Technology in Society 67 (2021) 101722

Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc

Automation in the future of public sector employment: the case of Brazilian Federal Government

Willian Boschetti Adamczyk [a,b,*], Leonardo Monasterio [b,c], Adelar Fochezatto [d,e]

[a] *Graduate Program in Economics at the Pontifical Catholic University of Rio Grande do Sul – PPGE/PUCRS, Av. Ipiranga, 6681, 50. CEP 90619-900, Porto Alegre, Brazil*
[b] *National School of Public Administration – ENAP, Brasília, Brazil*
[c] *Brazilian Institute of Education, Development and Research – IDP-DF, Brasília, Brazil*
[d] *Graduate Program in Economics at the Pontifical Catholic University of Rio Grande do Sul – PPGE/PUCRS, Porto Alegre, Brazil*
[e] *Research Fellow of the National Council for Scientific and Technological Development – CNPq, Brazil*

ARTICLE INFO

JEL classification:
JEL
J24
Human Capital
Skills
Occupational Choice
Labor Productivity

Keywords:
Automation
Machine learning
Public sector

ABSTRACT

What is the impact of automation on public sector employment? Using machine learning and natural language processing algorithms, this study estimates which occupations and agencies of the Brazilian Federal Government are most susceptible to automation. We contribute to the literature by introducing Bartik Occupational Tasks (BOT), an objective method used to estimate automation susceptibility that avoids subjective or *ad hoc* classifications. We show that approximately 20% of Brazilian public sector employees work in jobs with a high potential of automation in the coming decades. Government occupations with lower schooling and lower salary levels are most susceptible to future automation.

1. Introduction

Automation technology studies discuss future changes in the labor market reflecting general concern for the risk of unemployment caused by replacing humans with machines [1,25]. Despite the volume and quality of the literature, there is no record of studies that consider the distinction between the private sector and public sector occupations.

While the private sector has the necessary flexibility to adjust to technical changes by means of hiring, firing, and reallocating employees relying on price signal mechanisms, public sector rigidity hinders it from adjusting its labor force to address technological change. In the absence of a market mechanism, the allocation of workers is based on planning decisions made by government administrators. The Brazilian State's well-known and persistent problems reduce the rhythm of incorporation of new automation technology, contributing to the lag in public sector productivity compared to that of the private sector [50].

This research sets out to identify those occupations in which automation technology could be introduced to increase the productivity of public services and reduce costs. To that end, it presents predictive algorithms to determine the susceptibility to automation of occupations in the sphere of the Federal Executive branch in Brazil, enabling quantitative analysis of the impacts stemming from the introduction of automation by federal government agencies.

Regarding automation forecasts, the international literature [5,6,25,51] and the national literature [3,4,38] are based on experts' opinions in that area. This work innovates by applying an objective method inspired by the Bartik Instruments method [12], making it possible to estimate the occupational trajectory based on information observed in the labor market. The proposed method, known as Bartik Occupational Tasks (BOT), could be helpful not only for the federal government but also for other national and international federative entities.

Natural Language Processing (NLP) and Machine Learning algorithms have made it possible to extract quantitative information from texts that list the tasks involved in each one of the 2627 occupations listed in the Brazilian Classification of Occupations (*Classificação Brasileira de Ocupações* – CBO). This was the starting point for estimating the

* Corresponding author. Graduate Program in Economics at the Pontifical Catholic University of Rio Grande do Sul – PPGE/PUCRS, Av. Ipiranga, 6681, 50. CEP 90619-900, Porto Alegre, Brazil.
E-mail address: willianbadamczyk@gmail.com (W.B. Adamczyk).

https://doi.org/10.1016/j.techsoc.2021.101722
Received 23 March 2021; Received in revised form 17 July 2021; Accepted 16 August 2021
Available online 25 August 2021
0160-791X/

susceptibility to automation of the 398 occupations in the public sector. Access to the complete microdata of around 520,000 government employees registered in the Integrated Personnel Administration System (*Sistema Integrado de Administração de Recursos Humanos* – SIAPE) in December 2017, combined with the descriptive texts of Brazilian occupations listed in the CBO and information identified in the Annual Social Information Report (*Relação Annual de Informações Sociais* – RAIS) formed the database for this research.

This article's main contribution is to draw a detailed profile of the potential impact of automation in empirical terms. It shows that 20% of civil servants have occupations that are highly susceptible to being automated. Those occupations are commonly associated with lower schooling and lower salary levels.

2. Automation and employment: a literature review

Despite the undeniable long-term gains innovation can provide, studies on innovation have focused on private organizations in both private and public sectors. In the public sector, even small innovations can yield significant outcomes and spillovers to society [7]. Conditions vary across public organizations, but management and organizational strategies play essential roles in public sector innovation [22].

As defined by the OECD Oslo Manual [43], "an innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations". Since mechanisms of control and backup are in place, technological innovations, such as digitalization and Artificial Intelligence (AI), can provide better means of participation and legitimacy, a more efficient government [47], improved institutional quality, and less corruption [2].

Artificial Intelligence has been used as a broad concept with various definitions. It can be understood as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" [37]. Its constituent technologies include machine learning, reinforcement learning, artificial neural networks, deep learning, computer vision, and other techniques that are being rapidly developed [18].

As Artificial Intelligence technologies extended their applications to the labor market, fears over technological unemployment and social tensions reemerged [33]. The literature on technological unemployment[1] provides evidence that the introduction of new information technology has made it feasible to automatize tasks formerly carried out by workers. The horizon of tasks susceptible to automation moves progressively towards tasks of greater complexity. Tasks considered to be routine today are no obstacle to the advances of technology. Automation has the potential to progress from the replacement of the simplest tasks to the most abstract ones in a question of decades [25]. Against that background, it can be considered that a profession tends to be eliminated once the greater part of its activities has been replaced. Alternatively, the profession may transform itself into another, aggregating new tasks under a new denomination. There is no guarantee that new occupations will appear at the same rate or that the respective adjustment will be painless [26].

Therefore, society accompanies with apprehension the identification of occupations susceptible to automation [11,33]. According to Autor [11], journalists and the media at large tend to underestimate the outreach of replacement of human labor by machines and ignore the strong complementarity between labor and automation. It can increase productivity, raise salaries and increase the demand for labor on cheaper services and new occupations. Identifying the occupations most affected by automation and the latter's impacts on the labor market is no light task, and as of yet, it cannot rely on the support of consolidated literature and methodologies.

Recent research efforts are conflicting in regard to the magnitude of transformations and impacts that such technology can cause through jobs elimination. The discordancy in the results can be traced with the occupation-based [25] or task-based [5,6] estimation methodologies employed. The occupation-based approach to automation seeks to identify those professions with a tendency to disappear in the coming decades and identify the impacts of technological unemployment on salaries, inequality, and income polarization. Using this approach, Frey and Osborne [25] and McKinsey Global Institute [41] have estimated that around 47% of jobs in the United States are at high risk of disappearing in the next three decades due to automation of the respective professions.

Frey and Osborne [25] adopted a methodology that attributed automation probabilities to each profession according to their identification of bottlenecks characteristics and classified the activities as being at high or low risk of being automated. In the effort to develop predictive models, they estimated the possible results of automation based on the opinions of a group of machine learning experts. The 70 occupations for which the authors felt sure all the tasks could be automated were manually classified and used as information to train the model. Based on that training, the authors inferred the automation probability for over 900 professions listed in the Standard Occupational Classification (SOC).

In Brazilian formal labor market studies, Albuquerque et al. [3,4] adapted Frey and Osborne's methodology but relied on the assessments of 69 Artificial Intelligence experts. They estimated that 54.5% of the 45.9 million jobs are in occupations for which the probability of automation is high or very high. They considered the probability to be high for those occupations that lay in the third quartile of the automation probability distribution and very high for those that reside in the upper quartile.

Some criticism of these high percentages of workers in occupations at high risk of being automated has emerged in the literature. First, Arntz, Gregory and Zierahn [5,6] considered that occupation-based estimates provide overestimations of the degree of automation. Even in those jobs considered to be at high risk of automation, some of the tasks that workers carry out may be difficult to automate. Automation's impact on occupations is heterogeneous, and, furthermore, they are capable of being remodeled and taking on a new character, qualification, and denomination.

Second, subjective classification could confuse the potential for automation and the drop in employment levels in the occupations in question. In turn, automation may be technologically feasible but not economically viable [5]. Arntz, Gregory and Zierahn [5,6] proposed an alternative form of automation probability estimation based on the tasks involved in each occupation; a more disaggregated level that makes it possible to incorporate the heterogeneity among workers. Using this approach, they estimated that in the developed countries selected for their study, just 9% of the workers are in occupations that are at high risk of disappearing.

In Brazil's case, Kubota and Maciente [38] have estimated that 56.5% of the formal employment positions in the country are in occupations vulnerable to automation, based on a scenario of already consolidated technologies and others that could be implemented, in harmony with the regulatory framework, within the next five years. Using 19 thousand tasks described in the O*NET database, they constructed a dictionary of keywords associated with automation based on the classification of activities according to their routine requirements and cognition as proposed for Germany in Spitz-Oener [49].

Maciente, Rauen and Kubota [40] showed that, compared to developed countries, most of the labor activity in Brazilian occupations is in intense activities using routine skills and low cognitive levels. The

[1] Cross-cutting discussions of technological unemployment and automation, which however are not central issues for this study, are those addressing the yields of education, and technological change by skills [1,24,26,30], offshoreability [14,36], and inequality and polarization in the labor market [8,9,28,29,35].

literature indicates that operation and control tasks that are routine and depend on physical aptitudes tend to lose space in the labor market compared to those demanding greater qualifications. Making use of automation probabilities as in Albuquerque et al. [3], Maciente, Rauen and Kubota [40] consider that occupations highly likely to be automated represent 29% of the employment offer in Brazil.

It can be seen, therefore, that the current estimates are discordant about automation potential in Brazil. These studies fail to consider the specificities and dynamics of the public sector occupations compared to those in the private sector. Furthermore, they place their trust in the attribution of automation in *ad hoc* or subjective classifications.

3. Methodology and data

3.1. Estimation of automation's technological frontier

The objective of estimating the technological frontier is to identify automation tendencies in the private sector that could be replicated in the public sector, adopting a new empirical approach to classify occupations that could be automated. The appreciation of the method arises from the presupposition that public sector automation is already lagging behind the relevant technological frontier.

If this delay is maintained, then potentially, the public sector will reproduce the recent tendencies of the respective frontier. The selected frontier was that of the State of São Paulo in the period 2010 to 2018. São Paulo has the highest average income among the Brazilian states and concentrates investments in information and communications technologies.

However, it is not sufficient to observe which occupations presented variations in employment levels because factors other than automation can also lead to such variations. To address this aspect, the study uses a similar construction to the Bartik Instrument, also known as Shift-Share Instrument.

Shift-share analysis enables the decomposition of local employment growth in the light of national, structural and differential effects. Although it is a traditional method in the area of Regional Economics [23], the technique was revived based on the work of Bartik [12] and Blanchard and Katz [13], who made use of its components as an instrumental variable.[2] That decomposition of regional (or local) growth can be visualized as:

$$\begin{matrix}\textit{Local} \\ \textit{Growth}\end{matrix} = \begin{matrix}\textit{National} \\ \textit{Effect}\end{matrix} + \begin{matrix}\textit{Structural} \\ \textit{Effect}\end{matrix} + \begin{matrix}\textit{Differential} \\ \textit{Effect}\end{matrix} \tag{1}$$

In the regional Shift-Share method, the national effect represents that portion of the local employment growth that resulted from growth in employment in the country as a whole. The structural effect represents the proportion of the changes in local employment resulting from growth in employment at the national level. The differential effect represents the variation in employment due to factors internal to the localities that cannot be explained by national or structural effects.

In the context of automation of occupations, every effort was made to detect the effect of those variations on the occupations of local factors that cannot be explained by national or sector growth. That effect will be used as the dependent variable in training for automation susceptibility prediction, conditioned to the tasks of the occupations in accordance with the details set out in subsection 3.2.

In order to find the differential effect of automation, the Shift-Share method is adapted to control for the effects of sector employment growth instead of those of regional growth. Table 1 displays the modification of the Regional Shift-Share model to the Occupational one in response to the introduction of information on sectors and occupations.

Table 1
Modification of the Shift-Share method from regional to occupational model.

a) Regional Shift-Share			b) Occupational Shift-Share		
	Region 1	Region *j*		Occupation 1	Occupation *j*
Sector 1	a_{11}	a_{1j}	Sector 1	b_{11}	b_{1j}
Sector *i*	a_{i1}	a_{ij}	Sector *i*	b_{i1}	b_{ij}

Where a_{ij} indicates the employment of Region *i* in Sector *j*, in contrast to the modified method, whereby Sector replaces Region and considers the category Occupation instead of the former Sector. Thus, b_{ij} indicates the employment of Occupation *j* in Sector *i*.

These alterations mean that 'National effect' now represents the growth in employment as a whole; 'Structural Effect' represents the change in the sector employment as a result of the national growth in sector employment; and the Differential Effect represents the variation in sector employment due to internal changes in the occupations.

More specifically, the differential effect shows changes in the distribution of employment among the occupations after removing the effects of total growth and the growth in each sector. That change in composition of employment positions within the sectors is attributed to the production technologies which restructure the relations of replacement (or complementarity) between capital and labor stemming from the introduction of automation technologies in the labor market.

With the automation estimations for the technological frontier in hand, the next step is to attribute changes to the occupation tasks and generalize them to the occupations of the Federal Executive Branch.

3.2. Prediction of automation susceptibility

This study estimates the impact of automation on the tasks that compose each occupation rather than directly to the occupation itself. Thus, it uses the results of the preceding stage to construct the Bartik Occupational Tasks (BOT) method to predict susceptibility to automation based on the tasks of each occupation.

The tasks inherent to each occupation were obtained from the CBO 2002 Activities Matrix. The activities matrix displays the tasks that constitute each occupation. For example, the occupation 'Administrator' (252105) lists tasks such as 'Administer organizations', 'Elaborate organizational planning', 'Implement projects and programs', and so on.

Natural Language Processing (NLP) algorithms make it possible to weigh the terms of those activities that constitute each occupation and identify their contribution to the differential effect of automation estimated in the preceding stage. With that, it is possible to generalize the susceptibility to automation to all the other occupations based on the relative weights of the tasks.

The Term Frequency–Inverse Document Frequency (TF-IDF) technique is used to calculate the relative weights of the tasks. The value of the TF-IDF increases proportionally to the frequency of occurrence of a word in a given text extract in relation to the number of texts that contain the word. Ramos [46] proposed the following formula: given a set of documents *D*, a term or word *w*, and a single document $d \in D$, the calculation is:

$$w_d = f_{w,d} * log(|D| / f_{w,D}) \tag{2}$$

where $f_{w,d}$ is the number of times that *w* appears in *d*, $|D|$ is the size of the set of texts (*corpus*), and $f_{w,D}$ is the number of documents in *D* in which *w* appears. Whenever a high value is obtained for w_d it implies that *w* is an important word in *d*, but not very common in *D*. Therefore, the term *w* has discriminatory power in relation to the overall set of texts [46].

The main limitation of the TF-IDF method is that it is impossible to identify the grammatical flexions of the same term [44]. Spellings such as 'analyzed', 'analyze' or 'analyze' will each be treated as being

[2] For a critique of the use of the Bartik Instrument, see Goldsmith-Pinkham et al. [27] and Jaeger et al. [34]. In the case of Brazil, Dix-Carneiro [20]; Dix-Carneiro, Soares and Ulyssea [21]; Macedo and Monasterio [39], and others have applied the Bartik Instrument to the RAIS data.

different terms. To avoid that, before applying the TF-IDF method, a stemming procedure is carried out that links variations of terms to a single common root. In addition, another procedure discards stop-words, that is, phrase-connecting words such as conjunctions as well as numbers and special characters that have no semantic value for the analysis.

TF-IDF analysis of the CBO Activities Matrix offers a weighting scheme that makes it possible to diminish the importance of common terms to the tasks of other occupations and emphasize those singular tasks that differentiate them. Thus, in this case, D is the overall set of 20,003 activity descriptions of 2601 occupations resulting in 47 million words. Lastly, the measurement of the importance of the tasks within a given occupation is calculated as the sum of the w_d relative weights obtained. The resulting matrix has 2341 weighted terms for 2,601occupations.

The Bartik Occupational Tasks (BOT) method is then constructed based on the TF-IDF results with the differential effect. The BOT considers the relative importance of each task within the respective occupation as an explanatory variable for automation susceptibility obtained from the differential effect of the technological frontier. The model in (3) below is used to directly attribute the differential effect to the terms that make up the tasks of each occupation.

$$BOT = f(w_{d111} term_{111}, \ldots, w_{dijk} term_{ijk}) \tag{3}$$

where BOT attributes the estimated differential effect for the technological frontier on w_d weights found for each of the terms i, which constitute the tasks j of each occupation k obtained through the application of the TF-IDF method.

To find the best model for prediction and following the best practices in the field of machine learning, a sample of 2601 occupations is randomly divided into sub-samples, each with 70% of the original sample size, and separated into a training and test set. The models are constructed using the data obtained from the training set attributing a relative importance value to each of the explanatory variables. Then the model's performance is assessed using the test set. The selected model is then used to generalize the relative importance to all the occupations in the CBO database, based on the importance of the tasks.

Various econometric and machine learning methods were tested to identify which one gave the best general prediction performance with the test data. The Random Forest Regression method [17] was selected based on assessments of its accuracy and predictive power in comparison to linear regression and logistics models and other machine learning methods such as Support Vector Machines (SVM), Support Vector Regression (SVR), and Decision Trees.

Random Forest methods have become popular due to their good prediction performance especially in sparse matrix situations [8]. When the tasks are transformed by the TF-IDF analysis, the resulting matrix is sparse, that is, the number of regressor variables is greater than the number of observations. Thus, the Random Forest Regression proved to be better than all the other methods with accuracy around 95% based on a cross-referencing validation with 100 random sub-samples. It showed the lowest Root Mean Square Errors (RMSE) as well.

This procedure provided estimates of automation susceptibility for all the CBO-listed occupations, including private and public sectors. The numerical values ranged from 0 (least susceptibility) to 1 (greatest exposure to automation). These values were used to rank the occupations from the most to the least likely to be automated.

The classification of occupations as high or low susceptibility to being automated depends on their location in the quartile of the estimates of automation exposure. The label of 'high susceptibility' to being automated goes to those occupations above the 75th percentile; 'fairly-high susceptibility' goes to the occupations between the 50th and 75th percentiles; 'fairly-low susceptibility goes to the 25th and 50th percentiles; and lastly, 'low susceptibility' to those below the 25th percentile. Following the literature, those cut-off points are not static, but instead, the classification depends on the temporal horizon. The automation literature acknowledges that technology advances over the occupations if the focus shifts ahead one, two, or three decades [25].

3.3. Data sources

Among the most important databases for the execution of this study were the Brazilian Occupations Classification (*Classificação Brasileira de Ocupações* - 2002) [16], the Annual Social Information Report (*Relação Annual de Informações Sociais* - RAIS) [45] produced by the Brazilian Ministry of Economics and the Integrated Personnel Administration System (*Sistema Integrado de Administração de Recursos Humanos* – SIAPE) [48].

The SIAPE centralizes the processing of the monthly payrolls of all the bodies under direct government administration and government foundations and autarchies of the executive branch that depend on the National Treasury for their expenditure on personnel. It includes the payment of all civil servants contracted under the aegis of the Civil Service Statute, Law 8.112/90 (*Regime Jurídico Único Federal*), the Consolidation of Labor Laws *Consolidação das Leis do Trabalho* - CLT), all those on temporary contracts, trainee occupations, medical internships, and others. The database contains information on employees in active service and allocated in federal entities all over Brazil and pensioners and retirees [48].

According to their primary occupation, this study only considers civil government employees in active service, with a working week of 40 h or more. With the application of that filter, the information processed was of 521,701 employees out of a total of 627,284 persons registered in the SIAPE.

One of the first difficulties was the compatibilization of the nomenclatures of the work positions registered in the SIAPE, which do not follow a standardized pattern in codes, descriptions, or even spelling. This screening resulted in 1115 positions. In an endeavor to establish a standard, codes, and titles were assigned to all the Executive Branch positions according to the CBO Brazilian Occupations Classification [16] by cross-referencing the employees' tax registration numbers in the SIAPE and the RAIS-*Estatística Identificada* (Identified Statistics) for the year 2017. The model was refined by considering the CBO code for each position in the SIAPE[3] followed by a manual inspection of the correspondence between the description of the activities of the occupations in the CBO and the description of the activities corresponding to the positions. This compatibilization resulted in 389 clearly defined public occupations.

The Brazilian Occupations Classification describes the occupations and organizes them in a hierarchy, making it possible to systematize information regarding the workforce according to the occupational characteristics and the nature and content of the work involved. It describes the functions, duties, and tasks that make up each occupation and the content of the work in terms of the set of knowledge, skills, and training required for the performance of responsibilities for each occupation [19].

Estimating the technological automation frontier was gathered from the Annual Social Information Report (RAIS). It is widely recognized as one of the most reliable sources of Brazilian formal labor market statistics [20]. The micro-data constitute an administrative register that can be considered the equivalent of a census of the formal labor market [45]. The following section displays the results obtained for the automation technological frontier estimation, which is used to construct the measurement of the degree of susceptibility to automation based on the occupation tasks.

[3] We wish to thank Danilo Cardoso, Flávio da Vitoria and Pedro Masson, of the ENAP Data Science Department for their support in accessing and understanding the SIAPE database.

4. Technological frontier for automation

In order to calibrate technological change based on tendencies observed in the recent past, this study examined changes that took place in occupations in the private sector of the state of São Paulo from 2010 to 2018. São Paulo can be considered the technological frontier of Brazil because it has the highest average income among the Brazilian states [31]. It functions as a hub for international contact with many multinational corporations and companies in the technology sector. Furthermore, 69.5% of the total amount invested in research and technology by Brazilian states is concentrated in São Paulo [32] as are 42.9% of the workers engaged in hardware, software, IT services, Cloud, and IT production [32]. Based on the RAIS data, the state of São Paulo presented 18.5 million formal labor contracts in 2010 and 17 million in 2018, a variation of −5.6%. The decrease in the number of jobs may be related to the increase in unemployment observed in the Monthly Employment Survey (*Pesquisa Mensal do Emprego* - PME).

Variations in unemployment levels do not affect all sectors of the economy and their respective occupations in the same way. Thus, the study sought to extract the differential effects to understand the changes in employment composition within the sectors and attribute this to the introduction of automatable production technologies. This is done in Fig. 1, which compares the growth in employment in each sector (total effect) with the growth in employment independent of the overall sector growth (differential effect). The differential effect was estimated using the BOT method explained in item 3.1 in Methodology and Data. The results are aggregated by CBO Large Groups, which considers ten categories of occupations. Group 0 includes the armed forces, police and fire brigades, and was removed because the construction of the technological frontier differential effect only considers workers in the private sector.

Fig. 1 shows that the group of containing members of the higher echelons of public service, directors of organizations of public interest and companies, and managers (+28.6%) was the one in which the total number of jobs grew most, followed by science and arts professionals (+28.5%), workers in services provision, salespersons in shops and markets (+6.9%). The groups for which the most significant reduction of employment positions occurred were workers in cattle and crop farming, forestry and fishing (31.3%), industrial goods and services production workers disaggregated by discrete production (−28%), and continuous production (−18.9%).

In differential terms, the group with the members of the higher echelons of public service, directors of organizations of public interest and companies, and managers showed the highest growth (+31%) followed by the by science and arts professionals (+215%) and the workers in services provision, salespersons in shops and markets (+6.8). The groups with the greatest differential retractions were industrial goods and services production workers (discrete production) (−12%), workers in cattle and crop farming, forestry and fishing (−9.2%), and industrial goods and services production workers (continuous production) (−5.2%)

In alignment with the literature, the occupation groups most susceptible to automation aggregate those workers with lower qualification and lower remuneration – performing tasks with a higher possibility of being standardized and codified in algorithms [25]. Such occupations are most common in discrete and continuous industrial production of goods and services and farming, forestry, and fishing.

On the other hand, occupations that require the use of creativity to solve problems and social intelligence for communication and interaction in teams are indicated as the least susceptible to being automated [25]. That tendency is reflected in the differential growth detected in the groups that include members of the higher echelons of government, directors and managers of organizations in general, and professionals of the Arts and sciences.

The results obtained with the BOT method are in line with the reports of studies carried out in the United States and OECD member countries and based on experts' opinions. Thus, the estimated differential effect is used to classify the occupations most prone to be automated, attributing that effect to the various tasks carried out in the scope of each one of the occupations. To that end, the differential effect is calculated for the 49 main sub-groups of occupations according to the disaggregation given by the two-digit categorization of the CBO 2002. Fig. 2 is a schematic representation of the study's estimation procedure.

In short, the São Paulo RAIS is used to obtain the differential effect by means of the occupational modification of the Shift-Share method while the CBO Activities Matrix is subject to the TF-IDF analysis. The results of the differential effect and the TF-IDF are combined to construct the BOT method, which estimates the susceptibility of occupations to be automated using Random Forest Regression. The model is then used to identify the impact of automation on the occupations of the Federal Executive Branch which contains information from the SIAPE database.

5. Automation susceptibility results for occupations

This section presents the automation susceptibility estimations obtained in the preceding sections displaying the descriptive statistics of the Federal Executive Branch occupations (5.1), the general impact of automation on the occupations (5.2), and the impact of automation discriminated/disaggregated by government Ministries (5.3).

5.1. Descriptive statistics of Federal Executive Branch occupations

Table 2 displays the descriptive statistics of the qualitative and quantitative variables of the public service employees analyzed in this research. The mean age of the employees is 46.7 years old, ranging from 18 to 90, with a standard deviation of 11.65 years. 49.7% of civil servants are in the 39 to 50 age brackets, 41.9% in the 50 to 70 range, 8.1% are aged 18 to 30, and just 0.4% are over 70 years old.

The average number of schooling years is 15.3, ranging from a minimum of 8 (complete lower secondary education) to a maximum of 21 years (Doctorate). The schooling variable describes the highest qualification declared by the employee. The majority (36%) have a university degree, followed by 20% with complete higher secondary education, 19% with a Ph.D., 19% with a Master's degree or MBA and 5% with lower secondary education.

The average salary[4] is US$ 3,088, the median value is US$ 2517 and the standard variation US$ 1932. Most of the employees, 60.3%, receive up to US$ 3 thousand, 31.1% from US$ 3 thousand to US$ 6 thousand, 8.3% from US$ 6 thousand to US$ 9 thousand, 0.3% from US$ 9 thousand to US$ 12 thousand, and 0.003% over US$ 12 thousand. Among those with the highest salaries, 17 receive more than US$ 12, and they occupy the positions of Ministers, Doctors, and Federal Police Superintendents.[5]

5.2. Automation impacts on Federal Executive Branch occupations

The quantitative impact of automation on Federal Executive Branch occupations is based on the estimations of susceptibility to being automated, taking into account the automation technology frontier. The term 'high susceptibility' to be automated is attributed to occupations in the upper quartile of the distribution of public service occupations, that is, 96 of the 389 occupations.

The occupations are presented in groups according to their degree of susceptibility, quantities per group, schooling and remuneration of

[4] All monetary values were converted from Brazilian real (R$) to US dollar (US$) at July's 2017 exchange rate, 3.21 R$/US$.

[5] There are 6877 registrations with zero salaries, but they refer to employees whose payments are made by other systems, not SIAPE. Examples of such employees are resident doctors or physicians enrolled in special government programs.

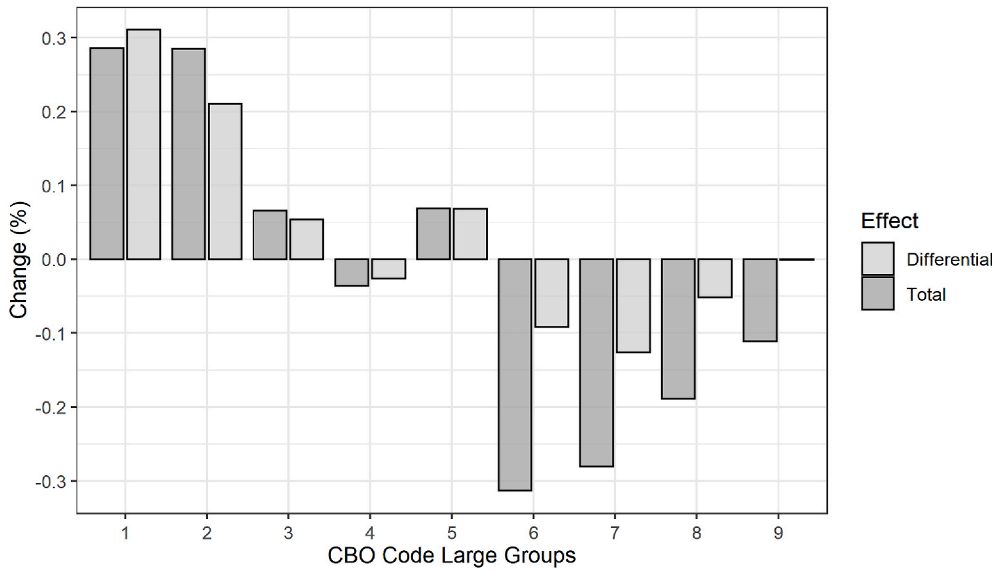

Fig. 1. Employment effects by CBO in the private sector in São Paulo - 2010 to 2018.

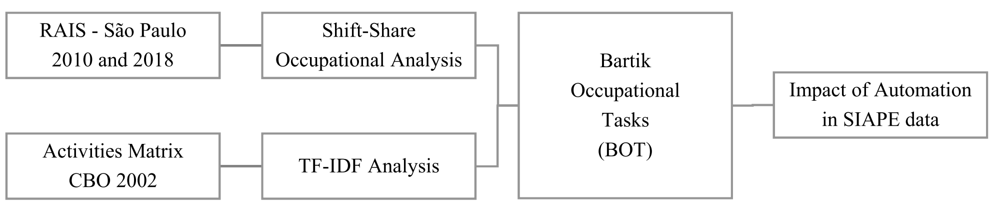

Fig. 2. Stages in the estimation of automation susceptibility in the Federal Executive Branch.

Table 2
Descriptive statistics of civil servants in the SIAPE database – December 2017.

Variable	Minimum	Median	Mean	Maximum	Stand. Var.
Schooling years	8.00	15.00	15.35	21.00	3.68
Age	18.00	47.00	46.70	90.00	11.65
Monthly pay (US$)	0.00	2517	3088	15,670	1932

Schooling	**Quant.**	**%**	**Age**	**Quant.**	**%**	**Remuneration (in 1000 US$)**	**Quant.**	**%**
Lower. 2ndary	28,396	5.4	18 |— 30	42,043	8.1	0 |— 3	314,593	60.3
Higher 2ndary	106,141	20.3	30 |— 50	259,273	49.7	3 |— 6	162,438	31.1
University	187,932	36.0	50 |— 70	218,346	41.9	6 |— 9	43,295	8.3
Master's or MBA	97,937	18.8	70 |—| 90	2039	0.4	9 |— 12	1358	0.3
PhD	101,295	19.4				12 |— 16	17	0.003
Total	521,701	100		521,701	100		521,701	100

employees. Exposure to automation for the Federal Executive Branch occupations should be understood as an ordinal rather than a cardinal scale. The susceptibility values do not represent a probability value but make it possible to establish a ranking of them from most susceptible to least susceptible to automation. Whenever there is a tie in values among the occupations, the difference in schooling levels is used to break the tie. The higher average value for the schooling levels was the differentiating criterion in consonance with the reports of other authors, whereby automation tends to have a greater impact on those professions associated with a lesser qualification [5,25].

Fig. 3a) shows the distribution of occupations according to their susceptibility to automation, that is, it represents the number of occupations situated in each score range. The occupations in dark blue, with scores higher than 0.87 represent the 96 occupation with the greatest susceptibility to being automated. To facilitate the visualization, 12 occupations with scores lower than 0.6 have been omitted.

Fig. 3b) shows the distribution of employees along the range of automation susceptibility scores. 20% (104,670) of the 521,701 employees analyzed by this research are in occupations considered to be highly susceptible to being automated, represented by the dark blue blocks in the graph.

There are 90,696 employees in 15 occupations in the range of exposure to automation between 0.87 and 0.88. In that bracket, there are some of the occupations with large numbers of employees such as administrative assistant (73,208 employees), office assistant (8,022), and typist (4,559). This explains the peak in the quantity distribution graph (2.b) For those occupations with a rating of 0.87–0.88, the average schooling years are 13.77, and the average remuneration is US$ 1783.

In the range 0.88–0.95, 9063 employees are engaged in 34 occupations. The outstanding occupations in this range are those of van or van-like vehicle drivers (4,703), cattle or crop farming workers in general (1,511), and library assistants (1,123). In the highest range, of 0.95 and over, there are 3932 employees in 41 occupations. Schooling years in

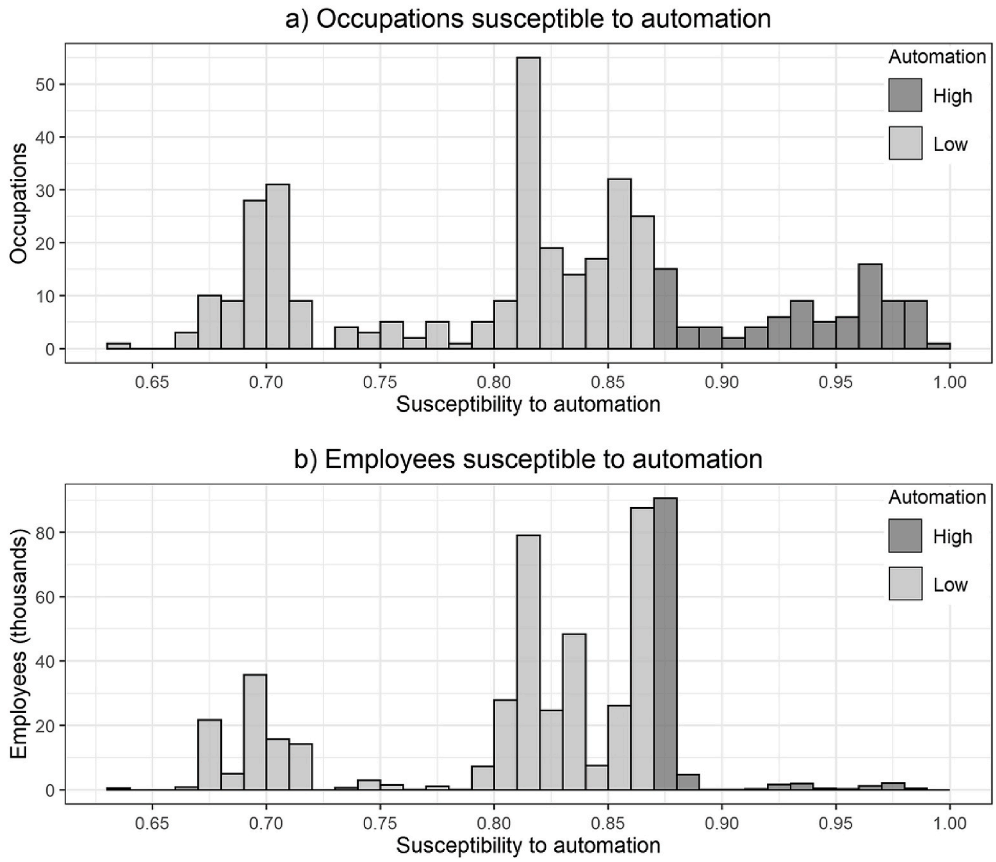

Fig. 3. Distributions of occupations and employees by degree of susceptibility to automation.

this range average 10.82 years, and the average salary is US$ 1643. The outstanding occupations are those of carpenter (687), construction worker (441), and bricklayer (306). Schooling years in this group average 11.67 years, and the average salary is US$ 1698.

Comparing the groups considered to be highly susceptible to automation with the others shows that the average schooling years of those at high risk is 13.42 years, while for the others it is 15.83. When looking at salaries, the average for the highly susceptible occupations is US$ 1,770, while for the others it is US$ 3419. Thus, occupations with high susceptibility to automation are those associated with lower levels of education and average salaries.

Table 3 displays those occupations with over 50 employees in decreasing order of susceptibility to automation. The cut-off point of 50 employees was adopted because there are various occupations with very low numbers of employees in them and therefore irrelevant for our purposes. The table also displays the number of employees in each Federal Executive Branch occupation and the employees' schooling and average salaries.

There is a visible predominance of technical occupations in graphics and audiovisual activities and employees in construction-related occupations such as steel fixer, bricklayer, painter, and carpenter. Generally speaking, they are all occupations with low schooling requirements and salaries well below the average for the total occupations of US$ 3088. This tendency is in line with the conclusions of other research reports that consider that automation's most significant impact is in professions with low qualification requirements and low salary levels [3,5,25].

Table 4 displays occupations with over 50 employees engaged in increasing order of susceptibility to automation and the numbers of Federal Executive Branch employees involved in them, their average schooling years, and their average salaries. In general, occupations with little risk of being automated require a high schooling level and receive high salaries, such as researchers in various fields, criminalistics experts, health services managers, and clinical psychologists.

Among the researchers, the most common activities involve

Table 3
Occupations in decreasing order of susceptibility to automation.

Occupation title	CBO Code	Automation Susceptibility	Quantity	Schooling (years)	Average Salary (US$)
Audiovisual systems technician	373130	0.9845	58	10.78	1555
Audiovisual operations assistant	373145	0.9845	88	11.28	1851
Audiovisual media operator	373105	0.9845	51	12.78	1442
Scenic arts technician (cinema, video, television, theatre and shows)	374205	0.9814	89	13.76	2351
Visual programming technician	371305	0.9780	291	16.35	2288
Graphic production technician	371310	0.9757	267	13.08	1895
Steel fixer (reinforced concrete)	715315	0.9736	285	12.29	2439
Bricklayer	715210	0.9732	306	9.91	1305
Painter (construction)	716610	0.9731	223	10.06	1353
Carpenter	715505	0.9727	687	10.37	1775

Note: Occupations with over 50 employees engaged in them.

Table 4
Occupations in increasing order of susceptibility to automation.

Occupation title	CBO Code	Automation Susceptibility	Quantity	Schooling (years)	Average alary (US$)
Electrical and electronic engineering researcher	203215	0.3966	453	18.99	4503
Engineering and technology researcher (other areas of engineering)	203210	0.3966	198	17.04	3660
Public health researcher	203320	0.4009	2675	19.23	5063
Social and human sciences researcher	203505	0.4060	4478	18.61	4759
Educational sciences researcher	203515	0.4060	252	17.52	3573
Meteorological researcher	201205	0.4184	518	17.95	4894
Criminalistics expert	204105	0.5590	1089	15.02	8446
Biologist	221105	0.6341	438	17.68	3276
Health services manager	131210	0.6638	800	17.38	4231
Clinical psychologist	251510	0.6704	1784	16.53	2698

Note: Occupations with over 50 employees engaged in them.

developing new materials, products, processes, or methods in line with the CBO's description. In addition, there are those associated with identifying opportunities, the execution of research projects, and the provision of technical consultancy. These activities are on the frontiers of knowledge and can not be standardized. They require highly complex tasks, including elements of creativity and innovation. Other activities central to research activities are the dissemination of knowledge through the orientation of research work, coordination of seminars, congresses, and courses in general to capacitate teams and future researchers.

Also, to be found among the occupations with little susceptibility to automation are those professionals in the knowledge areas of health and social sciences such as economists, sociologists, geographers, biologists, psychologists, and anthropologists. Others are in administration and communication, such as health services managers, public relations professionals, advertising specialists, and copywriters. These occupations comprise activities that are central to the development of automation technology itself. Insofar as they gain space in the labor market, they will demand greater investment to ensure the continuation of automation's evolution and dissemination in society.

5.3. Automation impacts by government ministries and agencies

In addition to the investigation of occupation at the disaggregated level, it is interesting to analyze the distribution of automation impacts among the Ministries and other superior bodies of the Federal Executive Branch. Given their differences in terms of the professionals employed in each of them with their focuses on areas such as education, health, or security, the impacts of automation on those bodies will inevitably be heterogeneous among them.

The SIAPE provides information on 27 higher bodies such as ministries to which 208 other bodies are subservient such as universities, federal institutes, social insurance and security agencies, intelligence agencies, federal police, agencies of control, regulation and inspection, research foundations and institutions, and other public administration entities. Corporative entities and the military are excluded from this analysis.

Ministries and occupations that have changed since December 2017 have been left out. An example of that is the occupation 'typist', an occupation with many employees engaged in it but was extinguished by the Decree nº 9.262/2018 [15].

Table 5 summarizes the automation impacts on the five Federal Executive Higher Entities that employ the greatest numbers of civil servants. The last line of the table displays the result for the entire set of employees showing that 104,670, that is 20.1% of the 521,701 analyzed by this research, are in occupations highly prone to being automated. In terms of remuneration in December 2017, these employees at high automation represent US$ 185 million of the total Federal Executive payroll of US$ 1611 million, that is, 11.5%.

The Ministry of Education is the High-Level body with the highest numbers of employees, and it also has the highest number of occupations highly susceptible to being automated: 78 of the 272 occupations it involves. Of the total number of its 252,272 employees, 47,296, that is, 18.8%, are engaged in those highly susceptible occupations. In terms of salaries, 9.5% of the total payroll goes to those automation-prone occupations.

Among the Ministry of Education occupations classified as highly prone to automation are administrative assistant (33,418), office assistant (6,189), library assistant (1,112), crop and cattle farmworkers in general (906), van (or similar vehicle) driver (701) and others with 4970 employees engaged in 73 occupations. In the other ministries, administrative assistant, office assistant and van (or similar vehicle) drivers stand out for being the occupations with the greatest numbers of employees engaged in them and susceptible to being automated.

The Ministry of Health has 26 of the 129 occupations highly exposed to automation which corresponds to 11,904 of its 66,465 employees. In terms of the total number of employees highly susceptible: the Ministry of Social Development comes next with 1727 out of a total of 32,358 employees (5%); the Ministry of Finance with 5295 employees out of a total 29,815 (17.8%); and the Ministry of Justice with 3816 out of a total of 29,273 employees (13%).

In general, the analysis of the quantitative aspect identifies those occupations most susceptible to automation with the highest numbers of employees engaged in them. The occupations classified as highly susceptible are associated with average schooling levels and salary levels below the average for the Federal Executive Branch as a whole: 15.4 years and US$ 3088. This analysis is consistent with the literature on automation which shows that the occupations with lower average

Table 5
Automation impacts by Federal Executive Branch Ministries and Other High-Level Bodies.

High-level body	Numbers Susceptible to Automation	Total Numbers	Susceptible Percentage	Payroll susceptible to Automation (US$ x 10^6)	Total payroll (US $ x 10^6)	Payroll percentage susceptible to automation
Ministry of Education	47,296	252,272	18.8%	69.2	728.0	9.5%
Ministry of Health	11,904	66,465	17.9%	22.4	147.7	15.3%
Ministry of social Development	1727	32,358	5.3%	5.3	103.1	5.1%
Ministry of Finance	5295	29,815	17.8%	9.3	152.3	6.1%
Ministry of Justice	3816	29,273	13.0%	7.2	124.6	5.7%
Others (22)	34,632	111,518	31.1%	72.0	355.5	20.2%
TOTAL	104,670	521,701	20.1%	185.4	1611.2	11.5%

schooling levels and smaller salaries are those most susceptible to being automated.

6. Discussion and conclusion

This research estimated the susceptibility to automation of occupations in the Federal Executive Branch in Brazil, making it possible to analyze the impacts of automation discriminated by occupations and government service entities. To our knowledge, this is the first study of automation with a focus on employment in the public sector. The methodology developed in this study can be applied to other spheres and branches of power in the Brazilian public sector or even to the public sectors of other countries.

In additions, this study has innovated insofar as it presents the method entitled Bartik Occupational Tasks – BOT, which makes it feasible to estimate susceptibility to automation without having to rely on subjective criteria. The method presupposes that the public sector, albeit lagging, is following the automation tendencies of the private sector in the country's technological frontier.

Based on the tendencies to automation identified in private sector occupations, the study shows that more than one hundred thousand of the 521,701 Federal Executive employees are in positions highly susceptible to automation. Thus, 20% of the total number of employees are in occupations that could potentially have their tasks attributed to automated systems in the near future.

The occupations most prone to automation are technical occupations associated with audiovisual and graphics systems and construction-related occupations such as steel fixer, bricklayer, painter, and carpenter. Generally speaking, these occupations have low schooling requirements, and salaries are below the average of US$ 3088. That conclusion is in harmony with the literature that considers that automation's greatest impact is on those professions with lower qualifications and salary levels [3,5,25].

The occupations with little susceptibility to automation generally require intense performance of analytical tasks or tasks that are not very repetitive. Among those in such occupations are researchers and professionals in natural, social and health sciences, such as engineers, economists, sociologists, geographers, biologists, psychologists, and anthropologists. In addition, there are professionals in the fields of administration and communication such as production managers and health services managers, public relations persons, advertising professionals and copywriters. Such occupations demand high qualifications and are highly paid.

The quantitative impact of automation on the public sector was shown to be expressive due to the large numbers of employees in occupations highly susceptible to being automated, such as administrative assistants, office assistants, library assistants and drivers. In budget terms, considering the situation in 2017, the employees in high risk of automation occupations received US$ 185 million of the Federal Executive Branch's total monthly payroll amount of US$ 1.6 billion.

Despite the improvements proposed in the automation estimation methodology, some limitations remain. Heterogeneities in activities performed by the same occupation are not accessed since we use formal standard descriptions of each occupation. The results do not evaluate the impact of new occupations on public sector employment. Advances in automation technologies may, as well, provide different activities in new occupational titles.

Automation has the potential to replenish the public labor force as employees grow older and retire. Future research lines may explore the relationship between automation and aging in the public sector, identifying opportunities to substitute or complement worker's activities and areas where workers' specialties are scarce. Therefore, our results can subsidize policymakers' decisions that aim at preparing the public sector workforce to dwell with future technological challenges.

Funding

This work was supported by National School of Public Administration – ENAP, as Basic Project 284/2019/CGDAD/DPP/ENAP and financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, with no direct involvement on the research. Monasterio thanks Federal District Research Foundation – FAP-DF, for financial support Grant 04/2017.

Authorship statement

Willian Boschetti Adamczyk: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft. Leonardo Monasterio: Conceptualization, Data curation, Methodology, Funding acquisition, Supervision, Validation, Writing – review & editing. Adelar Fochezatto: Methodology, Supervision, Validation, Writing – review & editing. All of the authors agree with the detailed description of their individual contributions presented above, and they agree to be accountable for any aspect of this work submitted.

Declaration of competing interest

None.

References

[1] D. Acemoglu, D. Autor, Skills, Tasks and Technologies: Implications for Employment and Earnings, Handbook of Labor Economics, Elsevier Inc, 2011, https://doi.org/10.1016/S0169-7218(11)02410-5.

[2] I.O. Adam, Examining E-Government development effects on corruption in Africa: the mediating effects of ICT development and institutional quality, Technol. Soc. 61 (2020) 101245, https://doi.org/10.1016/j.techsoc.2020.101245.

[3] P.H.M. Albuquerque, C.A.P.B. Saavedra, R.L. de Morais, P.F. Alves, P. Yaohao, Na Era Das Máquinas, O Emprego É De Quem? Estimação Da Probabilidade De Automação De Ocupações No Brasil. Texto para Discussão 2457, Instituto de Pesquisa Econômica Aplicada – IPEA, 2019.

[4] P.H.M. Albuquerque, C.A.P.B. Saavedra, R.L. De Morais, Y. Peng, The robot from ipanema goes working: estimating the probability of jobs automation in Brazil, Lat. Am. Bus. Rev. 20 (3) (2019) 1–22, https://doi.org/10.1080/10978526.2019.1633238.

[5] M. Arntz, T. Gregory, U. Zierahn, The risk of automation for jobs in OECD countries: a comparative analysis, OECD Soc. Employ. Migr. Work. Pap. 189 (2016).

[6] M. Arntz, T. Gregory, U. Zierahn, Revisiting the risk of automation, Econ. Lett. 159 (2017) 157–160, https://doi.org/10.1016/j.econlet.2017.07.001.

[7] B. Aschhoff, W. Sofka, Innovation on demand—can public procurement drive market success of innovations? Res. Pol. 38 (8) (2009) 1235–1247, https://doi.org/10.1016/j.respol.2009.06.011.

[8] S. Athey, G.W. Imbens, Machine learning methods that economists should know about, Annu. Rev. Econom. 11 (2019) 685–725, https://doi.org/10.1146/annurev-economics-080217-053433.

[9] D.H. Autor, L.F. Katz, M.S. Kearney, E. Berman, A. Chandra, The Polarization of the U.S. Labor Market, American Economic Review, 2006, https://doi.org/10.1257/000282806777212620.

[11] D.H. Autor, Why are there still so many jobs? The history and future of workplace automation, J. Econ. Perspect. 29 (2015) 3–30, https://doi.org/10.1257/jep.29.3.3.

[12] T.J. Bartik, Who Benefits from State and Local Economic Development Policies? Kalamazoo, W. E. Upjohn Institute for Employment Research, Michigan, 1991, https://doi.org/10.17848/9780585223940.

[13] O.J. Blanchard, L.F. Katz, Regional evolutions, Brookings Pap. Econ. Activ. 1 (1992) 1–75, https://doi.org/10.2307/2534556.

[14] A.S. Blinder, How many US jobs might Be offshorable? World Econ. 10 (2) (2009) 41–78.

[15] Brasil, Decreto nº 9.262, de 9 de janeiro de 2018. Extingue cargos efetivos vagos e que vierem a vagar dos quadros de pessoal da administração pública federal, e veda abertura de concurso público e provimento de vagas adicionais para os cargos que especifica, Diário Oficial da União 7 (2018) 5–42. Available at: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/D9262.htm.

[16] Brasil, Ministério do Trabalho e do Emprego, Classificação Brasileira de Ocupações (CBO). Portal Emprega Brasil, 2019, 2019. Available at: https://empregabrasil.mte.gov.br/76/cbo/.

[17] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[18] E. Brynjolfsson, A. Mcafee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, WW Norton & Company, 2014.

[19] Concla, Comissão Nacional de Classificação, Classificação Brasileira de Ocupações – CBO, Instituto Brasileiro de Geografia e Estatística, 2019. Available at: https://co

ncla.ibge.gov.br/classificacoes/por-tema/ocupacao/classificacao-brasileira-de-ocupacoes.html.
[20] R. Dix-Carneiro, Trade liberalization and labor market dynamics, Econometrica 82 (2014) 825–885, https://doi.org/10.3982/ECTA10457.
[21] R. Dix-Carneiro, R.R. Soares, G. Ulyssea, Economic shocks and crime: evidence from the Brazilian trade liberalization, Am. Econ. J. Appl. Econ. 10 (4) (2018) 158–195, https://doi.org/10.1257/app.20170080.
[22] M.A. Demircioglu, D.B. Audretsch, Conditions for innovation in public sector organizations, Res. Pol. 46 (9) (2017) 1681–1691, https://doi.org/10.1016/j.respol.2017.08.004.
[23] E.S. Dunn, A statistical and analytical technique for regional analysis, Pap. Reg. Sci. 6 (1) (1960) 97–112, https://doi.org/10.1111/j.1435-5597.1960.tb01705.x.
[24] S. Firpo, N.M. Fortin, T. Lemieux, Occupational Tasks and Changes in the Wage Structure, IZA Discussion Paper, 2011, p. 5542.
[25] C.B. Frey, M.A. Osborne, The future of employment: how susceptible are jobs to computerisation? Technol. Forecast. Soc. Change 114 (2017) 254–280, https://doi.org/10.1016/j.techfore.2016.08.019.
[26] C. Goldin, L.F. Katz, The origins of technology-skill complementarity, Q. J. Econ. 113 (3) (1998) 693–732, https://doi.org/10.1162/003355398555720.
[27] P. Goldsmith-Pinkham, I. Sorkin, H. Swift, Bartik Instruments: what, when, why, and how, Am. Econ. Rev. 110 (2020) 2586–2624, https://doi.org/10.1257/aer.20181047.
[28] M. Goos, A. Manning, Lousy and lovely jobs: the rising polarization of work in britain, Rev. Econ. Stat. 89 (1) (2007) 118–133, https://doi.org/10.1162/rest.89.1.118.
[29] M. Goos, A. Manning, A. Salomons, Explaining job polarization: routine-biased technological change and offshoring, Am. Econ. Rev. 104 (8) (2014) 2509–2526, https://doi.org/10.1257/aer.104.8.2509.
[30] Z. Griliches, Capital-Skill complementarity, Rev. Econ. Stat. 51 (4) (1969) 465–468, https://doi.org/10.2307/1926439.
[31] Instituto Brasileiro de Geografia e Estatística – IBGE, Pesquisa Nacional por Amostra de Domicílios Contínua – PNAD 2019, IBGE, Rio de Janeiro, 2019. Available at: https://www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html.
[32] Investesp, Por que SP, Available at: https://www.investe.sp.gov.br/por-que-sp/, 2020.
[33] S. Ivanov, M. Kuyumdzhiev, C. Webster, Automation fears: drivers and solutions, Technol. Soc. 63 (2020) 101431, https://doi.org/10.1016/j.techsoc.2020.101431.
[34] D. Jaeger, J. Ruist, J. Stuhler, Shift-Share Instruments and the Impact of Immigration, National Bureau of Economic Research Working Paper Series, 2018, https://doi.org/10.3386/w24285.
[35] N. Jaimovich, H.E. Siu, Job polarization and jobless recoveries, Rev. Econ. Stat. 102 (1) (2020) 129–147, https://doi.org/10.1162/rest_a_00875.
[36] J.B. Jensen, L.G. Kletzer, Measuring the task content of offshorable services jobs, tradable services and job loss. Labor in the New Economy, 2013, pp. 309–335, https://doi.org/10.7208/chicago/9780226001463.003.0009.
[37] A. Kaplan, M. Haenlein, Rulers of the world, unite! the challenges and opportunities of artificial intelligence, Bus. Horiz. 63 (1) (2020) 37–50, https://doi.org/10.1016/j.bushor.2019.09.003.
[38] L.C. Kubota, A.N. Maciente, Propensão à automação das tarefas ocupacionais no Brasil. Radar: tecnologia, produção e comércio exterior, v. 61, IPEA, Brasília, 2019.
[39] G. Macedo, L. Monasterio, Local multiplier of industrial employment: Brazilian mesoregions (2000-2010), Brazilian Journal of Political Economy 36 (4) (2016) 827–839, https://doi.org/10.1590/0101-31572015v36n04a09.
[40] A.N. Maciente, C.V. Rauen, L.C. Kubota, Tecnologias digitais, habilidades ocupacionais e emprego formal no Brasil entre 2003 e 2017. Mercado de Trabalho: conjuntura e análise, v. 66, IPEA, Brasília, 2019.
[41] Mckinsey Global Institute, A Future that Works: Automation, Employment, and Productivity, McKinsey & Company, 2017.
[43] Organisation for Economic Co-operation and Development – OECD, Oslo manual, in: Guidelines for Collecting and Interpreting Innovation Data, OECD/Euro-stat, Paris and Luxembourg, 2005.
[44] S. Qaiser, R. Ali, Text mining: use of TF-IDF to examine the relevance of words to documents, Int. J. Comput. Appl. 181 (1) (2018) 25–29, https://doi.org/10.5120/ijca2018917395.
[45] Relação Anual de Informações Sociais – RAIS, Manual de Orientação da RAIS – ano-base 2019, MTb, Brasília, 2019. Available at: http://www.rais.gov.br/sitio/rais_ftp/ManualRAIS2019.pdf.
[46] J. Ramos, Using TF-IDF to determine word relevance in document queries, Proceedings of the first instructional conference on machine learning 242 (2003) 133–142.
[47] H.S. Sætra, A shallow defence of a technocracy of artificial intelligence: examining the political harms of algorithmic governance in the domain of government, Technol. Soc. 62 (2020) 101283, https://doi.org/10.1016/j.techsoc.2020.101283.
[48] SIAPENET, Portal siapenet. Apresentação, Available at: www.siapenet.gov.br/portal/servico/Apresentacao.asp.
[49] A. Spitz-Oener, Technical change, job tasks, and rising educational demands: looking outside the wage structure, J. Labor Econ. 24 (2) (2006) 235–270, https://doi.org/10.1086/499972.
[50] World Bank, Gestão de Pessoas e Folha de Pagamentos no Setor Público Brasileiro: o que os dados dizem, World Bank, Brasília, 2019.
[51] World Economic Forum – WEF, The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution, 2016, 2016.

Willian Boschetti Adamczyk is a Ph.D candidate at the Business School of Pontifícia Universidade Católica do Rio Grande do Sul (Porto Alegre, Brazil). His main areas of research are labor, inequality and regional economics. His most recent research has been published in academic journals such as the Journal of Economic Studies.

Leonardo Monasterio is Head of Data Science at Escola Nacional de Administração Pública and professor at Instituto Brasiliense de Direito Público (Brasília, Brazil). His most recent research has been published in academic journals such as Regional Studies, Journal of Regional Science and Geographical Analysis.

Adelar Fochezatto is a full-time Professor at the Business School of Pontifícia Universidade Católica do Rio Grande do Sul (Porto Alegre, Brazil). His main area of research is regional economics. His most recent research has been published in academic journals such as: Journal of Economic Studies, The Annals of Regional Science, Journal of Agricultural Studies, Sustainable Water Resources Management, The Empirical Economics Letters, International Journal of Economics and Finance, EconomiA, Estudos Econômicos and Revista de Economia Aplicada.