race, might reduce these costs to promote more innovation and eventually development through productivity growth in the Brazilian private sector.

Aside this introduction, the rest of the paper is structured in other five sections. A literature review is presented in Section 2. Section 3 describes our empirical methodology. Data

description and descriptive statistics are presented in Section 4. We discuss the econometric results in Section 5, and we provide our concluding remarks in Section 6.

2. Literature Review

Workforce diversity has been used in different aspects in the literature, yet while some are chosen by individuals, others are not, as suggested by (Hewlett, Marshall, and Sherbin 2013). These authors classify workforce diversity into two categories: inherent; or acquired. Inherent resumes on those characteristics which individuals are not able to change: such as gender, age and race. Acquired characteristics comprehend those attributes that each individual is able to acquire after being born, such as the level of education or the occupation. Most of the literature on workforce does not make this distinction, which we consider essential especially because acquired attribute might be influenced by each individual inherent characteristic.

Despite this neglected aspect from the literature, benefits and costs of having a diverse workforce is discussed in the literature. For instance, workforce diversity might impact positively on firms performance due to different reasons, as reported by (Garnero, Kampelmann, and Rycx 2014). According to these authors, workforce diversity might promote higher performance in firms, for instance, if complementarities between workers are fostered (such as between those with different backgrounds) and/or spillovers are generated (knowledge transferred between workers). Whenever these circumstances occur, firms benefit from workforce diversity.

Another strand of the literature argues that workforce diversity harms firm's performance. Their argument is based on the fact that a diverse workforce leads to personal conflicts and communication problems, see (Becker, 1971; Akerlof and Kranton, 2000; Choi, 2007). For instance (Pfeffer 1985) states that diversity decreases social similarity, which reduces job satisfaction and eventually firm's performance. In terms of a specific diversity, age diversity raises communication costs as different generations experienced distinct environments when entering the labor market, see Lazear (1999). As for gender diversity, Haile (2012) points out that firms tend to underperform when gender diversity increases, especially when the labor force is mainly composed of men.

As observed, pro and cons arguments of workforce diversity exist and the literature does not provide clear pieces of evidence on whether workforce diversity works in favor or against firm's

performance. Thus, the net effect of workforce diversity on firm's performance depends on whether the gains surpass the costs. Different indicators of firm's performance exist, but our empirical literature focuses on how workforce diversity affects innovation. This literature expanded substantially during the last two decades and our literature review synthesizes the main findings on this subject.[6]

Although some authors, such as Harrison and Klein (2007), argue that empirical evidence of the impact of workforce diversity on firm's performance is inconclusive, the literature provides more robust evidence on the effects of workforce diversity on innovation. According to the literature review provided in Table I.a in Appendix A, diversity in terms of gender appears to affect positively innovation, while diversity other inherent attributes seem inconclusive.

This increasing literature suffers from a diverse range of drawbacks. First, diversity measurements differ across and within empirical works. For instance, while gender diversity is measured in terms of entropy index, age diversity is measured by the coefficient of variation, see Van der Vegt and Janssen (2003) as an example. Therefore, distinct metrics might lead to different outcomes. Second, when one measure of diversity is used for one attribute, no alternative is provided. In other words, when the paper uses an indicator for gender diversity, such as Blau Index, no other metric of diversity is explored, see Teruel, Robles and Blasco (2015) as an example. Third, while some papers use panel data to investigate this issue (Bolli, Renold and Wörter, 2018), others are still based on cross-section analysis, see Galia and Zenou (2012); Østergaard, Timmermans, and Kristinsson (2011); Díaz-García, González-Moreno, and Jose Sáez-Martínez (2013); Romero-Martínez, Montoro-Sánchez, and Garavito-Hernández (2017). Last, all papers are still from developed countries.

In the next sections, we will show how our paper addresses all these issues using a rich dataset spanning for a long period in a developing country, which provides much more demographic diversity compared to developed countries, especially in terms of race and educational level.

3. Methodology

Our empirical approach should encompass different aspects, starting from defining some crucial concepts and ending with econometric strategies to deal with shortcomings. First, the definition

[6] See Alsos et al.(2013) for the state of art and research agenda on gender diversity. In addition, a survey on workforce aging and innovation can be found in Frosch (2011).

of what kind of diversity and innovation should be described. Second, an empirical approach needs to be presented, including all econometric specifications. Third, all issues to estimate the CDM model need to be discussed. The following sections present how these aspects are considered in this paper.

3.1 Diversity measurement

Diversity measurement in the workforce within the firm is the first step for this investigation. We choose the Shannon and Weaver (1949) diversity index to measure how the workforce within the firm is diverse, because the attributes investigated in this research are categorical variables. This index is defined as:

$$s_i = -\sum_{r=1}^{R_i} p_{i,r} \ln(p_{i,r}) \quad (1)$$

Where s_i is the Shannon-Weaver diversity index of firm i, and $p_{i,r}$ is the proportion of the category or species r, of firm i, and R_i denotes the number of categories in firm i. Obviously, the diversity of categories is the highest when $p_{i,r} = 1/R_i, \forall \mathrm{r}$ or in some words, when the proportions of the different categories are equal. In addition, the higher is the number of existing categories within the firm i (i.e. R_i), the higher is the diversity index.

In this research, we use inherent characteristics, gender (*female* or *male*), age (*juniors* who are below 24 years old, *adults* who are between 24 and 40, and *seniors* for those after 40 years old) and race (*white* and *non-white*[7]).

The measurement of diversity is important in several sciences and it has evolved mainly within ecology. For robustness check, we also tried different sets of diversity measurements following the discussion presented in Ottaviano et al. (2003) and Haughton & Mukerjee (1995) who compare different measures in socio-economic topics.

Despite the criticism about Shannon-Weaver on ecological studies due to its inability to portrays species richness or species abundances across communities (see Sandoval et al, 2018), it does not seem to be a major problem in socioeconomic studies. Firstly, as the "species" groups that we constructed cover all the possibilities within firm, it is very likely for a firm to present most of

[7] The race of a person is self-reported according to his own identification following Brazillian race groups.

the categories. Secondly, its sensitivity to sample size is a minor problem since our dataset is large enough.

For robustness check, we replicated the results for Blau's index and the Morisita's index to see whether the results are sensitive to the diversity measurement chosen.[8] Blau's index is calculated as $1 - \sum_{r=1}^{R_i} p_r$, where p_r is the proportion of individuals in a category and R_i is the number of categories in firm i. As an alternative for a measure of dispersion of individuals in a population, we also computed the Morisita's index, which is $M_i = q \frac{\sum_{i=1}^{q} x_i(x_i-1)}{T(T-1)}$, where q is the number of sample units, and T is $\sum_{i=1}^{q} x_i$ and does not influence the index[9].

3.2 Innovation and CDM model

The literature on innovation has been using the Crepon-Douguet-Mairesse (CDM) approach [(Crepon, Duguet, and Mairesse, 1998)] as the standard model for investigating how innovation inputs, innovation outputs and productivity are linked. The CDM model states that firms invest in knowledge inputs, such as investment in R&D, training, machinery and others, which are transformed into innovation outputs according to the efficiency of their innovation function. These outputs, which are consequences of the innovation effort, are product, process, organizational and marketing innovations. In the last channel, these innovation outputs are able to promote productivity gains.

As described in the previous paragraph, three stages are observed in the CDM model: what drives innovation inputs, the innovation production function and last, the productivity equation. This allows us to control for simultaneity determination and causality problem between innovation and workforce diversity.

3.3 First Stage - Determinants of Innovation Inputs

The CDM model begins by estimating what drives innovation inputs. R&D investments are considered the main innovation input, but other expenses are important innovation inputs, especially in developing countries. According to (Crespi and Zuniga 2012), innovation in developing countries is much more assimilation and adaptation of existent technologies rather

[8] See Blau (1977) and Morisita (1962) for further information on these two indices.
[9] Although the results for Blau and Morisita indexes are not reported, they are available upon request.

than disruptive new technologies. Therefore, our measure of innovation input comprehends the total expenditure in innovation, which includes acquisition of machinery, patents, trademarks and others.

Different sets of attributes are relevant to explain how much a firm invests in innovation. We follow the literature and consider the following variables: firm characteristics (size, age and foreign capital ownership); firm's market conditions (competition, which can be measured by the Herfindahl-Hirshmann index)[10]; demand-pull factors (exports and cooperation with other agents); investment climate (obstacles to innovate and public incentives to innovation). [11]

As not every firm invests in innovation, our innovation input presents a substantial number of zero values and missing observations. Thus, a particular econometric technique to estimate the determinants of innovation inputs needs to consider this truncated distribution. The empirical literature uses the Tobit model by maximum likelihood for these cases. Our dependent variable described as innovation input, denoted by E_{it} from a firm i in time t, is related to another latent variable E^*_{it} such that:

$$E_{it}^* = \beta_1' x_{1it} + \varepsilon_{1it} \text{ when } E_{it}^* > 0 \text{ and zero otherwise} \qquad (2)$$

In any particular year t, x_{1it}corresponds to independent variables that explain knowledge inputs described above, such as firm's size, export status and others. Considering a truncated distribution, we expect that error term in equations (1) for our input innovation measure is assumed to be *iid* normally distributed.

3.4 Second Stage – Innovation Inputs, Diversity and Innovation Outcomes

Innovation input in the first stage is an important variable to influence innovation outputs when innovation is successful. Innovation materializing (innovation outputs) is when the impact of different ideas and knowledge are relevant in the innovation process.

For the latter, we follow the Oslo Manual [see (OECD, 2005) for further information] and define four types of innovation outputs as described below.

[10] As highlighted by Aghion et a. (2005), competition can be important for incentivizing investments in innovation inputs.
[11] See Canedo-Pinheiro et al (2019) for further details on which obstacles are relevant for innovation.

1. *Product Innovation* - A good or service that is new or significantly improved. This includes significant improvements in technical specifications, components and materials, software in the product, user friendliness or other functional characteristics.
2. *Process Innovation* - A new or significantly improved production or delivery method.

   This includes significant changes in techniques, equipment and/or software.
3. *Marketing Innovation* - A new marketing method involving significant changes in product design or packaging, product placement, product promotion or pricing.
4. *Organizational Innovation* - A new organizational method in business practices, workplace organization or external relations.

Basically, our econometric specification for each innovation output is presented in equation 3.

$$Innov_{it}^{*} = \beta_2' x_{2it} + \varepsilon_{2it} \quad (3)$$

Where $Innov_{it}^{*}$ is a latent variable (we observe only a dummy indicator of innovation), x_{2it} includes all explanatory variables used when estimating equation 1, as well as, the predicted value of innovation input, since how much a firm is expected to invest in innovation is a key determinant of innovation outputs. Additionally to those variables, workforce diversity is included in equation 2 as independent variables.[12] Our hypothesis is that exchange of ideas between different workers might be able to boost innovation outputs, in which creativity is vital.[13] Our innovation outputs are binary variables, Probit model is required. However, the decision to innovate in any aspect, either product, process, marketing or organizational, might be correlated to each other, our approach is to estimate a multivariate Probit to accommodate the correlation between the error terms.[14]

[12] Initially, we treat diversity measures as exogenous variables, yet some scholars might argue that they are also endogenous. In how to solve the model, we discuss the issue of endogeneity and propose a way to deal with the endogeneity problem of the diversity measures.

[13] Although creativity might be relevant to decide to whether innovate or not, our assumption is that once the decision has been made to innovate, workforce diversity plays a much more important role on product innovation rather than how much should be invested. Moreover, papers in the literature focus on innovation outcomes, as shown in Section 2.

[14] Additionally, we also estimate probits and biprobits for technological innovation (product and process) and non-techonlogical innovation (marketing and organizational) for robustness check, results are available upon request.

3.5 Third Stage – Productivity Equation

The final stage to estimate the impact of innovation on firm performance is to derive the productivity equation. Considering the production function as the starting point, the third stage econometric specification is represented by equation 4.[15]

$$log\left(Y_{it}/L_{it}\right) = \delta_0 + \delta log(H_{it}) + \alpha log\left(K_{it}/L_{it}\right) + \beta log(L_{it}) + \theta Z_{it} + \varepsilon_{it} \quad (4)$$

Where Y is firm's value added; K, capital; L, labor; H innovation outputs; and Z controls (sector and year dummies). However, labor productivity, as described in equation 4, is a short-term measure of productivity. Therefore, estimating the impact of innovation outputs in a long-term productivity measure, Total Factor Productivity (TFP), is essential and equation 5 shows our econometric specification.[16]

$$log(TFP_{it}) = \delta_0 + \delta log(H_{it}) + \theta Z_{it} + \varepsilon_{it} \quad (5)$$

3.6 Solving the model

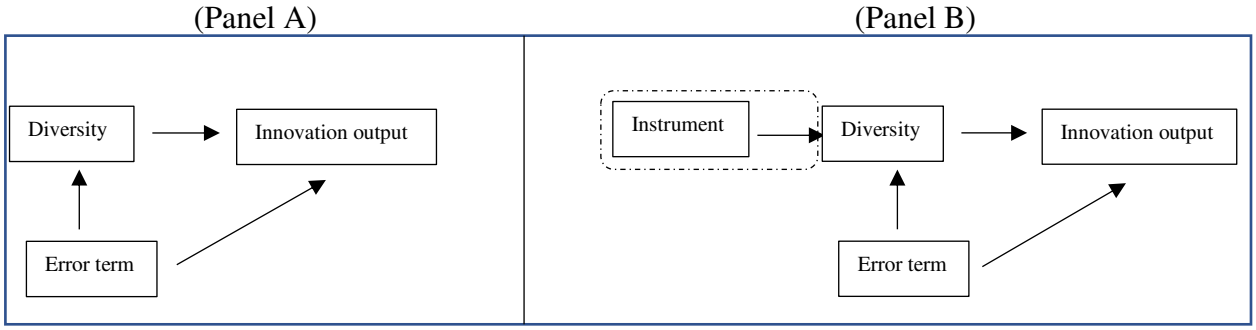
The main challenge to estimate the CDM model is the potential sources of reverse causality between inputs and outcomes, which means that more innovation can probably generate more knowledge investments, and between innovation outcomes and firm performance as the most productive firms could already be the most innovative ones. The literature has been using two different approaches to tackle this issue: either estimating the CDM by asymptotic least squares; or by a sequential approach, in which predicted values of endogenous variables are included in the estimation of the next stages, see Griffith et al. (2006) as an example. These two strategies are not exactly the same, yet outcomes from both approaches do not present substantial differences if endogeneity and selection are treated properly as suggested by Hall et al. (2009) and Musolesi and Huiban (2010). Since our econometric strategy in the second stage consists of estimating innovation outputs in a multivariate Probit, we estimate the model sequentially as in Griffith et al. (2006).

[15] See (Crepon, Duguet, and Mairessec 1998) for further details.
[16] TFP is estimated using Olley and Pakes (1996) methodology. Regressions outputs are available upon request.

3.7 Accounting for Endogeneity in Diversity Measures

If the pro-diversity firms are already the most innovative ones, our results are likely to be biased. In this case, the error term embodies all factors other than the diversity that determine innovation, such as management strategies. Therefore, estimations should be able to deal with endogenous decisions regarding innovation and workforce diversity. To account for this possibility, our approach to sort out the endogeneity of diversity measures consists of using the same strategy as used when incorporating innovation inputs in the second stage, which is estimating the model sequentially. In other words, a predicted value of innovation inputs is introduced as an independent value when determining what the main characteristics for innovation outcomes are. Therefore, our ideal instrument would be a variable that is strongly correlated with each type of workforce diversity (gender, age and ethnicity, for instance) but that it is not supposed to be related to innovation output (process, product, marketing and organizational) in the second stage.[17] The path diagram representing the association between diversity and the error term is:

Table 3.1: Instrument's Channel



Then, in Panel B we assume that changes in an instrument are associated with changes in workforce diversity but do not led to change in innovation output. The only way the instrument affects innovation is through the indirect diversity channel.

[17] We use all formal firms in the manufacturing sector to sort out the endogeneity issue. As the data from the first stage (300 thousand observations) has more information than those from the second stage (40 thousand observations), we are not able to implement any IV post-estimation test. Our stats rely on R-squared and F-Stat in the first stage.

Our strategy is exploring the main barriers hampering women labor force participation described in the literature as an instrument for gender diversity within firms. It relies on maternal health, which comprehends aspects related to maternity such as pregnancy, childbirth, and the period shortly after childbirth[18]; Fertility rate[19]; consumer durable (under the hypothesis that basic electrical appliances are labor-savings)[20]; and social and cultural factors[21]. Barriers to youth labor force participation, on the other side, are linked to the assumption that youth are less productive - mainly due to the fact that they have less (or none) work experience[22] or skill mismatch[23].

Table 3.2 summarizes the set of instruments chosen, while next sessions explain the rationale behind then. A primary test to assess its validity is to regress the proposed instrument against diversity, which need be significant and robust to a set of regional, sector and year controls, including region-year and sector-year dummies to capture any sector or geographical dynamics.

Table 3.2: Workforce Diversity and Instruments

Workforce diversity	Instrument
Gender diversity	• Maternity leave extension • Daycare coverage ratio • Divorce rates
Age diversity	• Number of vocational training workers
Racial diversity	• Sector and Regional Aspects

3.8 Maternity leave extension: the *Empresa Cidadã* Program (2008)

Participation of women in the Brazilian labor market is an upward trend, yet they face higher barriers, as highlighted by Gaddis and Pieters (2017). Countries are adapting their legislation to promote higher participation of women in the formal labor market. Among all the issues hampering women to engage in a professional career, the possibility to become mothers is a

[18] See Albanesi and Olivetti (2016)
[19] See Goldin and Katz (2002) and Lundborg, Plug and Rasmussen (2017)
[20] See Greenwood et al. (2005) show that married female labor force participation in the US can be explained by the consumer goods revolution, which introduced several time-saving products within households.
[21] See Fortin (2015) and Goldin and Katz (2016)
[22] See Pastore (2015)
[23] See Almeida, et al. (2012)

crucial one. In 2008, the Brazilian government created a program, named *Empresa Cidadã*, which is an exogenous and unanticipated policy aimed at extending the maternity leave from 120 days to 180 days. Women are then allowed to return to work only after their newborn child achieves 6 months; therefore, allowing mothers to exclusively breastfeed their babies during this recommended period by international standards.

Firms participate in the program on a voluntary basis by enrolling in an electronic form online. Enrolled firms in the program are allowed to deduct the expenditure on maternity leave on their income tax due. The firm's tax identifier of this program allows us to merge this data with micro level data from the Brazilian manufacturing survey (PIA, acronym in Portuguese) innovation survey (PINTEC, acronym in Portuguese) with employer-employee administrative record from the Ministry of Labor and Employment (named RAIS).

We consider all firms adopting this policy as firm's pro-women. In other words, even before this policy, these adopting firms might be considered as those interests to attract women to their workforce. Therefore, a dummy assuming value one for the whole period might be considered as a policy instrument which might be correlated to the gender diversity yet not with innovation outcomes. One drawback resides on the fact that all other firm fixed-effect over time should not be considered in the specification. In summary, we aim to estimate equation 6.

$$shannon_gender_{it} = policy_i\delta_1 + Sector\delta_2 + Region\delta_3 + \upsilon_t\delta_4 + \varepsilon_{it} \qquad (6)$$

Where *policy* indicates the target policy that extends maternity leave, $Sector$ and $Region$ are ISIC-2digits and regional specific dummies, τ_i is firm-especific, υ_t is time-variant to account for macroeconomic shocks and $\varepsilon_{it}$is the error term.

3.9 Daycare coverage ratio

The second approach for gender diversity is to explore exogenous variation of daycare coverage ratio within states. Preschool availability is expected to affect woman labor participation around the world, but it's a crucial aspect of childbearing in Brazil since trade liberalization has improved female labor force.

Exploring a natural experiment in the Canadian province of Quebec, Lefebvre and Merrigan (2008) provide evidence that childcare subsidies/supply has a positive effect on the labor supply

(result apply for mothers with young children). Although we don't know whether a woman has children in our employer-employee dataset, our rationale explores this female labor supply channel even for a single woman who expects to be a mother in the future.

For Brazil, to the best of our acknowledge, Barros et al. (2011) is the only study on childcare supply and female labor market participation. They explore a lottery assignment to childcare in Rio de Janeiro city carried out in 2008, which report an increase in mothers' employment by 36 to 46 percent and almost doubling (from 9 to 17 percent) for those who were not working before the lottery took place. These findings reinforce the argument for using this instrument, which is consolidated through the validity test running the following regression:

$$shannon_gender_{it} = daycarecoverage_{it}\delta_1 + Sector\delta_2 + Region\delta_3 + \tau_i\delta_4 + \upsilon_t\delta_5 + \varepsilon_{it} \qquad (7)$$

Where $daycarecoverage$ accounts for the daycare coverage rate in the municipality of firm i, while the remaning variables continue to have the same description as before.

3.10 Divorce Rates

It has been well recognized that household labor supply is affected by home and work demands that are placed on family time allocation, as seen in Becker (1973) and Chiaporri (1988). Therefore, marital status has been proved to determine long term income (Love, 2010), wage differentials (Korenman, 2004) and turnover (Ahituv and Lerman, 2011). But, more specifically, results may differ by gender as highlighted by Fernández and Wong (2014) and Eckstein et al. (2019). Brazilian main conclusions are presented in Madalozzo (2008) and Madalozzo & Gomes (2012).

Since 1988, the Brazilian Federal Constitution guarantee informal unions, such as cohabitation, the same legal rights as stable unions. On the other side, dissolution of marriage has become easier too: since 2007, Brazilian couples can request a divorce directly on notary's office by presenting basic documentation and paying a fee, provided that there's no minor children or mutual assets.

We argue that facilitate the divorce leads to a rise in Brazilian female labor force participation, releasing them from suboptimal marriages. This new facility, together with the fact of reducing the social stigma of being a woman (or man) divorced, lead women to invest in their own careers. In addition, women search for jobs after getting a divorce in order to provide resources

for their broken family, especially when children are present. Therefore, having information on divorce rates at each location might be correlated to the presence of women in the labor force and eventually to gender diversity within firms located at these regions. In order to investigate whether gender diversity can be explained by the divorce rate at each municipality in Brazil, we use specification 8.

$$shannon_gender_{it} = DivorceRate_{it}\delta_1 + Sector\delta_2 + Region\delta_3 + \tau_i\delta_4 + \upsilon_t\delta_5 + \varepsilon_{it} \quad (8)$$

Where $DivorceRate$ accounts for the divorce rate in municipality of firm i, while the remaning variables continue to have the same description as before.

3.11 Vocational training: the Brazilian Apprenticeship program (2005)

For age diversity we must seek for an exogenous variation affecting age composition within firm. Its widely recognized that young workers face major obstacles in the early stages of their careers, especially in school-job transition. Therefore, the unemployment rate of young workers is a fact, and incorporating them in the job market leads to social and economic benefits. In 2005, the Brazilian government established the rules of the: "*Jovem Aprendiz Act*", which requires that every medium or large firms should hire from 5% to 15% of their workforce in specific occupations composed by young workers (aging from 14 to 24 years old) and provide vocational training. According to Corseuil et al. (2014), the number of workers in each occupation is not correlated to firm's total employment, therefore we believe that the number of young professionals in each firm seems to be exogenous to workforce diversity in terms of age.

Since the firm is required to comply with the law, or otherwise will be subject to huge fines and other penalties, this program is eligible to instrument age diversity at workplace. To empirically test the validity of such instrument, we need to run the following regression:

$$shannon_age_{it} = apprenticeship_{it}\delta_1 + x_{3it}\delta_2 + \tau_i\delta_3 + \upsilon_t\delta_4 + \varepsilon_{4it} \quad (9)$$

Where $apprenticeship_{it}$accounts for the number of apprentices in firm i, $x_{3it}$is a set of controls, $\tau_i$is firm-especific, υ_t is time-variant to account for macroeconomic shocks and $\varepsilon_{it}$is the error term.

3.12 Racial Diversity: Sector and Region Aspects

We are not aware of any exogenous variation which might have affected racial diversity within firms in the Brazilian private sector, however we believe that regional and sector characteristics are relevant aspects to be correlated to firm's workforce diversity in terms of race, since a firm located in regions predominantly from one race tend to be less diverse than those in most racial balanced.[24] The same argument might be used for sector, as some might have some bias towards a specific race. Overall, we aim to estimate equation 9 to correct any potential bias in our racial diversity measure.

$$shannon_racial_{it} = Sector\delta_2 + Region\delta_3 + \tau_i\delta_4 + \upsilon_t\delta_5 + \varepsilon_{it} \quad (10)$$

4. Data and Descriptive Statistics

We manage a firm-level database with annual data for the period 1996 to 2014, creating an unbalanced panel with firms in the manufacturing industry, combining data from the manufacturing survey (PIA, acronym in Portuguese) and innovation survey (PINTEC, acronym in Portuguese) – both collected by the Brazilian Statistical Institute (IBGE) – with RAIS database, an employer-employee administrative record from the Ministry of Economy.

Data at the firm level in Brazil are merged on a unique firm identifier (CNPJ – National Registry of Legal Entities) in order to answer our main research questions. The three surveys are described as follows.

The Manufacturing Annual Survey (PIA) aims to identify the basic structural characteristics of the business segment of the manufacturing sector, providing information on production, revenue, costs, raw materials, and value added and shed some light on the firm's strategic orientation. The current PIA design uses a stratified sample of firms, collected all over the country in which all manufacturing firms with more than 30 employees are assigned to a take all strata. On the other hand, the majority of Brazilian firms is assigned to a basic sampling frame, which answers a simplified version of the survey questionnaire. On average, there are around 40 thousand firms in each year.

The second database is the Innovation Survey – PINTEC – that gathers information about innovative activities carried out by Brazilian firms in every two-three years. Currently, there are

[24] See Map A in the Appendix on how workforce diversity in terms of race differs geographically in Brazil.

six waves of PINTEC available in microdata (1998-2000, 2001-2003, 2004-2005, 2006-2008, 2009-2011, 2012-2014), where most of the qualitative variables refer to a period of three consecutive years. The quantitative variables refer to the last reference year of each wave. PINTEC is collected by Computer Assisted Telephone Interview System (CATI) and its sample design is based on tags for companies with higher probability in developing innovative activities. Because of the poor quality of PINTEC's first wave, which reports an unusual number of zero for critical variables, we restrict our analysis to the second onwards (2001-2014).

Finally, the employer-employee database – RAIS database – is an administrative record of the formally employed workers in Brazil. This registry is managed by the Ministry of Economy and its microdata is available from 1985-2017, including demographic, occupational and payroll characteristics of employees, which allows us to track the workforce composition of each firm through the firm identifier and, merge data with PIA and PINTEC.[25]

Table 4.1 shows whatever the source of the data, there are no substantial differences between the main estimated statistics. Precisely, this was the case by using the master and final dataset 2003's first wave of PINTEC (hereafter, first year) and 2014, the last wave of PINTEC (henceforth, last year). For instance, the major difference relies on Skilled/Non-skilled workers, because the PINTEC's sampling design assigns higher probability of being selected to the research to those firms with higher probability of being innovative. However, differences of all shares are not statistically different; therefore we conclude that reduction of the sample does not create any bias which might interfere in our investigation.

As for our instruments, different sources are utilized. For maternity leave, we use the information on each firm with their tax identifier provided by Ministry of Economy. Daycare coverage is provided by Teenagers and Child Observatory website ran by Abrinq Foundation, which is an institution responsible for mobilizing the Brazilian society on children rights. Data is available at State level from 2005 onwards. Divorce rates are provided by the Civil Statistics elaborated by the Brazilian Statistical Institute at Municipality level and records are from 2003 until 2017. Apprenticeship program information is available at RAIS.

[25] Description of all variables used in this paper as well as its source is presented in Appendix B.

Tabela 4.1 – Comparing main variable distributions across initial and final dataset

	2003		2014	
Atributes	PIA	PIA+PINTEC	PIA	PIA+PINTEC
Male	74,0%	76,1%	69,8%	70,7%
Female	26,0%	23,9%	30,2%	29,3%
Skilled	7,1%	8,8%	10,4%	12,5%
Unskilled	92,9%	91,2%	89,6%	87,5%
White	69,5%	69,0%	56,7%	56,1%
Non-white	30,5%	31,0%	43,3%	43,9%
#firms	28,451	7,632	38,591	9,995
#employees	5,840,015	3,745,151	5,967,359	3,960,502

Source: PIA, PINTEC and RAIS / Authors' Elaboration

4.1 Descriptive Statistics

Tables 4.2 presents the means and standard deviations of the main variables in this paper. On average, 50.2% of Brazilian firms in the sample reports at least product or process innovation in the period 2001-2014. A total of 54% of firms also indicated that they have been through an organizational or marketing change in this period. Also, a total of 52.8% of firms reports investment on product, process or incomplete/abandoned projects, which reveals the uncertainty/risk inherent in this type of investment since not every innovation project is realized. Table 4.2 also shows Shannon's index for diversity on a wide variety of categories, where maximum equality maximizes the index. Diversity in gender seems to be more balanced across workers (0.45) compared to race (0.40). When looking at different age categories, workforce diversity among senior workers (0.39) is lower than with younger workers (0.42). Racial diversity is more pronounced in male workers (0.40) than in female workers (0.35). As for age diversity, white workers seem more heterogeneous than non-white (0.92 versus 0.84, respectively). Nevertheless, it is also important to evaluate how these diversity indexes changed over time.

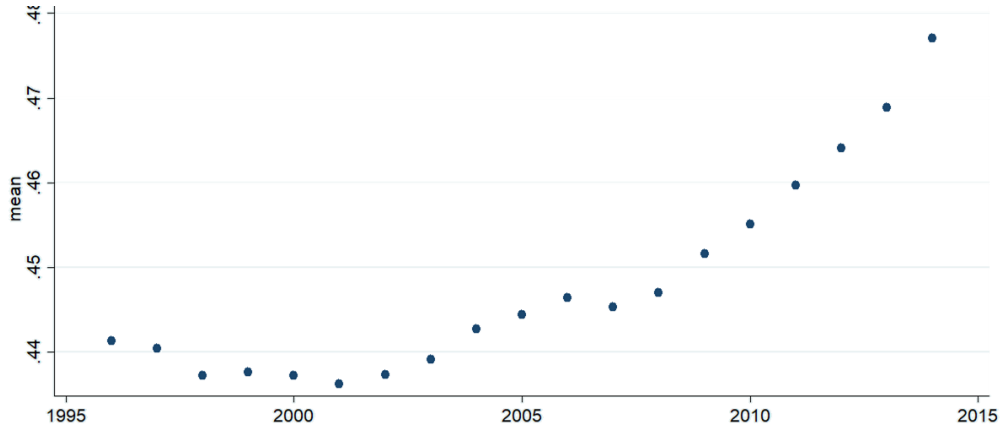
Tabela 4.2 – Statistics of the Main Variables

Variable description	*# obs.*	*Mean*	*sd*
Covariates			
ln(expenditure in innovative activities)	48,612	4671	55325
dummy for exporters	48,612	0.3848	0.4865
dummy for importers	48,612	0.4624	0.4986
herfindahl-Hirschman index	48,612	0.03095	0.08843
Firm's size (Number of Workers)	48,612	5120.1	1787
average age of workers	48,612	33.08	3679
Average employees schooling	48,094	8,792	1.932
Firm's internationalization	48,612	0.1143	0.3182
public incentives to innovation (dummy)	48,612	0.1788	0.3832
inter-firm cooperation for innovation (dummy)	48,612	0.1132	0.3168
obstacles to innovation faced by firm (dummy)	48,612	0.3285	0.4697
Shannon-Weaver entropy index for diversity on:**			
Gender	**48,612**	**0.4527**	**0.1877**
Gender within juniors	48,360	0.4241	0.2267
Gender within adults	48,595	0.4549	0.1957
Gender within seniors	48,529	0.396	0.2212
Age	**48,612**	**0.956**	**0.092**
Age within White	47,592	0.927	0.161
Age within Non-White	45,415	0.842	0.250
Race	**48,612**	**0.4046**	**0.2295**
Race within Male	48,582	0.402	0.2343
Race within Female	48,108	0.3556	0.2552
Innovation outputs			
Process	48,612	0.4279	0.4948
Product	48,612	0.3159	0.4649
Marketing	48,612	0.5419	0.4982
Organizational	48,612	0.5478	0.4977
Process, Product or Incomplete/Abandoned projects	48,612	0.5286	0.4992
Process or Product	48,612	0.5022	0.50
Any type of innovation	48,612	0.8051	0.3961
Overview (in US$)			
Labor productivity	48,612	20.768,65	67.868,39
Net sale revenues	48,612	51.147.581,35	646.854.764,77
Total revenues	48,612	57.041.231,09	762.304.265,28
Value added	48,612	14.257.355,44	235.987.763,73

Note: 1 US$ = 3,86R$ (June 2019); Overview variables are real values computed in Brazilian Reais, deflated by IPA-OG index published by Fundacao Getulio Vargas, and then converted in US$. ** In the case of two categories, the maximum Shannon index for gender is achieved at 0.69, while for three categories as age diversity that can be split into juniors, adults and seniors, the maximum is achieved at 1,10.
Source: PIA, PINTEC and RAIS / Authors' Elaboration

Figure 4.1 shows diversity on gender derived from employer-employee dataset from 1996-2014. Note that Shannon-Weaver index for the case of two categories, as gender, reaches its maximum at 0.69. Hence, in the late 90's, right after Brazil's trade openness, diversity within manufacturing was decreasing. In this period, Brazil was facing a financial crisis and, according to IBGE, the unemployment rate was 10.4% about in 1999. After that, in the early 2000's, the Brazilian economy recovered and experienced a new period of vigorous growth. According toWorld Bank, the GDP growth rate reached 5.74% in 2004 and 7.54% in 2010 (compared with the previous year). The key message is that gender diversity was in consonance with Brazilian economic growth, which may indicate that in periods of bonanza, firms start to hire those that were excluded from the labor market. When we plot the Shannon-Weaver index for gender within age groups (juniors, adults and seniors) we can infer that this diversity is driven by the juniors and adults individuals (See Figure 4.2).[26]

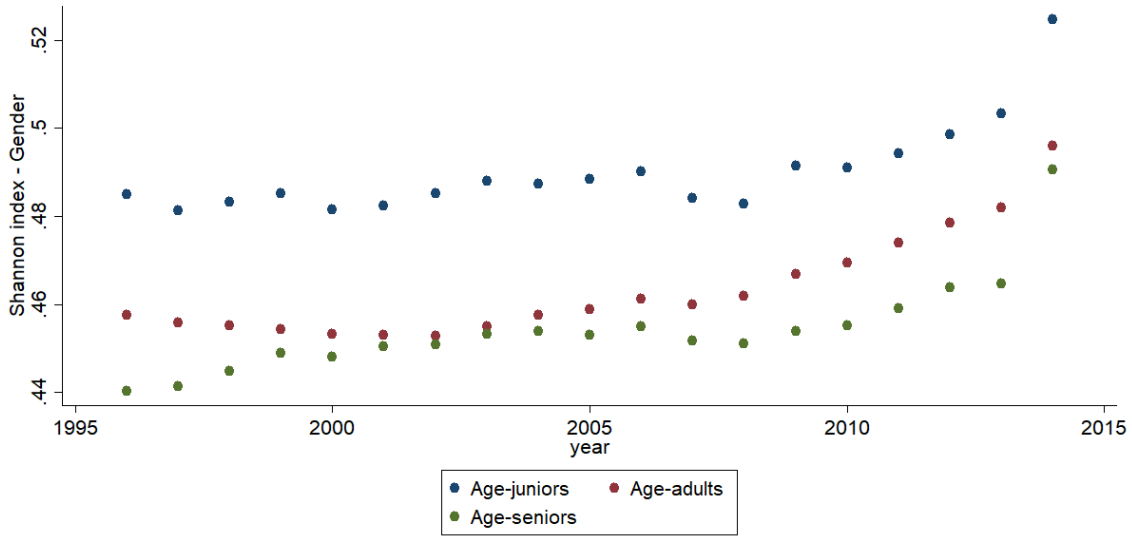
Figure 4.1: Shannon-Weaver entropy index for gender



Source: RAIS / Authors' Elaboration

[26] Shannon-Weaver entropy index for gender by manufacturing branch (ISIC two digits) is reported in the Appendix-A.

Figure 4.2: Shannon-Weaver entropy index for gender within age groups



Source: RAIS / Authors' Elaboration

Finally, in order to justify the study of the four distinct types of innovation, we plot the correlations between them. The highest correlation is between product and process innovation (0.46), which was expected according Damanpour and Gopalakrishnan (2001)[27]. In fact, the innovation survey is an interview made by telephone, so it is less likely that a firm report some type of innovation by mistake or misunderstanding, because any doubt can be remedied immediately during the interview. Therefore, each result on innovation output can be properly interpreted as independent without loss of generality.

Table 4.3: Correlation matrix – Innovation Outputs

	Product	*Process*	*Marketing*	*Organizational*
Product	1			
Process	**0,46**	1		
Marketing	0,24	0,21	1	
Organizational	0,22	0,24	0,26	1

[27] Since Product and Process innovation show correlation of 0.46, which can be interpreted as high, we apply a biprobit technique to account for simultaneity choice between both investments. For the sake of brevity we do not report this result, but it is available upon request.

5. Results

In general, previous research uses contemporaneous information on workforce diversity and innovation outcomes due to lack of previous information of labor force. As there might be a reverse causality there and our data enable us to use lag information, we introduce an extra cautious when estimating our econometric models. For all control variables derived from the manufacturing survey (PIA) and the employer-employee database (RAIS), we used them lagged in time. In other words, when using the information from the innovation survey (PINTEC), for example 2014, the controls from RAIS and PIA are lagged two years, which means from 2012 for those used in 2014. The use of lagged variables aims to reduce the potential problems of endogeneity from the regressors. Given that innovation information comprehends a period of three years, for instance product innovation considers the introduction of a new good or service in the last three years, then using the number of employees from the beginning of the period allows us to use a measure, which has not been interfered by the introduction of the new good. Nevertheless, we also present results using contemporaneous measures of workforce diversity in order to discuss our results with previous literature on this issue. [28]

As our focus is on whether diversity plays a role in innovation outcomes, this section focus on the results of the second stage from the CDM model.[29]

5.1 Results of Innovation Outputs' Determinants

As described in Section 3.4, different sets of innovation outcomes are considered in our second stage, in which workforce diversity might play a relevant factor as reported in the literature review. As discussed in our methodological section, the decision of one particular innovation output might be correlated to another, and then our results are based on a multivariate Probit considering all innovation outputs[30].

Our first outcome shows the gender diversity in Tables 5.1 until 5.4. The first table reports Shannon-Weaver diversity index on a naïve specification without accounting for endogeneity

[28] Besides the results presented with the independent variables lagged in time and contemporaneous, we also have outcomes using the average of the last three years for those characteristics. All results are in line with those presented in this paper and available upon request.
[29] Outcomes from the first and last stages are presented in Appendix C.
[30] We are restricting our analysis to manufacturing with 30 or more employees (which are at least medium firms on Brazilian context), then they are more likely to have a vertical organization and by taking overall diversification instead of board compositions can be a limitation.

issues. First four columns account for contemporaneous measure of diversity index, while the last four columns show lagged value of the diversity index.[31]

At first, outcomes using contemporaneous information on gender diversity show a positive association in all innovation measures, although marketing seems to be the most correlated. However, using gender diversity lagged two years; positive association vanishes in process and organizational innovation. These results suggest that product and marketing innovation present robust evidence on the positive association with workforce diversity in terms of gender. It also emphasizes the necessity to use lagged information to reduce the bias on assessing the importance of diversity measures on innovation outcomes.

By acknowledging potential bias for this naïve specification, that is likely to be influenced by firm's hiring policies as discussed in section 3, diversity measures might be endogenous and an instrument is needed. We explore three sets of instruments for gender diversity: maternity leave extension policy; daycare center supply; and divorce rates. When using maternity leave, results show a positive association only with marketing innovation in lagged and contemporaneous measures. However, a negative association is observed for organizational innovation when using lagged information of gender diversity measure.

Clearly there are other variables that can be used as instrument for gender diversity by easing women's labor force participation. Using daycare coverage as another approach, positive association remains on contemporaneous measure of diversity with marketing innovation, but using the lagged information, outcomes suggest that product and marketing innovation are positively related to gender workforce diversity. A negative association between gender diversity and organizational innovation is also present when using lagged information.

A third attempt for robustness check is to test whether divorce rates by municipality can be used as an instrument for gender diversity, following a growing literature stating how easiness on formal marriage dissolution affects woman labor participation, and, therefore, gender diversity within firms. Nearly all parameters become significant when using divorce rates. Considering this instrument, gender diversity is positively associated with product and marketing innovation, while negative for process and organizational innovation.

[31] First stage results fom each instruments are shown in Appendix C.

Main conclusions are twofold: first, taking into account the high correlation between product and process innovation (0.46), the gain of diversity on product innovation might be partially offset by the loss on process innovation; second, as Marketing innovation is more affected by diversity than Product innovation, this outcome suggests that higher levels of gender diversity tend to promote more intangible values, such as brand, than tangible ones. Therefore, increasing the number of women in the workforce enables firms to become more creative to add value in selling their products rather than creating new products.

Organizational innovation appeared either negative or non-significant. This definition of innovation involves decentralization of processes, changes in business strategy or search for best management practices, but since we only control for overall gender diversification, without specifying the composition of directorate board or decision makers, the result can be controversial. If we could be able to disentangle gender composition by occupation or main tasks within firms, the key message might be different.

Table 5.1: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index (Multivariate Probit)

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Gender Diversity Contemporaneous	0.365***	0.105**	0.0958**	0.756***				
	(-0.0847)	(-0.0473)	(-0.0418)	(-0.107)				
Gender Diversity Lagged Two Years					0.363***	0.0625	0.0553	0.751***
					(-0.0964)	(-0.0493)	(-0.0384)	(-0.103)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.2: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by the Maternity Leave extension policy (Multivariate Probit)

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Gender Diversity IV – Contemporaneous	0.363	0.126	-0.0598	1.191***				
	(0.505)	(0.318)	(0.298)	(0.385)				
Gender Diversity IV – Lagged Two Years					0.363	-0.266	-0.421*	0.706**
					(0.290)	(0.172)	(0.217)	(0.292)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.3: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by Daycare Coverage (Multivariate Probit)

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Gender Diversity IV – Contemporaneous	0.546	0.131	-0.130	1.141***				
	(0.538)	(0.323)	(0.317)	(0.383)				
Gender Diversity IV – Lagged Two Years					0.578*	-0.169	-0.493*	0.722**
					(0.302)	(0.248)	(0.279)	(0.331)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	37,984	37,984	37,984	37,984

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.4: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by Divorce Rate (Multivariate Probit)

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Gender Diversity IV – Contemporaneous	0.821*	-0.484*	-0.232	1.604***				
	(0.428)	(0.251)	(0.267)	(0.381)				
Gender Diversity IV – Lagged Two Years					0.727***	-0.525**	-0.581**	0.844**
					(0.280)	(0.235)	(0.242)	(0.330)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45,662	45,662	45,662	45,662	35,662	35,662	35,662	35,662

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Our second outcome shows the age diversity in Table 5.5 and 5.6, following the same rationale: the first table presents a naïve specification without accounting for endogeneity bias, while the second table presents the estimations with instrumental variables following the discussion on section 3. Following the structure from the previous tables, the first four columns account for contemporaneous value of the diversity index, while columns 5-8 includes diversity index lagged in time.

Contrary to what occurred in gender diversity, results without instruments for contemporaneous and lagged value of diversity are practically the same. In both cases, only process innovation is positively correlated to age diversity. Therefore, it seems that lagged information in terms of age does not interfere in our outcomes.

In order to account for potential endogeneity bias, we used instrumental variables approach by taking advantage of an exogenous variation of youngers labor market participation. The Brazilian Apprenticeship program is expected to affect age distribution among employees since it's mandatory for medium and large firms. Therefore, results on Table 5.6 shows that age diversity improves both Product and Marketing innovation, which reinforces the idea that diversity among generations promotes a creative and insightful environment that ultimately affects firm's ability to transform investment on innovation into effective outputs. By not mixing their employees, firms might be wasting resources.

Table 5.5: Second Stage - Innovation Outcomes Determinants Using Age Diversity Index (Multivariate Probit)

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Age Diversity – Contemporaneous	-0.0583	0.266***	0.0582	0.129				
	(0.122)	(0.0880)	(0.0882)	(0.0956)				
Age Diversity – Lagged Two Years					0.0307	0.211**	0.00595	0.102
					(0.131)	(0.0893)	(0.102)	(0.0930)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.6: Second Stage - Innovation Outcomes Determinants Using Age Diversity Index Instrumented by Apprenticeship Program (Multivariate Probit)

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Age Diversity IV – Contemporaneous	1.543***	0.622	0.529	1.411**				
	(0.396)	(0.395)	(0.336)	(0.604)				
Age Diversity IV – Lagged Two Years					1.296**	0.378	0.412	1.218**
					(0.512)	(0.411)	(0.342)	(0.583)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Our third outcome is a venue underexplored in the literature by considering individuals race as another diversity channel. We measured diversity by splitting employees into white and non-whites and calculated Shannon-Weaver diversity index for two groups. The key results are presented on Tables 5.7 and 5.8.

As before, the first four columns include contemporaneous information of racial diversity, while the last columns include racial diversity lagged in time. Contrary to the previous results, racial diversity consistently seems to undermine innovation outcomes, even when we use the information lagged in time.

On table 5.8, we control for potential endogeneity bias that comes from firms selecting the own employees based on its unobservable hiring policies. Thus, we minimize this pitfall by estimating the racial diversity controlling for geographical state-level dummies as well as sector-level dummies, which is intended to account for regional and sectoral differences on Brazilian racial composition. Additionally, sector and geographical dummies are also interacted with years, so that any time-varying information in terms of these two attributes is also considered in our first-stage estimation. Even by controlling for the set of geographic and sector variables, key interpretation remains practically unchanged. Thus, racial diversity hampers both Process and Organizational innovation, whether using lagged or contemporaneous information, which are innovation types effectively related to employees' interactions, decision making and

negotiations. Therefore, the cost of workforce diversity (miscommunication and background conflicts, for example) surpasses any benefit from innovating when it comes to racial diversity.

Table 5.7: Second Stage - Innovation Outcomes Determinants Using Racial diversity Index (Multivariate Probit)

Race	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Racial diversity – Contemporaneous	-0.0651**	-0.0930***	-0.0113	-0.0229				
	(0.0319)	(0.0280)	(0.0284)	(0.0382)				
Racial diversity – Lagged Two Years					-0.0904***	-0.131***	-0.0114	-0.0109
					(0.0283)	(0.0313)	(0.0315)	(0.0399)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.8: Second Stage - Innovation Outcomes Determinants Using Racel Diversity Index Instrumented by Region&Sector Dummies (Multivariate Probit)

Race	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Racial diversity IV – Contemporaneous	-0.323**	-0.317***	-0.240***	-0.252				
	(0.139)	(0.0786)	(0.0927)	(0.174)				
Racial diversity IV – Lagged Two Years					-0.252	-0.299***	-0.231**	-0.198
					(0.165)	(0.0947)	(0.0954)	(0.177)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Finally, Tables 5.9 and 5.10 put all measures together in order to account for any diversity, since it is arguable that the firms can explore more than one dimension at the same time. Table 5.9 includes gender diversity without instrument and Table 5.10 with instrument.[32]

[32] For parcimony, we present only results using lagged information. Outcomes with contemporaneous measures of diversity are available in Appendix C

Racial diversity remains strong enough to be significant on the same magnitude and sign, reinforcing its results on Process and Organizational innovation, even controlling for a full set of sector, year and sector-year fixed effects.

Product innovation is still enhanced by gender diversity without an instrument, even after controlling for other diversity dimensions, but it vanishes when considering an instrument. On the other side, Process and Organizational innovations are hampered by gender diversity, which reveals that the cost of diversifying can completely offset all the gains that a heterogeneous background provides. Therefore, the key policy message is that gender diversity might be a strategy contingent on the type of innovation implement by the firm. Marketing innovation becomes positive when considering maternity leave and divorce rates as instruments, which means that how ideas to sell the product flourishes when there is a diverse range of workers. Overall, marketing innovation seems to be positively associated with gender diversity, while process and organizational not.

Finally, age diversity was captured for the other workforce diversity sources, but remains boosting Marketing and Product innovation when using instruments, even controlling for other diversity measures. This result supports the hypothesis that workforce diversification can be seen as a strategy to enhance firm's innovation through employee's insights and creativeness across age cohorts.

Table 5.9: Second Stage - Innovation Outcomes Determinants Using All Diversity Indexes (Multivariate Probit) – Without Instrument

Diversity	(1)	(2)	(3)	(4)
Dependent Variable	Product	Process	Org.	Marketing
Gender Diversity Contemporaneous	0.363***	0.0713	0.0618	0.758***
	(0.0972)	(0.0484)	(0.0380)	(0.100)
Age Diversity – Contemporaneous	0.00463	0.158*	-0.0119	0.0594
	(0.132)	(0.0961)	(0.104)	(0.0936)
Racial Diversity – Contemporaneous	-0.0853***	-0.130***	-0.0141	-0.00985
	(0.0291)	(0.0314)	(0.0316)	(0.0410)
Sector Dummy	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes
Observations	44,600	44,600	44,600	44,600

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.9: Second Stage - Innovation Outcomes Determinants Using All Diversity Indexes (Multivariate Probit) – With Instrument

Diversity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Maternity Leave				Daycare Coverage				Divorce Rate			
Gender Diversity IV	0.101	-0.331*	-0.526**	0.516*	0.290	-0.368	-0.623**	0.436	0.458	-0.772***	-0.775***	0.571*
	(0.318)	(0.198)	(0.213)	(0.299)	(0.318)	(0.269)	(0.262)	(0.338)	(0.301)	(0.252)	(0.248)	(0.336)
Age Diversity IV	1.061**	0.163	0.386	0.832	0.949**	0.365	0.263	0.888**	0.992**	0.536	0.362	0.912**
by Apprenticeship program	(0.473)	(0.438)	(0.316)	(0.508)	(0.460)	(0.406)	(0.328)	(0.443)	(0.465)	(0.394)	(0.335)	(0.360)
Racial Diversity IV	-0.182	-0.284***	-0.204**	-0.136	-0.165	-0.321***	-0.214**	-0.148	-0.186	-0.342***	-0.239**	-0.144
by Sector and State level	(0.158)	(0.0922)	(0.0933)	(0.167)	(0.185)	(0.0910)	(0.0960)	(0.173)	(0.182)	(0.0889)	(0.0962)	(0.175)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	44,499	44,499	44,499	44,499	37,984	37,984	37,984	37,984	35,662	35,662	35,662	35,662

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6. Conclusion

Following the value-in-diversity literature, we tested the hypothesis to whether workforce diversity generates positive synergies between groups and increases firm capabilities to innovate, which could lead to productivity gains. Thus, this paper contributes to the literature investigating whether workforce diversity is a catalyzer for innovation in a developing country: Brazil.

Using a detailed and rich database from three different surveys, we are able to address the relationship between workforce diversity (using employer-employee dataset) with firm-level innovation (innovation survey) and productivity (manufacturing survey). Overall, our results suggest that even though there are some costs, the benefits of workforce diversity offset the costs in most innovation outcomes.

Generally, outcomes for gender diversity indicates that marketing innovation presents robust positive evidence, which means that firms more diverse on gender tend to introduce different marketing strategies over time. For other innovation measures, product innovation also seems to be positively related to gender diversity, yet not as robust as with marketing innovation.

In terms of age, results depend on whether instrument is considered or not. On one hand, controlling for potential endogeneity bias, both product and marketing innovation are positively related to age diversity. On the other hand, process innovation is positively associated with age diversity when no instrument is considered. Regardless the use of instrument, our results back up the idea that age diversity enables firms to innovate more, which is contrary to the international literature. A possible explanation is that age diversity measure similarly to gender diversity (entropy index rather than coefficient of variation) might present similar results, raising the question to what extent different measures lead to same results.

Finally, costs of workforce diversity in terms of race are still above of any benefit when it comes to provide insights for innovation, since most results show a negative association between innovation outcomes and racial diversity in the workforce. A feasible explanation might be the inexistence of direct policies to promote more non-white workers in the labor market as there are for women and young workers, such as Maternity Leave and Apprenticeship program. However, the Brazilian government has introduced quotas in some government universities providing more education for non-white population, which might eventually promote more understanding

between different race backgrounds and indirectly create an environment in the labor market with less miscommunication problems. Therefore, government policies might have an indirect effect which has not materialized yet.

In summary, our paper provides extremely heterogeneous effects of workforce diversity on innovation outcomes. This evidence shows that benefits and costs exist on having diversity among workers. The main policy message consists of promoting more integration of people with different backgrounds so that the economy benefits from its human assets.

7. References

Abowd, John M., Francis Kramarz, and David N. Margolis. (1999). "High Wage Workers and High Wage Firms." Econometrica 67 (2). Blackwell Publishing Ltd: 251–333.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. (2005). "Competition and Innovation: An Inverted-U Relationship*." *The Quarterly Journal of Economics* 120 (2): 701– 28.

Ahituv, A., & Lerman, R. I. (2011). Job turnover, wage rates, and marital stability: How are they related? *Review of Economics of the Household*, 9(2), 221–249.

Albanesi, S., & Olivetti, C. (2016). "Gender roles and medical progress". *Journal of Political Economy*, 124(3), 650-695.

Almeida, R., J. Behrman, and D. Robalino (2012), "The Right Skills for the Job? Rethinking Training Policies for Workers". Washington,DC: *World Bank Publications*, World Bank.

Alsos, G.A, Ljunggren, E., & Hytti, U. (2013). "Gender and innovation: State of the art and a research agenda". *International Journal of Gender and Entrepreneurship*, 5(3), 236–256.

Akerlof GA, Kranton RE (2000) "Economics and identity" Quarterly Journal of Economics 115, 715-753

Bantel, K.A., and S.E. Jackson, (1989), "Top Management and Innovation in Banking: Does the Composition of the Top Team Make a Difference?" Strategic Management Journal 10,107–124.

Barros, R. P., de Carvalho, M., Franco, S., Mendonça, R., & Rosalém, A. (2011). "Uma avaliação do impacto da qualidade da creche no desenvolvimento infantil." Pesquisa e planejamento econômico, 41, 213-232.

Becker, Gary. (1971) "The Economics of Discrimination" Second Edition, (Chicago, IL: University of Chicago Press).

Becker, G. S. (1973). "A theory of marriage: Part I". Journal of Political Economy, 81, 813–846

Blau, P. M. (1977). Inequality and heterogeneity: A primitive theory of social structure (Vol. 7). New York: Free Press.

Bolli, Thomas, Ursula Renold, and Martin Wörter. (2018). "Vertical Educational Diversity and Innovation Performance." *Economics of Innovation and New Technology* 27 (2). Routledge: 107–31.

Cânedo-Pinheiro, M, Lage de Sousa, F., and Barcelos, V. (2019). "Barriers to Innovation in the Brazilian Economy - Impacts and Lessons for Public Policy." IDB Discussion Paper, forthcoming.

Chiappori, Pierre-Andre´.(1988) "Rational Household Labor Supply." Econometrica 56, January: 63–90

Choi, Jin Nam. (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). John Wiley & Sons, 213–34

Cirera, Xavier, and William F. Maloney. (2017). *The Innovation Paradox: Developing-Country Capabilities and the Unrealized Promise of Technological Catch-Up*. The World Bank.

Cohen, Wesley M, and Daniel A Levinthal. (1990). "Absorptive Capacity: A New Perspective on Learning and Innovation." *Administrative Science Quarterly*. Vol. 35.

Corseuil, Carlos H, Miguel Foguel, Gustavo Gonzaga, and Eduardo Ribeiro. (2014). "Youth Turnover in Brazil: Job and Worker Flows and an Evaluation of a Youth-Targeted Training Program." CEDLAS, Universidad Nacional de La Plata.

Cox Jr., Taylor. (2001). *Creating the Multicultural Organization: A Strategy for Capturing the Power of Diversity. Creating the Multicultural Organization: A Strategy for Capturing the Power of Diversity.* University of Michigan Business School Management Series. San Francisco, CA, US: Jossey-Bass.

Crepon, Bruno, Emmanuel Duguet, and Jacques Mairessec. (1998). "Research, Innovation And Productivity: An Econometric Analysis At The Firm Level." *Economics of Innovation and New Technology* 7 (2). Routledge: 115–58.

Crespi, Gustavo, and Pluvia Zuniga. (2012). "Innovation and Productivity: Evidence from Six Latin American Countries." *World Development* 40 (2): 273–90.

Damanpour, F., & Gopalakrishnan, S. (2001). "The dynamics of the adoption of product and process innovations in organizations". *Journal of Management Studies*, 38(1), 45–65.

Díaz-García, Cristina, Angela González-Moreno, and Francisco Jose Sáez-Martínez. (2013). "Gender Diversity within R&D Teams: Its Impact on Radicalness of Innovation." *Innovation* 15 (2). Routledge: 149–60.

Eckstein, Z., Keane, M., & Lifshitz, O. (2019). Career and Family Decisions: Cohorts Born 1935-1975. *Econometrica*, 87(1), 217–253.

Fernández, R., & Wong, J. (2014). Unilateral Divorce, the Decreasing Gender Gap, and Married Women's Labor Force Participation. *American Economic Review*, 104(5), 342–347.

Fortin, N. M. (2015). "Gender role attitudes and women's labor market participation: opting-out, AIDS, and the persistent appeal of housewifery". *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (117/118), 379-401.

Frosch, Katharina H. (2011). "Workforce Age and Innovation: A Literature Survey." *International Journal of Management Reviews* 13 (4). Wiley/Blackwell (10.1111): 414–30.

Gaddis, I., & Pieters, J. (2017). "The Gendered Labor Market Impacts of Trade Liberalization". *Journal of Human Resources,* 52(2), 457–490

Galia, Fabrice, and Emmanuel Zenou. (2012). "Board Composition and Forms of Innovation: Does Diversity Make a Difference?" *European Journal of International Management* 6 (6). Inderscience Publishers: 630–50.

Garnero, Andrea, Stephan Kampelmann, and François Rycx. (2014). "The Heterogeneous Effects of Workforce Diversity on Productivity, Wages, and Profits." Industrial Relations: A Journal of Economy and Society 53 (3). John Wiley & Sons, Ltd (10.1111): 430–77. https://doi.org/10.1111/irel.12064.

Goldin, C., & Katz, L. F. (2002). "The power of the pill: Oral contraceptives and women's career and marriage decisions". Journal of political Economy, 110(4), 730-770.

Goldin, C., & Katz, L. F. (2016). "Women Working Longer: Facts and Some Explanations" (No. w22607). *National Bureau of Economic Research*.

Greenwood, J., Seshadri, A., & Yorukoglu, M. (2005). "Engines of liberation". *The Review of Economic Studies*, 72(1), 109-133.

Griffith, Rachel, Elena Huergo, Jacques Mairesse, and Bettina Peters. 2006. "Innovation and Productivity Across Four European Countries." *Oxford Review of Economic Policy* 22 (4): 483–98.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *The Bell Journal of Economics* 10 (1). RAND Corporation: 92.

Haile, Getinet Astatike. (2012). "Unhappy Working with Men? Workplace Gender Diversity and Job-Related Well-Being in Britain." Labour Economics 19 (3): 329–50.

Hall, Bronwyn H, Francesca Lotti, and Jacques Mairesse. (2009). "Innovation and Productivity in SMEs: Empirical Evidence for Italy." *Small Business Economics* 33 (1): 13–33

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). Academy of Management: 1199–1228. https://doi.org/10.5465/amr.2007.26586096.

Haughton, Dominique M.A., and Swati Mukerjee. (1995). "The Economic Measurement and Determinants of Diversity." Social Indicators Research 36 (3). Kluwer Academic Publishers: 201–25.

Hewlett, Sylvia Ann, Melinda Marshall, and Laura Sherbin. (2013) "How Diversity Can Drive Innovation." Harvard Business Review..

Horwitz, S.K. and I.B. Horwitz, (2007), "The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography," Journal of Management 33, 987–1015

Huber, George P. (1991). "Organizational Learning: The Contributing Processes and the Literatures." *Organization Science* 2 (1). INFORMS : 88–115.

Joshi, Amol M., Todd M. Inouye, and Jeffrey A. Robinson. (2018). "How Does Agency Workforce Diversity Influence Federal R&D Funding of Minority and Women Technology Entrepreneurs? An Analysis of the SBIR and STTR Programs, 2001–2011." *Small Business Economics* 50 (3). Springer New York LLC: 499–519.

Korenman, S. (2004). Does marriage really make men more productive? In: *Sex Differences in Labor Markets* (pp. 29–48). Routledge Taylor & Francis Group.

Lazear, Edward P. (1999). "Globalisation and the Market for Team-Mates." The Economic Journal 109 (454). John Wiley & Sons, Ltd (10.1111): 15–40.

Lefebvre, P., & Merrigan, P. (2008). "Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada." Journal of Labor Economics, 26(3), 519-548.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. *Review of Financial Studies*, 23(1), 385–432.

Lundborg, P., Plug, E., & Rasmussen, A. W. (2017). "Can Women Have Children and a Career? IV Evidence from IVF Treatments". *American Economic Review*, 107(6), 1611-1637.

Madalozzo, Regina, & Gomes, Carolina Flores. (2012). The impact of civil status on womencs wages in Brazil. Estudos Econômicos (São Paulo), 42(3), 457-487.

Madalozzo, Regina. (2008). An analysis of income differentials by marital status. Estudos Econômicos (São Paulo), 38(2), 267-292

Mayer, Roger C., Richard S. Warr, and Jing Zhao. (2018). "Do Pro-Diversity Policies Improve Corporate Innovation?" *Financial Management* 47 (3). John Wiley and Sons Inc.: 617–50.

McMahon, Anne M. (2011). "Does Workplace Diversity Matter? A Survey Of Empirical Studies On Diversity And Firm Performance, 2000-09." *Journal of Diversity Management (JDM)* 5 (2).

Montgomery, D. James. (1991). "Equilibrium Wage Dispersion and Interindustry Wage Differentials". The Quarterly Journal of Economics. 106. 163-79.

Morisita, M. (1962) Id--index, a measure of dispersion of individuals. Researches on Population Ecology, Tokyo, v. 4, n. 1, p. 1-7.

Musolesi, Antonio, and Jean-Pierre Huiban. (2010). "Innovation and Productivity in Knowledge Intensive Business Services." *Journal of Productivity Analysis* 34 (1). Springer US: 63–81.

OECD. (2005). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*.

Olley, G. Steven, and Ariel Pakes. (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263.

Østergaard, Christian R, Bram Timmermans, and Kari Kristinsson. (2011). "Does a Different View Create Something New? The Effect of Employee Diversity on Innovation." *Research Policy* 40 (3): 500– 509.

Ottaviano, Gianmarco I.P. and Pinelli, Dino and Maignan, Carole Juliette and Rullani, Francesco, (2003) "Bio-Ecological Diversity vs. Socio-Economic Diversity: A Comparison of Existing Measures". *FEEM Working Paper No. 13.2003*.

Pastore, F. (2015) "The Youth Experience Gap: Explaining National Differences in the School-to-Work Transition". *Heidelberg: Springer International Publishing* AG.

Pellegrino, Gabriele and Savona, Maria. (2017), No money, no honey? Financial versus knowledge and demand constraints on innovation, Research Policy, 46, issue 2, p. 510-521

Pfeffer, Jeffrey. (1985). "Organizational Demography: Implications for Management." California Management Review 28 (1). SAGE Publications Inc: 67–81.

Pfeifer, Christian, and Joachim Wagner. (2014). "Is Innovative Firm Behavior Correlated with Age and Gender Composition of the Workforce? Evidence from a New Type of Data for German Enterprises." *Journal for Labour Market Research* 47 (3): 223–31.

Richard, O.C., 2000, "Racial Diversity, Business Strategy, and Firm Performance: A Resource Based View," Academy of Management Journal 43, 164–177.

Romero-Martínez, Ana M., Ángeles Montoro-Sánchez, and Youseline Garavito-Hernández. 2017. "El Efecto de La Diversidad de Género Y El Nivel Educativo En La Innovación." *Revista de Administração de Empresas* 57 (2). Scielo: 123–34.

Ruiz-Jiménez, J. M., & Fuentes-Fuentes, M. del M. (2016). "Management capabilities, innovation, and gender diversity in the top management team: An empirical analysis in technology-based SMEs". *BRQ Business Research Quarterly*, 19(2), 107–121.

Sandoval, Luis, Gilbert Barrantes, and David R. Wilson. 2018. "Conceptual and Statistical Problems with the Use of the Shannon-Weiner Entropy Index in Bioacoustic Analyses." Bioacoustics, April 3

Shannon, C.E., and W. Weaver. 1949. *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana.

Talke, Katrin, Soeren Salomo and Katja Rost.(2010. "How top management team diversity affects innovativeness and performance via the strategic choice to focus on innovation fields." Research Policy 39, 907–918.

Teruel, M., M.D.P. Robles, and A.S. Blasco. 2015. "Gender Diversity and Innovation in Manufacturing and Service Firms."

Torchia, Mariateresa, Andrea Calabrò, and Morten Huse. 2011a. "Women Directors on Corporate Boards: From Tokenism to Critical Mass." *Journal of Business Ethics* 102 (2): 299–317.

Turner, Laure. 2009a. "Gender Diversity and Innovative Performance." *International Journal of Innovation and Sustainable Development* 4 (2/3): 123–34.

Vegt, Gerben S Van der, and Onne Janssen. 2003a. "Joint Impact of Interdependence and Group Diversity on Innovation." *Journal of Management* 29 (5). SAGE Publications Inc: 729–51.

8. Appendix A : Table I.a: Literature review for workforce diversity and Innovation

Authors	Title	Data description	Diversity measurement	Main findings
Bantel, K.A., and S.E. Jackson, (1989)	Top Management and Innovation in Banking: Does the Composition of the Top Team Make a Difference?	460 state chartered and national banks located in six midwestern states	Averages	Results indicate that more innovative banks are managed by more educated teams who are diverse with respect to their functional areas of expertise. These relationships remain significant when organizational size, team size, and location are controlled for.
Richard, O.C. (2000)	Racial Diversity, Business Strategy, and Firm Performance: A Resource Based View	Firm level data in Banking Industry. Total of 574 firms were interviewed	Blau's index of heterogeneity	Cultural diversity does in fact add value and, within the proper context, contributes to firm competitive advantage
Van der Vegt and Janssen (2003)	Joint impact of interdependence and group diversity on innovation	Database with 343 members of 41 work teams from a Dutch multinational in the financial services sector.	Shannon-Weaver entropy index for gender/ethnicity and coefficient of variation for age.	Their findings for demographic diversity do not indicate a direct relationship between group diversity and innovative behavior.
Horwitz, S.K. and I.B. Horwitz (2007)	The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography	Peer reviewed articles on the topic of team diversity published between 1985 and 2006	Diverse range	Support was found for the positive impact of task-related diversity on team performance although bio-demographic diversity was not significantly related to team performance. Similarly, no discernible effect of team diversity was found on social integration.
Turner (2009)	Gender Diversity and Innovative Performance	Project/individual data from 4 companies which participated in the European Commission's working group "Women in Science and Technology" in 2004: 3,812 workers who have participated in 272 projects	Share of female in each project	Gender balance at the team-level increases the performance not only at individual level but also at project level.
Talke, K., S. Salomo, and K. Rost (2010)	How Top Management Team Diversity Affects Innovativeness and Performance via	Mail survey to pre-identifield respodents (head of innovation or R&D). There are 122 responses,	Herfindahl-Hirschman Index	Top Management Team diversity has a strong impact on the strategic choice of firms to focus on innovation fields. Such focus then drives new product portfolio innovativeness and firm

	the Strategic Choice to Focus on Innovation Fields	coming from 10 manufacturing industries and originate from 17 countries.		performance.
Torchia, Calabrò, and Huse (2011)	Women directors on corporate boards: from tokenism to critical mass	Survey with nearly 3 thousand Norwegian firms in 2006 and, after, select only 317 of them, as they want to investigate firms having 6 to 12 board members in which woman are not minority group.	Share of female on corporate boards	The boards' contribution to the level of firm organizational innovation is higher in boards with 'at least three women'.
Østergaard et al. (2011)	Does a Different View Create Something New? The Effect of Employee Diversity on Innovation	Sample of 1,648 Danish firms from 2006	For workforce diversity, they use Shannon-Weaver Index for gender, ethnicity and educational status; for age they use standard deviation.	They found that employee diversity in terms of gender and education has a positive effect on the likelihood that firms innovate, yet age reveals a negative correlation and ethnicity none.
Galia and Zenou (2012)	Board Composition and Forms of Innovation: Does Diversity Make a Difference?	Cross-section of 176 French firms from 2008	For gender: the share of female; For age: the presence of five age groups in the board.	Gender diversity on board is positively correlated to marketing innovation yet negative to product innovation. As for age diversity, their outcomes show a positive association with product innovation and a negative with organizational innovation.
Díaz-García, González-Moreno, and Jose Sáez-Martínez (2013)	Gender Diversity within R&D Teams: Its Impact on Radicalness of Innovation	The sample consists of 4277 companies (the majority on manufacturing sector) regarding firms' innovation activities and employment characteristics in 2007	Blau index	Gender diversity is positively related to radical innovation, but not to the incremental
Pfeifer and Wagner (2014)	Is Innovative Firm Behavior Correlated with Age and Gender Composition of the Workforce? Evidence from a New Type of Data for German Enterprises	German firm-level database of 16,571 firms from 2003 to 2006	The share of female and people aged above 50 years old	Their findings reveal on average a negative correlation with older employees and a positive correlation with female employment.
Teruel et	Gender Diversity	Balanced panel of	Blau index	Gender-diverse teams increase the

al. (2015)	and Innovation in Manufacturing and Service Firms	5,392 of Spanish firms from 2007 to 2012		probability of innovating in all types of innovation, but this positive association is observed only in product and organizational innovation when considering gender diversity in R&D teams exclusively. Diversity is more evident in manufacturing firms compared to service firms.
Ruiz-Jiménez & Fuentes-Fuentes (2016)	Management capabilities, innovation, and gender diversity in the top management team: An empirical analysis in technology-based SMEs	A sample of 205 Spanish SMEs from technology sectors	Blau Index	Management capabilities have a greater influence on both product and process innovation when the management team is more balanced in number of men and women.
Joshi et al. (2017)	How does agency workforce diversity influence Federal R&D funding of minority and women technology entrepreneurs? An analysis of the SBIR and STTR programs, 2001–2011	52,126 Phase I SBIR/STTR awards granted by 11 Federal agencies (2001–2011).	Percentage of female employees in a Federal agency's workforce by year	Positive association between agency workforce diversity and Phase II funding for Phase I grantees, but minority and women technology entrepreneurs are less likely to receive this funding than their non-minority and male counterparts.
Romero-Martínez et al. (2017)	El Efecto de la diversidad de género y el nivel educativo en la innovación.	Sample of nearly eight thousand Spanish firms from 2012	Blau Index.	Gender diversity and educational level of R&D workers have a positive influence on product innovation, including radical innovation.
Bolli et al. (2018)	Vertical educational diversity and innovation performance	Six waves of the Swiss innovation survey	Herfindahl-Hirshman index.	Vertical educational diversity increases the extensive margin of R&D and product innovation. However, the relationship with process innovation, R&D intensity, and product innovation intensity is insignificant
Mayer et al; (2018)	Do Pro-Diversity Policies Improve Corporate Innovation	MSCI ESG STATS database	Two broad categories: Diversity Strengths and Diversity Concerns.	Corporate policies that promote more pro-diversity cultures, specifically treatment of women and minorities, enhance future innovative efficiency.

9. Appendix B – Description of the Variables

	Variable	Description
Input inovation	*Obstacles*	*dummy if firm receive some benefit from govermment*
	Cooperation	*dummy if firm cooperates each other to innovate*
	Government Support	*dummy if firm receive some benefit from govermment*
	Firm's internationalization	*dummy if firm receive some benefit from govermment*
	Firm's size (Number of Workers)	*log of Total employes december 31 plus 1 by firm*
	Average employees schooling	*Average education years by firm*
	Firm's age	*Age of firm estimated by oldest employee*
	ln (Herfindahl-Hirschman) (t-2)	*log of Herfindahl-Hirschman index in t-2*
	Import status (t-2)	*Dummy is firm import in t-2*
	Export status (t-2)	*Dummy is firm export in t-2*
	ln(expenditure in innovative activities)	*log o total expenditure in innovative activities plus 1*
Innovation and Index shannon	*Process*	*Dummy for Innovation in Process in last 3 years*
	Product	*Dummy for Innovation in Product in last 3 years*
	Marketing	*Dummy for Innovation in Marketing in last 3 years*
	Organizational	*Dummy for Innovation in Organizational*
	Process, Product or Incomplete/Abandoned projects	*Dummy of Process, Product or Incomplete/Abandoned pr years*
	Process or Product	*Dummy for Innovation in Process or Process in last 3 ye*
	Any type of innovation	*Dummy for Innovation in at lest one of 4 types above in l*
	Age Diversity	*Shannon weaver index calculated by age group*
	Racial diversity	*Shannon weaver index calculated by race group*
	Gender Diserviry	*Shannon weaver index calculated by gender group*
productivity	*Labor productivity*	*values of manufacturing diveded by total employees*
	Gross revenues	*Gross revenues*
	Net sale revenues	*Net sale revenues*
	Total revenues	*Total revenues*
	Value added	*Value added*
	Value of manufacturing	*Value of manufacturing*
instrument & overview	*Maternity Leave*	*Dummy starting by the first year of maternity leave polic*
	Daycare Coverage	*Tax of coverage daycare center in brasilian municipality*
	Divorce Rate	*Divorce yearly rate mesuare by municipality*
	Male	*Dummy if gender is male*
	Female	*Dummy if gender is female*
	Skilled	*Dummy if worker has college or more*
	Unskilled	*Dummy if worker has less than incomplete college*
	White	*Dummy if worker was declared was white race*
	Non-white	*Dummy if worker was declared was non-white race*

10. Appendix C – Accounting for Endogeneity in the CDM's Second Stage and Third Stage Results

10.1 Outcomes of Innovation Inputs' Determinants

Table C.1 shows the results for the first stage of the CDM model, which defines the determinants of innovation input, measured by the sum of R&D expenditure with training, trademarks, capital goods and others. We present outcomes for OLS and Tobit, just to elucidate about the bias when not considering a truncated distribution of the dependent variable.

Table C.1: First Stage - Determinants of Innovation Inputs

Dependent Variable Expenditure on Innovation	OLS	Tobit
	(1)	(2)
Obstacles	0.939***	2.013***
	(0.0425)	(0.102)
Cooperation	2.175***	3.293***
	(0.0675)	(0.167)
Government Support	3.346***	5.483***
	(0.0773)	(0.246)
Firm's internationalization	0.287***	0.190
	(0.0754)	(0.118)
Firm's size (Number of Workers)	0.533***	0.769***
	(0.0496)	(0.0710)
Average employees schooling	0.183***	0.337***
	(0.0215)	(0.0389)
Firm's age	0.00329**	0.00291
	(0.00122)	(0.00237)
ln (Herfindahl-Hirschman) (t-2)	0.0123	0.00457
	(0.0858)	(0.171)
Import status (t-2)	0.412***	0.906***
	(0.0578)	(0.107)
Export status (t-2)	0.305***	0.675***
	(0.0531)	(0.118)
Year dummy	yes	yes
Sector dummy	yes	yes
Year*Sector dummy	yes	yes
Constant	-2.390***	-11.84***
	(0.347)	(0.682)
#Observations	44,600	44,600
R squared	0.410	
Pseudo R-squared		0.104

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As discussed previously, all independent variables are lagged two years, which diminishes the potential issue of endogeneity. First, it is important to notice that the magnitudes of the parameters change when using Tobit compared to the OLS results. Outcomes suggest how important is to consider the truncated distribution in the estimation.

Overall, our outcomes suggest that larger firms tend to invest more in innovation as well as those foreign oriented (exporters and/or importers) and, on average, with more educated workers. Considering the variables of interaction with the business environment, firms cooperating with other economic agents and receiving government support tend to invest more on innovation inputs, which is an expected outcome. As for facing higher obstacles, firms tend to invest more, which seems a contradiction, yet our interpretation is that perhaps those facing higher barriers to innovate are those that actually spend more on innovation. As shown in (Cânedo-Pinheiro, Lage de Sousa, and Barcelos 2019), endogeneity is also an issue as we are including firms which do not innovate and results change when you consider only potential innovators rather than all available firms.[33] Since our focus here is on workforce diversity rather than on barriers itself, our strategy is to consider them in our estimation so that our predicted expenditure for all firms (not only potential innovators) does not neglect barriers faced by all firms in the survey.

Firm's internationalization (measured as those companies who shares foreign capital), Firm's age (as in year=2001) and market concentration (Herfindahl-Hirschman) are not significant. All results include sector, year and their jointly sector-year fixed effect on firm's outcome.

10.2 Instrumental Variable approach for Diversity

As discussed previously, there are some questions to whether our diversity measure is strictly exogenous. Therefore, we will address this issue by using policy, region and sector information to control for these characteristics which could be only related to innovation through diversity measures, as discussed in the methodology section. In this section, we will present the results of this first stage as well as how it affects our benchmark outcomes assuming that diversity indexes are exogenous.

[33] See (Pellegrino and Savona 2017) for further discussion on this issue.

10.3 Gender Diversity – Maternity Leave

Table C.2 shows whether maternity leave is associated to our gender diversity index. Initially, our gender diversity index is positively associated with firm's size (number of workers) and its workers' skills (workers' educational level). The younger is the firm, more diverse in terms of gender is. However, for maternity leave, our outcomes suggest that when a firm does not have this benefit, firms tend to be less diverse, which contradicts our initial assumption. Nevertheless, our R-square and F-Stat provide some reasonable results suggesting that our instrument is strongly correlated to our endogenous variable.

Table C.2: Maternity Leave extension policy and Gender Diversity

Maternity Leave	-0.00735***
	(0.00142)
Number of Employees	0.00457***
	(0.000327)
Workers' Education	0.0205***
	(0.000187)
Firm's Age	-0.000121***
	(3.08e-05)
Observations	393,057
R-squared	0.281
F-Stat	230.0
Year	yes
Sector	yes
State	yes
Year x Sector	yes
Year x State	yes
FE	no

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

10.4 Gender Diversity - Daycare coverage ratio

Table C.3 shows whether daycare coverage is positively correlated to our gender diversity index. Daycare coverage is positively correlated with gender index, as well as firm's size and workforce educational level. Firm's age remains negativeas expected. Moreover, outcomes also suggest that our instrument seems valid (R-Squared and F-Stat).

Table C.3: Daycare Coverage and Gender Diversity

Daycare Coverage	0.00105***
	(0.000216)
Number of Employees	0.00410***
	(0.000377)
Workers' Education	0.0202***
	(0.000231)
Firm's Age	-0.000250***
	(3.64e-05)
Observations	244,710
R-squared	0.297
F-Stat	153.9
Year	yes
Sector	yes
State	yes
Year x Sector	yes
Year x State	yes
FE	no

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

10.5 Gender Diversity - Divorce rate

Results using divorce rate as an instrument do not change previous results on the covariates (firm's size, workers' education and firm's age), see Table C.4. As for divorce rate, our outcomes suggest a positive correlation between divorce rate and gender diversity. Additionally, evidence on whether instrument is valid (R-Squared and F-Stats) shows that our strategy to use it as an instrument seems plausible.

Table C.4: Divorce Rate and Gender Diversity

Divorce Rate	0.00576***
	(0.000219)
Number of Employees	0.00316***
	(0.000388)
Workers' Education	0.0199***
	(0.000243)
Firm's Age	-0.000295***
	(3.78e-05)
Observations	224,141
R-squared	0.302
F-Stat	143.6
Year	yes
Sector	yes
State	yes

Year x Sector	yes
Year x State	yes
FE	no

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

10.6 Age diversity - Vocational training: the Brazilian Apprenticeship Program (2005)

Outcomes from the estimates from Apprenticeship Program is presented in Table C.5. Overall, this program is positively associated with age diversity in Brazilian firms, which means that the policy is working substantially well in order to promote a more heterogeneous and diverse workplace. As in the previous instruments, our stats show that Apprenticeship Program works reasonably well as an instrument.

Table C.5: Apprenticeship Program and Age Diversity

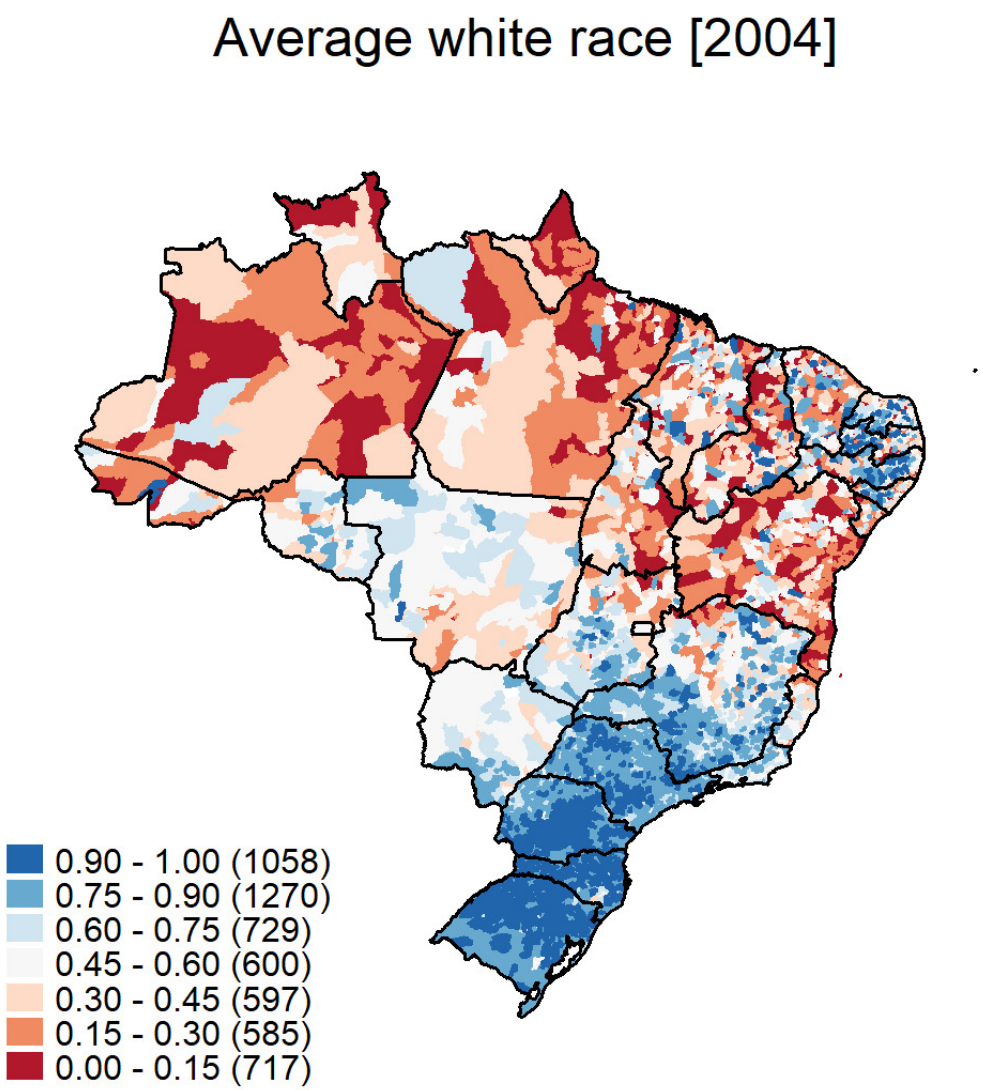
Divorce Rate	0.00459***
	(0.000355)
Number of Employees	0.00618***
	(0.000180)
Workers' Education	-0.00952***
	(9.83e-05)
Firm's Age	0.000230***
	(1.65e-05)
Observations	393,057
R-squared	0.123
F-Stat	82.52
Year	yes
Sector	yes
State	yes
Year x Sector	yes
Year x State	yes
FE	no

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

10.7 Racial diversity – Controlling for Geographical and Sector Variation

For racial diversity, we do not have an exogenous intervention which might interfere in the workforce diversity in terms of race. However, other exogenous information might be correlated to the racial diversity, such as firms geographic location. Figure C.1 shows the percentage of white workers across Brazilian municipalities.

Figure C.1: Percentage of White Workers at Municipality Level (2004)



As observed, while white workers are the majority in the South, non-white workers predominate in the North. Thus, racial diversity index should be controlled for geographical information as firms located in the South might have more white workers due to their regional labor supply.

Table C.6 presents the results for racial diversity index and its covariates. Outcomes suggest that larger firms tend to have higher levels of diversity in terms of race as well as older firms. Higher levels of workers' education tend to be associated with lower levels of racial diversity. As for the stats about the validity of the instrument, our

outcomes suggest that our covariates seem reasonably correlated to racial diversity since R-squared and F-Stats present adequate numbers.

Table C.6: Regional and Sector Dummies and Racial diversity

Number of Employees	0.0253***
	(0.000422)
Workers' Education	-0.00497***
	(0.000242)
Firm's Age	0.00130***
	(4.00e-05)
Observations	393,057
R-squared	0.185
F-Stat	133.4
Year	yes
Sector	yes
State	yes
Year x Sector	yes
Year x State	yes
FE	no

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

10.8 Second Stage of CDM Model

Table C.7: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index (Multivariate Probit) – Full Covariates

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Gender Diversity	0.365***	0.105**	0.0958**	0.756***	0.363***	0.0625	0.0553	0.751***
	(0.0847)	(0.0473)	(0.0418)	(0.107)	(0.0964)	(0.0493)	(0.0384)	(0.103)
Firm´s Size	-0.0258*	-0.0676***	0.0513***	-0.0272	-0.0268	-0.0679***	0.0522***	-0.0363**
	(0.0150)	(0.0200)	(0.00698)	(0.0181)	(0.0182)	(0.0197)	(0.00825)	(0.0169)
Education Level	0.0409***	-0.0621***	0.0416***	0.0289*	0.0425***	-0.0558***	0.0390***	0.0255*
	(0.0157)	(0.00482)	(0.00410)	(0.0156)	(0.0145)	(0.00550)	(0.00459)	(0.0154)
Firm´s Age	-0.000181	-0.00428***	-0.00285***	-0.000223	0.000889	-0.00341***	-0.00302***	-0.000265
	(0.000823)	(0.000526)	(0.000356)	(0.000902)	(0.000857)	(0.000579)	(0.000392)	(0.000910)
Obstacles	0.0796***	-0.107***	0.163***	0.114***	0.0922***	-0.0968***	0.159***	0.104***
	(0.0174)	(0.0318)	(0.00992)	(0.0149)	(0.0172)	(0.0299)	(0.0105)	(0.0154)
Cooperation	0.498***	0.0911***	0.193***	0.0790***	0.496***	0.0926***	0.185***	0.0825***
	(0.0359)	(0.0296)	(0.0273)	(0.0249)	(0.0340)	(0.0310)	(0.0258)	(0.0261)
Multnational Status	-0.0439	-0.0155	0.0135	-0.260***	-0.0588*	-0.0256	0.0400	-0.244***
	(0.0289)	(0.0358)	(0.0236)	(0.0306)	(0.0315)	(0.0384)	(0.0262)	(0.0300)
Import Dummy	0.0366	-0.0879***	-0.0464**	-0.0133	0.0717**	-0.0819***	-0.0453***	0.0135
	(0.0330)	(0.0205)	(0.0183)	(0.0331)	(0.0339)	(0.0232)	(0.0173)	(0.0239)
Export Dummy	0.0648	-0.115***	0.0276*	0.0871*	0.0526	-0.105***	0.00295	0.0671*
	(0.0494)	(0.0269)	(0.0166)	(0.0453)	(0.0461)	(0.0242)	(0.0101)	(0.0393)
Expenditure on Innovation	0.135***	0.230***	0.0699***	0.0661***	0.138***	0.228***	0.0712***	0.0670***
(Predicted)	(0.00242)	(0.0132)	(0.00341)	(0.00289)	(0.00285)	(0.0126)	(0.00345)	(0.00308)
Competition (HHI)	0.739	0.491	-0.285	-0.243	-0.0162	-0.0393	0.00342	-0.0555
	(0.534)	(1.016)	(0.490)	(0.710)	(0.0230)	(0.0241)	(0.0109)	(0.0350)
Constant	-16.19***	-0.422**	-0.886***	-1.207***	-7.408***	-0.663***	-1.233***	-1.125***
	(0.139)	(0.209)	(0.126)	(0.234)	(0.284)	(0.0800)	(0.0617)	(0.217)
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.8: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by Maternity Leave (Multivariate Probit) – Full Covariates

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Gender Diversity IV	0.363	0.126	-0.0598	1.191***	0.363	-0.266	-0.421*	0.706**
	(0.505)	(0.318)	(0.298)	(0.385)	(0.290)	(0.172)	(0.217)	(0.292)
Firm´s Size	-0.0282*	-0.0689***	0.0517***	-0.0356*	-0.0288	-0.0658***	0.0537***	-0.0428**
	(0.0153)	(0.0196)	(0.00742)	(0.0185)	(0.0185)	(0.0198)	(0.00879)	(0.0176)
Education Level	0.0407*	-0.0634***	0.0445***	0.0210	0.0432**	-0.0494***	0.0479***	0.0255
	(0.0220)	(0.00849)	(0.00642)	(0.0198)	(0.0185)	(0.00745)	(0.00592)	(0.0168)
Firm´s Age	-0.000372	-0.00433***	-0.00288***	-0.000527	0.000740	-0.00346***	-0.00296***	-0.000634
	(0.000856)	(0.000518)	(0.000363)	(0.000963)	(0.000915)	(0.000555)	(0.000400)	(0.000989)
Obstacles	0.0849***	-0.106***	0.166***	0.119***	0.0952***	-0.0982***	0.160***	0.110***
	(0.0172)	(0.0327)	(0.00997)	(0.0154)	(0.0172)	(0.0284)	(0.0105)	(0.0153)
Cooperation	0.495***	0.0901***	0.191***	0.0747***	0.491***	0.0882***	0.182***	0.0768***
	(0.0364)	(0.0307)	(0.0273)	(0.0256)	(0.0342)	(0.0309)	(0.0250)	(0.0267)
Multnational Status	-0.0489*	-0.0141	0.0142	-0.267***	-0.0674**	-0.0278	0.0342	-0.249***
	(0.0289)	(0.0368)	(0.0234)	(0.0314)	(0.0317)	(0.0381)	(0.0267)	(0.0302)
Import Dummy	0.0418	-0.0855***	-0.0454**	-0.00482	0.0738**	-0.0800***	-0.0400**	0.0199
	(0.0356)	(0.0209)	(0.0183)	(0.0362)	(0.0345)	(0.0229)	(0.0169)	(0.0267)
Export Dummy	0.0735	-0.113***	0.0312*	0.101**	0.0622	-0.104***	0.00533	0.0836**
	(0.0481)	(0.0270)	(0.0166)	(0.0470)	(0.0459)	(0.0243)	(0.0104)	(0.0408)
Expenditure on Innovation	0.135***	0.231***	0.0699***	0.0660***	0.138***	0.229***	0.0713***	0.0668***
(Predicted)	(0.00250)	(0.0131)	(0.00337)	(0.00279)	(0.00293)	(0.0126)	(0.00338)	(0.00301)
Competition (HHI)	0.571	0.376	-0.373	-0.268	-0.0125	-0.0396*	0.000308	-0.0479
	(0.504)	(0.994)	(0.493)	(0.681)	(0.0273)	(0.0223)	(0.0103)	(0.0380)
Constant	-1.518***	0.933***	-1.074***	-0.955***	-1.541***	0.768***	-1.015***	-0.965***
	(0.140)	(0.224)	(0.0917)	(0.141)	(0.124)	(0.149)	(0.0653)	(0.157)
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.9: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by Daycare Coverage (Multivariate Probit) – Full Covariates

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Gender Diversity IV	0.546	0.131	-0.130	1.141***	0.578*	-0.169	-0.493*	0.722**
	(0.538)	(0.323)	(0.317)	(0.383)	(0.302)	(0.248)	(0.279)	(0.331)
Firm´s Size	-0.0296**	-0.0686***	0.0519***	-0.0349*	-0.0369*	-0.0845***	0.0449***	-0.0391**
	(0.0150)	(0.0190)	(0.00711)	(0.0181)	(0.0201)	(0.0197)	(0.00934)	(0.0170)
Education Level	0.0366	-0.0623***	0.0465***	0.0216	0.0358**	-0.0522***	0.0518***	0.0245
	(0.0231)	(0.00886)	(0.00698)	(0.0205)	(0.0180)	(0.00747)	(0.00752)	(0.0170)
Firm´s Age	-0.000303	-0.00413***	-0.00287***	-0.000478	0.000701	-0.00337***	-0.00327***	-0.000540
	(0.000855)	(0.000549)	(0.000358)	(0.000951)	(0.000786)	(0.000609)	(0.000515)	(0.000935)
Obstacles	0.0821***	-0.108***	0.165***	0.116***	0.0719***	-0.111***	0.170***	0.115***
	(0.0171)	(0.0310)	(0.0100)	(0.0156)	(0.0166)	(0.0292)	(0.0134)	(0.0153)
Cooperation	0.494***	0.0904***	0.193***	0.0742***	0.488***	0.119***	0.182***	0.0884***
	(0.0358)	(0.0305)	(0.0265)	(0.0253)	(0.0321)	(0.0333)	(0.0258)	(0.0226)
Multnational Status	-0.0499*	-0.0198	0.0102	-0.273***	-0.0702**	-0.0514	0.0197	-0.256***
	(0.0292)	(0.0367)	(0.0241)	(0.0326)	(0.0344)	(0.0463)	(0.0290)	(0.0302)
Import Dummy	0.0403	-0.0880***	-0.0456**	-0.00521	0.0715*	-0.0801***	-0.0428**	0.0176
	(0.0355)	(0.0201)	(0.0183)	(0.0350)	(0.0423)	(0.0233)	(0.0176)	(0.0252)
Export Dummy	0.0738	-0.111***	0.0321**	0.103**	0.0756	-0.104***	-0.00225	0.0870**
	(0.0478)	(0.0253)	(0.0163)	(0.0462)	(0.0487)	(0.0280)	(0.0152)	(0.0371)
Expenditure on Innovation	0.135***	0.231***	0.0701***	0.0660***	0.139***	0.232***	0.0732***	0.0673***
(Predicted)	(0.00242)	(0.0131)	(0.00327)	(0.00270)	(0.00257)	(0.0114)	(0.00335)	(0.00333)
Competition (HHI)	0.655	0.413	-0.225	-0.128	-0.0216	-0.0359*	-5.92e-05	-0.0550
	(0.531)	(0.980)	(0.490)	(0.666)	(0.0295)	(0.0195)	(0.0130)	(0.0470)
Constant	-13.17***	-0.341	-0.881***	-1.252***	-16.67***	-0.494***	-1.143***	-1.356***
	(0.140)	(0.226)	(0.132)	(0.225)	(0.135)	(0.0997)	(0.0827)	(0.258)
Observations	48,612	48,612	48,612	48,612	37,984	37,984	37,984	37,984

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.10: Second Stage - Innovation Outcomes Determinants Using Gender Diversity Index Instrumented by Divorce Rate (Multivariate Probit) – Full Covariates

Gender	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Gender Diversity IV	0.821*	-0.484*	-0.232	1.604***	0.727***	-0.525**	-0.581**	0.844**
	(0.428)	(0.251)	(0.267)	(0.381)	(0.280)	(0.235)	(0.242)	(0.330)
Firm´s Size	-0.0298**	-0.0644***	0.0505***	-0.0358*	-0.0349*	-0.0770***	0.0452***	-0.0364**
	(0.0150)	(0.0205)	(0.00788)	(0.0184)	(0.0193)	(0.0198)	(0.0100)	(0.0180)
Education Level	0.0285	-0.0507***	0.0496***	0.00995	0.0309*	-0.0431***	0.0569***	0.0230
	(0.0194)	(0.00679)	(0.00588)	(0.0193)	(0.0165)	(0.00721)	(0.00678)	(0.0172)
Firm´s Age	-0.000310	-0.00415***	-0.00282***	-0.000411	0.000676	-0.00337***	-0.00328***	-0.000568
	(0.000874)	(0.000533)	(0.000397)	(0.000891)	(0.000839)	(0.000650)	(0.000507)	(0.000883)
Obstacles	0.0751***	-0.104***	0.160***	0.116***	0.0617***	-0.116***	0.164***	0.110***
	(0.0186)	(0.0337)	(0.00995)	(0.0159)	(0.0179)	(0.0302)	(0.0133)	(0.0163)
Cooperation	0.490***	0.0946***	0.190***	0.0641**	0.483***	0.120***	0.180***	0.0747***
	(0.0362)	(0.0303)	(0.0270)	(0.0262)	(0.0316)	(0.0326)	(0.0274)	(0.0220)
Multnational Status	-0.0489	-0.0179	0.00892	-0.270***	-0.0630*	-0.0473	0.0130	-0.256***
	(0.0299)	(0.0380)	(0.0233)	(0.0305)	(0.0340)	(0.0456)	(0.0298)	(0.0301)
Import Dummy	0.0390	-0.0739***	-0.0396**	-0.0120	0.0597	-0.0776***	-0.0469***	0.00535
	(0.0326)	(0.0201)	(0.0172)	(0.0337)	(0.0375)	(0.0222)	(0.0181)	(0.0229)
Export Dummy	0.0714	-0.108***	0.0273	0.0991**	0.0779*	-0.0992***	-0.00472	0.0906***
	(0.0478)	(0.0244)	(0.0181)	(0.0434)	(0.0464)	(0.0245)	(0.0154)	(0.0340)
Expenditure on Innovation	0.137***	0.227***	0.0701***	0.0656***	0.141***	0.229***	0.0738***	0.0672***
(Predicted)	(0.00245)	(0.0133)	(0.00369)	(0.00267)	(0.00270)	(0.0112)	(0.00374)	(0.00317)
Competition (HHI)	0.338	0.329	-0.470	-0.499	-0.0242	-0.0489***	-0.000688	-0.0493
	(0.658)	(0.926)	(0.530)	(0.505)	(0.0227)	(0.0178)	(0.0162)	(0.0413)
Constant	-1.442***	0.934***	-1.045***	-0.868***	-7.433***	-0.221**	-1.135***	-1.167***
	(0.146)	(0.221)	(0.101)	(0.135)	(0.291)	(0.0867)	(0.0797)	(0.247)
Observations	45,662	45,662	45,662	45,662	35,662	35,662	35,662	35,662

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.11: Second Stage - Innovation Outcomes Determinants Using Age Diversity Index (Multivariate Probit) – Full Covariates

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Age Diversity	-0.0583	0.266***	0.0582	0.129	0.0307	0.211**	0.00595	0.102
	(0.122)	(0.0880)	(0.0882)	(0.0956)	(0.131)	(0.0893)	(0.102)	(0.0930)
Firm´s Size	-0.0263*	-0.0718***	0.0504***	-0.0317*	-0.0283	-0.0680***	0.0508***	-0.0400**
	(0.0154)	(0.0205)	(0.00742)	(0.0184)	(0.0184)	(0.0191)	(0.00884)	(0.0174)
Education Level	0.0463***	-0.0564***	0.0450***	0.0449***	0.0497***	-0.0513***	0.0406***	0.0410***
	(0.0162)	(0.00541)	(0.00433)	(0.0159)	(0.0149)	(0.00594)	(0.00439)	(0.0156)
Firm´s Age	-0.000267	-0.00429***	-0.00288***	-0.000448	0.000736	-0.00356***	-0.00301***	-0.000589
	(0.000853)	(0.000525)	(0.000360)	(0.000945)	(0.000912)	(0.000573)	(0.000409)	(0.000951)
Obstacles	0.0840***	-0.106***	0.166***	0.120***	0.0948***	-0.0973***	0.160***	0.111***
	(0.0173)	(0.0315)	(0.0103)	(0.0150)	(0.0171)	(0.0289)	(0.0107)	(0.0155)
Cooperation	0.490***	0.0836***	0.192***	0.0717***	0.491***	0.0869***	0.186***	0.0752***
	(0.0356)	(0.0295)	(0.0265)	(0.0250)	(0.0347)	(0.0331)	(0.0247)	(0.0273)
Multnational Status	-0.0523*	-0.00605	0.0139	-0.268***	-0.0627**	-0.0189	0.0394	-0.248***
	(0.0289)	(0.0354)	(0.0244)	(0.0319)	(0.0314)	(0.0376)	(0.0283)	(0.0309)
Import Dummy	0.0421	-0.0885***	-0.0459**	-0.00194	0.0747**	-0.0827***	-0.0437***	0.0264
	(0.0344)	(0.0200)	(0.0180)	(0.0357)	(0.0353)	(0.0229)	(0.0160)	(0.0291)
Export Dummy	0.0752	-0.109***	0.0324**	0.109**	0.0622	-0.105***	0.00579	0.0846**
	(0.0484)	(0.0272)	(0.0165)	(0.0461)	(0.0459)	(0.0237)	(0.0102)	(0.0408)
Expenditure on Innovation	0.135***	0.230***	0.0698***	0.0658***	0.138***	0.228***	0.0711***	0.0664***
(Predicted)	(0.00247)	(0.0131)	(0.00345)	(0.00293)	(0.00288)	(0.0126)	(0.00332)	(0.00293)
Competition (HHI)	0.701	0.554	-0.289	-0.179	-0.0181	-0.0368	0.00518	-0.0562
	(0.502)	(1.021)	(0.486)	(0.729)	(0.0252)	(0.0245)	(0.0108)	(0.0382)
Constant	-9.143***	-0.410**	-0.913***	-1.063***	-7.324***	-0.879***	-1.198***	-1.074***
	(0.158)	(0.199)	(0.138)	(0.249)	(0.277)	(0.0990)	(0.103)	(0.203)
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.12: Second Stage - Innovation Outcomes Determinants Using Age Diversity Index Instrumented by Apprenticeship Program (Multivariate Probit) – Full Covariates

Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Age Diversity IV	1.543***	0.622	0.529	1.411**	1.296**	0.378	0.412	1.218**
	(0.396)	(0.395)	(0.336)	(0.604)	(0.512)	(0.411)	(0.342)	(0.583)
Firm´s Size	-0.0476***	-0.0762***	0.0445***	-0.0481**	-0.0442**	-0.0727***	0.0460***	-0.0553**
	(0.0168)	(0.0212)	(0.00953)	(0.0222)	(0.0200)	(0.0198)	(0.0115)	(0.0220)
Education Level	0.0631***	-0.0532***	0.0498***	0.0588***	0.0632***	-0.0499***	0.0443***	0.0510***
	(0.0148)	(0.00713)	(0.00599)	(0.0178)	(0.0131)	(0.00743)	(0.00566)	(0.0170)
Firm´s Age	-0.000501	-0.00428***	-0.00296***	-0.000683	0.000553	-0.00352***	-0.00311***	-0.000805
	(0.000839)	(0.000526)	(0.000363)	(0.000957)	(0.000890)	(0.000564)	(0.000426)	(0.00102)
Obstacles	0.0818***	-0.107***	0.165***	0.120***	0.0938***	-0.101***	0.158***	0.112***
	(0.0168)	(0.0319)	(0.00968)	(0.0151)	(0.0164)	(0.0284)	(0.0100)	(0.0154)
Cooperation	0.495***	0.0877***	0.192***	0.0711***	0.490***	0.0891***	0.184***	0.0773***
	(0.0352)	(0.0306)	(0.0274)	(0.0262)	(0.0343)	(0.0332)	(0.0257)	(0.0270)
Multnational Status	-0.0435	-0.0149	0.0168	-0.267***	-0.0642**	-0.0297	0.0362	-0.249***
	(0.0288)	(0.0373)	(0.0241)	(0.0318)	(0.0322)	(0.0374)	(0.0268)	(0.0314)
Import Dummy	0.0412	-0.0910***	-0.0473***	-0.00299	0.0763**	-0.0818***	-0.0423**	0.0263
	(0.0344)	(0.0208)	(0.0181)	(0.0345)	(0.0343)	(0.0227)	(0.0173)	(0.0272)
Export Dummy	0.0664	-0.116***	0.0270	0.0972**	0.0549	-0.105***	0.00281	0.0777*
	(0.0471)	(0.0267)	(0.0166)	(0.0487)	(0.0448)	(0.0225)	(0.0113)	(0.0417)
Expenditure on Innovation	0.135***	0.230***	0.0699***	0.0659***	0.138***	0.229***	0.0716***	0.0669***
(Predicted)	(0.00237)	(0.0132)	(0.00341)	(0.00286)	(0.00289)	(0.0124)	(0.00337)	(0.00286)
Competition (HHI)	0.694	0.569	-0.284	-0.338	-0.0130	-0.0308	0.00947	-0.0504
	(0.502)	(1.018)	(0.501)	(0.814)	(0.0262)	(0.0233)	(0.0106)	(0.0394)
Constant	-2.958***	0.356	-1.562***	-2.327***	-8.567***	-0.910***	-1.486***	-1.840***
	(0.353)	(0.374)	(0.304)	(0.564)	(0.359)	(0.330)	(0.263)	(0.518)
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.13: Second Stage - Innovation Outcomes Determinants Using Racial diversity Index (Multivariate Probit) – Full Covariates

Race	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Racial Diversity	-0.0651**	-0.0930***	-0.0113	-0.0229	-0.0904***	-0.131***	-0.0114	-0.0109
	(0.0319)	(0.0280)	(0.0284)	(0.0382)	(0.0283)	(0.0313)	(0.0315)	(0.0399)
Firm´s Size	-0.0248	-0.0651***	0.0512***	-0.0284	-0.0254	-0.0665***	0.0502***	-0.0393**
	(0.0153)	(0.0203)	(0.00678)	(0.0175)	(0.0183)	(0.0197)	(0.00833)	(0.0168)
Education Level	0.0479***	-0.0606***	0.0434***	0.0440***	0.0497***	-0.0534***	0.0401***	0.0398**
	(0.0165)	(0.00526)	(0.00375)	(0.0165)	(0.0156)	(0.00581)	(0.00406)	(0.0159)
Firm´s Age	-0.000235	-0.00419***	-0.00287***	-0.000445	0.000850	-0.00337***	-0.00300***	-0.000561
	(0.000881)	(0.000523)	(0.000365)	(0.000945)	(0.000928)	(0.000579)	(0.000414)	(0.000970)
Obstacles	0.0828***	-0.110***	0.163***	0.116***	0.0948***	-0.0963***	0.159***	0.111***
	(0.0170)	(0.0313)	(0.0102)	(0.0152)	(0.0169)	(0.0286)	(0.0103)	(0.0160)
Cooperation	0.494***	0.0886***	0.191***	0.0740***	0.494***	0.0861***	0.183***	0.0724***
	(0.0357)	(0.0309)	(0.0269)	(0.0255)	(0.0350)	(0.0329)	(0.0253)	(0.0270)
Multnational Status	-0.0507*	-0.0169	0.0128	-0.274***	-0.0664**	-0.0247	0.0371	-0.251***
	(0.0299)	(0.0370)	(0.0233)	(0.0319)	(0.0324)	(0.0389)	(0.0262)	(0.0302)
Import Dummy	0.0434	-0.0889***	-0.0451**	-0.00301	0.0758**	-0.0810***	-0.0443***	0.0245
	(0.0340)	(0.0209)	(0.0183)	(0.0353)	(0.0349)	(0.0240)	(0.0169)	(0.0285)
Export Dummy	0.0733	-0.113***	0.0304*	0.105**	0.0617	-0.105***	0.00605	0.0848**
	(0.0483)	(0.0263)	(0.0164)	(0.0467)	(0.0456)	(0.0245)	(0.0102)	(0.0411)
Expenditure on Innovation	0.135***	0.231***	0.0703***	0.0659***	0.137***	0.228***	0.0716***	0.0667***
(Predicted)	(0.00243)	(0.0132)	(0.00340)	(0.00267)	(0.00292)	(0.0124)	(0.00339)	(0.00291)
Competition (HHI)	0.682	0.642	-0.203	-0.0244	-0.0210	-0.0411*	0.00593	-0.0606
	(0.487)	(1.041)	(0.500)	(0.765)	(0.0253)	(0.0238)	(0.0108)	(0.0387)
Constant	-1.552***	0.933***	-1.065***	-0.994***	-8.795***	-0.625***	-1.193***	-1.009***
	(0.134)	(0.227)	(0.0888)	(0.148)	(0.286)	(0.0834)	(0.0625)	(0.214)
Observations	48,612	48,612	48,612	48,612	44,600	44,600	44,600	44,600

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.14: Second Stage - Innovation Outcomes Determinants Using Racial diversity Index Instrumented by Region&Sector Dummies(Multivariate Probit) – Full Covariates

Race	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Contemporaneous				Lagged Two Years			
Racial Diversity IV –	-0.323**	-0.317***	-0.240***	-0.252	-0.252	-0.299***	-0.231**	-0.198
	(0.139)	(0.0786)	(0.0927)	(0.174)	(0.165)	(0.0947)	(0.0954)	(0.177)
Firm´s Size	-0.0181	-0.0609***	0.0574***	-0.0233	-0.0207	-0.0582***	0.0577***	-0.0335**
	(0.0146)	(0.0202)	(0.00688)	(0.0173)	(0.0186)	(0.0201)	(0.00801)	(0.0167)
Education Level	0.0469***	-0.0612***	0.0437***	0.0434***	0.0499***	-0.0534***	0.0404***	0.0390**
	(0.0165)	(0.00554)	(0.00357)	(0.0165)	(0.0158)	(0.00575)	(0.00400)	(0.0160)
Firm´s Age	9.31e-05	-0.00379***	-0.00256***	-0.000127	0.00110	-0.00311***	-0.00272***	-0.000236
	(0.000951)	(0.000547)	(0.000379)	(0.000916)	(0.00102)	(0.000574)	(0.000442)	(0.000903)
Obstacles	0.0832***	-0.106***	0.167***	0.120***	0.0966***	-0.0948***	0.160***	0.113***
	(0.0175)	(0.0320)	(0.0101)	(0.0154)	(0.0168)	(0.0291)	(0.0107)	(0.0159)
Cooperation	0.494***	0.0834***	0.194***	0.0710***	0.496***	0.0916***	0.184***	0.0758***
	(0.0355)	(0.0299)	(0.0271)	(0.0261)	(0.0343)	(0.0306)	(0.0258)	(0.0265)
Multnational Status	-0.0407	-0.00586	0.0205	-0.265***	-0.0646**	-0.0245	0.0403	-0.253***
	(0.0297)	(0.0376)	(0.0239)	(0.0335)	(0.0324)	(0.0391)	(0.0271)	(0.0319)
Import Dummy	0.0385	-0.0924***	-0.0477***	-0.00620	0.0746**	-0.0842***	-0.0435**	0.0241
	(0.0344)	(0.0217)	(0.0182)	(0.0356)	(0.0352)	(0.0230)	(0.0173)	(0.0262)
Export Dummy	0.0705	-0.116***	0.0276*	0.104**	0.0588	-0.108***	0.00196	0.0817**
	(0.0471)	(0.0261)	(0.0160)	(0.0468)	(0.0450)	(0.0249)	(0.0107)	(0.0415)
Expenditure on Innovation	0.134***	0.230***	0.0694***	0.0656***	0.137***	0.227***	0.0708***	0.0664***
(Predicted)	(0.00238)	(0.0133)	(0.00350)	(0.00268)	(0.00280)	(0.0127)	(0.00343)	(0.00284)
Competition (HHI)	0.760	0.671	-0.266	-0.141	-0.0176	-0.0305	0.00834	-0.0566
	(0.504)	(1.104)	(0.451)	(0.740)	(0.0259)	(0.0241)	(0.0111)	(0.0395)
Constant	-12.38***	-0.200	-0.840***	-0.988***	-17.36***	-0.608***	-1.096***	-0.961***
	(0.293)	(0.228)	(0.111)	(0.245)	(0.139)	(0.110)	(0.0689)	(0.239)
Observations	48,612	48,612	48,612	48,612	44,499	44,499	44,499	44,499

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.15: Second Stage - Innovation Outcomes Determinants Using All Diversity Indexes Contemporaneous (Multivariate Probit) – Full Covariates

Diversity Dependent Variable	(1) Product	(2) Process	(3) Org.	(4) Marketing
Gender Diversity	0.365***	0.0995**	0.0958**	0.763***
	(0.0871)	(0.0489)	(0.0424)	(0.106)
Age Diversity	-0.0901	0.239***	0.0444	0.0720
	(0.123)	(0.0913)	(0.0896)	(0.0967)
Racial Diversity	-0.0637*	-0.0950***	-0.0119	-0.0201
	(0.0336)	(0.0275)	(0.0283)	(0.0398)
Firm´s Size	-0.0235	-0.0672***	0.0518***	-0.0270
	(0.0150)	(0.0205)	(0.00741)	(0.0180)
Education Level	0.0388**	-0.0604***	0.0417***	0.0282*
	(0.0151)	(0.00521)	(0.00439)	(0.0156)
Firm´s Age	-7.43e-05	-0.00412***	-0.00281***	-0.000152
	(0.000820)	(0.000528)	(0.000348)	(0.000903)
Obstacles	0.0794***	-0.107***	0.165***	0.112***
	(0.0172)	(0.0327)	(0.0102)	(0.0145)
Cooperation	0.499***	0.0919***	0.192***	0.0761***
	(0.0351)	(0.0284)	(0.0274)	(0.0255)
Multnational Status	-0.0453	-0.00480	0.0188	-0.254***
	(0.0285)	(0.0354)	(0.0244)	(0.0326)
Import Dummy	0.0357	-0.0909***	-0.0475***	-0.0188
	(0.0329)	(0.0208)	(0.0181)	(0.0328)
Export Dummy	0.0664	-0.110***	0.0284*	0.0913**
	(0.0488)	(0.0272)	(0.0164)	(0.0455)
Expenditure on Innovation	0.135***	0.230***	0.0699***	0.0662***
(Predicted)	(0.00234)	(0.0134)	(0.00343)	(0.00307)
Competition (HHI)	0.537	0.317	-0.355	-0.240
	(0.519)	(1.049)	(0.476)	(0.648)
Constant	-14.87***	-0.375*	-0.899***	-1.188***
	(0.281)	(0.199)	(0.129)	(0.226)
Observations	48,612	48,612	48,612	48,612

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.16: 2nd Stage : Innovation Outcomes Using All Diversity Indexes Contemporaneous IV (Multivariate Probit) Full Covariates

Diversity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Maternity Leave				Daycare Coverage				Divorce Rate			
Gender Diversity IV	-0.717	-0.331	-0.502*	0.631	-0.420	-0.156	-0.532	0.660	0.0412	-0.958***	-0.688**	1.349***
	(0.638)	(0.477)	(0.304)	(0.652)	(0.654)	(0.518)	(0.350)	(0.654)	(0.508)	(0.348)	(0.340)	(0.508)
Age Diversity IV	1.630***	0.538	0.619*	0.853	1.479***	0.379	0.581	0.803	1.383***	0.746	0.797**	0.482
by Apprenticeship program	(0.400)	(0.579)	(0.343)	(0.768)	(0.370)	(0.623)	(0.362)	(0.783)	(0.407)	(0.551)	(0.386)	(0.599)
Racial Diversity IV	-0.227*	-0.301***	-0.216**	-0.180	-0.235*	-0.289***	-0.207**	-0.174	-0.243*	-0.299***	-0.223**	-0.135
by Sector and State level	(0.138)	(0.0911)	(0.0888)	(0.156)	(0.139)	(0.0962)	(0.0887)	(0.154)	(0.129)	(0.0807)	(0.0913)	(0.159)
Firm´s Size	-0.0373**	-0.0639***	0.0522***	-0.0372*	-0.0387**	-0.0649***	0.0525***	-0.0385*	-0.0376**	-0.0618***	0.0485***	-0.0372*
	(0.0161)	(0.0232)	(0.00974)	(0.0220)	(0.0160)	(0.0237)	(0.00955)	(0.0219)	(0.0160)	(0.0234)	(0.0110)	(0.0211)
Education Level	0.0779***	-0.0483***	0.0599***	0.0410	0.0713***	-0.0533***	0.0602***	0.0391	0.0594***	-0.0325**	0.0676***	0.0206
	(0.0258)	(0.0156)	(0.00892)	(0.0298)	(0.0261)	(0.0173)	(0.0103)	(0.0311)	(0.0196)	(0.0127)	(0.0102)	(0.0248)
Firm´s Age	-0.000256	-0.00392***	-0.00264***	-0.000344	-0.000131	-0.00386***	-0.00267***	-0.000290	-0.000192	-0.00399***	-0.00263***	-0.000332
	(0.000918)	(0.000521)	(0.000376)	(0.000900)	(0.000914)	(0.000521)	(0.000398)	(0.000930)	(0.000902)	(0.000533)	(0.000408)	(0.000822)
Obstacles	0.0825***	-0.104***	0.167***	0.120***	0.0822***	-0.106***	0.165***	0.119***	0.0749***	-0.105***	0.159***	0.117***
	(0.0167)	(0.0319)	(0.00993)	(0.0148)	(0.0169)	(0.0320)	(0.00991)	(0.0156)	(0.0184)	(0.0337)	(0.00981)	(0.0163)
Cooperation	0.498***	0.0924***	0.195***	0.0795***	0.495***	0.0871***	0.194***	0.0743***	0.492***	0.0932***	0.189***	0.0659**
	(0.0357)	(0.0300)	(0.0267)	(0.0253)	(0.0362)	(0.0295)	(0.0265)	(0.0263)	(0.0356)	(0.0300)	(0.0274)	(0.0261)
Multnational Status	-0.0428	-0.0130	0.0165	-0.266***	-0.0416	-0.0122	0.0191	-0.265***	-0.0379	-0.00983	0.0155	-0.267***
	(0.0296)	(0.0367)	(0.0248)	(0.0346)	(0.0295)	(0.0376)	(0.0237)	(0.0345)	(0.0312)	(0.0361)	(0.0246)	(0.0337)
Import Dummy	0.0412	-0.0888***	-0.0460**	-0.00443	0.0395	-0.0942***	-0.0476***	-0.00739	0.0381	-0.0821***	-0.0433**	-0.0173
	(0.0348)	(0.0205)	(0.0185)	(0.0347)	(0.0351)	(0.0210)	(0.0181)	(0.0342)	(0.0320)	(0.0201)	(0.0177)	(0.0322)
Export Dummy	0.0659	-0.116***	0.0281*	0.0984**	0.0653	-0.114***	0.0278*	0.0998**	0.0650	-0.112***	0.0213	0.0946**
	(0.0469)	(0.0274)	(0.0167)	(0.0483)	(0.0468)	(0.0259)	(0.0164)	(0.0470)	(0.0464)	(0.0247)	(0.0178)	(0.0451)
Expenditure on Innovation	0.134***	0.229***	0.0686***	0.0645***	0.134***	0.230***	0.0692***	0.0652***	0.136***	0.226***	0.0696***	0.0655***
(Predicted)	(0.00236)	(0.0132)	(0.00342)	(0.00277)	(0.00233)	(0.0134)	(0.00351)	(0.00280)	(0.00238)	(0.0134)	(0.00388)	(0.00279)
Competition (HHI)	0.661	0.580	-0.289	-0.0636	0.609	0.446	-0.335	-0.232	0.482	0.597	-0.358	-0.181
	(0.530)	(0.957)	(0.499)	(0.677)	(0.516)	(1.056)	(0.500)	(0.672)	(0.651)	(0.884)	(0.555)	(0.553)
Observations	48,612	48,612	48,612	48,612	48,612	48,612	48,612	48,612	45,662	45,662	45,662	45,662

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.17: Second Stage - Innovation Outcomes Determinants Using All Diversity Indexes Lagged Two Years (Multivariate Probit) – Full Covariates

Diversity Dependent Variable	(1) Product	(2) Process	(3) Org.	(4) Marketing
Gender Diversity	0.363***	0.0713	0.0618	0.758***
	(0.0972)	(0.0484)	(0.0380)	(0.100)
Age Diversity –	0.00463	0.158*	-0.0119	0.0594
	(0.132)	(0.0961)	(0.104)	(0.0936)
Racial Diversity –	-0.0853***	-0.130***	-0.0141	-0.00985
	(0.0291)	(0.0314)	(0.0316)	(0.0410)
Firm´s Size	-0.0242	-0.0650***	0.0517***	-0.0371**
	(0.0176)	(0.0201)	(0.00877)	(0.0173)
Education Level	0.0427***	-0.0528***	0.0389***	0.0256*
	(0.0140)	(0.00560)	(0.00483)	(0.0146)
Firm´s Age	0.000930	-0.00342***	-0.00299***	-0.000341
	(0.000850)	(0.000564)	(0.000391)	(0.000885)
Obstacles	0.0917***	-0.101***	0.157***	0.101***
	(0.0172)	(0.0292)	(0.0101)	(0.0149)
Cooperation	0.497***	0.0928***	0.184***	0.0837***
	(0.0337)	(0.0323)	(0.0253)	(0.0261)
Multnational Status	-0.0577*	-0.0157	0.0404	-0.238***
	(0.0315)	(0.0377)	(0.0280)	(0.0314)
Import Dummy	0.0686**	-0.0883***	-0.0466***	0.00920
	(0.0332)	(0.0231)	(0.0162)	(0.0253)
Export Dummy	0.0529	-0.107***	0.00406	0.0694*
	(0.0459)	(0.0244)	(0.0100)	(0.0397)
Expenditure on Innovation	0.138***	0.229***	0.0715***	0.0678***
(Predicted)	(0.00284)	(0.0125)	(0.00336)	(0.00295)
Competition (HHI)	-0.0177	-0.0398	0.00639	-0.0524
	(0.0230)	(0.0249)	(0.0107)	(0.0347)
Constant	-1.524***	0.639***	-0.981***	-1.055***
	(0.158)	(0.169)	(0.0918)	(0.142)
Observations	44,600	44,600	44,600	44,600

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.18: 2nd Stage : Innovation Outcomes Using All Diversity Indexes Lagged 2 Years IV (Multivariate Probit) Full Covariates

Diversity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
	Maternity Leave				Daycare Coverage				Divorce Rate			
Gender Diversity IV	0.101	-0.331*	-0.526**	0.516*	0.290	-0.368	-0.623**	0.436	0.458	-0.772***	-0.775***	0.571*
	(0.318)	(0.198)	(0.213)	(0.299)	(0.318)	(0.269)	(0.262)	(0.338)	(0.301)	(0.252)	(0.248)	(0.336)
Age Diversity IV	1.061**	0.163	0.386	0.832	0.949**	0.365	0.263	0.888**	0.992**	0.536	0.362	0.912**
by Apprenticeship program	(0.473)	(0.438)	(0.316)	(0.508)	(0.460)	(0.406)	(0.328)	(0.443)	(0.465)	(0.394)	(0.335)	(0.360)
Racial Diversity IV	-0.182	-0.284***	-0.204**	-0.136	-0.165	-0.321***	-0.214**	-0.148	-0.186	-0.342***	-0.239**	-0.144
by Sector and State level	(0.158)	(0.0922)	(0.0933)	(0.167)	(0.185)	(0.0910)	(0.0960)	(0.173)	(0.182)	(0.0889)	(0.0962)	(0.175)
Firm´s Size	-0.0359*	-0.0591***	0.0549***	-0.0477**	-0.0431*	-0.0790***	0.0492***	-0.0437**	-0.0418*	-0.0734***	0.0486***	-0.0428**
	(0.0198)	(0.0208)	(0.0110)	(0.0206)	(0.0230)	(0.0210)	(0.0117)	(0.0200)	(0.0222)	(0.0211)	(0.0122)	(0.0197)
Education Level	0.0592***	-0.0454***	0.0539***	0.0375*	0.0516***	-0.0435***	0.0576***	0.0414**	0.0469***	-0.0315***	0.0648***	0.0373**
	(0.0188)	(0.0104)	(0.00681)	(0.0197)	(0.0194)	(0.0106)	(0.00816)	(0.0190)	(0.0172)	(0.0106)	(0.00827)	(0.0183)
Firm´s Age	0.000750	-0.00315***	-0.00277***	-0.000553	0.000752	-0.00292***	-0.00305***	-0.000509	0.000721	-0.00307***	-0.00303***	-0.000515
	(0.000978)	(0.000585)	(0.000446)	(0.000932)	(0.000888)	(0.000635)	(0.000519)	(0.000924)	(0.000929)	(0.000672)	(0.000528)	(0.000863)
Obstacles	0.0927***	-0.0991***	0.158***	0.108***	0.0706***	-0.111***	0.171***	0.116***	0.0602***	-0.116***	0.164***	0.109***
	(0.0172)	(0.0286)	(0.0104)	(0.0159)	(0.0161)	(0.0292)	(0.0134)	(0.0156)	(0.0184)	(0.0310)	(0.0131)	(0.0158)
Cooperation	0.491***	0.0892***	0.182***	0.0755***	0.486***	0.112***	0.184***	0.0836***	0.486***	0.122***	0.180***	0.0719***
	(0.0341)	(0.0320)	(0.0259)	(0.0265)	(0.0318)	(0.0330)	(0.0258)	(0.0219)	(0.0315)	(0.0324)	(0.0282)	(0.0216)
Multnational Status	-0.0616*	-0.0251	0.0399	-0.246***	-0.0638*	-0.0413	0.0216	-0.252***	-0.0563	-0.0357	0.0227	-0.247***
	(0.0327)	(0.0382)	(0.0263)	(0.0331)	(0.0356)	(0.0463)	(0.0294)	(0.0318)	(0.0356)	(0.0466)	(0.0302)	(0.0317)
Import Dummy	0.0732**	-0.0870***	-0.0465***	0.0172	0.0707*	-0.0853***	-0.0434**	0.0155	0.0581	-0.0848***	-0.0500***	0.00340
	(0.0348)	(0.0228)	(0.0169)	(0.0267)	(0.0421)	(0.0231)	(0.0179)	(0.0255)	(0.0376)	(0.0225)	(0.0181)	(0.0223)
Export Dummy	0.0547	-0.102***	0.00491	0.0844**	0.0677	-0.109***	-0.00737	0.0786**	0.0716	-0.104***	-0.00998	0.0829**
	(0.0440)	(0.0233)	(0.0114)	(0.0424)	(0.0471)	(0.0270)	(0.0165)	(0.0387)	(0.0441)	(0.0241)	(0.0166)	(0.0335)
Expenditure on Innovation	0.137***	0.227***	0.0709***	0.0659***	0.139***	0.231***	0.0723***	0.0668***	0.140***	0.227***	0.0727***	0.0669***
(Predicted)	(0.00278)	(0.0126)	(0.00327)	(0.00282)	(0.00251)	(0.0115)	(0.00335)	(0.00314)	(0.00266)	(0.0115)	(0.00388)	(0.00345)
Competition (HHI)	-0.0110	-0.0346	0.000950	-0.0466	-0.0250	-0.0334*	0.00177	-0.0551	-0.0258	-0.0519***	-0.00255	-0.0504
	(0.0272)	(0.0214)	(0.0104)	(0.0388)	(0.0298)	(0.0176)	(0.0137)	(0.0468)	(0.0236)	(0.0190)	(0.0168)	(0.0405)
Observations	44,499	44,499	44,499	44,499	37,984	37,984	37,984	37,984	35,662	35,662	35,662	35,662

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

10.9 Impact of Innovation on Productivity

For simplicity, predicted values using different types of innovation are selected to estimate the production function on tables C.8.a and C.8.b . This section reports results for the third stage of CDM model with the complet set of diversity indexes – gender, age and racial - as the following-up stage for Table 5.9.

Overall results implies that even when controlling for diversity on a wide variety of dimensions, the expected positive impact of innovation on productivity is positive. Moreover, difference according to which diversity index is utilized does not interfere substantially in the results, because distances between the estimated parameters are generally within the standard deviation.

Table C.19: Third Stage: Outcomes for Production Function – Age diversity

Dependent Variable: Productivity	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Ln(Number of Employees)	0.0583***	0.0849***	0.0468**	0.0926***	0.0542***	0.0841***	0.0450**	0.0776***	0.0450**	0.0737***	0.0425**	0.0654***	0.0438**	0.0727***	0.0418**	0.0640***
	(0.0182)	(0.0211)	(0.0181)	(0.0186)	(0.0184)	(0.0211)	(0.0182)	(0.0182)	(0.0184)	(0.0208)	(0.0185)	(0.0180)	(0.0179)	(0.0202)	(0.0179)	(0.0175)
Ln(Capital/Employees)	0.117***	0.121***	0.119***	0.121***	0.116***	0.121***	0.119***	0.118***	0.116***	0.121***	0.119***	0.118***	0.116***	0.121***	0.119***	0.118***
	(0.0105)	(0.0112)	(0.0106)	(0.0114)	(0.0103)	(0.0112)	(0.0105)	(0.0108)	(0.00986)	(0.0108)	(0.0102)	(0.0104)	(0.00978)	(0.0107)	(0.0101)	(0.0103)
Gender Diversity lagged two years	0.735***	0.370***	1.209***	0.698***												
	(0.0694)	(0.0462)	(0.126)	(0.126)												
Gender Diversity IV – by Maternity Leave					0.789***	0.378***	1.239***	1.238***								
					(0.0712)	(0.0468)	(0.126)	(0.138)								
Gender Diversity IV – by Daycare Coverage									0.752***	0.357***	1.119***	1.125***				
									(0.0646)	(0.0444)	(0.114)	(0.124)				
Gender Diversity IV – by Divorce Rate													0.749***	0.363***	1.125***	1.155***
													(0.0645)	(0.0443)	(0.118)	(0.128)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,157	43,157	43,157	43,157	43,058	43,058	43,058	43,058	36,852	36,852	36,852	36,852	34,594	34,594	34,594	34,594

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.20: Third Stage: Outcomes for Production Function – Age and Racial diversity

Dependent Variable: Productivity	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing	Product	Process	Org.	Marketing
Ln(Number of Employees)	0.0544***	0.0844***	0.0454**	0.0769***	0.0538***	0.0839***	0.0448**	0.0766***	0.0543***	0.0843***	0.0459**	0.0769***	0.0544***	0.0841***	0.0451**	0.0771***
	(0.0184)	(0.0211)	(0.0182)	(0.0181)	(0.0184)	(0.0211)	(0.0182)	(0.0181)	(0.0184)	(0.0211)	(0.0182)	(0.0180)	(0.0184)	(0.0211)	(0.0182)	(0.0182)
Ln(Capital/Employees)	0.116***	0.121***	0.118***	0.117***	0.116***	0.121***	0.118***	0.117***	0.116***	0.121***	0.119***	0.118***	0.116***	0.121***	0.118***	0.118***
	(0.0102)	(0.0112)	(0.0105)	(0.0107)	(0.0102)	(0.0112)	(0.0105)	(0.0107)	(0.0102)	(0.0112)	(0.0105)	(0.0107)	(0.0103)	(0.0112)	(0.0105)	(0.0108)
Age Diversity –	0.789***	0.376***	1.240***	1.262***												
lagged two years	(0.0710)	(0.0463)	(0.126)	(0.137)												
Age Diversity IV –					0.795***	0.381***	1.245***	1.268***								
by Apprenticeship program					(0.0709)	(0.0470)	(0.125)	(0.138)								
Racial Diversity –									0.789***	0.380***	1.235***	1.263***				
lagged two years									(0.0711)	(0.0468)	(0.126)	(0.138)				
Racial Diversity IV –													0.788***	0.379***	1.241***	1.241***
by Sector and State level													(0.0711)	(0.0468)	(0.125)	(0.137)
Sector Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,157	43,157	43,157	43,157	43,058	43,058	43,058	43,058	43,157	43,157	43,157	43,157	43,058	43,058	43,058	43,058

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1