

# Diversity to foster innovation:

## Using the lens of Brazilian Microdata

---

Filipe Lage de Sousa, Glaucia Ferreira, Leandro Veloso and **Synthia Santana**  
WIDER Development Conference, Transforming economies – for better jobs  
September 11th, 2019 – Bangkok, Thailand

# Contents

---

Context

Main questions and objective

Data and Empirical strategy

Results



# Context

## Labor Market and Growth Challenges

- Low Productivity Growth in Some Developing Countries (especially Brazil)
  - Innovation is a Key Driver for Productivity Growth (Syverson, 2011)
  - Trade Liberalization
  - Labor Market Facts
    - **Women** Participation is Increasing
    - High Levels of **Youth** Unemployment
    - **Racial** Discrimination



- Does Firms' **Workforce Diversity** Play Any Role in **Innovation**?
  - **Benefits:** Complementarities and Spillovers (Huber, 1991; Cox Jr., 2001; Garnero, Kampelmann, and Rycx, 2014);
  - **Costs:** Personal conflicts, communication problems, decreases social similarity & reduces job satisfaction (Becker, 1971; Akerlof and Kranton, 2000; Choi, 2007);
- Empirical Literature: Gender (+), Age (- ou 0) and Race (- ou 0);
- Focused in Developed Countries.



# Data

## Official Statistical Records – Three Sources

### Employer-employee data (yearly)

Worker x firm x year  
(Ministry of Employment and Labor)



reshuffle

firm x year

### Manufacturing survey (yearly)

firm x year  
(Brazilian Statistics Office)

### Innovation Survey - PINTEC

firm x wave  
(Brazilian Statistics Office)

### FINAL DATASET

Number of employers  $\geq 30$   
firm x wave (5)

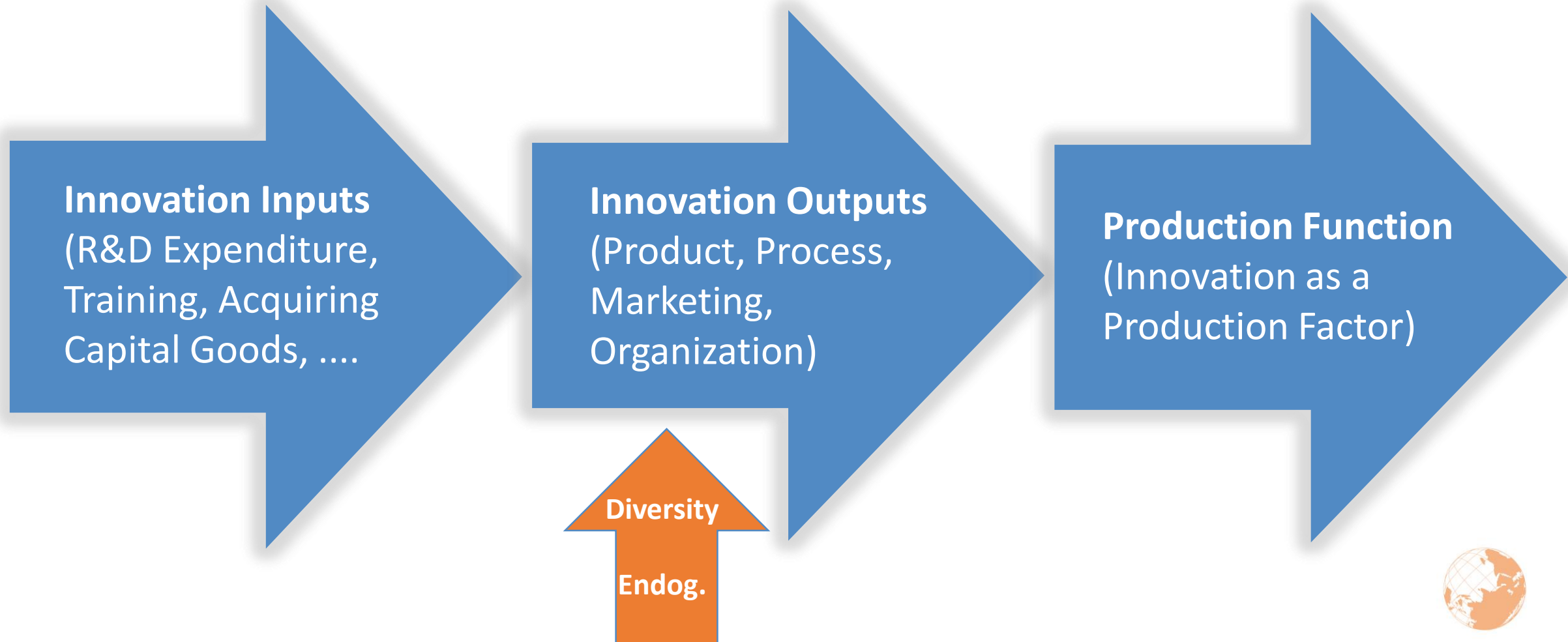
N=48,612 observations  
 $\approx 9,722$  firms by year



# Empirical Strategy

How To Estimate it?

We use the **CDM Model** proposed by Crepon, Duguet, and Mairesse (1998)



# Empirical Strategy

CDM model

Input (innov)  
(1st stage)

$$E_{it} = \begin{cases} E_{it}^* = x_{1it}\beta_1 + \tau\beta_2 + \upsilon\beta_3 + \varepsilon_{1it}, & \text{if } D_{it} = 1 \\ 0, & \text{if } D_{it} = 0 \end{cases}$$

Output(innov)  
(2nd stage)

$$Innov_{it} = \widehat{E}_{it}\gamma_1 + s_{it}\gamma_2 + x_{1it}\gamma_3 + \tau\gamma_4 + \upsilon\gamma_5 + \varepsilon_{2it}$$

Productivity  
(3rd stage)

$$Prod_{it} = \widehat{Innov}_{it}\delta_1 + x_{3it}\delta_2 + \tau\delta_3 + \upsilon\delta_4 + \varepsilon_{4it}$$



### Workforce Diversity and Instruments

- |                         |  |
|-------------------------|--|
| <b>Gender diversity</b> | <ul style="list-style-type: none"><li>• Maternity leave extension (Pro-Woman Firm)</li><li>• Daycare coverage ratio</li><li>• Marriage dissolution</li></ul> |
| <b>Age diversity</b>    | <ul style="list-style-type: none"><li>• Vocational training (Brazilian Apprenticeship Policy)</li></ul>  |
| <b>Racial diversity</b> | <ul style="list-style-type: none"><li>• Sector and Region Dummies</li></ul>  |

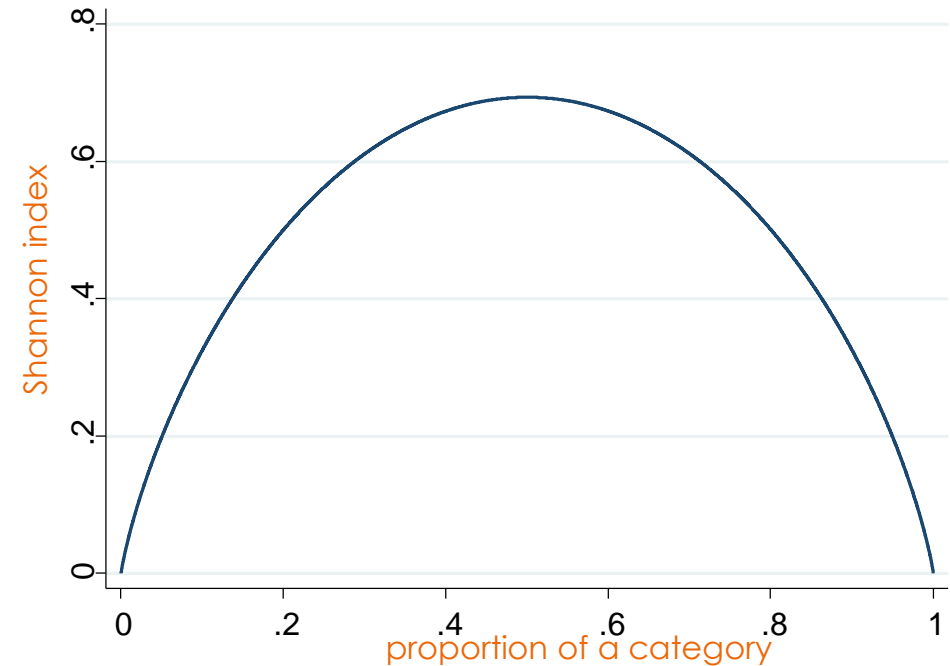




# Diversity measurement

## Shannon-Weaver index

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



Where  $s_i$  is the Shannon-Weaver (1949) diversity index of firm  $i$ , and  $p_{i,r}$  is the proportion of the category or species  $r$  of firm  $i$ . Obviously, the diversity of categories is the highest when  $p_{i,r} = \frac{1}{R}$ .



# Data and sources

## Description of variables

INPUT INNOVATION – 1ST STAGE

| Variable   | Description   | Source |
|--|---|--------|
| Obstacles  | dummy if the firm received some benefit from government     | PINTEC |
| Cooperation  | dummy if the firm cooperated with other company to innovate | PINTEC |
| Government Support                                 | dummy if the firm received some benefit from government     | PINTEC |
| Firm's internationalization                        | dummy if the firm shared foreign capital                    | PINTEC |
| Firm's size (Number of Workers)                    | log of Total #employees on December 31 plus 1 (by firm)     | PIA    |
| Average employees schooling                        | average workers' year of schooling (by firm)                | RAIS   |
| Firm's age   | age of the firm proxied by its oldest registered employee   | RAIS   |
| $\ln(\text{Herfindahl-Hirschman}) (t-2)$           | log of Herfindahl-Hirschman index in t-2                    | PIA    |
| Import status (t-2)                                | dummy if the firm import in t-2                             | SECEX  |
| Export status (t-2)                                | dummy if the firm export in t-2                             | SECEX  |
| $\ln(\text{expenditure in innovative activities})$ | log of total expenditure in innovative activities plus 1    | PINTEC |

# Data and sources

## Description of variables

### INSTRUMENTS AND OVERVIEW

| Variable               | Description  | Source                    |
|------------------------|--|---------------------------|
| Maternity Leave        | dummy from the first year of maternity leave policy onwards                                      | Federal Revenue of Brazil |
| Daycare Coverage ratio | ratio between the 'number of registrations' and 'the population aged 0 to 3 years'               | Abring Foundation         |
| Divorce Rate           | divorces granted at first instance without judicial appeals (by municipality)                    | IBGE                      |
| Male                   | dummy if the worker is male  | RAIS                      |
| Female                 | dummy if the worker is female  | RAIS                      |
| Skilled                | dummy if the worker holds at least a bachelor degree   | RAIS                      |
| Unskilled              | dummy if the worker does not hold a degree   | RAIS                      |
| White                  | dummy if the worker self-declared as white   | RAIS                      |
| Non-white              | dummy if the worker self-declared as non-white (black, indigenous, brown or other dark skinned ) | RAIS                      |

# Results – Maternity Leave

## Multivariate Probit Model – with instruments

| <b>Gender</b>  | (1)            | (2)              | (3)             | (4)              |
|--|----------------|------------------|-----------------|------------------|
| <b>Dependent Variable</b>                              | <b>Product</b> | <b>Process</b>   | <b>Org.</b>     | <b>Marketing</b> |
| <b>Gender Diversity IV (by <b>maternity leave</b>)</b> | <b>0.101</b>   | <b>-0.331*</b>   | <b>-0.526**</b> | <b>0.516*</b>    |
|  | (0.318)        | (0.198)          | (0.213)         | (0.299)          |
| <b>Age Diversity (by Apprenticeship program)</b>       | <b>1.061**</b> | <b>0.163</b>     | <b>0.386</b>    | <b>0.832</b>     |
|  | (0.473)        | (0.438)          | (0.316)         | (0.508)          |
| <b>Racial Diversity (by sector and region dummies)</b> | <b>-0.182</b>  | <b>-0.284***</b> | <b>-0.204**</b> | <b>-0.136</b>    |
|  | (0.158)        | (0.0922)         | (0.0933)        | (0.167)          |
| <b>Sector Dummy</b>                                    | <b>Yes</b>     | <b>Yes</b>       | <b>Yes</b>      | <b>Yes</b>       |
| <b>Year Dummy</b>                                      | <b>Yes</b>     | <b>Yes</b>       | <b>Yes</b>      | <b>Yes</b>       |
| <b>Other Controls</b>                                  | <b>Yes</b>     | <b>Yes</b>       | <b>Yes</b>      | <b>Yes</b>       |
| <b>Observations</b>                                    | <b>44,499</b>  | <b>44,499</b>    | <b>44,499</b>   | <b>44,499</b>    |

Robust standard errors in parentheses

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



# Results – daycare coverage ratio

## Multivariate Probit Model – with instruments

| <b>Gender</b>  | (1)                | (2)                   | (3)                  | (4)                |
|--|--------------------|-----------------------|----------------------|--------------------|
| <b>Dependent Variable</b>                              | <b>Product</b>     | <b>Process</b>        | <b>Org.</b>          | <b>Marketing</b>   |
| <b>Gender Diversity IV (by daycare cov ratio)</b>      | 0.290<br>(0.318)   | -0.368<br>(0.269)     | -0.623**<br>(0.262)  | 0.436<br>(0.338)   |
| <b>Age Diversity (by Apprenticeship program)</b>       | 0.949**<br>(0.460) | 0.365<br>(0.406)      | 0.263<br>(0.328)     | 0.888**<br>(0.443) |
| <b>Racial Diversity (by sector and region dummies)</b> | -0.165<br>(0.185)  | -0.321***<br>(0.0910) | -0.214**<br>(0.0960) | -0.148<br>(0.173)  |
| <b>Sector Dummy</b>                                    | Yes                | Yes                   | Yes                  | Yes                |
| <b>Year Dummy</b>                                      | Yes                | Yes                   | Yes                  | Yes                |
| <b>Other Controls</b>                                  | Yes                | Yes                   | Yes                  | Yes                |
| <b>Observations</b>                                    | <b>37,984</b>      | <b>37,984</b>         | <b>37,984</b>        | <b>37,984</b>      |

Robust standard errors in parentheses

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



# Results – Divorce rate

## Multivariate Probit Model – with instruments

| <b>Gender</b>  | (1)                | (2)                   | (3)                  | (4)                |
|--|--------------------|-----------------------|----------------------|--------------------|
| <b>Dependent Variable</b>                              | <b>Product</b>     | <b>Process</b>        | <b>Org.</b>          | <b>Marketing</b>   |
| <b>Gender Diversity IV (by divorce rate)</b>           | 0.458<br>(0.301)   | -0.772***<br>(0.252)  | -0.775***<br>(0.248) | 0.571*<br>(0.336)  |
| <b>Age Diversity (by Apprenticeship program)</b>       | 0.992**<br>(0.465) | 0.536<br>(0.394)      | 0.362<br>(0.335)     | 0.912**<br>(0.360) |
| <b>Racial Diversity (by sector and region dummies)</b> | -0.186<br>(0.182)  | -0.342***<br>(0.0889) | -0.239**<br>(0.0962) | -0.144<br>(0.175)  |
| <b>Sector Dummy</b>                                    | Yes                | Yes                   | Yes                  | Yes                |
| <b>Year Dummy</b>                                      | Yes                | Yes                   | Yes                  | Yes                |
| <b>Other Controls</b>                                  | Yes                | Yes                   | Yes                  | Yes                |
| <b>Observations</b>                                    | 35,662             | 35,662                | 35,662               | 35,662             |

Robust standard errors in parentheses

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



- Does Firms' **Workforce Diversity** Play Any Role in **Innovation**?
  - Yes!
    - Result is contingent on the type of innovation that the firm aim to invest on.
  - Generally, outcomes for **gender diversity** indicates that marketing innovation presents robust positive evidence;
  - **Gender diversity** seems to be more relevant to promote intangible values (such as brand) than tangible ones (new product).



- **Age diversity:** Both product and marketing innovation are positively related;
- **Racial diversity:** the cost of workforce diversity (miscommunication and background conflicts, for example) surpasses any benefit;
- **Policy implication:** promoting more integration of people with different backgrounds so that the economy benefits from its human assets.





# Thank you !

---

This research was carried out by with technical and financial support from the **Partnership for Economic Policy (PEP)**

Under the **PEP** research and capacity building initiative for  
**“Policy Analysis on Growth and Employment” (PAGE)**

Supported by:



**IDRC | CRDI**

International Development Research Centre  
Centre de recherches pour le développement international

Canada<sup>ca</sup>



**UKaid**  
from the British people

