Bruno Azuma Balzano

Computers and Jobs:
Evidence for the Brazilian Labor Market

São Paulo
November 23, 2023

Bruno Azuma Balzano

Computers and Jobs:
Evidence for the Brazilian Labor Market

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA

Supervisor: Renata del Tedesco Narita

São Paulo
November 23, 2023

Abstract

This project examines the impact of computer adoption on the Brazilian labor market, specifically focusing on the phenomenon of job polarization. Job polarization refers to the increasing concentration of high-skilled and low-skilled jobs, accompanied by a decline in medium-skilled jobs. By analyzing the effects of computer adoption on labor share and earnings across skill types (routine and nonroutine, manual and cognitive), this study aims to investigate the routinization hypothesis in the Brazilian context.

Using an industry-year measure of computer exposure, the study employs a fixed effects model to estimate the impact of computers on wage polarization. By utilizing data from the Brazilian Ministry of Economics' Annual Social Information Report (RAIS) and the World Bank's World Integrated Trade Solution (WITS), this research hopes to contribute to the understanding of job polarization in the Brazilian labor market and provide valuable insights for policymakers addressing the potential challenges posed by technological change.

Keywords: computer adoption, labor market, job polarization, wage polarization, wage inequality, skill-biased technological change, routinization hypothesis, Brazil.

JEL classification: J23 (Labor Demand), J24 (Human Capital; Skills; Occupational Choice; Labor Productivity), J31 (Wage Level and Structure; Wage Differentials), O33 (Technological Change: Choices and Consequences; Diffusion Processes).

Resumo

Este projeto tem como objetivo avaliar o impacto da adoção de computadores no mercado de trabalho no Brasil, com foco no fenômeno de polarização de empregos. A polarização de empregos se refere ao aumento relativo de empregos de alta e baixa qualificações, enquanto os empregos de média qualificação diminuem. Como é esperado que o uso de computadores aumente a produtividade dos trabalhadores de alta qualificação e substitua trabalhadores em tarefas de média qualificação, este estudo propõe analisar o impacto da adoção de computadores na renda dos trabalhadores por tipo de habilidade utilizada (cognitiva abstrata, cognitiva rotineira, manual rotineira e manual abstrata), a fim de examinar a hipótese de *routinization* para o caso brasileiro.

Para medir o emprego de computadores, será construída uma medida de exposição a nível de indústria-ano a computadores com base em dois índices: o índice de exportação, para representar o lado da oferta do mercado de computadores, e a participação de ocupações cognitivas rotineiras no nível da indústria, representando o lado da demanda. O estudo utilizará um modelo de forma reduzida baseado em efeitos fixos para estimar o efeito da exposição à tecnologia de computador na polarização salarial. Os resultados servirão para analizar as relações salariais entre categorias de ocupações, com as categorias manual rotineira e cognitiva rotineira agrupadas para se relacionarem com os trabalhadores de habilidade média.

O estudo utilizará dados do Relatório Anual de Informações Sociais (RAIS) do Ministério da Economia do Brasil para informações sobre salários e participação rotineira, e a *World Integrated Trade Solution* (WITS) do Banco Mundial para o índice de exportação. Espera-se que o estudo contribua para a análise da polarização de empregos no mercado de trabalho formal no Brasil, fornecendo *insights* sobre a relação entre adoção de computadores e polarização salarial por categoria de habilidade e tarefa. Os resultados tambem podem servir para informar decisões políticas destinadas a mitigar os potenciais efeitos negativos da polarização de empregos sobre os trabalhadores no Brasil.

Palavras-chave: adoção de computadores, mercado de trabalho, polarização de empregos, polarização salarial, desigualdade salarial, mudança tecnológica com viés de habilidade, hipótese de *routinization*, Brasil.

Códigos JEL: J23 (Demanda de trabalho), J24 (Capital humano; Habilidades; Escolha ocupacional; Produtividade do trabalho), J31 (Nível e estrutura salarial; Diferenciais salariais), O33 (Mudança tecnológica: escolhas e consequências; Processos de difusão).

List of Figures

List of Tables

Contents

1 Introduction

The impact of automation on jobs has got more attention than ever. Specialists are still trying to understand how the Future of Work and Artificial Intelligence (AI) may impact the economy (FRANK et al., 2019; AUTOR, 2019). Acemoglu (2021, p. 20-24) argue that there is something like excessive automation that could cause more negative effects, like unemployment and inequality, than good, via productivity gains. However, despite the exceptional evolution of AI in the last decade and the relevance of the subject for the present day, this phenomenon is not novel. In fact, it has been studied for centuries (FRANK et al., 2019, p. 6531).

Abramovay (2021) highlights that classic economists, like John Stuart Mill, Karl Marx, and John Maynard Keynes already expressed concerns related to the possibility of the end of jobs. After reviewing recent literature about this subject, however, Abramovay (2021) concludes that the concerns shouldn't be pointed to the end of jobs. Instead, the focus should be on the sharp inequality generated by the current technical changes.

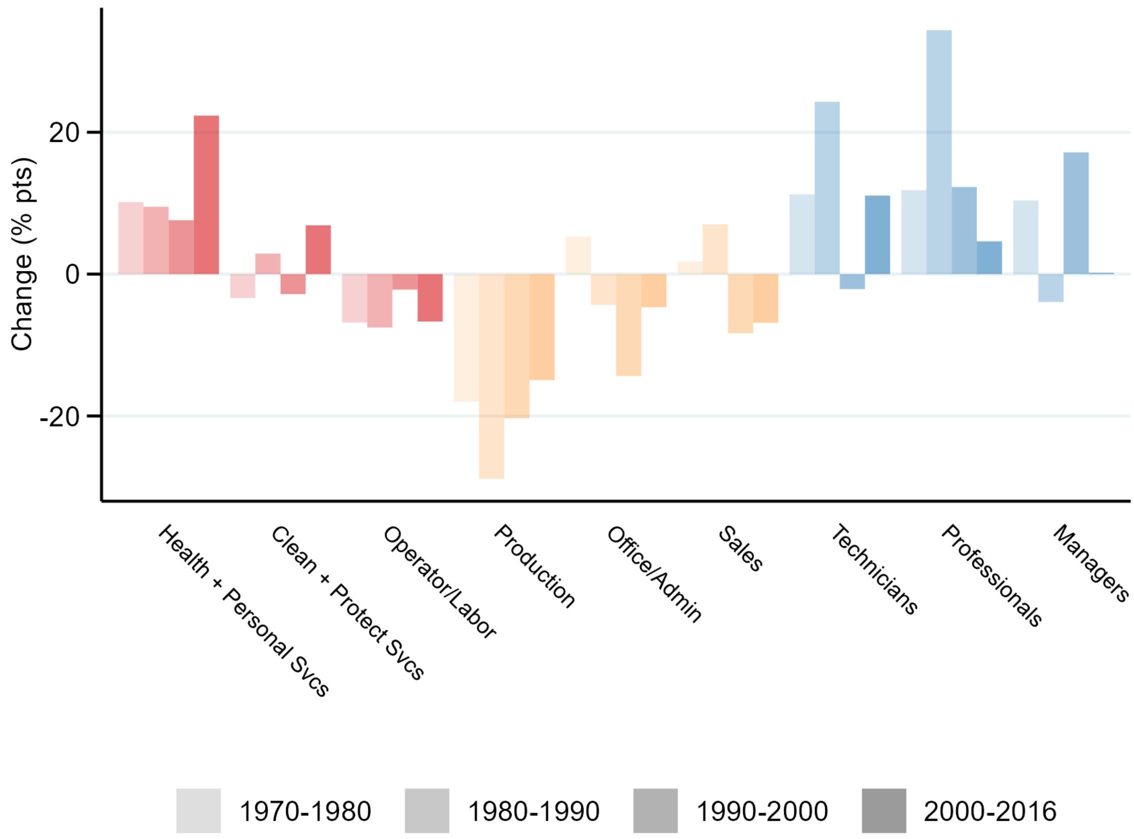
Again, this inequality is not contemporary. Decades ago, Katz and Murphy (1992) developed a supply and demand framework for the labor market to investigate the wage differentials between genders and education levels, concluding that "the growth of the supply of college graduates as a fraction of the labor force appear to play an important role in explaining these large differences in the behavior of the relative earnings of college graduates" (KATZ; MURPHY, 1992, p. 76). Besides, they note that Krueger (1993 apud KATZ; MURPHY, 1992) "suggests that the spread of computers in the workplace may be an important component of these within-sector changes in the composition of labor demand".

That suggested relation between computers and jobs is further examined by several posterior works. In their seminal paper, Autor, Levy, and Murnane (2003) introduce what begin to be called *routinization* or *routine biased technical change* (RBTC) hypothesis, which states that computers and its related technology complement workers in the execution of more generalized and abstract tasks (like medical diagnosis or janitorial services) — *non-routine tasks*, while substitutes workers on the execution of tasks that can be "described with programmed rules" (AUTOR; LEVY; MURNANE, 2003, p. 1332) — *routine tasks*. The dual effects of information and communication technologies (ICT) on the labor market cause what is going to be called a "hollowing out" effect: the investment in ICT capital would, simultaneously, reduce the labor input for routine tasks and increase the demand for non-routine tasks, increasing the employment in the occupations with most and least years of education required and reducing the employment in the occupations between them. This effect causes a U-shaped form in the chart of the percent employment change, as can be seen in Figure 1.

Building atop of that, Acemoglu and Autor (2011) further develop a model that directly links these changes in the demand for specific types of tasks to changes in the relative wage between groups of workers based on a classification of skills similar to the ordering of **??**.

Figure 1 – Changes in Occupational Employment Shares, 1970-2016

(Percent Change Over Decade)



Source: Autor (2019)

In summary, by classifying workers into an ordered set of three groups (low, medium, and high skill) it was possible to analytically deduce what would happen to each group's mean wage with the introduction of capital capable of substituting the labor input for some tasks (ACEMOGLU; AUTOR, 2011, p. 1118-1152). And, although the empirical application present in this work is simple, this model introduced a promising explanation capability for the wage and job polarization phenomena.

In fact, some recent papers expand this three-group model to a model based on an ordered set of any finite size and evaluate the direct and indirect effects of the introduction of labor substituting capital. The authors state that task displacement explains approximately 50% of the observed changes in U.S. wages while providing small productivity gains. (ACEMOGLU; RESTREPO, 2022, p. 2014) Yet, it's worth mentioning that one of the measures of substituting capital used by the authors is composed, amongst other things, of the adoption of specialized software across industries, which suggests that the adoption of Information and Communication Technologies is still relevant for the study this subject.

Now, whereas this theoretical models have been developed mostly focusing U.S. or other developed countries labor market, there is also a few studies testing these models for Brazil and developing countries. Adamczyk, Ehrl, and Monasterio (2022) indicate that the evidences about jobs and wages polarization in Brazil are mixed. While most of the studies observe some sign of polarization, some attribute the wage inequality to other factors, like the decline in returns to experience, and some encounter mixed trends with technology favouring middle skilled workers at first, and disfavouring later. The details of these differences will be explained in a little more detail in the chapter 2, dedicated for the literature review.

To contribute to this literature, this undergraduate thesis will evaluate the effects of computer adoption on the wages of three groups of workers based on the classification developed in Adamczyk, Ehrl, and Monasterio (2022). We use this classification to calculate a measure of exposure to computers in each industry sector based on the previous dependence of that sector in routine tasks. Then, an indicator of global export of computers is added to serve as an exogenous variable for computer prices, which also contributes to the exposure to computers. Finally, the exposure to computers is regressed on the wages of each combination of worker group and industry sector. This approach is mostly based on Boustan, Choi, and Clingingsmith (2022), with a minor adjustment to incorporate the skill based model.

The rest of this thesis is structured as follows. The chapter 2, as mentioned earlier, is dedicated for a detailed literature review, covering both the theoretical and the empirical works related to the subject. chapter 3 details the theoretical model adopted in this thesis and the reduced form that is used for the empirical analysis. chapter 4 provides an overview about the data sources and some basic exploratory analysis. chapter 5 discusses the results of the regressions utilized for empirical analysis. And, lastly, chapter 6 highlights the more important findings and suggests what can be done in future.

2 Literature review

For the main concern of this project, Katz and Murphy (1992) is the oldest work to get emphasis. It "provides a parsimonious framework for thinking about the skill premium and the determinants of the earnings distribution" (ACEMOGLU; AUTOR, 2011, p. 117) and apply it to analyze wage structure changes from 1963 to 1987. They found a sharp increase in wage inequality in this period and described what they called college wage premium, with college graduates assumed as more-skilled workers, so it became a starting point for further investigations of the "skill premium".

Looking to better understand this skill premium and its relation with technological change, Autor, Levy, and Murnane (2003) develop a "task model". In that context, a task is a sequence of procedures that generate some output. The tasks can be manual or cognitive, depending on the majority of the effort needed to fulfill the task, and routine or nonroutine. A task is routine if "it can be accomplished by machines following explicit programmed rules" (AUTOR; LEVY; MURNANE, 2003, p. 1283) and nonroutine if "the rules are not sufficiently well understood to be specified in computer code and executed by machines" (AUTOR; LEVY; MURNANE, 2003, p. 1283).

Autor, Levy, and Murnane (2003), then, assume the price for computer capital to be falling exogenously and concludes that this declining price of computer capital is the causal force by which computer technology affects skill demand. Yet, larger investments in computer capital would reduce labor input of routine tasks, which computer capital substitutes, and increase demand for nonroutine task input, which computer capital complements.

However, an exogenous assumption of the price of computer capital seems too restrictive. To address this shortcoming, Acemoglu and Autor (2011) develop a richer task-based model referred to as "Ricardian model of the labor market". This model combines a generalization of Acemoglu and Zilibotti (2001 apud ACEMOGLU; AUTOR, 2011), that was used to evaluate productivity differences across countries, with the skills and tasks interactions suggested by Autor, Levy, and Murnane (2003). A synthesis of the model can be seen in the following snippet:

> [...] A skill is a worker's *endowment* of capabilities for performing various tasks. This endowment is a stock, which may be either exogenously given or acquired through schooling and other investments. Workers apply their skill endowments to tasks in exchange for wages. Thus, the task-based approaches emphasize that skills are applied to tasks to produce output — skills do not directly produce output. Task models provide a natural framework for interpreting patterns related to occupations in the labor market, as documented above since we can think of occupations as bundles of tasks. In this light, the canonical model may be seen as a special case of the general task-based model in which there is a one-to-one mapping between skills and tasks. (ACEMOGLU; AUTOR, 2011, p. 1118-1119).

Also, this Ricardian model can more appropriately account for non-monotonic changes in the wage and employment changes, as the ones documented in Acemoglu and Autor (2011). So, the authors explain the theoretical implications of the model for job polarization, arguing that an increase in the supply of skills in a given group or complementary capital for that group expands the set of tasks performed by that group and contracts the set of tasks, decrease the skill wages and generate an endogenous change in technology that increases demand for that skill group (ACEMOGLU; AUTOR, 2011, p. 1131-1151).

Although there is no deep empirical application in Acemoglu and Autor (2011), several subsequent studies applied that framework. One of such is Adamczyk, Ehrl, and Monasterio (2022), an empirical study derived from Autor, Levy, and Murnane (2003) and Acemoglu and Autor (2011) that evaluate the routinization and job polarization hypothesis for the Brazilian labor market with data from 2003 to 2018. They note that the evidence for jobs and wages polarization in Brazil is mixed, with some studies observing some signs of polarization, some not observing polarization, and some observing other kinds of technological bias (ADAMCZYK; EHRL; MONASTERIO, 2022, p. 9). The result found in Adamczyk, Ehrl, and Monasterio (2022) is aligned with previous findings, indicating that routine manual workers are heavily impacted by economic downturns and evidence for routine-biased technological change, especially since 2014, corroborating with the computerization hypothesis and theoretical impacts of automation on the labor market. The impact of technology itself was not evaluated in that study, however, but has been studied in earlier works.

Almeida, Corseuil, and Poole (2017), for example, uses an interaction between the internet service provided in each city with the task intensity level of each industry sector based directly on the O*NET database and finds that technology adoption displaces both nonroutine and routine workers, but the displacement being larger on routine workers. Another example, Riva (2016), uses the 1991 market liberalization as a natural experiment for exogenous variation in technological changes, finding that the lower computer prices displaced labor from routine to nonroutine tasks, based on a classification by Funchal and Junior (2012 apud RIVA, 2016).

Finishing the evaluated studies with empirical applications of the Autor, Levy, and Murnane (2003) and Acemoglu and Autor (2011) framework for the Brazilian case, it is worth to mention Firpo et al. (2021). Among another set of analysis about inequality on the Brazilian labor market, they evaluate both job and wage polarization with data from formal and informal employees and self-employed workers, first using wage on initial period as a proxy for skill, then by using a Routine Task Intensity (RTI) index for each occupation and year, covering three periods: 2003/04, 2011/12 and 2018/19 (FIRPO et al., 2021). With the initial wage variable as a proxy, the study found strong evidence of wage polarization on all periods, with all the parameters significant at 1% level in all periods. Job polarization, on the other hand, is not so clear. There are mixed results, as the first period shows statistic significant signs of the exact inverse effect of polarization (i.e., more employment on the middle of the distribution) and the other periods showing signs of polarization, but not with statistic significance. When using the RTI index, consistent wage polarization signs are found, but mostly without statistical significance, and mixed effects on employment again without statistical significance too (FIRPO et al., 2021, p. 15-17).

These were all empirical applications of the Acemoglu and Autor (2011) framework, but there is a different theoretical approach on the subject of technological change emerging. Acemoglu and Restrepo (2020b) synthesizes this model, which also derives from Katz and Murphy (1992, p. 356 apud ACEMOGLU; RESTREPO, 2020b), but evolves the model to take into account that the mass of tasks performed by the workers also expands over time. The theoretical details of this model extrapolate the scope of this project, but there are some empirical works worth noting.

Based on this framework, Acemoglu and Restrepo (2020a) evaluates the impact of robots on the labor market by industry and commuting zone level. To do this, Acemoglu and Restrepo (2020a) develops an exposure to robots measure and use as an instrumental variable for robot adoption. The conclusion is that each additional robot per thousand workers reduces the local employment-to-population ratio by 0.39 percentage points and wages by about 0.77 percent, with a reduction to 0.2 and 0.42 on the aggregate level, due to trade linkages. Boustan, Choi, and Clingingsmith (2022) uses a similar approach to evaluate the impact of CNC adoption, using an intent-to-treat approach instead of an instrumental variable, and finds that CNC exposure is associated with a general increase in employment, but with fall in employments of workers without college, while also reducing wages of less than High school level workers.

Lastly, Acemoglu and Restrepo (2022) further develops the model presented in Acemoglu and Restrepo (2020b) and incorporates tasks and labor types to the model, but generalizing it to any finite size of types/groups of workers. The empirical analysis on this study indicates "that 50–70% of the changes in the U.S. wage structure between 1980 and 2016 are accounted for by the relative wage declines of worker groups specialized in routine tasks in industries experiencing rapid automation" (ACEMOGLU; RESTREPO, 2022).

3 Methodology

The theoretical framework for this project will be the one developed in Acemoglu and Autor (2011). That model is based on a continuum of tasks that can be performed by low, medium, and high skilled workers and capital, with each task having the following production function

$$y(i) = A_L\alpha_L(i)l(i) + A_M\alpha_M(i)m(i) + A_H\alpha_H(i)h(i) + A_K\alpha_K(i)k(i), \tag{1}$$

where A_L represents the general factor augmenting technology for the low skill, $\alpha_L(i)$ represents the productivity of low skilled workers on task i and $l(i)$ represents the number of workers allocated to task i, with analogous understandings for medium and high skilled workers and capital (on the K terms). This function is capable of giving us two types of capital use: complementary, represented by the A terms; and substitute, on the K terms. Additionally, that form of production function implies that a given i task could be supplied by any combination of low, medium, and high skilled workers or substitute capital, but the comparative advantage of each input differs across tasks, as captured by the α terms (ACEMOGLU; AUTOR, 2011, p. 1121).

Then, Acemoglu and Autor (2011) further analyzes the equilibrium noting that, given a continuum of tasks represented by the unit interval, [0,1], and the comparative advantage differences assumed, there will exist some I_L and I_H such that $0 < I_L < I_H < 1$ and all tasks $i < I_L$ will be performed (only) by low skilled workers and all tasks $i > I_H$ will be performed (only) by high skilled workers, with the intermediate tasks being performed by medium skilled workers. These intermediate tasks can be seen as the routine tasks performed by workers in many production, clerical, and administrative support occupations (ACEMOGLU; AUTOR, 2011, p. 1122).

The proposition 4 of Acemoglu and Autor (2011, p. 1140) allow the inference of the introduction of a substituting technology through the relative analysis between the mean wage of the skill groups. So, the first evaluation of this undergraduate thesis will be if there are signs of the introduction of such type of technological change through the period under analysis.

The second evaluation will be the existence of job and wage polarization in the period under analysis. For that evaluation, the methodology applied will be the same as Firpo et al. (2021), using the wages in the first period as a proxy for the skill index.

The third evaluation will be about the effectiveness of the task-based model to capture the wage and employment. To do that, the discrete classifications of occupations based on the predominant skill type of tasks assigned to that occupation will be used as explanatory variables for wage and number of employment relationships of each occupation. The discrete classification of occupations applied will be the ones available at the annex A3 of Adamczyk, Monasterio, and Fochezatto (2021, p. 39).

Finally, the main analysis of this thesis will be about the relation of these trends with computers (and ICT's in general). For that, a measure of *exposure to computers* will be constructed and used in conjunction of the discrete classifications of occupations to explain the changes on wages and employment. This approach is based on Boustan, Choi, and Clingingsmith (2022), which do the same analysis focused on CNC machines.

3.1 Hypothesis

There are some main hypotheses assumed that motivate the methodology applied in this project.

The first one is that technology can be simultaneously augmenting and substituting for workers, and is captured in the model by the fact that any share of the tasks executed by a given worker could be either augmented or substituted by new technology (ACEMOGLU; AUTOR, 2011). That hypothesis cannot be tested in this study, as it is fundamental for the theoretical framework used, but the model is general enough to also accommodates an alternative hypothesis that only one of the phenomena occurs.

The second hypothesis is that the task model is able to capture the interactions of the Brazilian formal labor market, specifically through the changes in wages and the number of employment relationships. This will be tested in the third evaluation along with the discrete classification of occupations present in Adamczyk, Monasterio, and Fochezatto (2021).

The last one is that computers, on average, work as augmenting technology for abstract cognitive tasks and substituting for routine tasks (both cognitive and manual), thus potentially increasing job polarization, in general (AUTOR, 2019). This hypothesis is supported by earlier evidence pointing to an effect of displacement of jobs from routine to nonroutine caused by computer adoption in the Brazilian context, as noted on chapter 2 (ALMEIDA; CORSEUIL; POOLE, 2017; RIVA, 2016).

That last hypothesis will be tested in some degree by our methodology approach. We can identify if that hypothesis holds for the Brazilian labor market, so to accumulate evidence of this hypothesis for the developing world. It is important to note, however, that the method and the measure used for that test will also count for the final result, so if the null hypothesis holds it does not mean necessarily that the theoretical hypothesis is false, just that or empirical approach needs some improvement.

3.2 Empirical Strategy

As there multiple evaluations being made, this section will be broke down for each set of models in order to better organize and explain the variables involved in each evaluation.

3.2.1 The existence of substituting technology

3.2.2 The existence of job and wage polarization

As stated earlier, this analysis will be based on the methodology applied in Firpo et al. (2021). It uses the following reduced model to evaluate changes in labor market outcomes:

$$\Delta \log(y_{j,t}) = \varphi_0 + \varphi_1 \log(x_{j,t-1}) + \varphi_2 \log(x_{j,t-1})^2 + \varepsilon_{j,t} \tag{2}$$

where $log(x_{j,t-1})$ represents the log of mean earnings of occupation j in the initial period $t-1$ and $\Delta \log(y_{j,t})$ represents either the change in log of employment share or the change in log mean earnings in occupation j for the period t (FIRPO et al., 2021, p. 4). Adapting this model to the data used in this thesis, Equation 2 will be applied with $\log(x_{j,t-1})$ representing the mean wage of occupation j in the initial period $t-1$ and $\Delta \log(y_{j,t})$ representing either the change in log of the number of employment relationships or the change in log of mean wage in occupation j for the period t.

3.2.3 Effectiveness of the task-based model for the Brazilian formal labor market

This is the most simple of all models. It is a simplification of the model described in subsection 3.2.4, but without the measure for exposure to computers. The model has the following form:

$$\log(y_{j,c,t}) = \lambda_c + \alpha_j + \theta_t + \varepsilon_{j,c,t} \tag{3}$$

where λ_c is a dummy vector for skill classification of occupations, α_j is an industry fixed effect, θ_t is a period fixed effect, and $\log(y_{j,c,t})$ represents either the log of the number of employment relationships or the log of mean wage in occupation j and skill group c for the period t. This will allow the analysis of the statistical significance of the skill classifications and, therefore, if there is some degree of relevance of the task-based model for explaining the changes in wages and employment level in the Brazilian formal labor market.

The classification of nonroutine manual, routine manual, routine cognitive, and nonroutine cognitive is based on the definition of discrete skill groups in Adamczyk, Monasterio, and Fochezatto (2021, p. 39), which comes from a new method of skill measurement across occupations: they build a dictionary of verbs associated with each skill group from the O*NET data adapted by Acemoglu and Autor (2011), apply machine learning techniques to translate and derive weights for these verbs in Portuguese and apply these weights to tasks descriptions of the occupations in the Brazilian Catalog of Occupations (Catálogo Brasileiro de Ocupações — CBO) (ADAMCZYK; EHRL; MONASTERIO, 2022, p. 13).

According to the authors, this method has three advantages over the previous ones:

> [...] The first advantage is that we do not need ad hoc lists of words related to each type of task. Second, the method does not assume that the occupations perform the same activities across countries. Third, our method can measure the skills even of occupations that are not listed in a database such as O*NET. A large number of Brazilians work in occupations that cannot be found in O*NET, such as elevator operators in commercial buildings (CBO 5141-05 - Ascensorista) or bus ticket/money collectors (CBO 5112-15 - Cobrador). (ADAMCZYK; EHRL; MONASTERIO, 2022, p. 13).

3.2.4 Relation of computer adoption and the labor market

Lastly, this model will allow the evaluation of the computerisation hypothesis for the Brazilian formal labor market. With the Acemoglu and Autor (2011) model in mind, the empirical strategy of this project is based on a combination of the Acemoglu and Restrepo (2020a) and Boustan, Choi, and Clingingsmith (2022) approach with the discrete classification of occupations in Adamczyk, Monasterio, and Fochezatto (2021). It consists of the examination of whether and how the diffusion of computer technology was associated with the wage polarization effect. Following Boustan, Choi, and Clingingsmith (2022, p. 8-11), an industry-year exposure to computers measure will be constructed to infer the computer adoption with the combination of two indexes. The first is an export index, to proxy the supply side of the computer market, based on the global export of computers from the top 60% exporters, as it is not likely to be endogenous with local-level wages. The second, to proxy the demand side is the share of routine cognitive occupations at the industry level before the periods used in the regression. So, the measure takes the form

$$Exposure_{j,t} = \underbrace{\frac{\sum_i X_{i,t}}{\sum_i X_{i,t=0}}}_{Export\_index} \times \underbrace{\frac{L_{c=RC,t=0}}{\sum_c L_{c,t=0}}}_{RC\_share_j} \tag{4}$$

where $X_{i,t}$ is the total export of computers of the exporter i on period t, $L_{c,t}$ stands for the number of employment relationships in category $c \in C\{NRM, RM, RC, NRC\}$. The skill classification of the occupations will be again based on the definition of discrete skill groups in Adamczyk, Monasterio, and Fochezatto (2021, p. 39).

Finally, we make use of the exposure measure with the following reduced-form model:

$$\ln(wage_{j,c,t}) = \beta Exposure_{j,t} + \alpha_j + \lambda_c + \theta_t + \varepsilon_{j,c,t}, \tag{5}$$

where $Exposure_{j,t}$ is the industry-year measure of exposure to computers, α_j is an industry fixed effect, λ_c is a skill classification fixed effect, and θ_t is a period fixed effect. Thus, β will identify the effects of the changes in exposure to computers within each outcome. The outcome here is the log wage, where $wage_j, t, c$ stands for the average wage for occupations in the industry j and category c on year t. Because data on actual computer usage is not applied, we instead run the "intent to treat" specification in Equation 5, estimating the effect of exposure to computer technology on wage polarization, so we cannot interpret the magnitude of the coefficient of interest β as estimating the effect of changes in tool usage directly.

Lastly, the findings of Acemoglu and Autor (2011, p. 1131-1151) — concerning the expansion of tasks performed by each group of workers and the wage premium between these groups, also assuming that an increase in the supply of a group of skills will lead to an endogenous change in technology that increases the demand for that skill group — will be used to analyze the results found on the estimators.

4 Data Sources and initial evaluation

4.1 Wages and employment data

The jobs and wages data will be extracted from the Annual Social Information Report (Relação Anual de Informações Sociais - RAIS) produced by the Brazilian Ministry of Economics. This database contains administrative data about all of the employment relationships throughout the year with mean wages, agreed weekly hours, occupation, industry sector of the employer, and several demographic characteristics of the employee, as scholarship level, race/color, gender, age, *etc.*

This data is complemented with the discrete classification of occupations applied will be the ones available at the annex A3 of Adamczyk, Monasterio, and Fochezatto (2021, p. 39), which provides a classification for each occupation in the Brazilian Catalog of Occupations (Catálogo Brasileiro de Ocupações — CBO) in the categories of non-routine manual, routine manual, routine cognitive and non-routine cognitive. This combination allow the extraction of some initial insights from the data.

Table 1 shows the share of employment relationships for each group of skill classification of occupations and the mean wages and number of employment relationships both for each group and in general for the first year of the sample and Table 2 shows the same for the last year. There is also some employment relationships which does not have a group classification, either because the occupation was not available on RAIS database, or because the occupation was not classified in Adamczyk, Monasterio, and Fochezatto (2021) — which excluded police and military professionals from the analysis (ADAMCZYK; EHRL; MONASTERIO, 2022, p. 17). The unclassified employment relationships are excluded from the analysis for the remainder of this thesis, but it does not seem to affect the analysis as they are a very small share of the employment relationships[1].

Table 1 – Mean wages, number and share of employment relationships, 2006

Skill group	**Mean hourly wage**	**Employment relationships**	
		Number	**Share**
NRM	R$ 9.30	7,735,410	15.26%
NRC	R$ 31.95	7,995,661	15.77%
RC	R$ 16.19	16,154,345	31.86%
RM	R$ 11.04	18,025,205	35.55%
Non classified	R$ 31.09	790,399	1.56%
All groups	R$ 19.91	50,701,020	100.00%

Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

[1] The number of relationships without occupation code and the share of them by year can be seen in Appendix A.

Table 2 – Mean wages, number and share of employment relationships, 2019

Skill group	Mean hourly wage	Employment relationships	
		Number	Share
NRM	R$ 11.65	11,774,724	17.66%
NRC	R$ 35.95	13,219,736	19.83%
RC	R$ 17.32	22,764,576	34.15%
RM	R$ 13.82	18,334,787	27.50%
Non classified	R$ 54.78	573,594	0.86%
All groups	R$ 26.70	66,667,417	100.00%

Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

Now, it is interesting to see the general trends of employment relationships and wages for the period in analysis, as well as the trends for each group of skill classification. Starting with wages, Figure 2 shows the general patterns of mean wages[2] in relation the first year of the sample, while Figure 3 does the same for each group of skill classification. Figure 2 shows a general rising trend, with three noticeable exceptions: 2009-2010, related to the impacts of 2008 international financial crisis; and 2013-2015 and 2018-2019, related to Brazilian internal political crises. The details of each crisis are outside of the scope of this thesis, but it is worth to see how these downturns affected each group of workers.

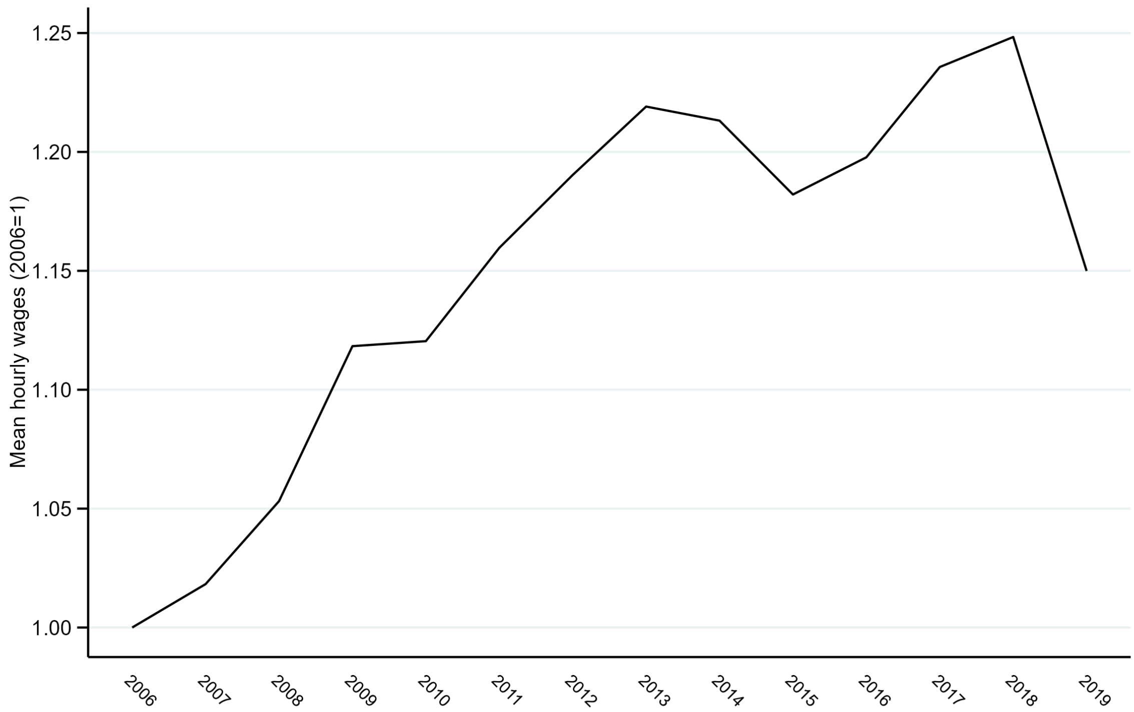
Figure 3 shows approximately the same general rising trend as Figure 2, but the patterns on the crisis periods differ for each group. For the 2008 financial crisis, non-routine manual workers seem to be unaffected, with wages rising somewhat steadily between 2008-2011; routine cognitive workers was the only group to face a decline in wages between 2009-2010; and routine manual workers face a stagnation on wages until 2011, while other groups recovered the rising pattern in 2010. For the 2013-2015 crisis, non-routine manual workers was the least affected, with a decline in wages only between 2014-2015 and fully recovered between 2015-2016, while routine manual and non-routine cognitive workers recover the wage levels only in 2017 and routine cognitive workers never recovered the 2013 wage level. The importance of this observations are less for the actual behavior of each group and more for the notice of remarkably distinct behaviors of each group, showing that the skill classification of the occupations seem to be relevant.

Yet, the most important pattern to be observed in Figure 3 are the comparison between non-routine and routine groups. Non-routine manual workers have the smallest mean wage in 2006, and non-routine cognitive workers have the biggest mean wage in 2006, leaving the routine manual and routine cognitive workers in the middle of the distribution. If it happens to be an ongoing wage polarization, it would be expected to be seen a prominent rise in both non-routine groups in relation to the routine groups. However, the group with the biggest rise in wages was the routine manual, followed by non-routine manual, non-routine cognitive, and

[2] For the remainder of the analysis, the wages are converted to 2021 real hourly values using the contract weekly hours and Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Figure 2 – Mean wages, 2006-2019

(Index based on 2006 level)



Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

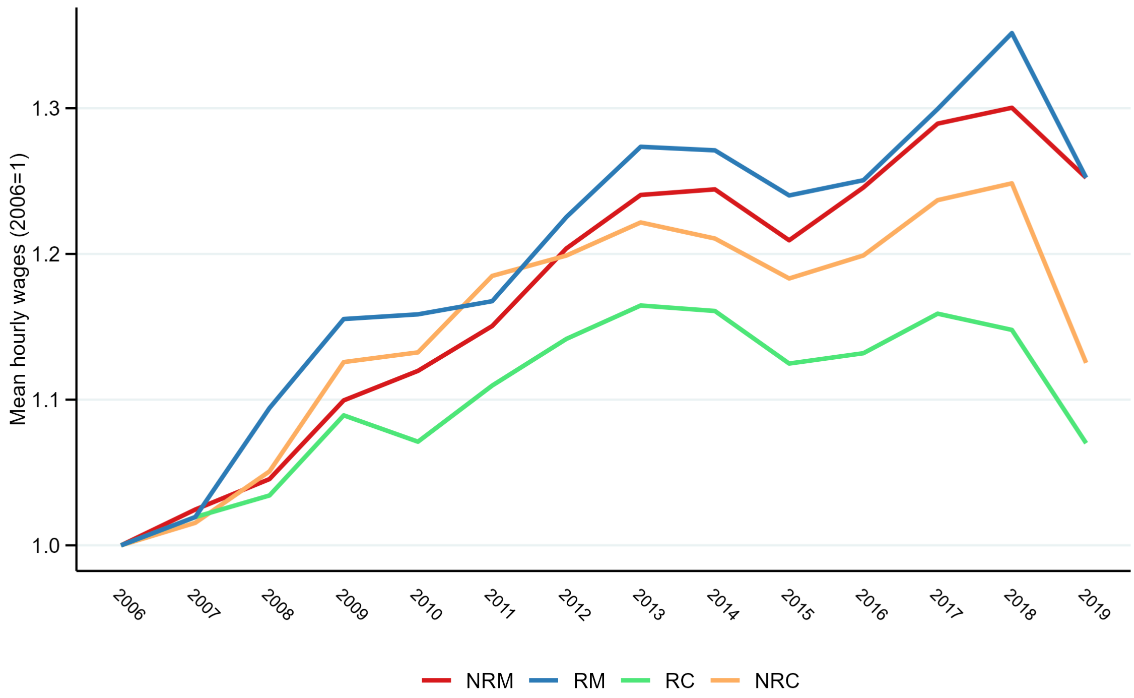
routine cognitive. So, it seems to be the case of an appreciation of manual workers in general rather than a wage polarization trend.

Now, for employment relationships, Figure 4 shows the general trends for the number of employment relationships for the period of 2006-2019, while Figure 5 does the same for the groups of skill classification. Here, the trends differ from the wages pattern. Figure 4 shows that the 2008-2009 period suffered a considerable stagnation on the employment relationship rising, recovered almost fully in 2009-2010 — an increase from 1.08 to 1.18 in 2007-2008, to 1.21 in 2009 and 1.32 in 2010. The 2014-2016 political crisis, however, has a much more pronounced effect than it had on wages, with the number of employment relationships falling almost to the level of 2010 in 2016 and falling bellow it in 2017, when it raised to approximately the same level of 2010 and 2016. In short, a rise in the employment relationships that took 4 years, fell off in two years and never fully recovered.

As for each group of workers, again, the impact suffered from each crisis was different. Figure 5 show that non-routine manual and routine cognitive workers had the same behavior as the general trend, with a huge reduction on the rising step between 2008-2009, but recovering on 2009-2010. On the other hand, the routine manual workers had a much more prominent stagnation (the index number goes from 1.159 in 2008 to 1.158 in 2009), but also recovers the rising step in 2009-2010, and the non-routine cognitive workers had not suffered any effect from the 2008 financial crisis.

Figure 3 – Mean wages by skill classification, 2006-2019

(Index based on 2006 level)



Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

For the 2014-2016 political crisis, the non-routine manual and routine manual workers show the same behavior and somewhat similar to the general trend: a noticeable decline between 2014-2016, a smooth in the decline in 2016-2017 and a slow recovery from 2017-2019, with 2019 levels being close to 2012 and 2011, respectively. Non-routine cognitive workers, however, suffered much less from this crisis, with a slight decline between 2014-2016 (with 2016 level still close to 2013 level) and a rise from 2016-2019 that leaves the non-routine cognitive workers in 2019 with the highest level of employment relationships of the sample. The routine manual workers, though, suffer quite the opposite, as the decline from 2014-2016 leaves the level of employment relationships below the 2007 level, and declining a little more from 2016-2017, almost to the level of 2006, and stagnating until 2019.

Moreover, Figure 5 do shows some signs of job polarization. Although the routine cognitive workers did not suffer as much as the routine manual workers on both of the crises, non-routine workers become the groups with most prominent rise on employment relationships approximately since 2012, with a bigger gap since 2016. This can either be a sign of a rise on the demand for non-routine workers (which corroborates with the routinization hypothesis), or a shortage in the supply of routine workers. In addition, Figure 6 shows the percent change on the employment relationships of each group side by side, ordered by the mean wages in 2006 (as a proxy for the skill index). This figure do shows some resemblance of a U-shaped form since 2009, which corroborates with the routinization hypothesis, even though there is some

Figure 4 – Employment relationships, 2006-2019

(Index based on 2006 level)

Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

bias towards the non-routine cognitive workers.

4.2 Export data about computers

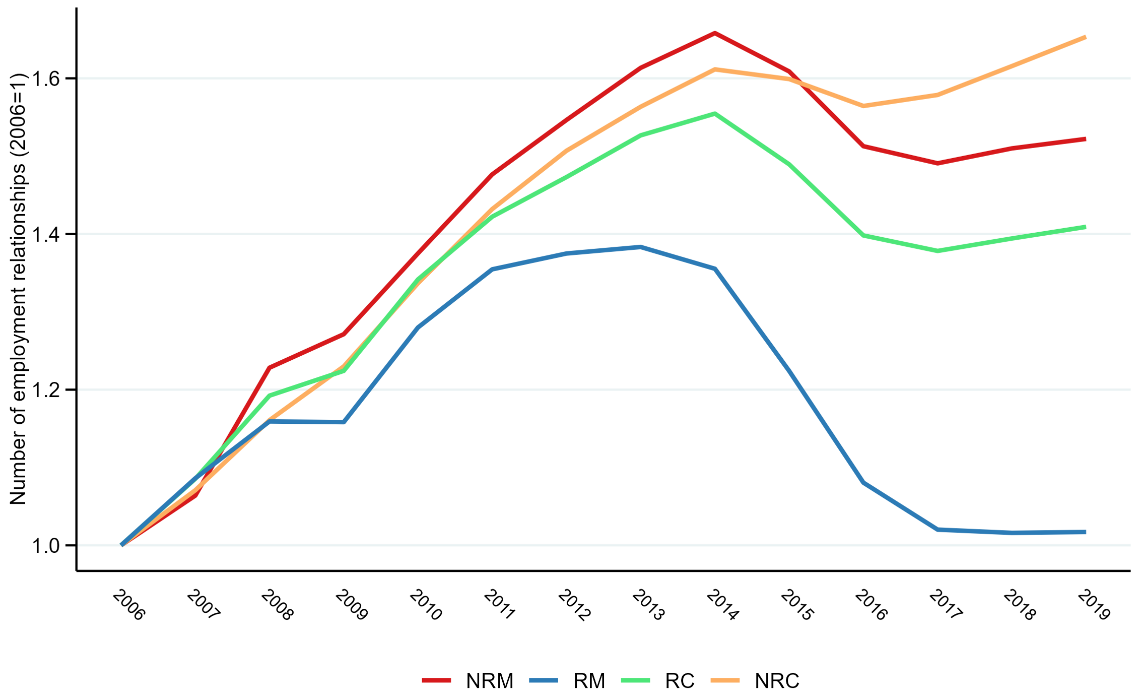
The export data comes from the from the World Integrated Trade Solution (WITS), a software developed by the World Bank that compiles data from several databases about international trade. It provides data about gross exports at the level of Harmonized System (HS) 6 digit Product for several countries from 1992 to 2022. As it does not presents data for the 847130 HS 6 digit code, which comprises personal computers, the export index for the Equation 4 is built with data for the whole 8471 heading of the HS. It also seems reasonable to do that because the 8471 heading comprises other ICTs components that are correlated with professional usage of computers, like network devices, server storage devices, *etc*.

Figure 7 shows the exported value in billions of U.S. dollars for the selected countries and Figure 8 shows data of quantity[3] of items exported by the same countries. The charts use logarithmic scale on the vertical axis due to the huge difference between the top exporter and the other ones. Only five exporters were kept in the final sample as a mean to mitigate the

[3] Some years had missing values for quantity and were treated as follow: countries with quantity reported for less than half of the years were dropped from the sample; the missing values for the other countries were replaced by the nearest previous reported value.

Figure 5 – Employment relationships by skill classification, 2006-2019

(Index based on 2006 level)



Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

possibility of an influence from the Brazilian domestic demand for computers (and other ICT components), thus compromising the exogeneity of the measure, as done by Boustan, Choi, and Clingingsmith (2022). The selected exporters were kept as they figured as top 3 exporters in some of the years in the sample[4].

The unit prices do seem to fall to some degree between 2012-2016, as the value exported fell and the quantity exported remained somewhat constant, which corroborates with the hypothesis of exogenous falling in the prices of computers. However, the trend does not seem to be happening either before 2012 or after 2017. Nevertheless, the construction of a measure related to the supply of computers prevents the model in Equation 5 from being affected by some erroneous assumption about the prices. Lastly, as the export index intends to serve as a proxy for the supply of computers, it is constructed with the quantity of items exported.

4.2.1 Data Basis project

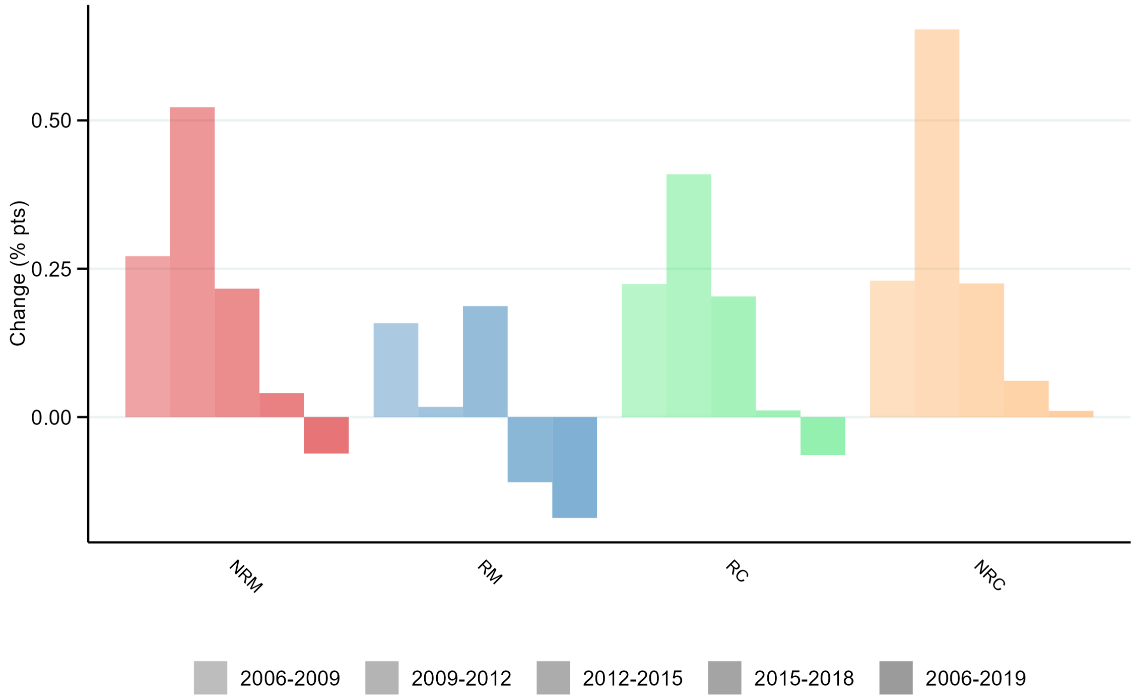
With the exception of WITS, all the other databases used in this undergraduate thesis were obtained from the Data Basis project[5], a project from a non-profit organization that makes available, among other components, a data lake with hundreds of structured, normalized, and up-to-

4 Thailand also figured in top 3 exporters for some year, but was removed from the sample do to insufficient data.

5 The organization's website in Portuguese is available at <https://basedosdados.org/>.

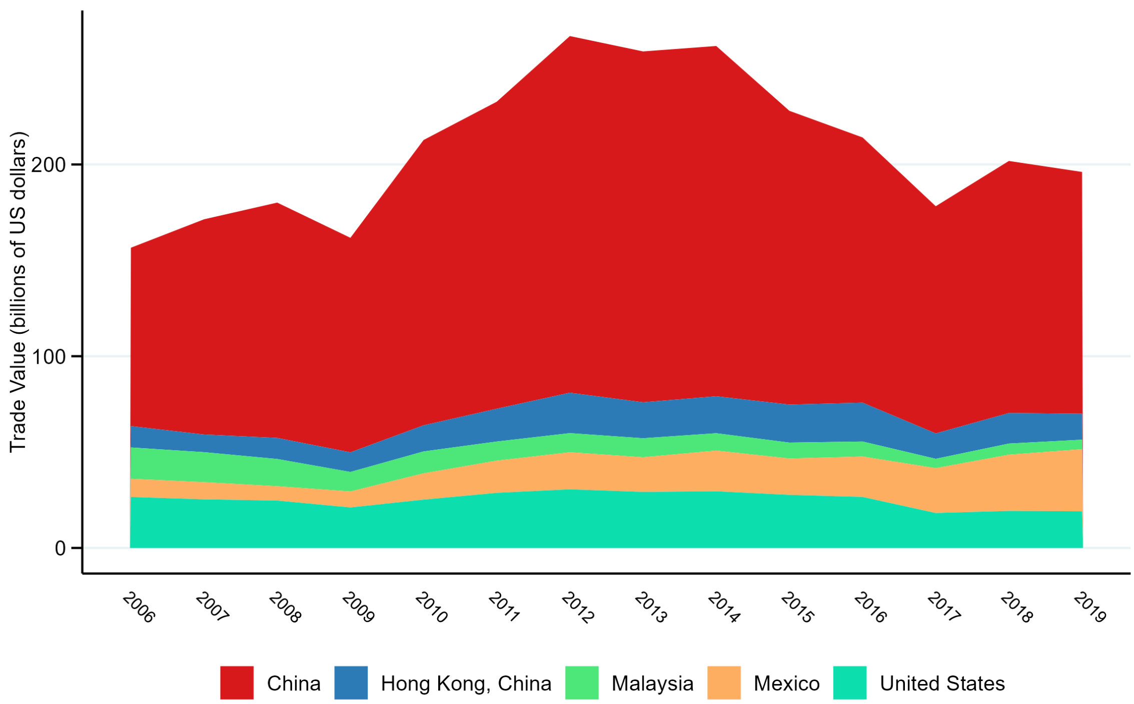
Figure 6 – Changes in employment relationships, 2006-2019

(Percent change over the intervals)



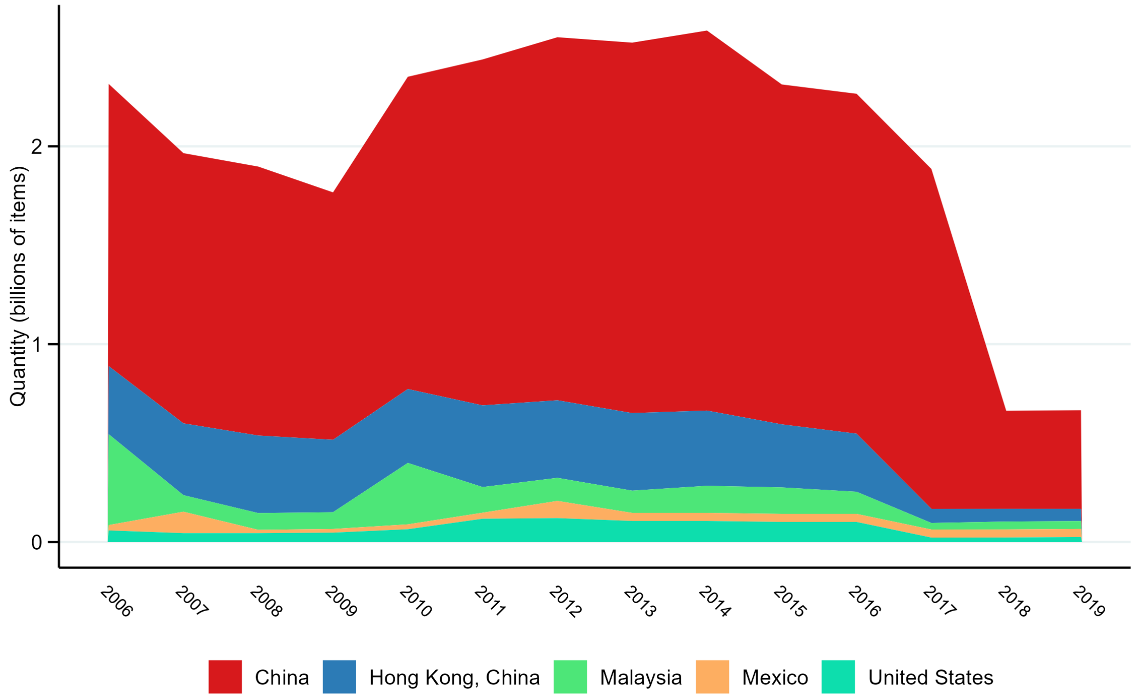
Source: Own elaboration based on RAIS data and Adamczyk, Monasterio, and Fochezatto (2021).

Figure 7 – Exports of HS 8471 heading, 2006-2019



Source: Own elaboration based on WITS data.

Figure 8 – Exports of HS 8471 heading, 2006-2019



Source: Own elaboration based on WITS data.

date tables (DAHIS et al., 2022). Several databases from the Brazilian Federal Government — and related institutes, such as the Brazilian Institute of Geography and Statistics (IBGE) and the Institute of Applied Economic Research (IPEA) — are made available on this project.

5 Results

5.1 Effectiveness of the task-based model for the Brazilian formal labor market

To assess the relevance of the task-based model in explaining the trends of the Brazilian formal labor market, three slight variations of the model described in Equation 3 were evaluated. They were all based on the One Way Fixed Effects Model described in Baltagi (2021, p. 16-17), which demonstrates that estimating the time demeaned model gives the same estimators as the original model. So, the estimated models were

$$\left[\log(y_{j,c,t}) - \overline{\log(y_t)}\right] = (\lambda_c - \overline{\lambda}) + (\varepsilon_{j,c,t} - \overline{\varepsilon_t}) \tag{6}$$

$$\left[\log(y_{j,c,t}) - \overline{\log(y_t)}\right] = (\lambda_c - \overline{\lambda}) + (\alpha_j - \overline{\alpha}) + (\varepsilon_{j,c,t} - \overline{\varepsilon_t}) \tag{7}$$

$$\left[\log(y_{j,c,t}) - \overline{\log(y_t)}\right] = (\alpha_j - \overline{\alpha}) + (\varepsilon_{j,c,t} - \overline{\varepsilon_t}) \tag{8}$$

where

$$\overline{\log(y_t)} = \frac{1}{C}\sum_c \frac{1}{J}\sum_j \log(y_{j,c,t})$$

$$\overline{\lambda} = \frac{1}{C}\sum_c \lambda_c$$

$$\overline{\alpha} = \frac{1}{J}\sum_j \alpha_j$$

with C and J being the number of skill classification groups and industry sectors, respectively. The model 8, however, was estimated only for the evaluation of a possible bias in the model 7. This evaluation is available on the Appendix B.

Table 3 shows the results for the regressions of the models 6 and 7 with wages as the explained variable. Although these models do not represent any causal relationship and can be biased by omitted variables, the estimates do have some interesting results to be interpreted. Almost all of the variables are extremely significant and the inclusion of the industry sector effects did not change the coefficients of the estimators for the parameters for skill group dummies, which is a good indicator of the nonexistence of omitted variable bias. Again, further specifications could be thought to investigate this bias, but that would also require additional theoretical investigation to compose the model.

Now, the interpretation of the coefficients is very simple: a worker with an occupation composed mostly of tasks belonging to the RM group is expected to receive, on average, 14.7%

Table 3 – Results for the regressions of models 6 and 7 for wages

Characteristic	(6)		(7)	
	Coef. (SE)[1,2]	p-value	Coef. (SE)[1,2]	p-value
Skill group				
NRM	—		—	
RM	0.1472 (0.095)	0.12	0.1472*** (0.041)	<0.001
RC	0.3985*** (0.103)	<0.001	0.3985*** (0.041)	<0.001
NRC	1.0368*** (0.113)	<0.001	1.0368*** (0.048)	<0.001
Industry sector effects			X	
R^2	0.527		0.918	
Adjusted R^2	0.521		0.916	
No. Obs.	1,176		1,176	

[1] * p<0.05; ** p<0.01; *** p<0.001
[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.

more for his hourly wage than a worker with an occupation composed mostly of tasks belonging to the NRM group. For the RC and NRC groups, the premium is 39.9% and 103.7%, respectively. However, considering that the order of the groups in relation to the first period (2006) mean wages is, from lowest to highest, { NRM, RM, RC, NRC }, there is no sign of wage polarization effect, as it would require a negative sign on RM and RC coefficients. Although the task model exhibits an excellent adherence to the Brazilian formal labor market, the values encountered for the coefficients indicate a monotone change in wages.

Table 4 shows the analogous results for the regressions of the models 6 and 7 in relation to employment relationships. Again, is important to stress that these models do not represent any causal relationship and can be biased by omitted variables, although there was no sign of the existence of omitted variable bias. In opposition to what was seen on the wages regressions, this time almost all of the variables are not significant, but the inclusion of the industry sector effects did not change the coefficients of the estimators for the parameters for skill group dummies, which is an indicator of consistency. Again, further specifications could be thought to investigate this bias, but that would also require additional theoretical investigation to compose the model.

The interpretation of the coefficients is very similar: an occupation composed mostly of tasks belonging to the RM group is expected to present, on average, 27.1% more employment relationships than an occupation composed mostly of tasks belonging to the NRM group. For the RC and NRC groups, the difference is 65.4% and 12.7%, respectively. However, it is not statistically possible to assert that the RM or NRC coefficients differ from zero.

Table 4 – Results for the regressions of models 6 and 7 for employment relationships

Characteristic	(6)		(7)	
	Coef. (SE)[1, 2]	p-value	Coef. (SE)[1, 2]	p-value
Skill group				
NRM	—		—	
RM	0.2705 (0.667)	0.7	0.2705 (0.319)	0.4
RC	0.6507 (0.665)	0.3	0.6507** (0.230)	0.005
NRC	0.1265 (0.657)	0.8	0.1265 (0.264)	0.6
Industry sector effects			X	
R^2	0.012		0.827	
Adjusted R^2	-0.001		0.822	
No. Obs.	1,176		1,176	

[1] $p<0.05$; $p<0.01$; $p<0.001$
[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.

5.2 Relation of computer adoption and the labor market

For the main evaluation of this undergraduate thesis, three variations of the model specified in Equation 5 were estimated. All of them were based on the Two Way Fixed Effects Model described in Baltagi (2021, p. 48-49). As in the One Way Fixed Effects Model, Baltagi (2021) demonstrates that estimating the time demeaned and individual model gives the same estimators as the original model. So, the estimated models were

$$\widetilde{\log(y_{j,t})} = \beta \widetilde{Exp}_{j,t} + \widetilde{\varepsilon}_{j,t} \tag{9}$$

where

$$\begin{aligned}
\widetilde{\log(y_{j,t})} &= \left[\log(y_{j,t}) - \overline{\log(y_t)} - \overline{\log(y_j)} + \overline{\log(y)}\right] \\
\overline{\log(y_t)} &= \frac{1}{J}\sum_j \log(y_{j,t}) \\
\overline{\log(y_j)} &= \frac{1}{T}\sum_t \log(y_{j,t}) \\
\overline{\log(y)} &= \frac{1}{T}\sum_t \frac{1}{J}\sum_j \log(y_{j,t})
\end{aligned}$$

$$\begin{aligned}
\widetilde{Exp_{j,t}} &= \left[Exp_{j,t} - \overline{Exp_t} - \overline{Exp_j} + \overline{Exp)}\right] \\
\overline{Exp_t} &= \frac{1}{J}\sum_j Exp_{j,t} \\
\overline{Exp_j} &= \frac{1}{T}\sum_t Exp_{j,t} \\
\overline{Exp} &= \frac{1}{T}\sum_t \frac{1}{J}\sum_j Exp_{j,t}
\end{aligned}$$

And

$$\widetilde{\log(y_{j,c,t})} = \gamma \widetilde{Exp}_{j,t} + \widetilde{\varepsilon}_{j,c,t} \tag{10}$$

$$\widetilde{\log(y_{j,c,t})} = \delta \widetilde{Exp \times} \lambda_{j,c,t} + \widetilde{\varepsilon}_{j,c,t} \tag{11}$$

where

$$\begin{aligned}
\widetilde{\log(y_{j,c,t})} &= \left[\log(y_{j,c,t}) - \overline{\log(y_t)} - \overline{\log(y_{j,c})} + \overline{\log(y)}\right] \\
\overline{\log(y_t)} &= \frac{1}{C}\sum_c \frac{1}{J}\sum_j \log(y_{j,c,t}) \\
\overline{\log(y_{j,c})} &= \frac{1}{T}\sum_t \log(y_{j,c,t})
\end{aligned}$$

$$\begin{aligned}
\widetilde{Exp_{j,t}} &= \left[Exp_{j,t} - \overline{Exp_t} - \overline{Exp_j} + \overline{Exp}\right] \\
\overline{Exp_t} &= \frac{1}{C}\sum_c \frac{1}{J}\sum_j Exp_{j,t} = \frac{1}{J}\sum_j Exp_{j,t} \\
\overline{Exp_j} &= \frac{1}{T}\sum_t Exp_{j,t} \\
\overline{Exp} &= \frac{1}{T}\sum_t \frac{1}{C}\sum_c \frac{1}{J}\sum_j Exp_{j,t} = \frac{1}{T}\sum_t \frac{1}{J}\sum_j Exp_{j,t}
\end{aligned}$$

$$\begin{aligned}
\widetilde{Exp \times \lambda}_{j,c,t} &= \left[Exp \times \lambda_{j,c,t} - \overline{Exp \times \lambda_t} - \overline{Exp \times \lambda_{j,c}} + \overline{Exp}\right] \\
\overline{Exp \times \lambda_t} &= \frac{1}{C}\sum_c \frac{1}{J}\sum_j Exp \times \lambda_{j,c,t} \\
\overline{Exp \times \lambda_j} &= \frac{1}{T}\sum_t Exp \times \lambda_{j,c,t} \\
\overline{Exp} &= \frac{1}{T}\sum_t \frac{1}{C}\sum_c \frac{1}{J}\sum_j Exp \times \lambda_{j,c,t}
\end{aligned}$$

with C, J, and T being the number of skill classification groups, industry sectors, and periods, respectively. The difference between models 9 and 10 is that the last one takes account of skill classification groups' fixed effects, while the former ignores it. In addition, The difference between models 10 and 11 is that the last one also takes account of distinct effects of exposure to each skill classification group

Table 5 shows the results for the regressions of models 9, 10, and 11 with wages as the explained variable. On the contrary of what was hypothesized, it is not possible to state that exposure to computers has any effect on wages. In all the models the estimated coefficient for the *Exposure to computers* measure and its interactions were not statistically different from zero when the explained variable was $\log(wages)$.

Table 5 – Results for the regressions of models 9, 10, and 11 for wages

Dependent variable: $\log(wages)$	(9) Beta (SE)[1,2]	(10) Beta (SE)[1,2]	(11) Beta (SE)[1,2]
Exposure to computers	-0.0117 (0.109)	-0.0015 (0.064)	—
Exposure to computers x Skill type			
NRM	—	—	-0.0748 (0.079)
RM	—	—	-0.1070 (0.080)
RC	—	—	0.0825 (0.075)
NRC	—	—	0.0023 (0.080)
R²	0	0	0.008
Adjusted R²	-0.131	-0.09	-0.092
No. Obs.	294	1,176	1,092

[1] * p<0.05; ** p<0.01; *** p<0.001
[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.

Table 6 shows the analogous results for the log of employment relationships as the explained variable. Now, most of the results were statistically significant at least at a 5% level.

Table 6 – Results for the regressions of models 9, 10, and 11 for employment relationships

Dependent variable: $\log(jobs)$	(9)	(10)	(11)	
	Beta (SE)[1,2]	Beta (SE)[1,2]	Beta (SE)[1,2]	Beta (SE)[1,2]
Exposure to computers	-0.8675* (0.347)	-0.5762** (0.205)	-0.5641* (0.224)	
Exposure to computers x Skill type				
NRM				-0.5641* (0.224)
RM			0.5519* (0.223)	-0.0122 (0.214)
RC			-0.0443 (0.196)	-0.6084** (0.200)
NRC			-0.439* (0.203)	-1.0031*** (0.207)
R^2	0.042	0.016	0.045	0.045
Adjusted R^2	-0.083	-0.072	-0.05	-0.05
No. Obs.	294	1176	1092	1092

[1] * $p<0.05$; ** $p<0.01$; *** $p<0.001$

[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.

6 Concluding remarks

References

ABRAMOVAY, Ricardo. O fim do trabalho. Entre a distopia e a emancipação. **Estudos Avançados**, v. 35, p. 139–150, 101 Apr. 2021.

ACEMOGLU, D.; ZILIBOTTI, F. Productivity Differences. **The Quarterly Journal of Economics**, v. 116, p. 563–606, 2 May 2001.

ACEMOGLU, Daron. **Harms of AI**. [S.l.: s.n.], Sept. 2021.

ACEMOGLU, Daron; AUTOR, David. Skills, Tasks and Technologies: Implications for Employment and Earnings. In: [s.l.: s.n.], 2011. v. 4, p. 1043–1171.

ACEMOGLU, Daron; RESTREPO, Pascual. Robots and Jobs: Evidence from US Labor Markets. **Journal of Political Economy**, University of Chicago Press, v. 128, p. 2188–2244, 6 June 2020.

________. Tasks, Automation, and the Rise in U.S. Wage Inequality. **Econometrica**, The Econometric Society, v. 90, p. 1973–2016, 5 2022.

________. Unpacking Skill Bias: Automation and New Tasks. **AEA Papers and Proceedings**, American Economic Association, v. 110, p. 356–361, May 2020.

ADAMCZYK, Willian; EHRL, Philipp; MONASTERIO, Leonardo. **Skills and employment transitions in Brazil**. [S.l.]: ILO, 2022. v. 65.

ADAMCZYK, Willian Boschetti; MONASTERIO, Leonardo; FOCHEZATTO, Adelar. Automation in the future of public sector employment: the case of Brazilian Federal Government. **Technology in Society**, Elsevier Ltd, v. 67, p. 101722, Nov. 2021.

ALMEIDA, Rita K; CORSEUIL, Carlos H L; POOLE, Jennifer P. **The Impact of Digital Technologies on Routine Tasks Do Labor Policies Matter?** [S.l.: s.n.], 2017. Available from: <http://hdl.handle.net/10986/28364>.

AUTOR, D. H.; LEVY, F.; MURNANE, R. J. The Skill Content of Recent Technological Change: An Empirical Exploration. **The Quarterly Journal of Economics**, v. 118, p. 1279–1333, 4 Nov. 2003.

AUTOR, David H. Work of the Past, Work of the Future. **AEA Papers and Proceedings**, v. 109, p. 1–32, May 2019.

BALTAGI, Badi H. **Econometric Analysis of Panel Data**. [S.l.]: Springer International Publishing, 2021.

BOUSTAN, Leah Platt; CHOI, Jiwon; CLINGINGSMITH, David. Automation After the Assembly Line: Computerized Machine Tools, Employment and Productivity in the United States. [S.l.], Aug. 2022. Available from: <https://www.nber.org/papers/w30400>.

DAHIS, Ricardo et al. Data Basis: Universalizing Access to High-Quality Data. [S.l.], 2022. DOI: `https://doi.org/10.31235/osf.io/r76yg`.

FIRPO, Sergio et al. WIDER Working Paper 2021/162-The changing nature of work and inequality in Brazil (2003–19): A descriptive analysis, 2021. Available from: <`https://doi.org/10.35188/UNU-WIDER/2021/102-0`>.

FRANK, Morgan R et al. Toward understanding the impact of artificial intelligence on labor. **Proceedings of the National Academy of Sciences**, v. 116, p. 6531–6539, 14 2019.

FUNCHAL, Bruno; JUNIOR, Jadir Soares. **Understanding demand for skills after technological trade liberalization**. [S.l.: s.n.], 2012.

KATZ, L. F.; MURPHY, K. M. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. **The Quarterly Journal of Economics**, v. 107, p. 35–78, 1 Feb. 1992.

KRUEGER, A. B. How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989. **The Quarterly Journal of Economics**, v. 108, p. 33–60, 1 Feb. 1993.

MACKINNON, James G; WHITE, Halbert. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. **Journal of Econometrics**, v. 29, p. 305–325, 3 Sept. 1985.

RIVA, Flavio Luiz Russo. **Computerization, occupational tasks and the labor market: evidence from a natural experiment in Brazil**. 2016. Fundação Getulio Vargas, São Paulo. Available from: <`http://hdl.handle.net/10438/16233`>.

ANNEX A – Availability of occupation codes in RAIS

Year	CBO availability	Number of relationships	Share of relationships by year
2006	Available	50,700,755	99.999%
2006	Unavailable	265	0.001%
2007	Available	54,649,128	100.000%
2007	Unavailable	1	0.000%
2008	Available	59,706,419	100.000%
2009	Available	61,126,589	99.999%
2009	Unavailable	307	0.001%
2010	Available	66,747,268	100.000%
2010	Unavailable	34	0.000%
2011	Available	70,956,582	99.980%
2011	Unavailable	14,543	0.020%
2012	Available	73,286,517	99.945%
2012	Unavailable	39,968	0.055%
2013	Available	75,332,652	99.910%
2013	Unavailable	67,858	0.090%
2014	Available	76,041,075	99.913%
2014	Unavailable	66,204	0.087%
2015	Available	72,102,638	99.900%
2015	Unavailable	72,464	0.100%
2016	Available	67,073,222	99.894%
2016	Unavailable	71,376	0.106%
2017	Available	65,585,781	99.893%
2017	Unavailable	70,101	0.107%
2018	Available	66,154,789	99.910%
2018	Unavailable	59,903	0.090%
2019	Available	66,664,399	99.995%
2019	Unavailable	3,018	0.005%
2020	Available	65,665,754	99.613%
2020	Unavailable	255,440	0.387%
2021	Available	70,268,833	99.641%
2021	Unavailable	253,132	0.359%

Source: Own elaboration based on RAIS data.

ANNEX B – Evaluation of bias on industry sector parameters of models 7 and 8

The main interest in this results is the possibility of bias in the industry sector coefficients, which did not appear, as the majority of the coefficients kept the same with and without the skill group variable — 95% of the coefficients for the wage models, and 80% of the coefficients for the employment relationships model.

Characteristic	7		8	
	Coef. (SE)[1,2]	p-value	Coef. (SE)[1,2]	p-value
CNAE industry sector				
Administração pública, defesa e seguridade social	—		—	
Agricultura, pecuária, produção florestal, pesca e aqüicultura	-0.3656*** (0.064)	<0.001	-0.3656 (0.256)	0.2
Alojamento e alimentação	-0.5785*** (0.102)	<0.001	-0.5785** (0.207)	0.005
Artes, cultura, esporte e recreação	-0.2582** (0.086)	0.003	-0.2582 (0.227)	0.3
Atividades administrativas e serviços complementares	-0.3831*** (0.065)	<0.001	-0.3831 (0.242)	0.11
Atividades financeiras, de seguros e serviços relacionados	0.2592 (0.150)	0.085	0.2592 (0.355)	0.5
Atividades imobiliárias	-0.3316*** (0.066)	<0.001	-0.3316 (0.260)	0.2
Atividades profissionais, científicas e técnicas	-0.0792 (0.076)	0.3	-0.0792 (0.297)	0.8
Comércio; reparação de veículos automotores e motocicletas	-0.4707*** (0.074)	<0.001	-0.1077903	0.04
Construção	-0.2066*** (0.057)	<0.001	-0.2066 (0.269)	0.4
Educação	0.1874 (0.134)	0.2	0.1874 (0.364)	0.6
Eletricidade e gás	0.7208*** (0.102)	<0.001	0.7208** (0.262)	0.006
Água, esgoto, atividades de gestão de resíduos e descontaminação	0.0743 (0.101)	0.5	0.0743 (0.317)	0.8
Indústrias de transformação	-0.1696** (0.055)	0.002	-0.1696 (0.271)	0.5
Indústrias extrativas	0.4979*** (0.079)	<0.001	0.4979 (0.307)	0.11
Informação e comunicação	0.0297 (0.074)	0.7	0.0297 (0.250)	¿0.9
Organismos internacionais e outras instituições extraterritoriais	0.3386*** (0.070)	<0.001	0.3386 (0.273)	0.2
Outras atividades de serviços	-0.2250*** (0.065)	<0.001	-0.2250 (0.282)	0.4
Saúde humana e serviços sociais	-0.2189*** (0.059)	<0.001	-0.2189 (0.249)	0.4
Serviços domésticos	-0.6639*** (0.117)	<0.001	-0.6639*** (0.197)	<0.001
Transporte, armazenagem e correio	-0.0712 (0.066)	0.3	-0.0712 (0.258)	0.8
R^2	0.918		0.391	
Adjusted R^2	0.916		0.374	
No. Obs.	1,176		1,176	

[1] * $p<0.05$; ** $p<0.01$; *** $p<0.001$

[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.

Characteristic	7		8	
	Coef. (SE)[1,2]	p-value	Coef. (SE)[1,2]	p-value
CNAE industry sector				
Administração pública, defesa e seguridade social	—		—	
Agricultura, pecuária, produção florestal, pesca e aqüicultura	-2.0903** (0.798)	0.009	-2.0903** (0.798)	0.009
Alojamento e alimentação	-1.0275023	0.023	-0.9785007	0.017
Artes, cultura, esporte e recreação	-3.2638*** (0.489)	<0.001	-3.2638*** (0.480)	<0.001
Atividades administrativas e serviços complementares	-0.3698 (0.565)	0.5	-0.3698 (0.578)	0.5
Atividades financeiras, de seguros e serviços relacionados	-3.0181*** (0.882)	<0.001	-3.0181** (0.946)	0.001
Atividades imobiliárias	-3.9409*** (0.440)	<0.001	-3.9409*** (0.491)	<0.001
Atividades profissionais, científicas e técnicas	-1.9080*** (0.466)	<0.001	-1.9080*** (0.531)	<0.001
Comércio; reparação de veículos automotores e motocicletas	0.2371 (0.483)	0.6	0.2371 (0.564)	0.7
Construção	-1.2120 (0.679)	0.075	-1.2120 (0.691)	0.08
Educação	-1.2436146	0.015	-1.2610566	0.016
Eletricidade e gás	-4.2553*** (0.582)	<0.001	-4.2553*** (0.616)	<0.001
Água, esgoto, atividades de gestão de resíduos e descontaminação	-3.0543*** (0.526)	<0.001	-3.0543*** (0.498)	<0.001
Indústrias de transformação	-0.2372 (0.624)	0.7	-0.2372 (0.660)	0.7
Indústrias extrativas	-3.6409*** (0.582)	<0.001	-3.6409*** (0.598)	<0.001
Informação e comunicação	-2.1590*** (0.526)	<0.001	-2.1590*** (0.558)	<0.001
Organismos internacionais e outras instituições extraterritoriais	-7.1360*** (0.397)	<0.001	-7.1360*** (0.419)	<0.001
Outras atividades de serviços	-1.6148*** (0.431)	<0.001	-1.6148*** (0.437)	<0.001
Saúde humana e serviços sociais	-1.0261791	0.022	-1.027713	0.022
Serviços domésticos	-7.3112*** (0.682)	<0.001	-7.3112*** (0.643)	<0.001
Transporte, armazenagem e correio	-0.6656013	0.039	-0.6879054	0.045
R^2	0.827		0.815	
Adjusted R^2	0.822		0.809	
No. Obs.	1,176		1,176	

[1] * p<0.05; ** p<0.01; *** p<0.001
[2] SE = MacKinnon and White (1985) HC1 standard error

Source: Own elaboration.