Do Innovative Firms Pay Higher Wages? Micro Level Evidence from Brazil*

Xavier Cirera and Antonio Soares Martins-Neto†

Abstract

A number of studies have documented a positive and causal relationship between innovation inputs - R&D - and outputs - product or process innovation - and labor productivity. Given the links between labor productivity and wages, a likely implication of this positive relationship is the fact that innovation is associated with higher wages of more productive firms. This paper explores the relationship between innovation and wages using the Brazil employer-employee census (RAIS) and a novel measure of innovation derived from the share of technical and scientific occupations of workers in the firm. The results show a robust and positive wage premium associated to innovative firms. The decomposition of this innovation related wage premium suggests a series of important stylized facts: i) the innovation wage premium is larger for manufacturing but also positive and significant for agriculture and services; ii) it is larger for large firms, but also positive and significant for all firm size categories including micro firms; and iii) it is larger for medium and low skill occupations, although this depends on the use of firm fixed effects. More importantly, the paper explores the causality between innovation and wages and finds empirical support to the idea of both "self-selection" - firms that innovate pay already higher wages before becoming innovator - and increases in wages associated with starting innovation activity; which are persistent over 3 years after firms start innovating.

JEL: O30, O54, J31

Keywords: Innovation; Wage premium; Skills; Brazil

*We thank the participants of the World Bank's FIE Knowledge Sharing Seminar for their useful comments and suggestions.

†Senior Economist and consultant, Finance, Competitiveness and Innovation Global Practice, World Bank. Corresponding author: xcirera@worldbank.org

1 Introduction

Innovation is the engine of creative destruction (Schumpeter et al., 1942) and a key ingredient of economic growth (Romer, 1990; Aghion and Howitt, 1992). At the micro/firm level, a growing empirical literature has documented the impact of innovation on productivity (see Hall (2011) and Mohnen and Hall (2013) for a survey of a large sample of studies in the OECD) and on employment (see Dosi and Mohnen (2018)). Less is known, however, about the relationship between innovation and wages, especially in developing economies where innovation intensity is lower than in advanced economies despite the potential gains from technological catch up (Cirera and Maloney, 2017).

A significant number of studies have analyzed the impact of skill-biased technological change on wages and employment in developed economies (see Acemoglu, 2002; Acemoglu and Autor, 2011). More recently, in the vain of Schumpeterian theory, some studies have analyzed more directly the relationship between innovation, invention and patenting on wages (Reenen, 1996; Aghion et al., 2019, 2018; Akcigit et al., 2017). These studies focus on the impact on wages via the extent of appropriation of the rents created by innovation. An additional channel through which innovation affects wages is via an increased performance of innovative firms. When firms keep some degree of wage-setting in imperfectly competitive labor markets (Mortensen, 2003) and innovation increases firm productivity, workers may benefit from higher wages in innovative firms. Dickens and Katz (1986) find significant sector effects explaining wage setting in the US. Card et al. (2018), reviewing different studies, find that the elasticity of wages to value added per worker is in the range of 0.05-0.15 and, more importantly, firm specific wage premiums can explain up to 20% in total wage variation.[1] So, in summary, if innovation creates improvements in productivity or rents that can be appropriated by workers, one should expect a wage premium associated to innovation.

One important empirical question, however, is to what types of workers and for what types of innovation is there a wage premium. The scant literature on innovation in developing countries finds lower innovation activity in these countries (Cirera and Maloney, 2017), in some cases as a result of lower returns to innovation (Goñi and Maloney, 2017). If these returns are lower, it is unclear what rents or productivity gains can be extracted by workers. Also, if different types of workers have different bargaining powers, one may expect an increase in within firm wage inequality (Aghion, 2002) but not necessarily a positive wage premium in lower skill occupations. Therefore, it is important to understand if this wage premium applies to firms in middle and low-income countries, and to what types of skills

[1] Also, a large literature has analyzed the "decoupling" of wages with productivity (see Feldstein, 2008), which measures how much wage changes depart from labor productivity changes.

and occupations.

A final important element when analyzing the relationship between innovation and wages is the direction of causality. A particular strand of literature has focused on examining the links between incentive systems within firms and employee innovation activities. Although large-scale quantitative evidence is sparse, it appears that performance-based reward systems have a positive influence on the creativity and "innovative behavior" of individual employees (De Jong and Den Hartog (2007); Ederer and Manso (2013); Gibbs, Neckermann, and Siemroth (2017)) and, in some cases, on the firm's overall innovation-related activities and outcomes (Leiblein and Madsen, 2009), including for example its patenting propensity (Chen et al., 2016) and quality (Mao and Weathers, 2019). In addition, human resource management and incentives system play an important role in explaining performance in general (Bloom and Van Reenen, 2011; Bender et al., 2018). Therefore, it is possible that firms increase wages to attract more talented workers and incentivize innovation within firms. This is also consistent with efficiency wage theories (Katz, 1986). Entorf et al. (1999) find for the US and computer-related occupations that these enjoy a wage premium, although these workers were already better paid before introducing the new technologies. Thus, there can be some "self-selection" of firms that pay higher wages to provide incentives to workers that can implement innovative projects.

Thus, in addition to estimate whether there is a wage premium associated with innovation, it is critical to understand whether this is the outcome of firms that "self-select" into innovation by paying higher wages already or, on the other hand, whether these higher wages of workers capturing part of the rents and firm performance gains associated to innovation. The objective of this paper is to shed some light on these questions using a large national employer-employee census (RAIS) in Brazil, and employing a novel measure of innovation based on the share of technical and scientific occupations of the firm (PoTec).

The paper contributes to the literature in three dimensions. First, it contributes to narrow the knowledge gap in relation to estimating the innovation wage premium for a middle-income country like Brazil. Second, it empirically tests the hypothesis that firms pay higher wages before becoming innovative. Third, it measures the extent to which workers appropriate rents created by innovative activities. In the first part of the paper, we estimate the innovation wage premium in Brazil by sector, type of firm and type of skill during the period 2014-2017. To advance some of the main findings, we find a positive innovation wage premium that is robust across different definitions of skills and also across different firm sizes. The premium is higher for manufacturing compared to agriculture and services, and although larger in large firms, it also positive in all firm size types. In the second part of the paper, using the empirical tools of the "self-selection" and "learning-by-exporting" literature

(Bernard and Bradford Jensen, 1999), we find evidence of "self-selection" on innovation activities; innovative firms pay higher wages already in the three previous years to become innovative. In addition, we also find evidence of increasing in the wage premium after becoming innovative; effect that persists three years later.

The paper is structured as follows. The next section briefly describes the literature on wage premiums and the links to the innovation premium. The third section describes the data and methodology used. Section 4 shows the results for the estimated wage premium and its decomposition by type of firm and worker. Section 5 focuses on the causality between innovation and wages. The last section concludes.

2 The Innovation wage premium

The availability in the last two decades of microdata linking firms and workers characteristics has spurred a significant number of empirical studies looking at different wage premiums associated to key workers or firms characteristics. In a perfectly competitive labor market, sector and firms' differences are expected to result in a different composition of workers, but should not influence wages. However, in imperfectly competitive labor markets, firms and workers have some ability to influence wages and capture some of the rents associated with improved performance. Evidence that firms and workers have some ability for wage-setting (Mortensen, 2003) has triggered a large literature trying to estimate what types of workers and firms have larger wage premiums, since these premiums can account for a significant share of wage inequality. For example, Alvarez et al. (2018) document a large decrease in earnings inequality in Brazil from 1992 to 2012 and find that firm premiums account for 40 percent of the total decrease in inequality. This section summarizes some of the most relevant studies that focus on premiums associated with firm size and age, international activity - exporters and FDI - and technology and innovation.

One of the first analysis of wage premium goes back to Moore (1911), who finds that the mean rate of wages in the textile industries increases with the establishments' size. Over the decades, the large-firm wage premium (LFWP) passed the test of many empirical exercises (see Oi and Idson (1999) for a literature review and Bloom et al. (2018) for a recent description of the LFWP).[2] A more recent empirical strand has focused on the wage premium by age group, especially the young-firm pay premium. Initial findings argued that older firms pay more, even when controlling for size and industry (Dunne and Roberts, 1990). Heyman (2007), using matched employer-employee data for Sweden, also finds a positive premium

[2] Berlingieri et al. (2018) argue that, in the service sector, the stylized fact is a "productivity-wage premium" rather than "size-wage premium".

associated with firm's age. Conversely, Adrjan (2018) finds that young firms pay a small premium for new hires. This positive young-firm wage premium is more recently supported by Babina et al. (2019) that show a positive and significant young-firm pay premium for the US that only appears when one controls for the necessary both worker (sorting) and firm fixed effects (employer heterogeneity) (Abowd et al., 1999).

A related literature has focused on the wage premium associated to international activity and find that multinational enterprises and exporting firms pay higher wages than their counterparts. Lipsey and Sjoholm (2003) find that wages in domestically-owned Indonesian manufacturing plants taken over by foreign firms increased sharply after the takeover. Hijzen et al. (2013) find a large and positive wage effect following a foreign acquisition of a domestic firm in Brazil, Indonesia, Germany, Portugal and UK. The authors also find that wage effects are higher in developing countries and that effects on wages after foreign acquisition is restricted to high-skilled workers. In the case of exporting firms, an empirical literature has focused on the impacts of exporting on workers' wages. Brambilla et al. (2017), using comparable data for 61 developing and low-income countries, show a prevalent wage premium in exporting firms compared to their non-exporting counterparts. Schank et al. (2008), using German employer-employee data, also find a wage premium in exporting firms, but argue that it is largely explained by the self-selection of firms that are already more productive and pay higher wages.

Explaining the dynamics and causality of these wage premiums is an important gap in some of the literature, especially regarding the firm size premium, which has primarily focused on identifying the premium size but not on how the premium is created. On the one hand, rent sharing models of the labor market argue that the premium is associated with firms applying more advance technology equipment, innovation projects, or other investments; thus enabling workers to capture some of the rents associated with innovation and other competitive advantages (Reenen, 1996; Dunne and Schmitz, 1995). On the other hand, efficiency-wage arguments suggest large and "better" firms will pay a premium because it is harder to monitor their workers (Krueger and Summers, 1988), provide better working conditions, reduce worker turnover and minimize the risk of their productivity advantage to be captured by their counterparts (Fosfuri et al., 2001). Understanding whether the wage premium is due to rent-sharing or the fact that some firms have already larger wages is important to understand the impact of innovation on income distribution.

A related literature to this paper has focused on measuring the premium associated with technology and innovation. Some of these studies have analyzed the impact of skill-biased technological change on wages and employment in developed economies (see Acemoglu, 2002; Acemoglu and Autor, 2011). In the skill-biased technological change framework, new technologies, changes in the production process, or changes in the organization of work are more complementary to skilled workers, so that shifts in the level of technological capabilities of the economy increases the demand for skilled labor relative to unskilled labor and the wages of skilled labor. While the skill bias technological change impact on wages has received some empirical backup, there is no full consensus on the importance of this effect (Card and DiNardo, 2002), and some studies looking at the use of particular technologies, for example, computers, also find that the premium can be exaggerated if not controlled for the sorting of higher-paid workers. For example, Entorf et al. (1999) studies the effects of computer use on wages and find that workers that use computers are better paid than non-users; however, computer users were already better paid before, with only a small wage increase afterwards.

Looking more directly at the impact of innovation on the wage premium, Reenen (1996) uses a panel of British manufacturing firms to analyze the wage differential between innovative and non-innovative firms. Innovating firms are found to have higher wages due to the rent sharing associated with the R&D projects. Martinez-Ros (2001) also find for a panel of manufacturing firms in Spain a positive wage premium associated with product and process innovation and that, in fact, innovation and wages are jointly determined. Cirillo (2014) estimates the wage premium for Chile using both factor and cluster analysis, as well as panel data techniques, and finds that the innovation premium is not significant when including firms fixed effects. However, as these papers lacked information on workers' characteristics, it is not clear whether the observed premium is due to the sorting of high (low) paid workers into high (low) paying firms or due to firms paying higher (lower) wages to seemingly similar workers. Aghion et al. (2019) are the first to use matched employer-employee data for larger firms in the UK to investigate the innovation wage premium. The authors find a positive innovation premium and larger for low-skilled workers, which in a further theoretical and empirical analysis, the authors associate with increasing complementarity between workers in high-skilled occupations and those in some low-skilled occupations in innovative firms.

Overall, the literature documents several wage premiums associated with firms characteristics. The few existing studies looking at innovative firms suggest a positive wage premium associated with innovation. However, what is less clear from the evidence is what part of this premium is associated to sorting of workers into more innovative firms, what part is explained by innovative firms paying higher wages before turning innovative and what is the change in the wage premium once firms become innovative. These questions are explored empirically in the next sections.

3 Data and Empirical strategy

3.1 Data

To estimate the innovation premium in Brazil, we use the RAIS database (Relação Anual de Informações Sociais) from 2014 to 2017. This is an administrative database from the Brazilian Ministry of Economy considered a high-quality census of the Brazilian formal labor market. The census includes all establishments nationwide with at least one registered worker. We use data on over 40 million employees per year, matched with firm and establishment information, including location and industry. At the individual level, the database includes information on workers' gender, age, education, employment status, type of contract, tenure, data of hiring, among others. We excluded observations where the data was clearly miscoded or missing, as well as individuals younger than 18 or older than 65 years old. We also restricted our sample to one observation per worker-year. We impose this restriction by choosing the highest-paying in any given year. Finally, we have opted for using data at the firm level to account for firm's characteristics.

Due to the large size of the database, we focus on the period 2014 to 2017. This period represents a period of stagnation in the economy with recessions in 2015 and 2016 when GDP growth dropped 3.5% and 3.2%, respectively. Thus, this is a period of stagnation of wages, and it is possible that since the wage premium oscillates with the business cycle, the estimates for the period provide a lower bound of the premium.

3.2 Classification of occupations

The RAIS database also includes information on the occupation of each worker, coded according to the Brazilian Code of Occupations (CBO). To improve comparability, we matched CBO with two other codes, the International Standard Classification of Occupations – ISCO-08 and a match between the National Qualification Framework (NQF) from the UK and the Standard Occupation Code (SOC). These two codes resulted in two different classifications of skills. We follow Autor and Dorn (2013) and Dorn (2009) and classify occupations according to Table 1.

For robustness of the analysis we also follow Aghion et al. (2019), that use the NQF from the UK. The NQF defines 8 levels of skill based on the required qualification from PhD (level 8) to Entry level (less than GCSE grade D-G). The current UK immigration rules use 6 groups (see Table A1). One crucial distinction between these classifications is related to the definition of medium-skill workers. While the previous classification includes most of production and clerical support workers as medium-skill, many manufacturing and other

Table 1: Classification of skills

Skill classification	Occupations	ISCO-08 (2 digits)
High-Skill	Managers	11-14
	Professionals	21-26
	Technicians	31-35
Medium-Skill	Clerical support	41-44
	Sales	52
	Skilled agricultural workers	61-62
Low-Skill	Agricultural workers	63 & 92
	Protective services	54
	Food prep., cleaning, transp.	91 & 93-96
	Personal care/services	51 & 53

maintenance basic occupations are defined as low-skill occupations in the NQF classification. This distinction is essential when interpreting the results since some of the changes in the premium for the same skill category between both classifications just capture differences in how some occupations enter the low skill category.

3.3 Definition of innovative firms

A critical step for the empirical analysis is the definition of innovative firms. One key challenge when working with employer-employee dataset is the lack of firm-level information, which in the case of RAIS is restricted to employment, wages, occupations, and geographical location. Innovation information is often found in other surveys. In the case of Brazil, the national innovation survey PINTEC (Industrial Research on Technological Innovation) is a representative survey for more than 17,000 firms every three years that includes information on firms' expenditure on R&D. The survey, however, only includes manufacturing, services and electricity and does not cover all states of the country. Also, given the probabilistic sampling design and the three-year gap in data collection, the longitudinal dimension is not well defined in the dataset.

Given the existing limitations in the data, a proxy of innovative activity can be developed using the share of occupations dedicated to research and research-related activities. There is an extensive theoretical and empirical tradition looking at innovation and technology from the angle of human capital. Endogenous growth models, such as Romer (1990) and Aghion and Howitt (1992) or more recently Bloom et al. (2020), assume that total factor productivity growth depends on the amount of human capital devoted to research. A number of empirical applications of this tradition has used some measures of human capital as a measure of innovation inputs. For example, OECD (2018) uses the number of scientists and engineers as an indicator for measuring a country's technological human resources. Maloney and Caicedo (2017) present a historical analysis of the impact of engineers density on technological intensification and structural change. de Rassenfosse and van Pottelsberghe (2009) use the number of full-time scientists and engineers as a measure of innovation effort in a patent knowledge function and show evidence of a positive relationship between research effort and patents fillings. Huang and Yu (2011) also use the number of R&D personnel as a measure of in-house R&D, and analyze the effect of competitive and non-competitive R&D collaboration. Finally, Cirillo et al. (2017) use the number of technical-scientific workers as a proxy of innovation when looking at the impact of innovation on within-firm wage inequality in Europe.

We follow this literature and use the measure proposed for Brazil by Araújo et al. (2009) and Taveira et al. (2019) based on the firm's share of technical-scientific employees (PoTec) as a proxy for R&D investments (see Table 2 for definition). Araújo et al. (2009), using RAIS and PINTEC databases, show that PoTec has a high-correlation - between 0.8 to 0.9 - to intramural and extramural R&D expenditure. We also compare the R&D expenditures by states, according to the latest PINTEC survey in 2017, with our PoTec firms by state in 2017 and find a correlation of 0.81. Thus, in the absence of direct data on innovation expenditure for the firms in the census, the PoTec approach appears as a good proxy of R&D expenditure intensity. More importantly, it allows using the employer-employee census to follow firms and workers over time and to match worker's characteristics with those of the innovative an non-innovative firms. Exploiting these dynamics allows exploring transitions from non-innovative to innovative status, which is critical in the causality analysis of innovation and wages.

3.4 Methodology

3.4.1 Measuring the innovation premium

To identify the wage premium associated with innovation a canonical Mincer-type wage equation is estimated:

$$ln(w_{ifjt}) = \alpha_0 + \alpha_1 I_{jt} + + \sum_{n=1}^{n=f} \beta_f X_i t + \beta_9 firmsize_{ft} + \beta_{10} publicservant_{ift} + \theta_t + \sigma_j + \epsilon_{ijft} \qquad (1)$$

where i indexes individual, j region, f firm and t years; w_{ijft} is the hourly wage, which is regressed against our proxy of innovation (innovative) - the share of PoTec workers -and a set of personal characteristics: gender (male = 0), age, tenure, measures of education

Table 2: Classification of PoTec

Occupational group	Professionals
Researchers	Researchers
Engineers	Mechanical and electrical engineers Civil engineers, among others Agronomists and fishing engineers
R&D directors and managers	R&D directors R&D managers
Scientific professionals	Biotechnologists, geneticists, metrology researchers, and specialists in meteorological calibrations Mathematicians, statisticians, among others IT and computer professionals Physicists, chemists, among others Biologists and other professionals

Source: (Araújo et al., 2009)

(respectively, primary school, middle school, high school and college degree or higher), a dummy for whether the employee has a temporary contract, firm's number of employees (firmsize), and a dummy for public servants. θ_t and σ_j represent common time and region effects. The parameter of interest in this study is α_1, the semi-elasticity or wage premium of working in an innovative firm.

One challenge when estimating Equation 1 is that if worker characteristics are not captured adequately by the control variables, and more productive or higher ability workers self-select into innovative firms, then OLS estimates of the innovation wage premium are likely to be biased. In order to control for these unexplained worker factors, we use worker fixed effects:

$$ln(w_{ift}) = \alpha_0 + \alpha_1 I_{jt} + \; + \sum_{n=1}^{n=f} \beta_f X_i t + \beta_9 firmsize_{ft} + \beta_{10} publicservant_{ift} + \lambda_i + \theta_t + \epsilon_{ijft} \tag{2}$$

where λ_i represents time-invariant worker effect. The error term is now decomposed into a time-invariant, person-specific quality component (λ_i) and a purely random element (ϵ_{ijft}). Since we control for individuals and year effects, we omit information on some of workers' characteristics due to additive worker and year fixed effects, remaining with age squared and a dummy for temporary contracts.

A final challenge is the fact that there may be unobserved firm characteristics that are not controlled and are correlated with the decision of becoming an innovator and, therefore, bias the coefficient on the wage premium. The problem, however, is that given that identification of the wage premium with workers and firm fixed effects occurs using information from firms that change innovation during the period - i.e., increase or decrease the number of PoTec workers, firms fixed effects will also absorb part of the innovation premium for those firms that do not change innovation intensity. As a result, in the empirical analysis we compare the estimates of Equation 2 with the results of augmenting it to include firm fixed effects.

$$ln(w_{ift}) = \alpha_0 + \alpha_1 I_{jt} + \ + \sum_{n=1}^{n=f} \beta_f X_i t + \beta_9 firmsize_{ft} + \beta_{10} publicservant_{ift} + \lambda_i + \delta_f + \theta_t + \epsilon_{ijft} \quad (3)$$

where δ_j represents time-invariant firm effects. As discussed in more detail below, the use of worker and firm effects will produce unbiased estimates of the wage premium under the assumption of *exogenous mobility* or lack of sorting in firm-worker pairs.

3.4.2 The causal relationship between innovation and wages

As discussed above, there are two hypotheses on why innovative firms pay higher wages. The first is related to self-selection, as those firms can pay higher wages prior to becoming innovative to, for example, incentivize workers. The other hypothesis assumes that there is a sharing of the quasi-rents from innovation between workers and firms (Reenen, 1996). To test these hypotheses, we build on the empirical literature on foreign wage premia (Hijzen et al., 2013) and the large literature on learning-by-exporting (Haidar, 2012; Yang and Mallick, 2010; Temouri et al., 2013).

To test whether high-paying firms self-select into innovation, we compare workers' of innovative firms with workers in non-innovative firms years before there is a change in their innovative-status, defined by starting to employ a at least a PoTec worker.[3] In this framework, innovative firms are firms that did not have a PoTec employee in time t but have at least one in time $t + l$. Non-innovative in t are firms that neither have PoTec workers in t nor in $t + l$. Specifically, we estimate:

$$ln(w_{ijft}) = \alpha_0 + \alpha_1 I_{t+l} + + \sum_{n=1}^{n=f} \beta_f X_i t + \beta_9 firmsize_{ft} + \theta_t + \sigma_j + \epsilon_{ijft} \quad (4)$$

where I is a dummy for the innovation-status (1 if firm i hires at least one PoTec worker in year $t + l$, where $l \in$ [1,3], and 0 otherwise. w_{ijft} is the log hourly wage, which is regressed against a set of personal and firm characteristics, as in Equation 1.

[3]While when estimating the innovation wage premium in the previous section we used the share of PoTed workers, in the causality framework we use as the innovation variable "treatment" whether the firm has at least one PoTec worker. We also perform some sensitivity analysis increasing the number of PoTec workers necessary to be considered as an innovative firm to two PoTec workers.

In order to test for the possibility that becoming an innovator increases wages by allowing some rent sharing we use a different methodological framework. While the wage premium estimated in Equation 3 is unbiased if both worker and firm effects control for potential endogeneity of innovation, a robust alternative methodology used in the learning-by-exporting literature is the use of propensity score matching (PSM) tools in combination with difference-in-differences (DiD) methods (see Heckman et al., 1997). Specifically, the propensity score matching (PSM) estimates the probability of receiving a given treatment, in this case, becoming an innovator, conditional on the pre-entry characteristics of workers and firms. The probability of becoming an innovator is estimated in Equation 5, where $D = \{0, 1\}$ is a dummy variable indicating whether worker i is employed in a firm that becomes innovative for the first time at t, and X is the multidimensional vector of pre-treatment covariates include workers (wage, gender, age, tenure, type of contract, education, and skill level) and firms (number of workers, average wage, sector,and region) characteristics.

$$p(X) = Pr\left\{D = 1|X\right\} = E\left\{1|X\right\} \tag{5}$$

In the model, we also include the change in wages and employment between $t-2$ and $t-1$ to ensure that matched workers have similar trends in terms of wages and employment growth. Matching is then performed using a single index (the propensity score) that captures all information from the (observable) characteristics. Pairs of treated and control individuals are selected using one-to-one radius matching without replacement.[4] To avoid composition effects, the dataset is balanced, so that individuals are present in each year of our sample - from $t-2$ to $t+2$ - and remain in the same firm for all periods. Also, firms that were innovative in time $t-1$ were excluded. Matching is then complemented with the difference-in-differences estimator. Specifically, the estimated difference-in-differences equation takes the form:

$$ln(w_{it}) = \beta_0 + \beta_1 D_i + \beta_3 T_{it} + \lambda_i + \theta_t + \delta_s + \sigma_j + \epsilon_{ijst} \tag{6}$$

where w_{it} is the hourly wage. D is a dummy for being in the treatment group and T is a treatment indicator that is equal to one if an individual i is treated at time t and zero otherwise. λ_i and θ_t represent individual and time fixed effects, while σ_j and δ_s represent common region and sector effects. It is the significance or otherwise of β_3 that is of primary interest, measuring the average treatment effect on wages from working on a firm that

[4]The matching could result in some poorly matched observations if the closest neighbor is not a good match. One solution is to impose a propensity score distance requirement, namely, the caliper. The value entered in caliper draws a maximum distance of matched firms in the treated and control groups that is closest in terms of the propensity score. In our estimations, the caliper was set to 0.001.

becomes innovative.

3.5 Descriptive Statistics

Table 3 reports summary statistics of key variables comparing innovative - those with at least one PoTec worker - with non-innovative firms in 2017 for all sectors of the economy and excluding public sector institutions. Innovative firms represent only 3% of total firms in Brazil but 43% of private employment in 2017. Innovative firms are significantly larger than non-innovative. The median innovative firms have more than eight times the size of the median non-innovative firm; which is mainly a micro firm. Innovative firms have younger and more educated workers, have a larger share of male workers and also employ a larger share of high-skill occupations.

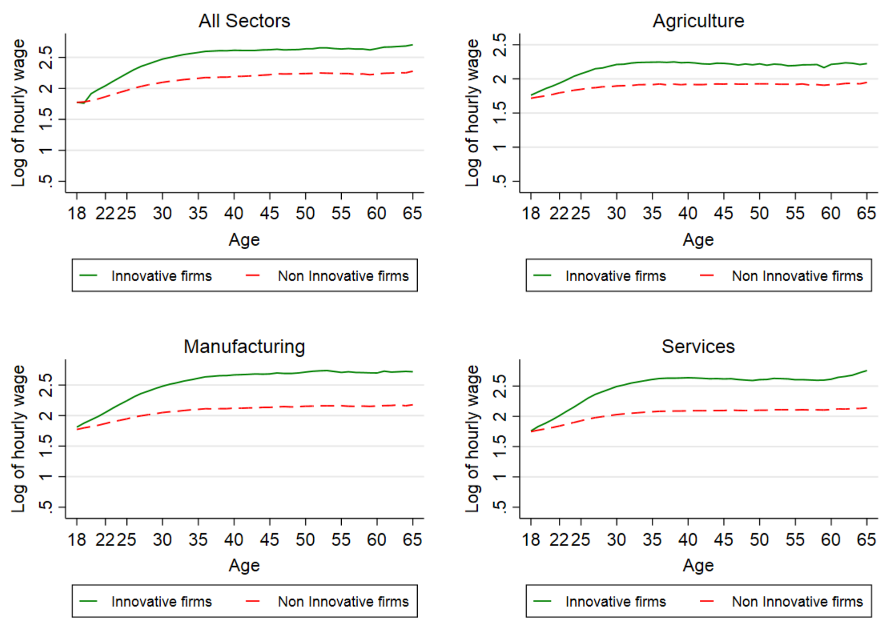
Table 3: Summary statistics of private sector firms by innovative status, 2017

	Innovative	**Non-Innovative**
Log of hourly wage	2,4	1,9
Employment (mean)	187,4	7,3
Employment (median)	25,0	3,0
Age	35,4	35,9
Share of female workers	34,1%	45,4%
Share of high-skill workers	67,4%	40,2%
Share of college graduate	33,2%	7,3%
Average worker tenure	45,1	42,4

Innovative firm is defined as those with at least one PoTec workers in 2017. Employment is the number of workers by innovative status. Tenure is measured in months, and age shows the average wage of workers.

These basic descriptive statistics highlight the fact that innovative firms tend to be much larger than non-innovative firms and employ a more educated labor force. Thus, given the stylized facts established by the LFWP summarized above and the extensive mincerian literature on the returns to education, one should expect that innovative firms also pay higher wages. This is confirmed in Figure 1, which plots the average log of wage per hour on workers' age for innovative firms (green light) and non-innovative firms (red dotted line). The top left panel includes all sectors, also the public sector, and shows higher wages for all age groups for innovative firms, and a wage difference of around 50 log points for workers above 35 years old. Similar results are shown in the panels below for manufacturing and services firms, with also larger wages of workers in innovative firms and a similar wage gap across experience groups. In the case of agriculture (top right panel), the wage gap is also positive but much narrower.

Figure 1: Average wage by innovative status and sector, 2017



Since these results may be simply correlated to the employment structure of those firms, characterized by more educated workers and high-skill occupations, the next section investigates whether the wage premium does still exist when controlling for other factors, such as firm size or workers' characteristics.

4 Results

4.1 The innovation wage premium

We start estimating Equation 1 to measure the innovation wage premium (Table 4) for the period 2014-2017. Columns (1)-(3) show baseline estimates that include a set of mincerian workers' characteristics and year, sector, and region effects. The columns also differ in terms of sample used: column (1) use the complete sample, while columns (2) and (3) restrict the sample to the private and public sector, respectively. The main finding is the existence of positive premium, that is larger in the public sector. Workers in innovative firms earn over 71% more than workers in non-innovative firms, 53% in the private sector and 207% in public organizations.[5] Also, the mincerian variables suggest positive returns to experience,

[5]The impact is measured as the exp(predicted wage at max(innovative)) - exp(predicted wage at mix(innovative)) / exp(predicted wage at min(innovative)). The minimum value is zero (non-innovative firms) and we assume the maximum value is equal to the p99 of the distribution of innovative firms – 0.43

tenure, and education, especially on tertiary education. Finally, consistent with the LFWP literature, workers in larger firms have larger wage premiums.

Table 4: Innovation Premium, 2014-2017

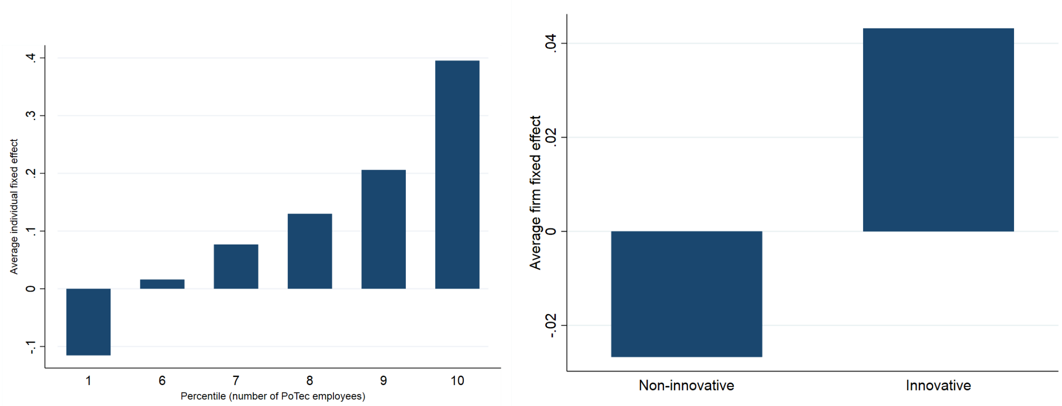
	Dependent variable: $ln(w_{ijkft})$						
	(1) All	(2) Private	(3) Public	(4) All	(5) Private	(6) Public	(7) Private
Innovation	1.254*** (0.107)	0.996*** (0.0496)	2.616*** (0.312)	0.189*** (0.0118)	0.178*** (0.0118)	0.366*** (0.0905)	0.0792*** (0.0246)
Age	0.303*** (0.0128)	0.262*** (0.00289)	0.365*** (0.0618)				
Age Squared				0.249*** (0.0176)	0.235*** (0.00403)	0.486*** (0.143)	0.224*** (0.00383)
Gender	-0.214*** (0.00352)	-0.198*** (0.00191)	-0.233*** (0.0114)				
Primary school graduate	0.0386*** (0.00394)	0.0326*** (0.00269)	0.0457** (0.0210)				
Middle school graduate	0.0837*** (0.00939)	0.0730*** (0.00317)	0.0891* (0.0542)				
High school graduate	0.218*** (0.00774)	0.173*** (0.00391)	0.405*** (0.0342)				
College graduate or higher	0.839*** (0.0181)	0.752*** (0.0116)	1.004*** (0.0466)				
Tenure	0.0930*** (0.00254)	0.0736*** (0.000832)	0.176*** (0.0103)	0.0263*** (0.000579)	0.0292*** (0.000447)	0.00423 (0.00451)	0.0286*** (0.000540)
Temporary	-0.0325 (0.0216)	-0.0117 (0.0140)	-0.171*** (0.0598)	-0.0165*** (0.00588)	-0.0258*** (0.00506)	0.0895*** (0.0273)	0.0537*** (0.0102)
Public sector	0.0953*** (0.0193)			0.0164* (0.00891)			
Firm size	0.0509*** (0.00320)	0.0513*** (0.00207)	0.0521*** (0.0164)	0.0258*** (0.000914)	0.0264*** (0.000813)	0.00498 (0.00494)	0.0311*** (0.00113)
Geo-Occupation-Year	Yes	Yes	Yes				
Individual-Year				Yes	Yes	Yes	
Individual-Year-Firm							Yes
Observations	172519611	136163626	36355985	157682620	122814877	33708892	122340110
R-squared	0.562	0.518	0.504	0.942	0.927	0.954	0.938

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Columns (7)-(9) exclude the public sector. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

The innovation premium in columns (1) and (3) could be explained by the lack of controls for workers sorting and employer heterogeneity. Abowd et al. (1999) show how worker effects can be very important in explaining wage differentials and the fact that more capable workers could sort into better and potentially innovative firms. Columns (4)-(6) estimate Equation 2 and control for worker and year effects; thus providing better control for workers' characteristics. The new results show that the coefficients are reduced significantly to 0.18, 0.17, and 0.36, and the premium on workers earnings is thus reduced to 8%, 8%, and 17%, respectively; between 5 times for private firms and 7 times for public organizations lower premium when controlling for worker effects. This suggests significant sorting of workers to innovative firms, especially in the public sector, but also demonstrates that there is still a positive and statistically significant wage premium associated with innovation.

It is possible that part of the premium is also explained by employer heterogeneity and the fact that we are not able to control for key elements of wage setting in the firm such as the quality of the entrepreneur, the firms' ability to capture rents, or their productivity level. As a result, column (7) estimates Equation 3 with both firms and workers fixed effects as in Abowd et al. (1999). A challenge when also adding firm fixed effects is the fact that identification of the innovation wage premium is only possible when there is some variation in innovation status and the number of PoTec workers, and there is significant persistence over the four year period. As a result, the innovation premium is also captured by the firm fixed effects, and the "true" innovation premium in column (7) is the sum of the innovation wage premium coefficient and the impact of innovation on firm fixed effects.[6] Comparing the estimates with column (5) show that firm fixed effects reduce in more than half the innovation premium, but even when controlling for firms and worker effects, the premium is still positive and statistically significant at 1% confidence level.

Figure 2: Average fixed effect



(a) Individual fixed effect

(b) Firm fixed effect

Note: Figures show the mean of individual and firm fixed effects by percentile of firms' number of PoTec employees. In practice, non-innovative firms correspond to the first nine deciles of firms distribution. Fixed effects are estimated from the regression in Table 5, column 7.

Using the estimated worker effects in column (7) we can establish how large is the sorting of workers to innovative firms. Figure 2a plots for each decile of innovative firms the average

[6] We regress firm fixed effects from column (7) in Table 4 on the mean share of PoTec for each firm. The estimated coefficient using a total of 3.132.403 firm observations is 0.108 with standard error of 0.0029327.

worker effects. Given the large number of non-innovators the first decile represents in practice the first five deciles. The figure suggests that workers in the most innovative firms have 50 log points higher wages than in non-innovative firms. Thus, there is strong sorting of more capable and highly paid workers to more innovative firms that tend to pay higher wages.

Figure 2b plots the average estimated firm fixed effects for innovative and non-innovative - no PoTec workers. The average firm effects are larger for innovative than non-innovative firms by around 7 log points. This suggests that there are unexplained factors at the level of the firm - for example higher productivity, higher rents, better entrepreneurs or that already innovate - that make innovative firm pay higher wages.

The results suggest the existence of a innovation wage premium, which is robust to the inclusion of worker and firm fixed effects. However, obtaining unbiased estimates depends on the estimation of the additive form of firm and workers' fixed effects, which requires *exogenous mobility.* If the mobility of workers that joined innovative firms is due to match and personal specific characteristics beyond what is captured by the worker and firm fixed effects, then estimates in column (7) in Table 4 are likely to be biased. To recover unbiased estimates, Card et al. (2013) show that a sufficient condition is that the assignment of workers across firms obeys a strict exogeneity condition:

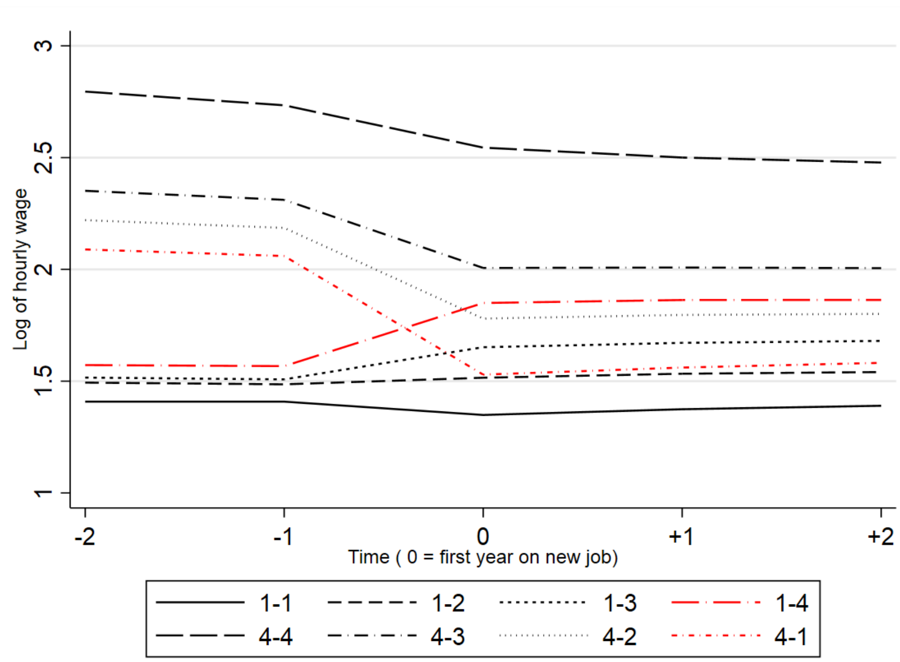
$$\mathbb{P}[J(i,t) = j|\epsilon_{it}] = \mathbb{P}[J(i,t) = j] = G_{jt}\left(\alpha_i, \{\psi_j\}_j\right) \forall i, t \tag{7}$$

We follow Card et al. (2013) and Alvarez et al. (2018) and use an event study framework to test Equation 7 using employees who transition between firms. If the variation in wage across firms is due to sorting, then workers moving across employers should not experience systemic wage changes. On the other hand, if workers moving to high-paying firms experience, on average, a wage gain, while workers moving to low-paying firms experience symmetric wage loss, then we can assume that column (7) provides a reasonable characterization of the mean wages resulting from different matching of workers and firms. We focused on workers who changed employers in year zero (2015) and held the previous job for two or more years and the new job for two or more years in order to investigate the impacts of changing jobs across high- and low-paying firms. Firms were classified based on the mean wage of coworkers. We classify the old job based on the quartile of coworkers in the last year at that job, and the new job based on the first year at that job. We end up with 16 groups based on the quartiles of mean coworker earnings at the origin and destination firms around a move between firms. Figure 3 shows movements from the first quartile to higher quartiles, as well as from the fourth quartile to lower quartiles.[7] The lines in red clearly show that workers moving to

[7]For the sake of clarity, Figure 3 showed only half of the possible transitions. Other transitions showed the same pattern and are available upon request.

low-paying firms experience wage loss and workers moving to high-paying firms present a significant and symmetric wage increase. This symmetry suggests that even without pair match effects, our specification with individual- and firm fixed effects provides a reasonable approximation to the average worker transition in the Brazilian labor market.

Figure 3: Mean wage of job changers classified by quartile of mean wage of coworkers at origin and destination establishment, 2012-2017



Note: Figure shows the mean wage of workers who changed employers in year zero (2015), and held the previous job for two or more years and the new job for two or more years. Each quartile is defined based on the mean wage of coworkers. We classify the old job based on the quartile of coworkers in the last year at that job, and the new job based on the first year at that job. Wages are log normalized to real 2012 Reais. Each number describe the wage quartile; thus 1-1 refers to workers that remained in first quartile, while 1-4 refers to workers that transitioned from the first to the forth quartile.

4.2 The innovation premium by sector

Table 5 shows the results of estimating the wage premium for different sector samples - agriculture, manufacturing, and service.[8] Odd columns use worker fixed effects, while even columns also add firm fixed effects. The results show a larger innovation wage premium for manufacturing than agriculture and services, although when adding firm fixed effects, only the coefficient on the premium for workers in services firms remain statistically significant.

[8]Given the strong persistence in our innovation measure and the fact that as a result firm fixed effects absorb a significant share of the innovation premium, we use worker and year fixed effects as our preferred specification. We also exclude the public sector.

Some of these differences in innovation premium across sectors could be related to having different technological and market structures. For example, there are potentially lower rents created in agriculture due to the production of more homogeneous good or lower bargaining power of workers. But also differences could reflect the fact that the PoTec maybe a less adequate proxy of innovation outside manufacturing, since the PoTec occupations have been selected in relation to their correlation with manufacturing R&D.[9]

Table 5: Private Sector: Innovation Premium per sector, 2014-2017

	Dependent variable: $ln(w_{ijkft})$					
	(1) Agriculture	(2) Agriculture	(3) Manufacturing	(4) Manufacturing	(5) Services	(6) Services
Innovation	0.285***	0.0632	0.401***	0.0838	0.109***	0.0686**
	(0.0515)	(0.0638)	(0.0754)	(0.109)	(0.0121)	(0.0282)
Age Squared	0.117***	0.123***	0.319***	0.317***	0.228***	0.211***
	(0.0115)	(0.0127)	(0.0100)	(0.0102)	(0.00553)	(0.00527)
Tenure	0.213***	0.0227***	0.0384***	0.0367***	0.0230***	0.0239***
	(0.000616)	(0.000690)	(0.000649)	(0.000714)	(0.000646)	(0.000829)
Temporary	0.0242	-0.0776	0.143	0.170	-0.00704	0.0479***
	(0.0629)	(0.0497)	(0.124)	(0.140)	(0.00612)	(0.0104)
Firm size	0.0234***	0.0176***	0.0309***	0.0433***	0.0321***	0.0265***
	(0.00112)	(0.00327)	(0.00124)	(0.00326)	(0.00102)	(0.00146)
Individual-Year	Yes		Yes		Yes	
Individual-Year-Firm		Yes		Yes		Yes
Observations	4625738	4618266	22898875	22870154	82711558	82330700
R-squared	0.915	0.923	0.945	0.948	0.930	0.940

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.3 The innovation premium and the large firm wage premium (LFWP)

A strong implication of the descriptive statistics in Table 3 is the significant difference in terms of firm's size between innovative and non-innovative firms, which raises the concern that previous results may be only capturing the well-established large-firm wage premium (Oi and Idson, 1999; Bloom et al., 2018). Even though in Table 4 there are controls for firm's size, as an additional robustness test, we re-estimate the innovation premium dividing the sample according to firm size groups (see Table 6).

[9]In the case of services for example, the role of design is as or more important than of R&D. See for example Cox (1990) for a discussion of the importance of design on R&D and innovation.

The results confirm the positive and statistically significant innovation wage premium, and the fact that this exists irrespective of size. The premium is larger for large firms, but positively and significant at 1% level for all size groups, even after controlling for within size group firm size and worker effects. Workers in micro firms, with less than 10 workers, also have a wage premium when working in innovative firms. The results, therefore, provide evidence of both an innovation premium and large firm wage premium.

Table 6: Private Sector: Innovation Premium per firm size, 2014-2017

	Dependent variable: (w_{ijkft})			
	(1)	(2)	(3)	(4)
Firm size	Micro (<10)	Small (10-49)	Medium (50-99)	Large (>99)
Innovation	0.0713***	0.132***	0.139***	0.171***
	(0.00476)	(0.00710)	(0.0198)	(0.0258)
Age Squared	0.106***	0.171***	0.242***	0.358***
	(0.00214)	(0.00258)	(0.00796)	(0.0113)
Tenure	0.0167***	0.0229***	0.0244***	0.0294***
	(0.000101)	(0.000132)	(0.000384)	(0.000989)
Temporary	0.00796	-0.00490	-0.0372***	-0.0325***
	(0.0138)	(0.0247)	(0.00688)	(0.00540)
Firm size	0.0242***	0.0368***	0.0588***	0.0172***
	(0.000224)	(0.000498)	(0.00265)	(0.00166)
Individual-Year	Yes	Yes	Yes	Yes
Observations	20453784	24511458	7462266	57445724
R-squared	0.914	0.919	0.940	0.940

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

4.4 Innovation premium by skill level

The literature on skill-biased technical change has emphasized the role of technology in increasing labor polarization and depressing the incomes of low-skilled workers (Acemoglu and Autor, 2011). This empirical literature suggests that the returns on high skill workers should be larger on average than for other skills. One exception to this evidence is Aghion et al. (2019), who develop a model where low and medium-skilled workers that possess soft skills that are complementary to high skill workers in innovative firms. The authors using data from the UK find empirical support to the hypothesis of large complementarities, and a larger wage premium associated with working in an innovative firm for low and medium-skilled workers. Thus, identifying how the wage premium is related to the worker's skill level is critical to understand the distributional impacts of innovation, especially in a country like

Brazil that has experienced a large decrease in earnings inequality in the last decade Alvarez et al. (2018).

In order to identify the skill premium associated with innovation, Table 7 shows the results of estimating Equation 2 but interacting our innovation PoTec variable and the workers' skill level[10] - column (1) - and then by restricting the sample for each skill level – columns (3), (5) and (7).[11] The main result is that, similarly to Aghion et al. (2019), the innovation premium is higher for low- and medium-skill workers. The innovation premium that a medium-skilled worker premium gets is 23 log points higher than the innovation premium of a high-skilled worker; 8 log points higher for low-skilled workers. Columns (3), (5) and (7) confirm these results splitting the sample; low and medium-skilled workers innovation premium remain high, and larger than the innovation premium of high-skilled workers.

These results seem to support the complementarity hypothesis that Aghion et al. (2019) found using UK data and also using a different skill classification. Table A2 shows the results using the same skill classification than in Aghion et al. (2019). The results show some differences in the size of the coefficients, with low-skilled workers, on average, obtaining a higher positive innovation wage premium. This is likely the result that the NQF classification defines many manufacturing and other maintenance basic occupations as low-skill occupations - in Table 7, about 21 million workers were classified as low-skill, against more than 73 million in Table A2.[12] The results, nevertheless, are aligned with Aghion et al. (2019) in supporting a potential complementarity of low and medium skill workers with high skills workers in innovative firms, which imply a larger innovation premium for these lower skills occupations.

One caveat of these findings, however, is that we are not controlling for potential unobserved firm factors. While, as discussed above, these firm fixed effects also capture part of the premium, they also allow controlling for these unobserved firm fixed effects. The estimates using both firm and worker fixed effects are shown in column (2). As before, the innovation wage premium is reduced, highlighting the larger firm fixed wage effect in innovative firms, which are better-paying firms. However, the relative distribution of the premium across skills changes, and now high skilled workers are getting a larger innovation premium. Although these estimates may suffer from limited mobility bias (see Bonhomme et al. (2020)), they suggest the need to look more in-depth on how different skill groups share the innovation

[10] Occupations were classified following Autor and Dorn (2013), as discussed in section 3.B.

[11] Table B1 and Table B2 re-estimates the premium per skill level including both private and public sectors.

[12] Note that the NQF classification is mainly used for immigration purposes, which in many cases is intended to prevent workers to obtaining their right to work in the UK. With this in mind, our preferred definition of skills is the one developed in Autor and Dorn (2013), as it takes into account workers' tasks and wages.

wage premium.

Table 7: Private Sector: Innovation Premium per skill level, 2014-2017

	Dependent variable: $ln(w_{ijkft})$							
	(1) All	(2) All	(3) Low	(4) Low	(5) Medium	(6) Medium	(7) High	(8) High
Innovation	0.154***	0.0964***	0.379***	-0.00315	0.397***	-0.000373	0.115***	0.0827***
	(0.0107)	(0.0244)	(0.0289)	(0.0224)	(0.0561)	(0.0549)	(0.0106)	(0.0257)
X Low-Skill	0.0856***	-0.336***						
	(0.0233)	(0.0241)						
X Medium-Skill	0.226***	-0.0949***						
	(0.0467)	(0.0266)						
Age Squared	0.233***	0.220***	0.0365***	0.0275***	0.169***	0.143***	0.393***	0.384***
	(0.00404)	(0.00380)	(0.00745)	(0.00683)	(0.00418)	(0.00466)	(0.00724)	(0.00723)
Tenure	0.0287***	0.0281***	0.0171***	0.0154***	0.0259***	0.0258***	0.0270***	0.0259***
	(0.000444)	(0.000539)	(0.000283)	(0.000311)	(0.000440)	(0.000520)	(0.000969)	(0.00126)
Temporary	-0.0231***	0.0519***	-0.000112	0.0654***	-0.0487***	0.0186	-0.0269***	0.0417***
	(0.00511)	(0.00987)	(0.00410)	(0.0114)	(0.00691)	(0.0126)	(0.00673)	(0.0102)
Firm size	0.0264***	0.0309***	0.00989***	0.0182***	0.0276***	0.0304***	0.0363***	0.0352***
	(0.000810)	(0.00112)	(0.000817)	(0.00114)	(0.00123)	(0.00142)	(0.000746)	(0.00178)
Low-Skill dummy	-0.0641***	-0.0756***						
	(0.00150)	(0.00132)						
Medium-Skill dummy	-0.0239***	-0.0138***						
	(0.00218)	(0.00206)						
Individual-Year	Yes		Yes		Yes		Yes	
Individual-Year-Firm		Yes		Yes		Yes		Yes
Observations	122814877	122340110	21667647	21481418	40800393	40492276	52325435	52029478
R-squared	0.928	0.938	0.874	0.893	0.891	0.909	0.945	0.953

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

5 The impact of becoming an innovator on wages

The analysis above shows robust evidence of an innovation wage premium in Brazil. The mechanics, however, about how this premium is generated and what happens to wages when firms become innovators is the focus of this section. First, we analyze whether firms pays higher wages before becoming innovators - analogous to the self-selection in the export literature (see Bernard and Bradford Jensen (1999)). Second, we estimate what is the impact of becoming an innovator on wages and how lasting over time is this effect. In what follows, we explore each of the questions separately.

5.1 Do innovative firms pay higher wages before becoming innovators?

The size of the firm fixed effects averages across innovator and non-innovators in Figure 2b suggests that innovative firms are the ones that pay higher wages due to unobserved characteristics and also because of the impact of innovation on these firms. We explore this more formally using the framework used in the learning by exporting literature (see Bernard and Bradford Jensen (1999) and applications by Haidar (2012); Temouri et al. (2013)) and the foreign wage premium (Hijzen et al., 2013). As described in section 3.4.2, the procedure requires to estimate the wage premium for firms before becoming innovators.

Table 8 presents the main results of estimating Equation 4 before firms becoming innovative, controlling for workers' and firms' characteristics. Each column in Table 8 takes a year of transition and compares workers' wages in three different periods, $t-1$, $t-2$, and $t-3$.[13] The results for the innovation wage premium are positive and statistically significant for all years and lagged values. Firms that innovate in t exhibit a positive wage premium even three years before becoming an innovator. The coefficients are not directly comparable to those in previous estimations given that we estimate the effect of a dummy measuring future change in innovative status, but the results suggest that future innovators have around 5 log points higher wages up to $t-3$.

As a robustness test, Table D1 in the Appendix shows the Kolmogorov-Smirnov test for stochastic dominance. The test compares the distribution of wages between future innovators and persistent non-innovators. The $p-values$ reject the hypothesis of a difference in wages favoring non-innovative firms and do not reject the hypothesis of a difference favoring innovative firms, thus indicating that the previous wage distribution of workers in future innovative firms stochastic dominates that of non-innovative firms. Table D2 in the Appendix performs an additional robustness test by using a more stringent innovation status when the firm has at least two workers in PoTec occupations in time t. The results do not change qualitatively and reinforce the previous findings of higher wages in firms that will become innovators.

Summing up, workers employed in firms that become innovative already earn higher wages than those in firms that remain non-innovative. These results show support to the hypothesis that firms may pay higher wages to workers to incentivize performing innovation activities. What is surprising is the strong degree of persistence up to three years before becoming an innovator and the robustness to more stringent measures of innovation. Firms

[13] From 2015 to 2017, between 10.000 and 12.000 firms transitioned from non-innovative to innovative per year. In terms of workers, this represents on average more than 639 thousand workers transitioning each year.

prepare and self-select to innovation using higher wages as incentives.

Table 8: Self-selection

	Treated (t-1)		
Year of transition	**2015**	**2016**	**2017**
Treated	0.0533*** (0.00902)	0.0462*** (0.00954)	0.0520*** (0.0104)
	Treated (t-2)		
Treated	0.0674*** (0.0107)	0.0540*** (0.0138)	0.0664*** (0.0152)
	Treated (t-3)		
Treated	0.0552*** (0.0111)	0.0471** (0.0157)	0.0657*** (0.0156)

Note: Estimation using OLS, controlling for region, and sector effects. The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is a PoTec workers for firm f in year t+1. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

5.2 What is the increase in wages associated with firms becoming innovative?

A critical question is whether workers earn positive premiums after firms become innovative. Regardless of being firms that pay higher wages, it is possible that innovation creates rents that are shared between firms and workers. To test for the appropriation of these rents, we combine PSM and DiD methods as described in section 3.4.2, and explore the wage difference of becoming an innovator in t, $t+1$ and $t+2$.

PSM allows us to match similar workers employed in firms with different innovation statuses, and following the matching, we use DiD and estimate Equation 6 for controlling for any remaining unobservable effects. Table 9 shows the results using the more straightforward definition of transition, moving from zero to one PoTec worker, while Table 10 restricts transitions to at least two PoTec employees. The sample used allows for reverse transitions, firms that dismiss PoTec worker,s and are re-categorized as non-innovator. Removing these reverse transitions could bias the results since they could reflect failed innovation strategies that could result in reduced wages.[14]

The results show evidence of a wage premium for workers once the firm becomes innovative. Table 9 shows that workers earn 2 log points more after their firms becoming innovators compared with workers in similar firms that remain non-innovators. Table 10 shows even

[14] Table E1 and Table E2 shows the results when the sample does not include reverse transitions

larger effects, consistent with using a more stringent innovation measure. Workers capture part of the rents of firms.

The last two columns of both Table 9 and Table 10 show the dynamic effect of innovation on wages. The effect on wages of a firm becoming innovative is positive and lasting, around 2 log points and constant for t and $t+2$. This suggests that workers also appropriate some of the rents or competitive advantages associated to innovation in a way that is persistent over time. Table E1 and Table E2 re-estimates the same models but without allowing transitions from innovators to non-innovators. The results are qualitatively identical.

Table 9: Do firms pay higher wages after becoming innovative?

	y: log of wage			
	(1) Mean	(2) Mean	(3) Dynamic	(4) Dynamic
Mean effect	0.0169*** (0.00327)	0.0168*** (0.00328)		
Effect at t			0.0195*** (0.00345)	0.0193*** (0.00345)
Effect at t+1			0.0181*** (0.00355)	0.0180*** (0.00355)
Effect at t+2			0.0197*** (0.00347)	0.0195*** (0.00347)
Individual-Year	Yes	Yes	Yes	Yes
Region-Sector		Yes		Yes
Observations	949896	949896	949896	949896
R-squared	0.946	0.946	0.946	0.946

Note: The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is a PoTec workers for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

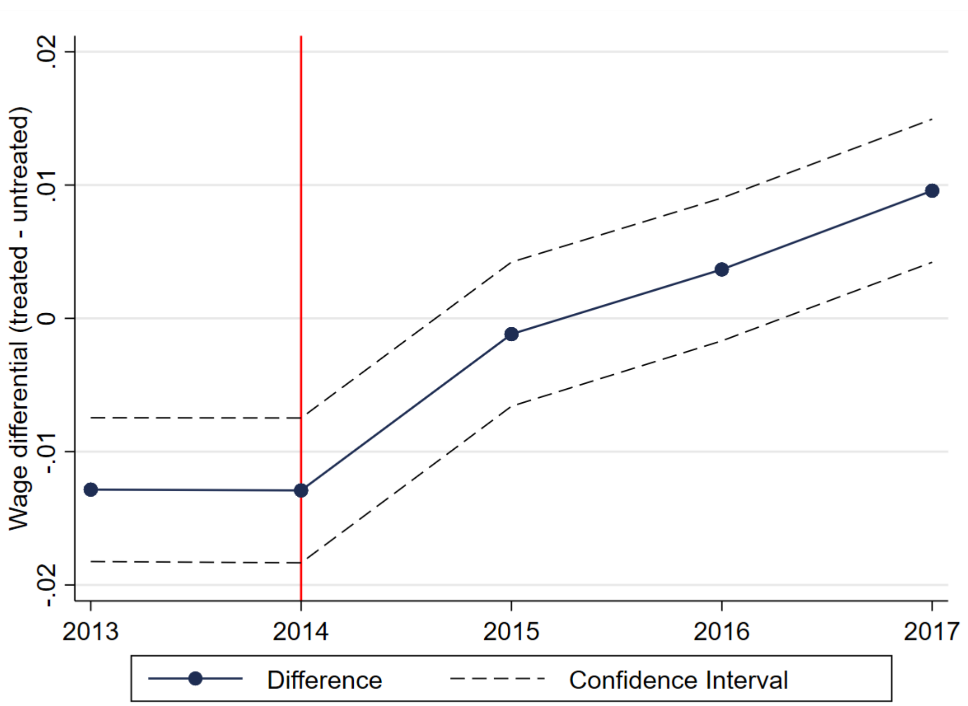
As a robustness check, Figure 4 plots the wage differences between treated and untreated workers of firms that become innovative after 2014. The figure provides some visual reassurance that changes in the wage premium between the two groups start after firms become innovative and last on time. Specifically, this positive effect persist over time on about 2 log points more wages. The figure shows a higher effect on wages on the year of becoming an innovator, and then a lower increase but constant over time. While the period used for the analysis does not allow to analyze the dynamic effect over a longer horizon, it is likely that the persists even longer than $t+2$. Overall, the results provide strong evidence for a causal effect of innovation on a wage premium, supporting other studies in developed countries (Reenen (1996);Martinez-Ros (2001))

Table 10: Do firms pay higher wages after becoming innovative?

	y: log of wage			
	(1) Mean	(2) Mean	(3) Dynamic	(4) Dynamic
Mean effect	0.0219**	0.0216**		
	(0.0105)	(0.0106)		
Effect at t			0.0324***	0.0318***
			(0.0104)	(0.0105)
Effect at t+1			0.0279***	0.0276***
			(0.0105)	(0.0106)
Effect at t+2			0.0278***	0.0275***
			(0.0103)	(0.0103)
Individual-Year	Yes	Yes	Yes	Yes
Region-Sector		Yes		Yes
Observations	156480	156477	156480	156477
R-squared	0.951	0.951	0.951	0.951

Note: The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is more than 1 PoTec workers for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Figure 4: Change in wage after treatment



Note: Figure shows wage differential between treated and untreated workers. Treated workers are those employed in firms that have at least one PoTec employee in 2015, but no PoTec employee in the previous years.

5.3 Sensitivity analysis

As a final exercise to better identify causality in the innovation wage premium, we use an exogenous shock to workers to re-estimate the premium. Specifically, we reduce our data to those workers in establishments that were not innovative in 2014 and experienced closure (following Dahl and Sorenson (2010); Schmieder and von Wachter (2010); Rege et al. (2019)), and follow these workers over time; since some workers are reallocated to other firms that are innovative while most of the remaining workers change companies but continue in non-innovative firms. To minimize temporary workers reallocation we keep workers that remain employed during the whole period of analysis, so once they change firms they remain employed. We match workers using initial firms' and workers' characteristics and use initial firm size, initial wages, average firm wage growth in the past, and average firm growth in 2013-2014.

Table 11 shows the results equivalent to Table 10 but using only workers that have experienced layoffs and found work in another firm. The results support the findings in the previous section of a positive, between 7 and 8 log points, and persistent innovation wage premium. Workers that move to firms that are innovative earn between 5 and 8 log points higher wages than other workers that reallocate to non-innovative firms. The effects are significant and persistent in magnitude in t+2 and t+3.

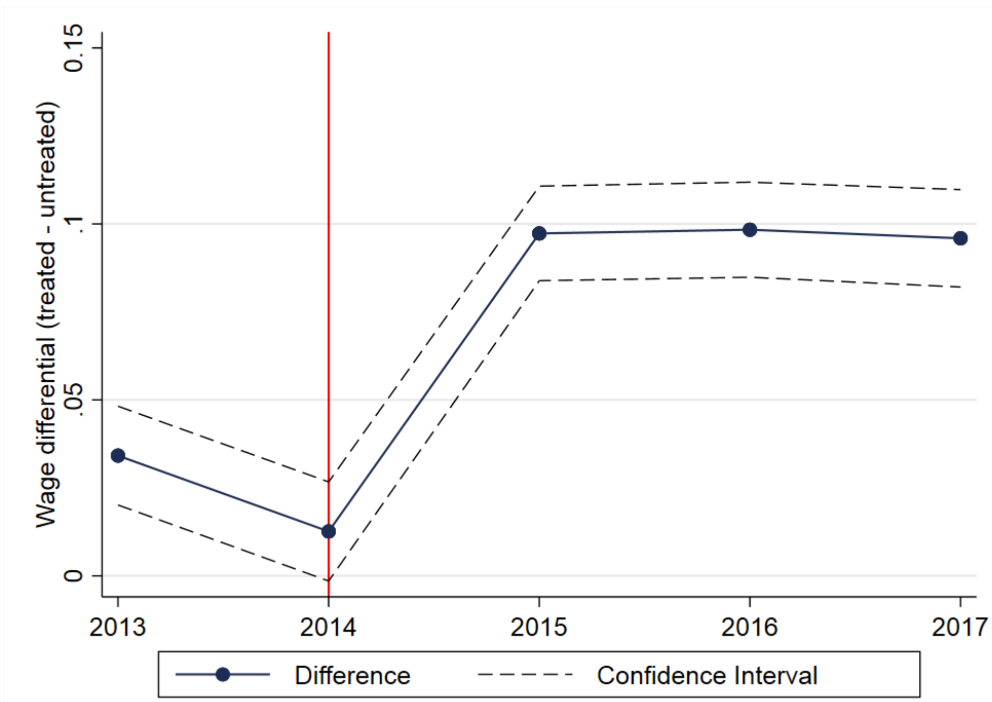
Table 11: Do reallocated workers get higher wages after moving to innovative firms?

	y: log of wage			
	(1) Mean	(2) Mean	(3) Dynamic	(4) Dynamic
Mean effect	0.0846*** (0.00641)	0.0669*** (0.00632)		
Outcome t+1			0.0790*** (0.0102)	0.0565*** (0.0101)
Outcome t+2			0.0774*** (0.0102)	0.0539*** (0.0100)
Outcome t+3			0.0786*** (0.0103)	0.0550*** (0.0101)
Individual-Year	Yes	Yes	Yes	Yes
Region-Sector		Yes		Yes
Observations	79144	79142	79144	79142
R-squared	0.838	0.840	0.837	0.840

Note: The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is more than 1 PoTec workers for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Figure 5 shows the wage trajectory of workers wage differentials, treated versus nontreated, and provides visual evidence of the wage premium. Before plant closure treated workers earned higher wages. However, this wage differential increased significantly after workers were reallocated to other firms in 2015, and the difference persists in 2016 and 2017.

Figure 5: Change in wage after treatment in reallocated workers



Note: Figure shows wage differential between treated and untreated workers. Treated workers are those reallocated to firms that have at least one PoTec employee in 2015.

6 Conclusions

This paper has used a novel measure of innovation to measure the wage premium associated with innovation in Brazil during the period 2014-2017. The results show a positive and statistically significant premium associated with innovation, that is robust across sectors, firm size, and skill groups. While there is strong evidence that innovative firms pay already higher wages before becoming innovative, there is also evidence that the wage premium increases after firms become innovators. Both effects, higher wages before becoming an innovator and the effect after becoming an innovator, are very persistent over time. Overall, working in an innovative company in Brazil pays off.

These results shed some light on some of the hypotheses of different pieces of literature. First, it suggests that firms "prepare" for innovation by paying higher wages prior to becoming an innovator. Given that workers are critical in developing and bringing to market innovation, firms may decide to reward and incentivize workers to have more innovative behavior. Second, it also supports the literature on the appropriation of rents from innovation. While most of the evidence focuses primarily on high-income countries, this paper shows evidence of this rent sharing in a large middle-income country like Brazil. Third, the results are inconclusive concerning the impact across skill groups and skill-biased technological change. While we find evidence of higher premium in medium and low skilled occupations, this distribution of the wage premium across skills disappear when controlling for firm fixed effects. Forth, the paper also contributes to the use of human capital and occupation proxies fro innovation.

More research is needed in regard to the distribution of the premium across skilled groups. Especially a more nuanced view of these groups and the tasks they perform is needed to fully understand whether they are to benefit from innovation. This is critical to understand the impact of innovation on income distribution.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Acemoglu, D. (2002, March). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature 40*(1), 7–72.

Acemoglu, D. and D. Autor (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4 of *Handbook of Labor Economics*, Chapter 12, pp. 1043–1171. Elsevier.

Adrjan, P. (2018, June). Risky Business? Earnings Prospects of Employees at Young Firms. Economics Series Working Papers 852, University of Oxford, Department of Economics.

Aghion, P. (2002). Schumpeterian growth theory and the dynamics of income inequality. *Econometrica 70*(3), 855–882.

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hemous (2019). Innovation and Top Income Inequality. *Review of Economic Studies 86*(1), 1–45.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2018, May). On the returns to invention within firms: Evidence from finland. *AEA Papers and Proceedings 108*, 208–12.

Aghion, P., A. Bergeaud, R. Blundell, and R. Griffith (2019, December). The Innovation Premium to Soft Skills in Low-Skilled Occupations. CEP Discussion Papers dp1665, Centre for Economic Performance, LSE.

Aghion, P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica 60*(2), 323–351.

Akcigit, U., J. Grigsby, and T. Nicholas (2017, January). The Rise of American Ingenuity: Innovation and Inventors of the Golden Age. NBER Working Papers 23047, National Bureau of Economic Research, Inc.

Alvarez, J., F. Benguria, N. Engbom, and C. Moser (2018, January). Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics 10*(1), 149–89.

Araújo, B. C., L. R. Cavalcante, and P. Alves (2009). Variáveis proxy para os gastos empresariais em inovação com base no pessoal ocupado técnico- científico disponível na relação anual de informações sociais (rais). *Radar: Tecnologia, Produção e Comércio Exterior* (5), 16–21.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

Babina, T., W. Ma, C. Moser, P. P. Ouimet, and R. Zarutskie (2019). Pay, Employment, and Dynamics of Young Firms . Working papers, U.S. Census Bureau, Center for Economic Studies.

Bender, S., N. Bloom, D. Card, J. V. Reenen, and S. Wolter (2018). Management Practices, Workforce Selection, and Productivity. *Journal of Labor Economics 36*(S1), 371–409.

Berlingieri, G., S. Calligaris, and C. Criscuolo (2018, May). The productivity-wage premium: Does size still matter in a service economy? *AEA Papers and Proceedings 108*, 328–33.

Bernard, A. B. and J. Bradford Jensen (1999, February). Exceptional exporter performance: cause, effect, or both? *Journal of International Economics 47*(1), 1–25.

Bloom, N., F. Guvenen, B. S. Smith, J. Song, and T. von Wachter (2018, May). The Disappearing Large-Firm Wage Premium. *AEA Papers and Proceedings 108*, 317–322.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020, April). Are ideas getting harder to find? *American Economic Review 110*(4), 1104–44.

Bloom, N. and J. Van Reenen (2011). Human Resource Management and Productivity. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4 of *Handbook of Labor Economics*, Chapter 19, pp. 1697–1767. Elsevier.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020, June). How much should we trust estimates of firm effects and worker sorting? Working Paper 27368, National Bureau of Economic Research.

Brambilla, I., N. Depetris Chauvin, and G. Porto (2017). Examining the export wage premium in developing countries. *Review of International Economics 25*(3), 447–475.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D. and J. E. DiNardo (2002, October). Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics 20*(4), 733–783.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Chen, C., Y. Chen, P.-H. Hsu, and E. Podolski (2016). Be nice to your innovators: Employee treatment and corporate innovation performance. *Journal of Corporate Finance 39*(C), 78–98.

Cirera, X. and W. F. Maloney (2017, December). *The Innovation Paradox*. Number 28341 in World Bank Publications. The World Bank.

Cirillo, V. (2014, December). Patterns of innovation and wage distribution. Do "innovative firms" pay higher wages? Evidence from Chile. *Eurasian Business Review 4*(2), 181–206.

Cirillo, V., M. Sostero, and F. Tamagni (2017). Innovation and within-firm wage inequalities: empirical evidence from major european countries. *Industry and Innovation 24*(5), 468–491.

Cox, P. J. (1990). Research and development — or research design and development? *International Journal of Project Management 8*(3), 144 – 150.

Dahl, M. S. and O. Sorenson (2010). The migration of technical workers. *Journal of Urban Economics 67*(1), 33 – 45. Special Issue: Cities and Entrepreneurship.

De Jong, J. and D. Den Hartog (2007). How leaders influence employees' innovative behaviour. *European Journal of Innovation Management 10*(1), 977–1002.

de Rassenfosse, G. and B. van Pottelsberghe (2009, June). A policy insight into the R&D-patent relationship. *Research Policy 38*(5), 779–792.

Dickens, W. T. and L. F. Katz (1986, September). Interindustry wage differences and industry characteristics. Working Paper 2014, National Bureau of Economic Research.

Dorn, D. (12009). *Essays on inequality, spatial interaction, and the demand for skills*. Ph. D. thesis, University of St. Gallen.

Dosi, G. and P. Mohnen (2018, 12). Innovation and employment: an introduction. *Industrial and Corporate Change 28*(1), 45–49.

Dunne, T. and M. J. Roberts (1990, August). Wages and The Rist of Plant Closings. Working Papers 90-6, Center for Economic Studies, U.S. Census Bureau.

Dunne, T. and J. A. Schmitz (1995). Wages, employment structure and employer size-wage premia: Their relationship to advanced-technology usage at us manufacturing establishments. *Economica 62*(245), 89–107.

Ederer, F. and G. Manso (2013). Is pay for performance detrimental to innovation? *Management Science 59*(7), 1496–1513.

Entorf, H., M. Gollac, and F. Kramarz (1999, July). New Technologies, Wages, and Worker Selection. *Journal of Labor Economics 17*(3), 464–491.

Feldstein, M. (2008). Did wages reflect growth in productivity? *Journal of Policy Modeling 30*(4), 591–594.

Fosfuri, A., M. Motta, and T. Ronde (2001, February). Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics 53*(1), 205–222.

Gibbs, M., S. Neckermann, and C. Siemroth (2017). A field experiment in motivating employee ideas. *The Review of Economics and Statistics 99*(4), 577–590.

Goñi, E. and W. F. Maloney (2017). Why don't poor countries do R&D? Varying rates of factor returns across the development process. *European Economic Review 94*(C), 126–147.

Haidar, J. I. (2012). Trade and productivity: Self-selection or learning-by-exporting in India. *Economic Modelling 29*(5), 1766–1773.

Hall, B. H. (2011, June). Innovation and Productivity. NBER Working Papers 17178, National Bureau of Economic Research, Inc.

Heckman, J., H. Ichimura, and P. E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies 64*(4), 605–654.

Heyman, F. (2007, June). Firm Size or Firm Age? The Effect on Wages Using Matched Employer–Employee Data. *LABOUR 21*(2), 237–263.

Hijzen, A., P. Martins, T. Schank, and R. Upward (2013). Foreign-owned firms around the world: A comparative analysis of wages and employment at the micro-level. *European Economic Review 60*(C), 170–188.

Huang, K.-F. and C.-M. Yu (2011, August). The effect of competitive and non-competitive R&D collaboration on firm innovation. *The Journal of Technology Transfer 36*(4), 383–403.

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation. *NBER Macroeconomics Annual 1*, 235–276.

Krueger, A. B. and L. H. Summers (1988). Efficiency wages and the inter-industry wage structure. *Econometrica 56*(2), 259–293.

Leiblein, M. J. and T. L. Madsen (2009). Unbundling competitive heterogeneity: incentive structures and capability influences on technological innovation. *Strategic Management Journal 30*(7), 711–735.

Lipsey, R. E. and F. Sjoholm (2003, January). Foreign firms and indonesian manufacturing wages: An analysis with panel data. Working Paper 9417, National Bureau of Economic Research.

Maloney, W. F. and F. V. Caicedo (2017). Engineering Growth: Innovative Capacity and Development in the Americas. CESifo Working Paper Series 6339, CESifo.

Mao, C. X. and J. Weathers (2019). Employee treatment and firm innovation. *Journal of Business Finance & Accounting 46*(7-8), 977–1002.

Martinez-Ros, E. (2001). Wages and innovations in spanish manufacturing firms. *Applied Economics 33*(1), 81 – 89.

Mohnen, P. and B. Hall (2013). Innovation and productivity: An update. *Eurasian Business Review 3*(1), 47–65.

Moore, H. L. (1911). *Laws of Wages: An essay in statistical economics.* McMaster University Archive for the History of Economic Thought.

Mortensen, D. T. (2003, 08). *Wage Dispersion: Why Are Similar Workers Paid Differently?* The MIT Press.

OECD (2018). *OECD Science, Technology and Innovation Outlook 2018.* OECD Publishing.

Oi, W. and T. Idson (1999). Firm size and wages. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 3, Part B, Chapter 33, pp. 2165–2214. Elsevier.

Reenen, J. V. (1996). The creation and capture of rents: Wages and innovation in a panel of u.k. companies. *The Quarterly Journal of Economics 111*(1), 195–226.

Rege, M., T. Skardhamar, K. Telle, and M. Votruba (2019). Job displacement and crime: Evidence from norwegian register data. *Labour Economics 61*, 101761.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy 98*(5), S71–S102.

Schank, T., C. Schnabel, and J. Wagner (2008). Higher wages in exporting firms: self-selection, export effect, or both? First evidence from German linked employer-employee data. Discussion Papers 55, Friedrich-Alexander University Erlangen-Nuremberg, Chair of Labour and Regional Economics.

Schmieder, J. F. and T. von Wachter (2010, July). Does wage persistence matter for employment fluctuations? evidence from displaced workers. *American Economic Journal: Applied Economics 2*(3), 1–21.

Schumpeter, J. A., I. T. S, and J. A. Schumpeter (1942). Capitalism, socialism and democracy.

Taveira, J. G., E. Gonçalves, and R. D. S. Freguglia (2019). The missing link between innovation and performance in brazilian firms: a panel data approach. *Applied Economics 51*(33), 3632–3649.

Temouri, Y., A. Vogel, and J. Wagner (2013). Self-selection into export markets by business services firms – Evidence from France, Germany and the United Kingdom. *Structural Change and Economic Dynamics 25*(C), 146–158.

Yang, Y. and S. Mallick (2010). Export premium, self-selection and learning-by-exporting: Evidence from chinese matched firms. *The World Economy 33*(10), 1218–1240.

Appendix A Alternative skill definition

Table A1: Classification of skills - second definition

Skill category	Description
Low-skill	
Skill cat 1	Lowest skill occupations, includes many manufacturing basic occupations, child-care related education, housekeeping, telephone, salespersons.
Skill cat 2	Corresponds to a NQF below 3, but not considered as an entry level. Occupations such as pharmaceuticals dispensers, greenkeepers, aircraft maintenance technician.
Medium-skill	
Skill cat 3	Requires a NQF of 3 which corresponds to a Level of Advanced GCE (A-level). This category spans many different occupations from Fitness instructors to Legal associate professionals.
Skill cat 4	Requires a NQF of 4 and above which corresponds to a Certificate of Higher Education. It includes many technical occupations like Medical technicians or IT operations technicians and some managerial occupations.
High-skill	
Skill cat 5	Includes most managerial and executive occupations as well as engineers. These occupations require at least a NQF of 6 which corresponds to a Bachelor's degree or a Graduate Certificate.
Skill cat 6	Corresponds to occupational skilled to PhD-level and include most scientific occupations like Chemical scientists, Biological scientists, Research and Development manager but Higher education teaching professionals.

Table A2: Private Sector: Innovation Premium per skill level – Skill definition as in the NQF from the UK, 2014 to 2017

	Dependent variable: $ln(w_{ijkft})$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	Low	Low	Medium	Medium	High	High
Innovation	0.0582***	0.0305	0.365***	-0.000475	0.250***	0.0736***	-0.000347	0.0605
	(0.0109)	(0.0248)	(0.0364)	(0.0235)	(0.0284)	(0.0183)	(0.0168)	(0.0386)
X Low-Skill NQF	0.235***	-0.0233						
	(0.0185)	(0.0144)						
X Medium-Skill NQF	0.218***	0.0861***						
	(0.0164)	(0.0126)						
Age Squared	0.229***	0.216***	0.160***	0.150***	0.235***	0.209***	0.557***	0.519***
	(0.00399)	(0.00378)	(0.00432)	(0.00425)	(0.00664)	(0.00792)	(0.0191)	(0.0199)
Tenure	0.0291***	0.0285***	0.0273***	0.0260***	0.0251***	0.0254***	0.0193***	0.0200***
	(0.000447)	(0.000536)	(0.000355)	(0.000442)	(0.000569)	(0.000753)	(0.00218)	(0.00280)
Temporary	-0.0259***	0.0512***	-0.00641	0.0436***	-0.0624***	0.0489**	-0.0632***	0.0238*
	(0.00511)	(0.0101)	(0.00419)	(0.00758)	(0.00744)	(0.0229)	(0.00905)	(0.0133)
Firm size	0.0265***	0.0306***	0.0207***	0.0267***	0.0344***	0.0358***	0.0400***	0.0346***
	(0.000798)	(0.00113)	(0.000811)	(0.00106)	(0.000880)	(0.00165)	(0.00127)	(0.00452)
Low-Skill NQF	-0.117***	-0.112***						
	(0.00254)	(0.00268)						
Medium-Skill NQF	-0.0739***	-0.0646***						
	(0.00208)	(0.00218)						
Individual-Year	Yes		Yes		Yes		Yes	
Individual-Year-Firm		Yes		Yes		Yes		Yes
Observations	122814877	122340110	73371134	72988821	29657823	29423259	12500861	12402900
R-squared	0.928	0.939	0.895	0.911	0.931	0.942	0.947	0.955

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Appendix B All Sectors

Table B1: All sectors: Innovation Premium per skill level, 2014-2017

	Dependent variable: $ln(w_{ijkft})$			
	(1)	(2)	(3)	(4)
	All	Low	Medium	High
Innovation	0.164***	0.374***	0.404***	0.125***
	(0.0108)	(0.0298)	(0.0554)	(0.0107)
X Low-Skill	0.104***			
	(0.0274)			
X Medium-Skill	0.241***			
	(0.0461)			
Age Squared	0.247***	0.0313***	0.167***	0.401***
	(0.0178)	(0.00795)	(0.00501)	(0.0238)
Tenure	0.0258***	0.0166***	0.0257***	0.0208***
	(0.000575)	(0.000328)	(0.000448)	(0.00102)
Temporary	-0.0144**	-0.000457	-0.0488***	0.00123
	(0.00589)	(0.00426)	(0.00683)	(0.0103)
Firm size	0.0258***	0.0105***	0.0276***	0.0316***
	(0.000911)	(0.000815)	(0.00122)	(0.00177)
Public sector	0.0113	0.0249**	-0.0305***	-0.0337***
	(0.00889)	(0.0100)	(0.00529)	(0.0105)
Low-Skill dummy	-0.0672***			
	(0.00201)			
Medium-Skill dummy	-0.0288***			
	(0.00223)			
Individual-Year	Yes	Yes	Yes	Yes
Observations	157682620	25965671	42366117	77979166
R-squared	0.943	0.908	0.895	0.949

Note: The dependent variable is log of hourly wage. Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B2: All Sectors: Innovation Premium per skill level, 2014-2017

	Dependent variable: $ln(w_{ijkft})$			
	(1) All	(2) Low	(3) Medium	(4) High
Innovation	0.0607***	0.374***	0.278***	-0.0000693
	(0.0111)	(0.0231)	(0.0206)	(0.0165)
X Low-Skill NQF	0.257***			
	(0.0192)			
X Medium-Skill NQF	0.233***			
	(0.0177)			
Age Squared	0.243***	0.160***	0.220***	0.516***
	(0.0180)	(0.00644)	(0.0314)	(0.0454)
Tenure	0.0262***	0.0263***	0.0221***	0.0159***
	(0.000568)	(0.000475)	(0.000763)	(0.00199)
Temporary	-0.0159***	-0.00331	-0.0357***	0.0137
	(0.00591)	(0.00481)	(0.0111)	(0.0260)
Firm size	0.0257***	0.0206***	0.0326***	0.0358***
	(0.000897)	(0.000829)	(0.00160)	(0.00153)
Public sector	0.00689	0.0423***	0.0342**	-0.191***
	(0.00905)	(0.00676)	(0.0174)	(0.0150)
Low-Skill NQF	-0.125***			
	(0.00360)			
Medium-Skill NQF	-0.0794***			
	(0.00339)			
Individual-Year	Yes	Yes	Yes	Yes
Observations	157682620	86389697	39234013	18175210
R-squared	0.943	0.923	0.945	0.946

Note: The dependent variable is log of hourly wage. Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Appendix C Dummy variable

Table C1: Innovation Premium, 2014-2017

	Dependent variable: $ln(w_{ijkft})$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Private	Public	All	Private	Public	Private
Dummy innovation	0.0973***	0.157***	0.0293	0.0258***	0.0280***	0.00309	-0.00225
	(0.0155)	(0.00706)	(0.0412)	(0.00336)	(0.00127)	(0.0157)	(0.0178)
Age	0.305***	0.262***	0.369***				
	(0.0127)	(0.00288)	(0.0608)				
Age Squared				0.250***	0.236***	0.486***	0.433***
				(0.0177)	(0.00405)	(0.143)	(0.141)
Gender	-0.217***	-0.197***	-0.244***				
	(0.00368)	(0.00188)	(0.0118)				
Primary school graduate or middle school dropout	0.0397***	0.0350***	0.0447**				
	(0.00396)	(0.00267)	(0.0212)				
Middle school graduate or high school dropout	0.0875***	0.0790***	0.0938*				
	(0.00951)	(0.00308)	(0.0530)				
High school graduate or college dropout	0.225***	0.179***	0.418***				
	(0.00770)	(0.00387)	(0.0339)				
College graduate or higher	0.862***	0.768***	1.030***				
	(0.0183)	(0.0122)	(0.0462)				
Tenure	0.0930***	0.0725***	0.181***	0.0263***	0.0291***	0.00424	0.0122**
	(0.00253)	(0.000859)	(0.0101)	(0.000579)	(0.000447)	(0.00451)	(0.00484)
Temporary	-0.0350*	-0.0113	-0.172***	-0.0182***	-0.0249***	0.0895***	0.0478*
	(0.0218)	(0.0146)	(0.0606)	(0.00623)	(0.00519)	(0.0273)	(0.0270)
Firm size	0.0405***	0.0320***	0.0500**	0.0225***	0.0228***	0.00436	0.0277
	(0.00460)	(0.00289)	(0.0179)	(0.00106)	(0.000914)	(0.00464)	(0.0346)
Geo-Occupation-Year	Yes	Yes	Yes				
Individual-Year				Yes	Yes	Yes	
Individual-Year-Firm							Yes
Observations	172519611	136163626	36355985	157682620	122814877	33708892	33708299
R-squared	0.558	0.518	0.491	0.942	0.927	0.953	0.955

Note: Innovative firms is measured by the share of PoTec workers in total employment for firm f in year t. Columns (7)-(9) exclude the public sector. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Appendix D Self selection - Stochastic Dominance test and different definition of transition

Table D1: Stochastic Dominance: Kolmogorov-Smirnov test

Year of transition	**t-1**		**t-2**		**t-3**	
2015	**D**	**P-value**	**D**	**P-value**	**D**	**P-value**
	0.1105	0.000	0.1125	0.000	0.0977	0.000
	-0.0005	0.727	-0.0000	0.999	-0.0008	0.634
	0.1105	0.000	0.1125	0.000	0.0977	0.000
2016	**D**	**P-value**	**D**	**P-value**	**D**	**P-value**
	0.1230	0.000	0.1102	0.000	0.1187	0.000
	-0.0006	0.708	-0.0006	0.796	-0.0022	0.057
	0.1230	0.000	0.1102	0.000	0.1187	0.000
2017	**D**	**P-value**	**D**	**P-value**	**D**	**P-value**
	0.1216	0.000	0.1368	0.000	0.1283	0.000
	-0.0009	0.390	-0.0001	0.997	-0.0000	1.000
	0.1216	0.000	0.1368	0.000	0.1283	0.000

Table D2: Self selection: different definition of transition

	Treated (t-1)		
Year of transition	2015	2016	2017
Treated	0.0606*** (0.0224)	0.0811*** (0.0280)	0.0972*** (0.0294)
		Treated (t-2)	
Treated	0.0740*** (0.0282)	0.0573 (0.0561)	0.106*** (0.0406)
		Treated (t-3)	
Treated	0.0859*** (0.0243)	0.0427 (0.0654)	0.0923* (0.0526)

Note: Estimation using OLS. The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is more than 1 PoTec worker for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Appendix E Causality

Table E1: Do firms pay higher wages after becoming innovative?

	y: log of wage			
	(1) Mean	(2) Mean	(3) Dynamic	(4) Dynamic
Mean effect	0.0214***	0.0211***		
	(0.00393)	(0.00395)		
Effect at t			0.0194***	0.0191***
			(0.00426)	(0.00426)
Effect at t+1			0.0199***	0.0196***
			(0.00430)	(0.00430)
Effect at t+2			0.0208***	0.0205***
			(0.00433)	(0.00433)
Individual-Year	Yes	Yes	Yes	Yes
Region-Sector		Yes		Yes
Observations	770016	770016	770016	770016
R-squared	0.946	0.946	0.946	0.946

Note: The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is a PoTec workers for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table E2: Do firms pay higher wages after becoming innovative?

	y: log of wage			
	(1) Mean	(2) Mean	(3) Dynamic	(4) Dynamic
Mean effect	0.0224**	0.0227**		
	(0.0113)	(0.0114)		
Effect at t			0.0350***	0.0353***
			(0.0123)	(0.0123)
Effect at t+1			0.0318***	0.0321***
			(0.0119)	(0.0119)
Effect at t+2			0.0306**	0.0307**
			(0.0124)	(0.0124)
Individual-Year	Yes	Yes	Yes	Yes
Region-Sector		Yes		Yes
Observations	144064	144061	144064	144061
R-squared	0.949	0.949	0.949	0.949

Note: The dependent variable is log of hourly wage. Innovative firms is measured by a dummy = 1 if there is more than 1 PoTec workers for firm f in year t but none in $t-1$. Heteroskedasticity robust standard errors clustered at the firm level are reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.