DOCUMENTOS DE TRABAJO

A Deep Learning Model for Classifying Job Positions

Diego Donoso
Roberto Gillmore
Jesica Olivares
Dagoberto Quevedo

N° 1089 Septiembre 2026
BANCO CENTRAL DE CHILE

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 38822311

Documento de Trabajo N° 1089

Working Paper N° 1089

A Deep Learning Model for Classifying Job Positions*

Diego Donoso
Central Bank of Chile

Roberto Gillmore
Central Bank of Chile

Jesica Olivares
Central Bank of Chile

Dagoberto Quevedo
Central Bank of Chile

Resumen

Los registros administrativos de empleo y los avisos de empleo en línea constituyen fuentes de información relevantes para comprender la dinámica del mercado laboral, con implicaciones directas para proyectar la rotación laboral, anticipar presiones salariales y orientar la política monetaria. Los avances recientes en procesamiento de lenguaje natural (NLP, por sus siglas en inglés) permiten procesar grandes volúmenes de texto no estructurado a escala y con alta granularidad, lo que abre la puerta al análisis sistemático de fenómenos que tradicionalmente han sido difíciles de medir. En este trabajo se desarrolla un clasificador de aprendizaje profundo basado en BETO (un modelo BERT preentrenado en español) que asigna cerca de dos millones de avisos de empleo en línea y seis millones de contratos de trabajo a la Clasificación Internacional Uniforme de Ocupaciones en su adaptación chilena (CIUO-08.CL). El clasificador alcanza un F1-score ponderado de 0,85 en el nivel más desagregado (grupo primario). A continuación, este clasificador se combina con un indicador de trabajo desde el hogar basado en palabras clave para examinar, por ocupación, el impacto de la pandemia de COVID-19 sobre la oferta de empleos con esta modalidad y sus efectos de mediano plazo. Utilizando avisos de empleo a nivel de empresa y un enfoque de diferencias en diferencias, se encuentra que la pandemia elevó la proporción de avisos que ofrecen trabajo desde el hogar en 2,0 puntos porcentuales en las industrias con mayor propensión a esta modalidad respecto de las menos expuestas, lo que equivale a un aumento de 78 % frente al nivel prepandemia. El efecto es particularmente pronunciado en ocupaciones susceptibles de realizarse a distancia, como las de directivos, profesionales y personal de apoyo administrativo. En cambio, no se observan efectos significativos en las ocupaciones con baja propensión al trabajo desde el hogar. Los resultados, respaldados por un análisis de tendencias paralelas previas al tratamiento, evidencian el valor del aprendizaje automático para comprender las transformaciones del mercado laboral, especialmente en el contexto de una crisis global como la pandemia de COVID-19.

Abstract

Administrative employment records and online job postings together constitute one of the richest sources of information for understanding labor market dynamics, with direct implications for forecasting labor turnover, anticipating wage pressures, and informing monetary policy. Recent advances in natural language processing (NLP) make it possible to process this large volume of unstructured text at scale and with high granularity, opening the door to systematic analysis of phenomena that have traditionally been hard to measure. In this paper, we develop a deep learning classifier based on BETO (a Spanish BERT model) that maps approximately two million online job postings and six million job contracts into the International Standard Classification of Occupations (ISCO-08), using its Chilean adaptation (CIUO 08.CL). The classifier achieves an F1-score of 0.85 at the most granular (Unit Group) level. We combine this classifier with a keyword-based indicator of working from home (WFH) to examine the impact of the COVID-19 pandemic on WFH job postings across occupations and its medium-term effects. Using job posting data at the firm level and a difference-in-differences (DiD) approach, we find that COVID-19 increased WFH job offers in industries more likely to offer WFH by 2.0 percentage points relative to less exposed industries (a 78% increase compared to the pre-pandemic period). The effect is particularly pronounced in occupations amenable to remote work, such as managers, professionals, and administrative assistants. However, we observe no significant effects in occupations less likely to offer WFH options. Our findings, supported by a pre-trend analysis, demonstrate the value of machine learning in understanding labor market trends, especially in the context of a global crisis like the COVID-19 pandemic.

* The views expressed are those of the authors and do not necessarily represent the views of the Central Bank of Chile (CBCh) or its board members. This study was developed within the scope of the research agenda conducted by the CBCh in economic and financial affairs of its competence. The CBCh has access to anonymized information from various public and private entities, by virtue of collaboration agreements signed with these institutions. To secure the privacy of workers and firms, the CBCh mandates that the development, extraction, and publication of the results should not allow the identification, directly or indirectly, of natural or legal persons. Officials of the CBCh processed the disaggregated data.
E-mail addresses: ddonoso@bcentral.cl, rgillmore@bcentral.cl, jolivareso@bcentral.cl, dquevedo@bcentral.cl.

1 Introduction

Administrative employment records have become an increasingly valuable source of information for analyzing labor market dynamics, their impact on the economy, and monetary policy decisions. Unlike traditional statistics, these sources enable the analysis of high-frequency and granular information on job characteristics, required skills, and working conditions. A key challenge, however, is that this information is embedded in unstructured text, which complicates large-scale analysis.

Recent advances in artificial intelligence (AI) techniques and specifically in natural language processing (NLP) allow the extraction of meaningful insights from unstructured labor market data. Sources like job postings and employment contracts contain rich information that can be systematically processed, enhancing the depth and robustness of the analyses conducted on the labor market.

Specific cases of application could be given, such as detecting skill shortages based on the analysis of online job postings (Dawson et al., 2019), or transforming job postings into labor market intelligence, such as identifying shifts in job demands, skill requirements, and types of employment (Tzimas et al., 2024), or measuring the impact of AI on the current job market (Gmyrek et al., 2025; Muro et al., 2019), using the trends in job contracts.

In this paper, we propose a deep-learning-based approach to classify online job postings and contracts into a standard international classification of occupations (ISCO-08 – CIUO 08.CL for Chile)[1]. To demonstrate the utility of this classification tool, we investigate the impact of the COVID-19 pandemic on jobs offering the option of working from home (WFH) and its medium-term effects. WFH postings are identified through remote-work keywords in the posting text, while the classifier assigns each posting to an occupational group, allowing us to trace how the response varies across occupations. Specifically, we examine whether the lockdown changed firms' labor demand for WFH positions. To address this, we use firm-level job posting data and a difference-in-differences (DiD) approach: pre-COVID industry trends in offering WFH jobs classify industries into control and treatment groups, allowing us to compare flexible industries against those less likely to offer remote options.

Our findings indicate that the COVID-19 pandemic led to a differential increase of 2.0 percentage points (pp) in the share of job postings advertising WFH arrangements in industries with higher pre-existing flexibility, corresponding to a 78% increase relative to the pre-pandemic baseline. This result highlights a substantial reallocation of advertised labor demand toward remote-compatible positions, despite the low initial prevalence of formally posted WFH vacancies. These findings are consistent with recent Big Data approaches that use online vacancies to capture structural adjustments in labor demand (Hansen et al., 2023; Adrjan et al., 2021; Bamieh and Ziegler, 2022). Using an event study framework, we show that the effect does not materialize immediately but becomes statistically significant approximately six months after the onset of the pandemic. This adjustment lag provides empirical support for organizational learning and friction mechanisms, whereby firms require time to internalize the persistence of a large shock before translating it into durable changes in hiring strategies (Barrero et al., 2021).

The aggregate effect is driven by occupations with high teleworkability, such as Directors, Managers, and Administrators, Professionals, Scientists, and Intellectuals, Technicians and Mid-level Professionals, Administrative Assistants, and Service Workers and Shop and Market Sales Workers. In contrast, we find no statistically significant effects for occupations with inherently low remote feasibility, including Craftsmen

[1] ISCO codes are proposed by the International Labour Organization (ILO) and standardize occupations across all countries. Codes are organized into four hierarchical levels, each grouping similar activities.

and Trade Workers, Plant and Machine Operators and Assemblers, and Elementary Occupations. This sharp occupational gradient provides a strong internal validity check for both the classification algorithm and the empirical strategy, as the response aligns closely with task-based constraints rather than reflecting a uniform change in posting behavior. The results are consistent with the Skill-Biased Technological Change framework and the teleworkability measures of Dingel and Neiman (2020), and extend the evidence to an emerging economy context. Moreover, the absence of effects for manual and elementary occupations reinforces concerns raised by Mindell et al. (2023) regarding the unequal distributional consequences of workplace digitalization. Pre-trend analyses further support the plausibility of the parallel trends assumption underlying the DiD estimates.

Beyond the specific application to WFH, this exercise illustrates how machine-learning-based classification of online job postings can uncover heterogeneous and economically meaningful labor demand responses to major aggregate shocks in the Global South, where traditional data sources are limited or slow to adjust. Chile is a particularly informative case: the stabilization of the WFH effect coincides with the enactment of Law No. 21,220, which provided legal certainty for remote work arrangements, consistent with evidence on the role of regulatory frameworks in formalizing labor flexibility (Goldin, 2023).

2 Classification Approach

An important challenge in this task is the unstructured nature of the data: postings are not standardized and often consist of free text, so each employer describes offers in its own way, yielding noisy inputs. Thus, an algorithm to automate the classification of job postings requires a language model capable of interpreting possible occupations from free-text inputs. This strategy requires:

1. Build a training dataset by labeling job postings with codes from the CIUO 08.CL classifier.
2. Train a contextual model to correctly classify the job posting text.
3. Carry out an exhaustive performance evaluation using the selected metrics and repeated stratified k-fold cross-validation on the entire labeled training dataset.

This section presents our classification approach. We first review the related literature and describe the data used in the study; we then introduce the models—FastText and the different Transformer-based models—and the dataset sampling strategy; finally, we describe the experimental setup and report the results.

2.1 Related Work

A growing body of literature has explored the use of natural language processing (NLP) techniques to analyze job postings, with applications ranging from keyword extraction to occupational classification and job-resume matching. We organize the most relevant contributions along three methodological lines and then position our work with respect to the closest precedent for the Chilean labor market.

A first strand focuses on *keyword and skill extraction* from unstructured job descriptions. Mahdi et al. (2021) leverage BERT-based representations to identify salient terms that effectively characterize job postings, providing a foundation for downstream tasks such as labor demand monitoring and skill-gap analysis.

A second strand develops *supervised classifiers* that map free-text job descriptions to standardized occupational taxonomies. Maitra et al. (2024) propose an enhanced BERT classifier to assign job titles from descriptions, improving on generic transformer baselines through task-specific fine-tuning. Rosenberger et al. (2025) introduce CareerBERT, which projects both résumés and ESCO occupational descriptions into a shared embedding space, recasting recommendation as a similarity-search problem rather than a pure classification task. In a related vein, Dahl et al. (2026) introduce OccCANINE, fine-tuning the character-level model CANINE on 15.8 million description-code pairs from 29 sources in 13 languages to map occupational descriptions to HISCO codes—the historical counterpart of the ISCO/CIUO classifications—reaching around 96 percent accuracy, precision, and recall. While their tool targets multilingual historical sources, we focus on contemporary Spanish descriptions, comparing a Spanish-specific encoder (BETO) against lightweight alternatives across the full CIUO hierarchy.

A third strand exploits *large language models (LLMs)* to bypass the need for large labeled corpora. Li et al. (2023) present LLM4Jobs, a two-stage pipeline that first generates embeddings of job descriptions via an LLM and then ranks candidate ESCO/ISCO codes by similarity, demonstrating that zero- or few-shot setups can be competitive with fully supervised classifiers when labeled data are scarce.

The closest precedent to our work is Cornejo et al. (2025), who fine-tune a Spanish BERT model (BETO) to classify job postings into the Chilean adaptation of ISCO-08 (CIUO 08.CL) and deploy it in a national labor market analytics platform. Our paper differs from theirs in two respects. First, we broaden the scope of the classification task to cover both online job postings and job contracts, and we construct the training sample through an active learning strategy that iteratively selects the most informative examples for labeling, reducing annotation costs without sacrificing model performance. Second, we apply the resulting classifier to a substantive economic question, the medium-term impact of the COVID-19 pandemic on work-from-home job postings, through a difference-in-differences design. In addition, given the methodological proximity, we compare our model's quality metrics against those reported by Cornejo et al. (2025) in Section 2.3, while noting that the two models are evaluated on different datasets and validation protocols.

2.1.1 Transformer Models

The Transformer architecture (Vaswani et al., 2017) replaced recurrent models—which process text sequentially and struggle to retain long-range information (Hochreiter and Schmidhuber, 1997; Hu, 2019)—with a self-attention mechanism that lets every token attend directly to every other token in the sequence, capturing long-range relationships in a single computation and enabling parallel processing at scale.

Based on the original Transformer architecture, modern models are classified as encoder-only (e.g., BERT (Devlin et al., 2019)), decoder-only (e.g., GPT (Radford et al., 2018)), or encoder-decoder (e.g., T5 (Raffel et al., 2020)), based on the transformer component used. In this study, we focus on encoder-only models from the BERT family, as these models can effectively learn an embedding space and perform well in classification tasks.

Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2019) was introduced, fundamentally changing how language models are trained. BERT uses only the encoder part of the Transformer architecture, allowing it to learn deep bidirectional representations by conditioning on both left and right context simultaneously in all layers.

BERT Pretraining

The key practical advantage of these models is transfer learning. Before any supervised training, they are pre-trained in an unsupervised manner on massive text corpora—through tasks such as Masked Language Modeling, in which the model learns to predict randomly hidden words—which endows them with rich representations of the statistical structure of the language. Fine-tuning can then adapt these representations to a specific downstream task using a comparatively small labeled dataset, which is what makes the approach viable in a setting like ours, where labels are costly to produce.

Models such as BETO adopt the BERT architecture but modify the pre-training dataset by using large Spanish corpora to adapt the model to the Spanish language. As a result, BETO outperforms the multilingual BERT model across multiple Spanish NLP tasks.

BERT Fine-Tuning

Once a Transformer has been pre-trained, it can be adapted to specific downstream tasks using a relatively small and specialized labeled dataset. In this work, the downstream task is text classification: given $\{(x_i, y_i)\}_{i=1}^{N}$, where x_i is an input text and y_i its class label, the goal is to learn $f(x) \approx y$.

To adapt the model, a classification head is added on top of the final hidden state of the `[CLS]` token, $h_{\texttt{CLS}}$.The head consists of a dropout layer followed by a linear projection:

$$z = \text{Dropout}(h_{\texttt{CLS}})\, W_c + b_c, \tag{1}$$

where W_c and b_c are the learnable weights and bias of the classifier.

The model is trained by minimizing the cross-entropy loss between the predicted logits z and the one-hot target y, which reduces to

$$\mathcal{L}(z, y) = -z_y + \log\left(\sum_k e^{z_k}\right). \tag{2}$$

This loss penalizes mass placed on incorrect classes and encourages the correct logit z_y to dominate. Because pre-training already endows the model with rich linguistic representations, fine-tuning typically requires only a few epochs on the labeled dataset.

2.2 Methodology

2.2.1 Data

We draw on two data sources to obtain job positions:

- *Trabajando.com*: Job postings were obtained from `trabajando.com`, one of the largest employment platforms in Chile. Founded in 1999, Trabajando.com serves as a comprehensive job portal that connects job seekers with employers across a wide range of industries. Each job posting typically includes a title, a detailed description of the position, and specific requirements.

- *Labor Directorate*: Job contracts were obtained from the Labor Directorate (*Dirección del Trabajo*), an institution that manages the electronic employment contracts in Chile. It allows employers to register detailed information about contracts. Each job contract includes a title of the position and a description of the main functions.

In total, our data comprise approximately two million job postings and six million employment contracts. The two sources play complementary roles: job postings feed the labeled training sample and the WFH application of Section 3, while employment contracts, beyond contributing to the training sample, provide the validation of the classifier against representative labor market statistics (Section 2.3).

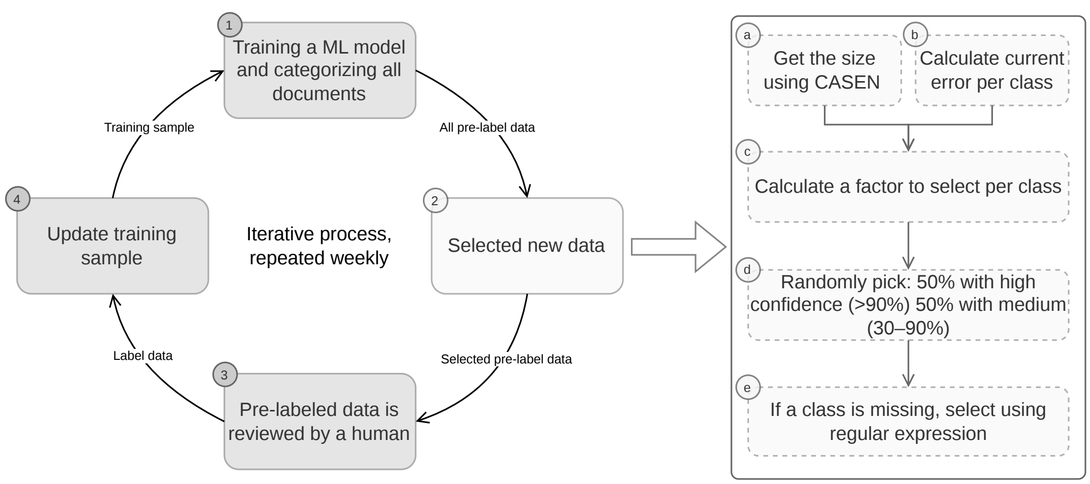
2.2.2 Preprocessing

For consistency, we standardized the textual content of the job positions following Gentzkow et al. (2019). The cleaning pipeline consists of the following operations: (i) conversion of all text to uppercase; (ii) whitespace normalization; (iii) standardization of programming-language tokens; (iv) removal of accents; (v) standardization of parentheses; (vi) removal of web-content tags; and (vii) processing of numerical patterns to remove residual numeric noise.

2.2.3 Training Dataset Creation

An effective method for sampling the data to be labeled is required to train the model. In this approach, an active learning strategy is used: the idea is that an algorithm can achieve better performance with a smaller labeled dataset if it is allowed to choose the data from which it learns (Settles, 2009). This is crucial in our setting, where no labeled data is available and the dataset must be generated from scratch. A specific application is uncertainty sampling, an iterative cycle in which examples are manually labeled, a classifier is trained on them, and the classifier then identifies new examples with ambiguous class membership for further labeling (Lewis and Gale, 1994). We adapt this approach to sample the job positions to be labeled and used to train the model; Figure 1 illustrates the process.

Figure 1: Dataset generation cycle using active learning



Notes: The learner iteratively selects samples from the unlabeled pool U using a CASEN-based sorting strategy. Labeled data is added to L and used to retrain the model. Source: Own elaboration based on Settles (2009), modified to include a more customized sampling strategy.

The process starts with a model trained with a small dataset, which is used to predict the entire unlabeled pool of job positions. A ranking of importance of the labels is generated using the National Socioeconomic Characterization Survey 2022 (de Desarrollo Social y Familia, 2022) (CASEN by its Spanish name), which provides the participation percentage for every job category in the current national dataset. The idea is to have a dataset that reflects the real proportions of the country.

Additionally, the current accuracy of the model is calculated for the already evaluated set. For those categories with a number of labeled job positions larger than a defined minimum (10), the proportion of samples given by CASEN is multiplied by one minus the percentage of correct predictions. This makes it possible to stop sampling categories that are already large enough given their importance and are being correctly labeled by the model, and instead focus on those that still need improvement. If this procedure yields no candidates for a category, texts are retrieved directly from the unlabeled pool using regular expressions tailored to that category.

The selected sample of pre-labeled job positions is then reviewed by a classifying technician who will define the correct label. This new labeled dataset is incorporated into the next training of the model, and the process is repeated.

The training dataset comprises 120 different CIUO codes with more than five records, representing 70% of the universe according to CASEN. Figure 2 shows the composition of the dataset, according to the first-level CIUO code and its origin: job postings from *Trabajando.com* or contracts from the *Labor Directorate*.

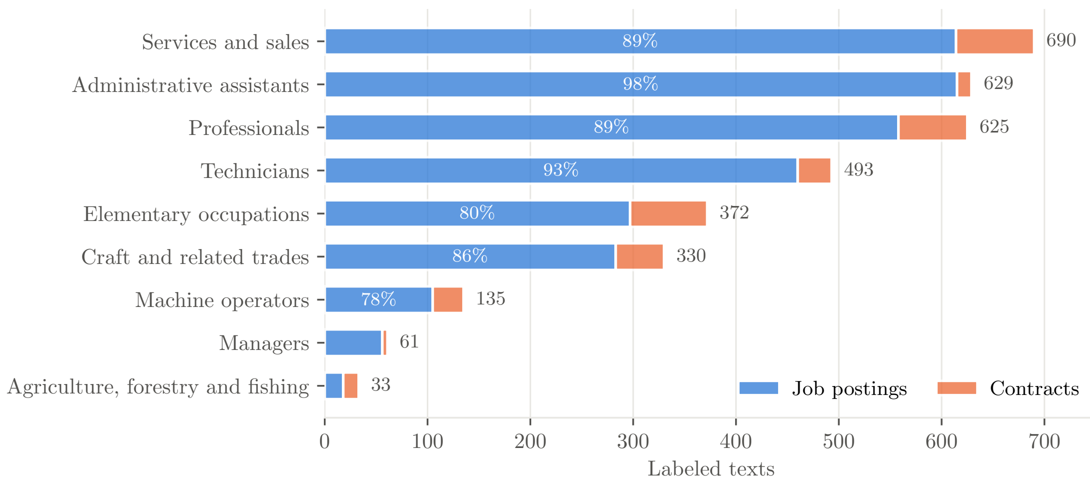


Figure 2: Training dataset composition by CIUO level 1 and data source

Notes: Composition of the labeled training dataset by CIUO level-1 group and data source (n = 3,006 job postings from *Trabajando.com* and 362 contracts from the *Labor Directorate*). Percentages inside the bars show each source's share within the group; the figure at the end of each bar is the group total. Bars are sorted by total count.

2.3 Evaluation

2.3.1 Experimental Setup

In this study, we evaluate the performance of two model families in text classification tasks. The first is an adaptation of the Continuous Bag of Words model (FastText (Joulin et al., 2017)), while the second consists of Transformer-based models (Devlin et al., 2019).

Using the labeled dataset, we train FastText (FT) as the baseline model using only the job title. FastText resembles the Continuous Bag of Words model (Mikolov et al., 2013) but predicts a label rather than a central word: because it relies on averaging word vectors—capturing no positional information or interactions between tokens—it performs well on shorter texts but becomes less effective on longer ones, where token-level features may be lost. This choice is further supported by its success in similar contexts, such as electronic invoices (Luttini et al., 2023; Albagli et al., 2026). On the other hand, we fine-tune BERT-like models using the title, description, business sector, and occupation description from the labeled data. We adapt the models by adding a classification layer on top of the [CLS] token (the first token of the sequence). Specifically, we evaluate the following three pre-trained BERT-like models; details of the hyperparameter search are provided in Appendix A.2.

1. **BETO** (Cañete et al., 2020): BETO is a pre-trained BERT model trained on Spanish corpora (Cañete, 2019).

2. DistilBETO (Cañete et al., 2022): DistilBETO is the Spanish version of DistilBERT (Sanh et al., 2020), trained on Spanish corpora as well. It is a faster variant of BERT that uses knowledge distillation during pre-training to reduce model size while retaining its performance.

3. ALBETO (Cañete et al., 2022): ALBETO is the Spanish version of ALBERT (Lan et al., 2020), trained on Spanish corpora. This model is a lightweight variant of BERT designed to improve parameter efficiency through factorized embedding parametrization and cross-layer parameter sharing, reducing model size while maintaining competitive performance.

All models are trained to predict the most granular, four-digit Unit Group code; because the taxonomy is hierarchical, performance at the coarser levels (Minor, Sub-major, and Major) is obtained by aggregating each predicted code to its parent group, rather than by training separate models per level. We then evaluate the performance of all models, reporting the precision, recall, and F1-score metrics after carrying out a repeated stratified k-fold cross-validation with 5 repeats and 5 folds, since the classes in our data are imbalanced (Figure 2).

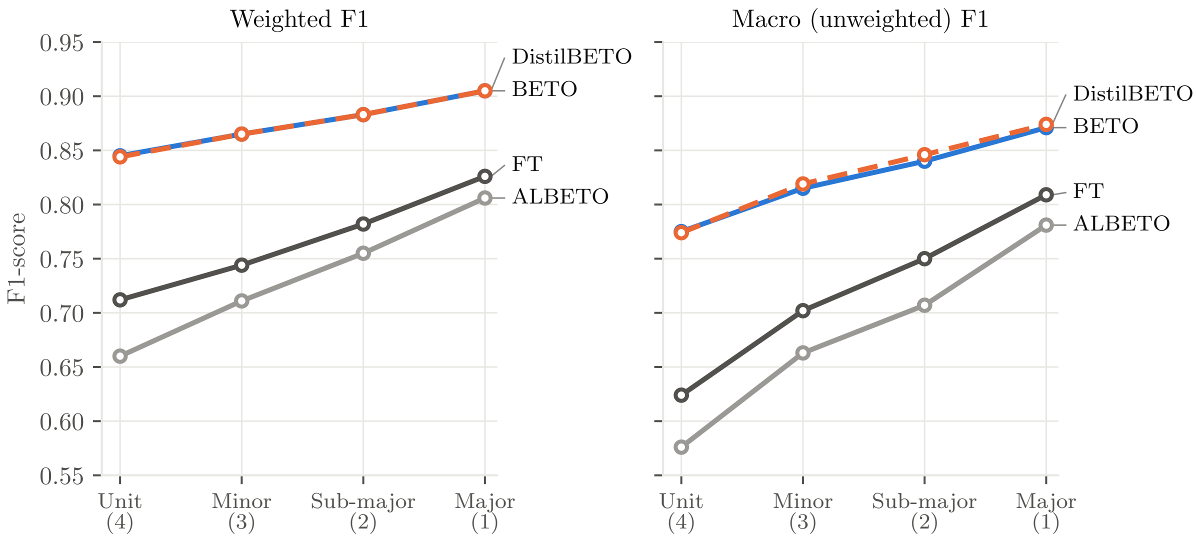
2.3.2 Results

Figure 3 compares the F1-scores of the four candidate models across all four CIUO code levels; Table 8 in the Appendix reports the full set of evaluation metrics. Performance is assessed via repeated stratified k-fold cross-validation with 5 folds and 5 repeats, and the best result for each (level, metric) pair is highlighted in bold.

Three findings emerge from Figure 3; all numerical comparisons below refer to weighted F1. First, BETO and DistilBETO consistently outperform the FastText (FT) baseline and ALBETO across every CIUO level, with the performance gap widening at finer granularities: at the Unit Group level, BETO exceeds FT by 13 F1 points and ALBETO by approximately 18 F1 points. Second, DistilBETO matches BETO across all four levels—the largest gap being 0.1 F1 points at the Unit Group level—despite having roughly half the parameters; this makes it an attractive option in production settings where inference cost is binding. Third, ALBETO underperforms even FastText across all levels, and increasingly so at finer granularities, suggesting that its lighter architecture is insufficient for the long-tailed label distribution characteristic of detailed occupational taxonomies.

Given these results, we adopt BETO as the primary classifier for the remainder of the paper, while noting that DistilBETO offers a competitive lightweight alternative. For reference, our BETO model attains a weighted F1-score of 0.85 and a macro (unweighted) F1-score of 0.78 at the Unit Group level (4), while Cornejo et al. (2025) report a weighted F1-score of 0.66 (macro 0.64) at the same four-digit level. The two

Figure 3: Model comparison: F1-scores across CIUO aggregation levels



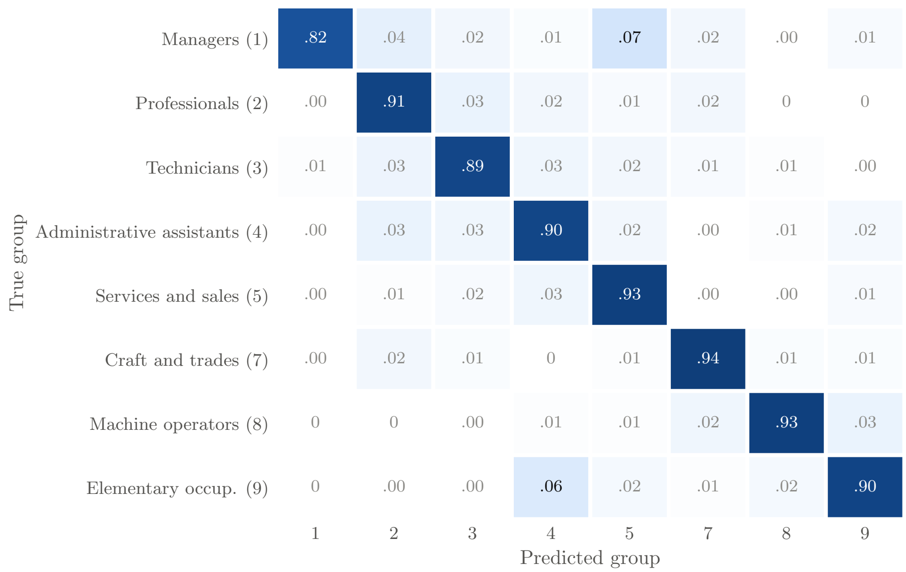
Notes: F1-scores from repeated stratified k-fold cross-validation (5 folds, 5 repeats). Models predict four-digit Unit Group codes; coarser levels aggregate each prediction up the hierarchy. FT = FastText; BETO = Spanish pre-trained BERT; ALBETO = Spanish ALBERT; DistilBETO = Spanish DistilBERT. Table 8 in the Appendix reports the full set of metrics (precision, recall, and F1).

sets of figures are not evaluated on a common benchmark and reflect different design objectives: Cornejo et al. (2025) use a balanced sample of 104 four-digit occupations (capped per class) with a single train/test split, a design well suited to their production setting, whereas we use a CASEN-representative sample of 120 codes with repeated stratified k-fold cross-validation. With that caveat in mind, our higher figures under both metrics are consistent with the gains expected from the active learning sampling strategy and the broader training corpus described in Section 2.2.

Figures 4 and 5 confirm the strong performance of BETO at the major-group level (CIUO Level 1). The confusion matrix is sharply diagonal-dominant: correct-classification rates range from 0.82 to 0.94, with most groups exceeding 0.90. Misclassifications are sparse and never exceed 0.07 for any off-diagonal cell, indicating that the model rarely confuses one major group for another. The lowest diagonal value corresponds to Group 1, of which 7% is misclassified as Group 5; the only other notable confusion is Group 9, of which 6% is assigned to Group 4. The ROC analysis corroborates these results: the macro-averaged curve attains an AUC of 0.956, sitting well above the random-classifier diagonal and reflecting a high true-positive rate across the full range of false-positive rates. Together, the two diagnostics indicate that the classifier is accurate and discriminates well across classes at the coarsest level of the taxonomy.

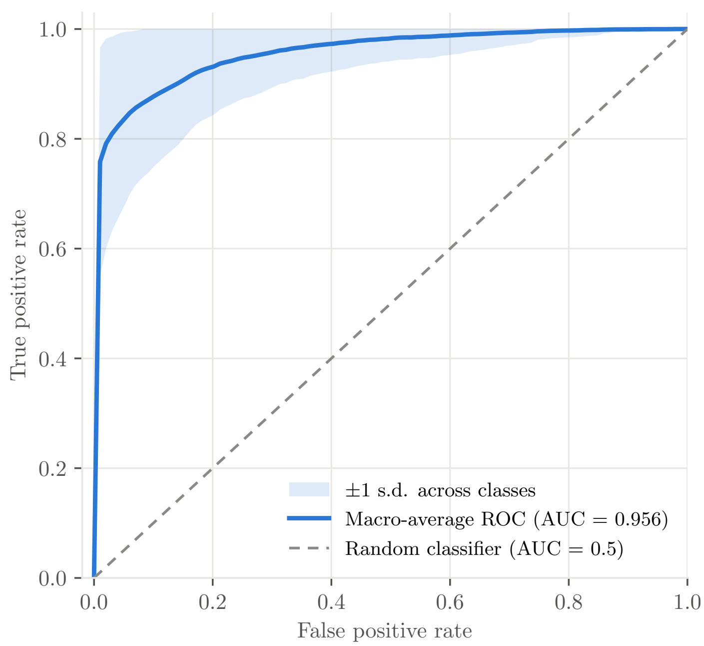
As an additional exercise, we evaluate whether the composition generated by the classifier is consistent with representative labor market statistics, using the CASEN survey (Section 2.2) as the benchmark distribution. We then classify job contracts provided by the Labor Directorate for every period and compare the compositions. Figure 6 shows that the aggregate distribution of occupations obtained from the classified contracts is broadly consistent with the CASEN benchmark over time, providing additional evidence that the model captures meaningful occupational structures beyond the labeled training sample. Figure 10 in the Appendix disaggregates this comparison by occupational group; the remaining gaps reflect the coverage

Figure 4: BETO confusion matrix at CIUO Level 1



Notes: Row-normalized confusion matrix of the BETO classifier at CIUO Level 1: each cell reports the share of a true group's observations assigned to each predicted group, so rows sum to one; "0" denotes an exact zero.

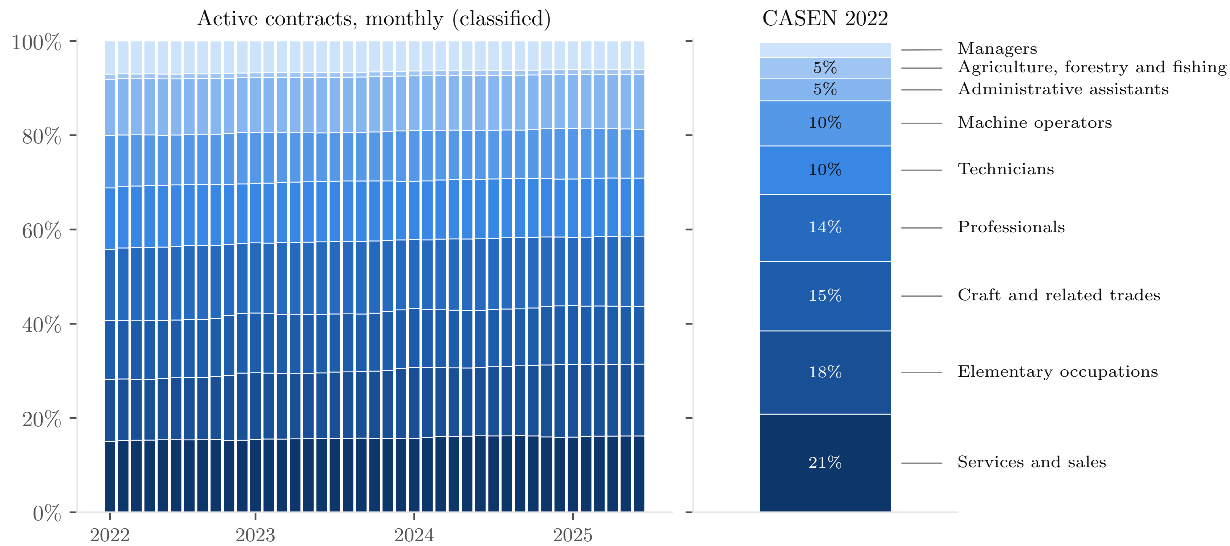
Figure 5: BETO mean ROC curve at CIUO Level 1



Notes: Macro average of the per-class ROC curves of the BETO classifier at CIUO Level 1. The shaded band is ±1 standard deviation across classes.

of electronic contracts, which capture formal salaried employment only, whereas CASEN measures total employment, including informal work.

Figure 6: Job contracts versus CASEN distribution



Notes: The left panel shows the monthly composition of active job contracts classified by the model into CIUO level-1 groups (February 2022–June 2025). The right panel shows the occupational distribution of employment in the CASEN 2022 survey. Bands are stacked in the same order in both panels, from the largest to the smallest CASEN share.

3 The Effects of the COVID-19 Pandemic on the Working-from-Home Modality: Evidence from Job Posting Classification

To demonstrate the utility of the classification algorithm, we examine the heterogeneous effects of the COVID-19 pandemic on work-from-home (WFH) job offers, a work arrangement that was previously uncommon between employers and employees (Dingel and Neiman, 2020; Bloom et al., 2013; Brynjolfsson et al., 2020), across different occupations in Chile. We employ a difference-in-differences (DiD) approach, utilizing both a two-way fixed effects model and an event study design to identify the impact. Our analysis is based on job postings published by firms on the website `trabajando.com` from January 2018 to June 2021. We classify job postings as WFH positions if the job description includes specific keywords related to remote work. The final classification model is then applied to assign the appropriate CIUO code. Importantly, our analysis focuses on changes in advertised work-from-home arrangements in online job vacancies, rather than on realized employment contracts or the overall prevalence of telework in the labor force. This distinction is central to our contribution, as online job postings provide a high-frequency and forward-looking measure of labor demand that is not observable in standard household or firm surveys.

3.1 Descriptive Data

Job postings from `trabajando.com` come with a firm ID that is used to obtain the sector of the firm. We follow a 42-sector classification generated by National Accounts and used within the Central Bank. The job postings are grouped by sector to create an indicator that assigns each sector to the control or treatment group. The definition is based on the percentage of WFH postings from the period before the pandemic, between January 2018 and February 2020. Sectors with a pre-pandemic WFH share below the median are assigned to the control group, and those above it to the treatment group. The full details of the sectors are provided in Appendix B.

The control group comprises mostly labor-intensive sectors, such as agriculture, fruit and livestock farming, mining, and some manufacturing industries such as wine, textile and wood. The treatment group includes sectors such as financial and business service activities, communications and education, and also considers some sectors in manufacturing industries such as chemical and rubber.

Table 1: Firm workers summary statistics

Sector	Pandemic	Average workers	25th percentile	50th percentile	75th percentile
Control	Pre-pandemic	274.57	15	66	241
Control	Pandemic	257.83	14	58	220
Treatment	Pre-pandemic	248.17	8	32	142
Treatment	Pandemic	226.85	7	27	121

A characterization of the firms in both groups is provided in Table 1, detailing the number of workers in the control and treatment groups. This data was obtained from the Pension Superintendence database (*Superintendencia de Pensiones*). On average, control firms are larger than treatment firms. Firm size also decreased in both groups during the pandemic period, by 6.1% for the control group and 8.59% for the treatment group. To analyze company sales, we use an internal classification that calculates firms' net sales and assigns a firm stratum from lowest to highest sales: 1, 2, 3, and 4. Control firms have a larger composition of stratum 4 as compared with the treatment group, as detailed in Table 2.

Table 2: Distribution of firms by size

Firm Size	Control Sector (%)	Treatment Sector (%)
1	9.8	15.16
2	21.7	28.6
3	18.5	21.6
4	50.0	34.6
Total	100.0	100.0

When examining the behavior of job postings in terms of occupations before and after COVID-19 (Table 3), it is noticeable that the number of monthly WFH job postings increased after COVID-19 for most CIUO 08.CL major groups, except "Elementary Occupations" where the value remained stable.

Table 3: Descriptive average of monthly job postings by firm

	Monthly AVG Total Postings		Monthly AVG WFH Postings	
Occupation	Before COVID	After COVID	Before COVID	After COVID
(1) Directors, Managers, and Administrators	2.293	2.367	0.005	0.033
(2) Professionals, Scientists, and Intellectuals	6.404	4.490	0.019	0.271
(3) Technicians and Mid-level Professionals	3.778	3.549	0.022	0.095
(4) Administrative Assistants	3.673	3.193	0.012	0.138
(5) Service Workers and Shop and Market Sales Workers	4.875	4.669	0.022	0.149
(7) Craftsmen and Trade Workers	2.540	2.446	0.003	0.010
(8) Plant and Machine Operators and Assemblers	3.426	3.809	0.002	0.003
(9) Elementary Occupations	3.545	3.896	0.004	0.004

3.2 Empirical Strategy

To estimate the effect of the COVID-19 pandemic on WFH job posting, we employ a difference-in-differences (DiD) strategy, comparing the job offers for WFH arrangements between economic sectors with varying levels of flexibility. We specifically choose to define treatment at the sector level (42 groups) rather than the firm level to mitigate endogeneity and selection bias. At the firm level, the choice to offer teleworking is often endogenous; firms that adopted it early were likely already more digitalized or financially stable, creating a selection bias. By shifting the analysis to the sector, we capture the technical feasibility of the tasks themselves, the natural capacity for remote work inherent to an industry, which is exogenous to an individual firm's management choices. We estimate the following equation:

$$r_{fst} = \alpha + \beta \times Treat_{fs} \times Post_t + \lambda_f + \gamma_t + \varepsilon_{fst} \tag{3}$$

where $r_{fst} = \eta_{ft}/N_{ft}$ is the share of WFH job postings (η_{ft}) in total job postings (N_{ft}) of firm f, belonging to economic sector s, in month t; since each firm belongs to a single sector, tables and figures abbreviate it as r_{ft}. $Treat_{fs} = 1$ if the firm belongs to a sector s that is flexible for WFH, and 0 otherwise. This approach serves as an Intention-to-Treat (ITT) framework, accounting for the fact that firms are exposed to the competitive labor pressures of their industry regardless of their individual take-up. While many firms stopped hiring during the pandemic and would otherwise drop out of a firm-level model (creating survivor bias), the sector-level grouping retains this information. This reduces idiosyncratic noise, where a single firm's hiring might jump for random reasons, and helps isolate the underlying economic signal.

$Post_t = 1$ for periods after the start of the COVID-19 shock (March 2020). This period also coincides with the enactment of Law No. 21,220 in March 2020, which formalized teleworking and remote work into the Chilean Labor Code. Because this law established universal regulations and protections, it affected all firms equally. Any common effects of this regulatory change are captured by the time fixed effects (γ_t, specified as month-year fixed effects), while β remains the parameter of interest, capturing the differential impact of the shock on sectors with higher pre-existing flexibility.

We also estimate a more flexible model that visually provides evidence of the parallel trends assumption and captures the dynamics of the effect, as in Alexander and Karger (2021); Hu et al. (2021). The specification is as follows:

$$r_{fst} = \alpha + \sum_{g \in G} \beta_g \, Treat_{fs} \mathbb{1}\left\{g(t) = g\right\} + \lambda_f + \gamma_t + \varepsilon_{fst} \tag{4}$$

where G is the set of months surrounding the shock $\{-10, ..., 0, ..., 15\}$, and all estimated effects β_g are normalized relative to February 2020 ($\beta_{-1} = 0$). By assigning treatment at the sector level, we ensure the stability of our groups, since firms cannot easily change their economic sector. This strengthens both the validity of our parallel trends assumption and our confidence that the results reflect plausible structural shifts in job posting rather than temporary firm-level pivots. The assumption states that in the absence of the COVID-19 pandemic, the demand for WFH jobs in sectors with more pre-COVID flexibility would have trended similarly to the demand in sectors with less flexibility. Although this assumption is not directly testable, we provide evidence by examining the pre-treatment coefficients in the event study plots.

A potential concern when using online vacancies is that estimates may partly reflect changes in firms' recruitment strategies or disclosure practices. However, such adjustments are not a limitation of our approach but rather a key margin through which labor demand responds to large economic shocks. In the context of the COVID-19 pandemic, decisions regarding whether to post vacancies online and how explicitly to advertise remote work arrangements constitute economically meaningful responses that are observable only through job posting data. Therefore, the estimated DiD effects should be interpreted as capturing the differential adjustment of advertised labor demand toward WFH-compatible arrangements across sectors, combining genuine changes in job design and increased explicitness in vacancy descriptions. Both margins are central to understanding how firms adapted to prolonged uncertainty.

3.3 Results

In Table 4, columns 1 and 2 report the within-group change in WFH job offers after the pandemic for each group. For the less flexible economic sectors, there is an average increase of 1.3 percentage points after the pandemic, whereas in the more flexible group, there is an increase of 3.3 percentage points. Column 3 shows the DiD estimate, which represents a 2.0 percentage-point increase in the offers for WFH jobs in the more flexible economic sectors relative to the less flexible ones.

In Table 5, we present the DiD estimates for different occupations based on the classification algorithm. For occupations more likely to engage in WFH activities, such as Directors, Managers, and Administrators, Professionals, Scientists, and Intellectuals, Technicians and Mid-level Professionals, Administrative Assistants, and Service Workers and Shop and Market Sales Workers, we find an increase of approximately

Table 4: Effects on WFH job posting

	Dependent variable: Ratio of postings: $r_{ft} = \frac{\eta_{ft}}{N_{ft}}$		
	Control Group	Treatment Group	Full Sample
	(1)	(2)	(3)
Post	0.013^{***}	0.033^{***}	
	(0.003)	(0.003)	
Treat x Post			0.020^{***}
			(0.005)
Firm Effects	✓	✓	✓
Time Effects			✓
AVG r_{ft} (Pre)	0.0001	0.0052	
% Change			3.8276
Observations	7027	30689	37716
N. of groups	1460	5660	7120
R^2	0.008	0.021	0.001

Notes: Summarizes the results of the regression defined in Equation 3, shown for each treatment group defined by $Treat_{fs}$. The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. The reported % Change is the Treat×Post estimate expressed relative to the treatment group's pre-pandemic mean ratio (AVG r_{ft} (Pre)), a base that differs from the one underlying the relative increase reported in the text. Standard errors are clustered at the firm level and are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

2.3 percentage points in the availability of WFH options. Conversely, we observe no significant effects for occupations less likely to offer WFH options, such as Craftsmen and Trade Workers, Plant and Machine Operators and Assemblers, and Elementary Occupations. The absence of significant effects for manual and elementary occupations provides a strong internal validity check for both the classification algorithm and the empirical design. If the estimated effects were driven purely by changes in posting behavior, we would expect to observe similar increases across all occupations, which is not the case.

We next examine the dynamic effects. Figure 7 presents the event study estimates. We observe that in the first few months following the pandemic outbreak, the trends in job offers with the option of WFH were similar between more and less flexible economic sectors. However, starting in the sixth month after the onset of COVID-19, job offers with WFH options increased by about 3 to 4 percentage points in more flexible sectors, with these effects being both statistically significant and stable through to the end of the sample, 15 months later. These later effects align with the idea that adjustments to WFH options in job postings likely occurred as businesses gained more certainty about the long-term impact of the pandemic. In other words, as it became clearer that the effects of the pandemic were not transitory, firms became more willing to embrace permanent changes, including the option for remote work.

Figure 7 also provides important information regarding the parallel trends assumption. In the months leading up to the COVID-19 outbreak, the trends in the treatment and control groups followed parallel paths, further validating the previous estimates.

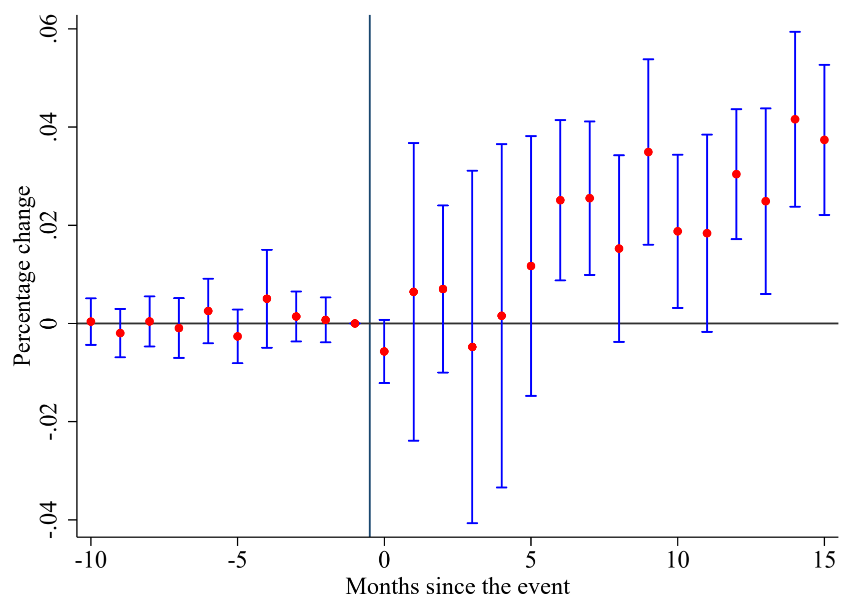
Further, we divide the event study analysis by occupation. In Figures 8 and 9, the event study plots provide visual evidence of the heterogeneous impact of the COVID-19 shock across different occupational categories, reinforcing the teleworkability framework and the hypothesis of Skill-Biased Technological Change. For high-skill and cognitive groups, specifically Professionals, Scientists, and Intellectuals, Technicians and

Table 5: Effects on WFH job posting by CIUO major group classification

	Dependent variable: Ratio of postings: $r_{ft} = \frac{\eta_{ft}}{N_{ft}}$							
	k=1	k=2	k=3	k=4	k=5	k=7	k=8	k=9
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat x Post	0.027***	0.025***	0.018***	0.028***	0.018**	0.001	-0.001	0.003
	(0.010)	(0.007)	(0.007)	(0.009)	(0.007)	(0.004)	(0.001)	(0.005)
Firm Effects	✓	✓	✓	✓	✓	✓	✓	✓
Time Effects	✓	✓	✓	✓	✓	✓	✓	✓
AVG $r_{ft} = \frac{\eta_{ft}}{N_{ft}}$ (Pre)	0.0026	0.0049	0.0089	0.0043	0.0043	0.0031	0.0018	0.0014
% Change	10.5366	5.1238	1.9666	6.5884	4.0995	0.373	-0.7988	1.9275
Observations	6146	23280	18055	13671	9632	9645	4738	5033
N. of groups	1081	3639	3171	2029	1705	2255	1033	1234
R^2	0.002	0.001	0.001	0.002	0.001	0.000	0.000	0.000

Notes: Summarizes the results of the regression defined in Equation 3 by CIUO code k, where each major group corresponds to the following label: (1) "Directors, Managers, and Administrators". (2) "Professionals, Scientists, and Intellectuals". (3) "Technicians and Mid-level Professionals". (4) "Administrative Assistants". (5) "Service Workers and Shop and Market Sales Workers". (6) "Farmers and Skilled Agricultural, Forestry, and Fishery Workers". (7) "Craftsmen and Trade Workers". (8) "Plant and Machine Operators and Assemblers". (9) "Elementary Occupations". The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. Standard errors are clustered at the firm level and are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

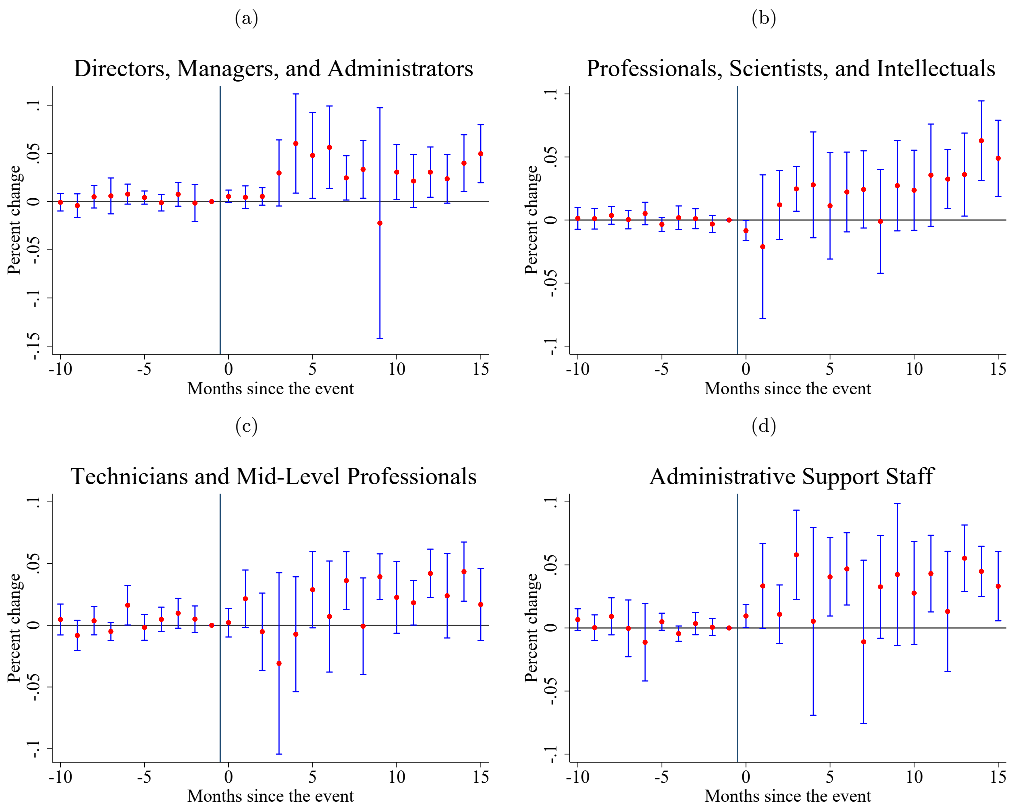
Figure 7: Event Study on the COVID-19 impact on WFH job posting



Notes: The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

Mid-level Professionals, and Directors and Managers, there is a clear and sustained structural shift. In the months leading up to the shock, months (-10 to -1), the coefficients are flat and statistically insignificant, which supports the parallel trends assumption necessary for our DiD identification.

Figure 8: Event studies on the COVID-19 impact on WFH job posting by occupation: higher-teleworkability groups

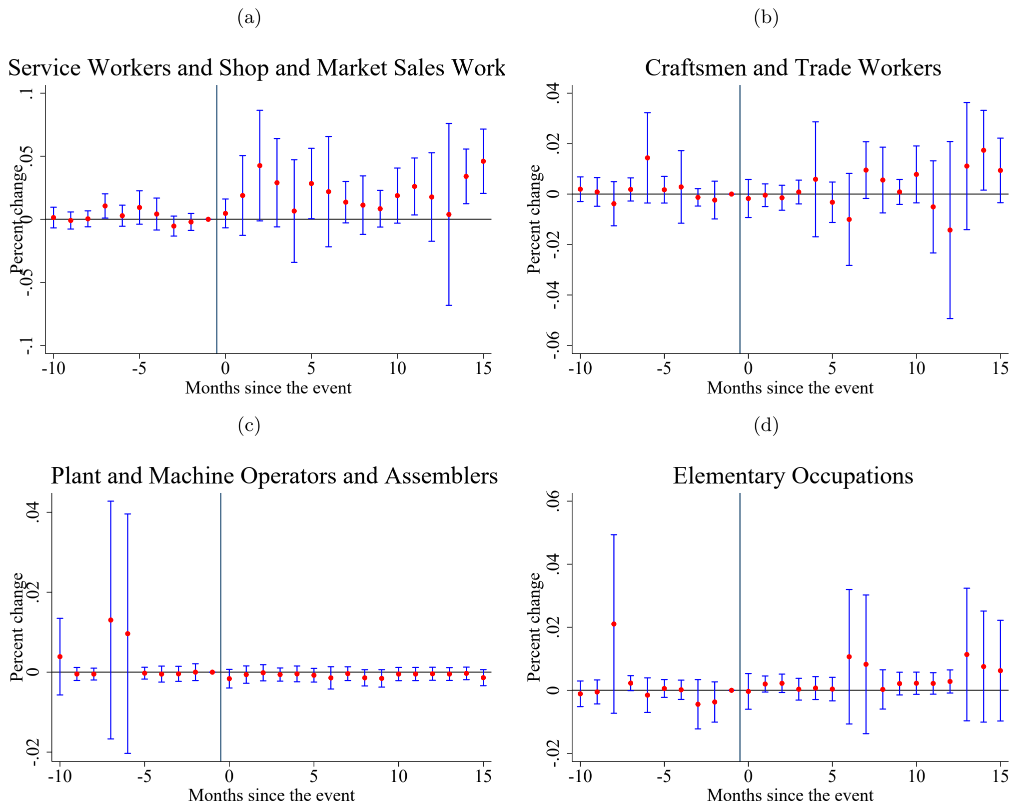


Notes: Groups (1) Directors, Managers, and Administrators; (2) Professionals, Scientists, and Intellectuals; (3) Technicians and Mid-level Professionals; (4) Administrative Assistants. The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

Consistent with the organizational learning theory of Barrero et al. (2021), the impact on these high-skill occupations was not instantaneous. The plots reveal a noticeable friction period of approximately five months characterized by volatile coefficients. Only afterward, around the sixth month, do these groups show a significant and persistent increase in WFH job postings that stabilizes at a higher plateau through month 15. This suggests that for high-skill roles, the pandemic was not a temporary disruption but a permanent catalyst for labor flexibility, likely cemented by the legal certainty provided by Law No. 21,220.

In contrast, manual and elementary occupations, such as Craftsmen, Plant and Machine Operators, and Elementary Occupations, show virtually null effects. The coefficients for these groups remain near zero and statistically insignificant throughout the entire post-event period. This lack of response highlights the rigid technical constraints of these roles, which typically require physical presence or specialized machinery. These findings provide empirical evidence for how workplace digitalization can exacerbate labor inequality in an emerging economy like Chile, as the gains in flexibility accrued almost exclusively to occupations at the upper end of the skill distribution.

Figure 9: Event studies on the COVID-19 impact on WFH job posting by occupation: service, manual, and elementary occupations



Notes: Groups (5) Service Workers and Shop and Market Sales Workers; (7) Craftsmen and Trade Workers; (8) Plant and Machine Operators and Assemblers; (9) Elementary Occupations. The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

Moreover, the intermediate groups, including Administrative Assistants and Service and Sales Workers, show positive but more volatile shifts with wider confidence intervals. This indicates that while these roles were capable of adapting to remote work, their adoption was likely more sensitive to specific firm-level digital maturity or fluctuating sector-specific lockdown measures. Collectively, the event studies indicate that the aggregate 2.0 pp increase in WFH demand is concentrated in occupations whose tasks and prior human capital permit remote work.

This application demonstrates that machine-learning-based classification of online job postings can recover economically meaningful and theoretically consistent labor demand responses to large aggregate shocks, even in emerging economies where traditional labor market statistics are limited or lagged.

A natural concern with the ratio outcome is that its denominator, total postings per firm, also moved during the pandemic, so a rising WFH share could partly reflect a contraction in overall posting volume rather than more WFH advertising. To address this, in Tables 10 and 11 and Figures 11 and 12 we re-estimate the average effects and the event studies using the log of the number of job postings offering teleworking, a specification that does not depend on the denominator (and is defined for all firm-months, including those without postings, which is why its estimation samples are larger). The effect also appears in this specification and increases over time, for both the aggregate sample and specific occupations, indicating that our results are not an artifact of denominator movements.

Finally, in Figure 13, we re-estimate the model by classifying treatment at the firm level[2], where $C_f = 1$ if a firm offered any WFH positions before the pandemic. We acknowledge that this firm-level definition may be subject to endogeneity and selection bias, as it captures individual management choices rather than the inherent technical feasibility of the industry. Furthermore, unlike our sector-level approach, it does not account for survivor bias or ensure the same level of group stability. While this specification introduces some minor fluctuations and slight pre-treatment drift, likely reflecting firm-specific anticipation effects or idiosyncratic noise, the post-treatment trajectory remains qualitatively similar to our primary results in Figure 7, confirming that these pre-existing variations do not threaten the validity of our main findings.

4 Conclusion

This study makes two contributions. First, we develop and validate a deep learning classifier that maps Spanish-language job postings and employment contracts into the CIUO 08.CL occupational taxonomy. Trained on a sample built through an active learning strategy that reduces annotation costs, the classifier attains a weighted F1-score of 0.85 at the most granular four-digit level, higher than the figures reported in the closest comparable exercise (Cornejo et al., 2025), and the occupational composition of the employment contracts it classifies is broadly consistent with the representative CASEN benchmark. Second, we use this infrastructure to study how advertised labor demand responded to a major aggregate shock.

Our findings reveal that COVID-19 led to a substantial increase in job offers with the option of working from home, especially in industries and occupations where WFH was more feasible. The increase in WFH job offers—78% higher than pre-pandemic levels—reflects a notable shift in labor market dynamics, with the effect emerging around six months after the pandemic's onset and concentrated in high-teleworkability occupations, while manual and elementary occupations show no response.

These results carry three caveats. First, our WFH indicator captures remote-work arrangements advertised in online postings, not realized telework, and relies on keyword matching; measurement error of this kind would tend to attenuate the estimates rather than generate the sharp occupational gradient we document. Second, classification error from the model introduces noise into the occupational breakdowns, although the aggregate difference-in-differences result does not depend on the classifier. Third, our postings come from a single—albeit large—platform, and the electronic employment contracts cover formal salaried employment only, which limits external validity for informal segments of the labor market.

Beyond this application, the classifier is validated down to the four-digit level, whereas the present analysis only required major occupational groups; this opens the door to finer-grained uses such as occupation-level teleworkability measures, skill-demand tracking, or the study of wage dynamics. More broadly, our results illustrate how machine learning applied to naturally occurring text data can support the timely, granular labor market monitoring that central banks and research institutions require.

[2] We do not show the estimates for the firm-level classification by occupations since they are too noisy; many firms rarely post job offers in certain categories, which makes the estimates statistically underpowered and highly sensitive to idiosyncratic hiring shocks.

References

P. Adrjan, G. Ciminelli, A. Judes, M. Koelle, C. Schwellnus, and T. Sinclair. Will it stay or will it go? analysing developments in telework during covid-19 using online job postings data. *OECD Productivity Working Papers*, 2021.

E. Albagli, F. Grigoli, E. Luttini, D. Quevedo, and M. Rojas. Beyond costs: The dominant role of strategic complementarities in pricing. *Journal of Monetary Economics*, page 103966, 2026. ISSN 0304-3932. .

D. Alexander and E. Karger. Do Stay-at-Home Orders Cause People to Stay at Home? Effects of Stay-at-Home Orders on Consumer Behavior. *The Review of Economics and Statistics*, pages 1–25, 09 2021. ISSN 0034-6535. .

O. Bamieh and L. Ziegler. Are remote work options the new standard? evidence from vacancy postings during the covid-19 crisis. *Labour Economics*, 76:102179, 2022.

J. M. Barrero, N. Bloom, and S. J. Davis. Why working from home will stick. Technical report, National Bureau of Economic Research, 2021.

N. Bloom, J. Liang, J. Roberts, and Z. J. Ying. Does working from home work? evidence from a chinese experiment. Working Paper 18871, National Bureau of Economic Research, March 2013.

E. Brynjolfsson, J. J. Horton, A. Ozimek, D. Rock, G. Sharma, and H.-Y. TuYe. Covid-19 and remote work: An early look at us data. Working Paper 27344, National Bureau of Economic Research, June 2020.

J. Cañete, S. Donoso, F. Bravo-Marquez, A. Carvallo, and V. Araujo. ALBETO and DistilBETO: Lightweight Spanish language models. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. European Language Resources Association, June 2022.

J. Cañete. Compilation of large spanish unannotated corpora, May 2019.

J. Cañete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang, and J. Pérez. Spanish pre-trained bert model and evaluation data. In *PML4DC at ICLR 2020*, 2020.

D. Cornejo, F. Vera, B. Gamboa-Labbé, R. B. Ruiz, B. Villena-Roldán, and J. D. Velásquez. A transformer-based intelligent system for hierarchical occupational classification in the labor market. *IEEE Open Journal of the Computer Society*, 6:1908–1919, 2025. ISSN 2644-1268. . Publisher Copyright: © 2020 IEEE.

C. M. Dahl, T. Johansen, and C. Vedel. Breaking the hisco barrier: Automatic occupational standardization with occcanine, 2026.

N. Dawson, M.-A. Rizoiu, B. Johnston, and M.-A. Williams. Adaptively selecting occupations to detect skill shortages from online job ads. In *2019 IEEE International Conference on Big Data (Big Data)*, page 1637–1643. IEEE, Dec. 2019. .

M. de Desarrollo Social y Familia. Encuesta de caracterización socioeconómica nacional (casen) 2022, 2022. Online, accessed 1 August 2024.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. .

J. I. Dingel and B. Neiman. How many jobs can be done at home? *Journal of public economics*, 189:104235, 2020.

M. Gentzkow, B. Kelly, and M. Taddy. Text as data. *Journal of Economic Literature*, 57(3):535–574, 2019.

P. Gmyrek, J. Berg, K. Kamiński, F. Konopczyński, A. Ładna, B. Nafradi, K. Rosłaniec, and M. Troszyński. *Generative AI and jobs: A refined global index of occupational exposure*. Number 140. ILO Working Paper, 2025.

C. Goldin. Why women won. Technical report, National Bureau of Economic Research, 2023.

S. Hansen, P. J. Lambert, N. Bloom, S. J. Davis, R. Sadun, and B. Taska. Remote work across jobs, companies, and space. Technical report, National Bureau of Economic Research, 2023.

S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

D. Hu. An introductory survey on attention mechanisms in nlp problems. In *Proceedings of SAI intelligent systems conference*, pages 432–448. Springer, 2019.

J. Hu, H. Xu, Y. Yao, and L. Zheng. Is working from home here to stay? evidence from job posting data and the covid-19 shock. *SSRN*, 2021. .

A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov. Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 2017.

Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut. Albert: A lite bert for self-supervised learning of language representations, 2020.

D. D. Lewis and W. A. Gale. A sequential algorithm for training text classifiers. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '94)*, pages 3–12, 1994.

N. Li, B. Kang, and T. De Bie. Llm4jobs: Unsupervised occupation extraction and standardization leveraging large language models. *arXiv preprint arXiv:2309.09708*, 2023.

E. Luttini, M. Pizarro, D. Quevedo, and M. Rojas. Invoices rather than surveys: Using machine learning to monitor the economy. In *IFC–Bank of Italy Workshop on "Data Science in Central Banking: Enhancing the Access to and Sharing of Data"*. Bank for International Settlements / Bank of Italy, 2023.

H. F. Mahdi, R. Dagli, A. Mustufa, and S. Nanivadekar. Job descriptions keyword extraction using attention based deep learning models with bert. In *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pages 1–6. IEEE, 2021.

M. Maitra, S. Sinha, and T. Kierszenowicz. An improved bert model for precise job title classification using job descriptions. In *2024 IEEE 12th International Conference on Intelligent Systems (IS)*, pages 1–6. IEEE, 2024.

T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. In Y. Bengio and Y. LeCun, editors, *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, 2013.

D. A. Mindell, E. Reynolds, et al. *The work of the future: Building better jobs in an age of intelligent machines*. Mit Press, 2023.

M. Muro, J. Whiton, and R. Maxim. What jobs are affected by ai? better-paid, better-educated workers face the most exposure. *Metropolitan Policy Program Report*, 2019.

A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever. Improving language understanding by generative pre-training. , 2018. Technical report.

C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67, 2020.

J. Rosenberger, L. Wolfrum, S. Weinzierl, M. Kraus, and P. Zschech. Careerbert: Matching resumes to esco jobs in a shared embedding space for generic job recommendations. *Expert Systems with Applications*, 275:127043, 2025.

V. Sanh, L. Debut, J. Chaumond, and T. Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2020.

B. Settles. Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences, 2009.

G. Tzimas, N. Zotos, E. Mourelatos, K. C. Giotopoulos, and P. Zervas. From data to insight: Transforming online job postings into labor-market intelligence. *Information*, 15(8):496, 2024.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

A Experimental Details

A.1 Package Environment

We used Python 3.9.13 to run the experiments. Training, hyperparameter tuning, and evaluation of BERT-like models were performed using PyTorch 1.10.2, Hugging Face Transformers (version 4.45.2), Datasets (version 3.0.1), and Optuna 4.1.0. For FastText, we used the FastText Wheel package version 0.9.2.

A.2 Hyperparameter Tuning

Table 6: Search space of values for hyperparameter tuning of Transformer models

Parameter	Value
Learning Rate	$[1e-5, 5e-5]$
Epochs	$[3, 8]$
Batch Size	$\{4, 8, 16\}$
Weight Decay	$[1e-5, 1e-1]$

Notes: Search space of values for each parameter, applied to all transformer models during hyperparameter optimization with Optuna.

Table 7: Best Parameters for evaluated Transformer models

Parameter	BETO	DistilBETO	ALBETO
Learning Rate	2.8e-5	4.6e-5	4.9e-5
Epochs	8	8	8
Batch Size	8	8	8
Weight Decay	0.016	6.1e-5	1.0e-5

Notes: During hyperparameter optimization, a total of 200 trials were executed for each model. The best-trial configurations were subsequently used for all experiments reported in this study.

A.3 Detailed Evaluation Metrics

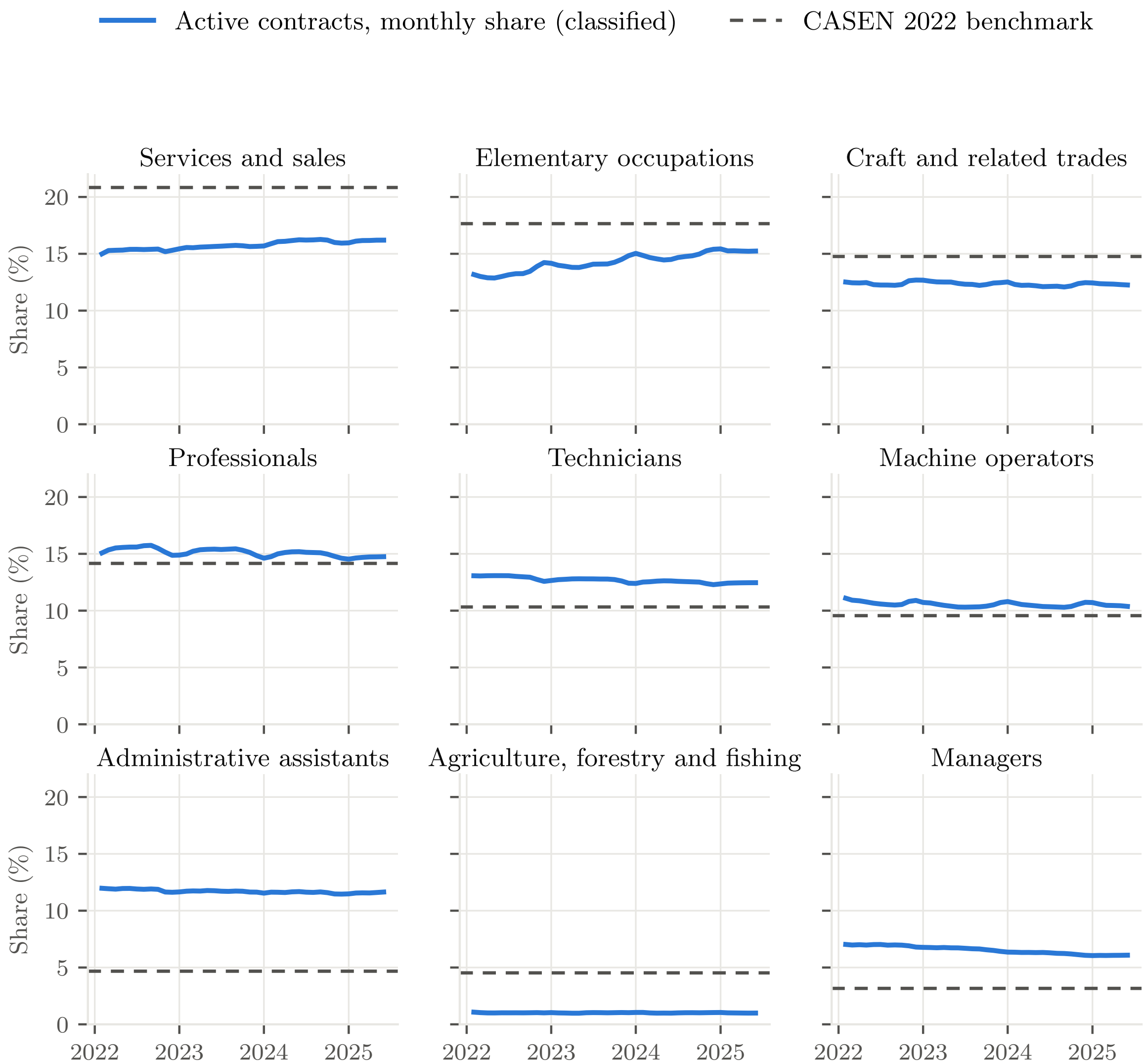
Table 8: Model evaluation across the four CIUO code levels

Level	Metric	FT	BETO	ALBETO	DistilBETO
		Model			
		Weighted			
Major Group (1)	Precision	0.829	**0.907**	0.812	0.906
	Recall	0.826	**0.905**	0.808	**0.905**
	F1-Score	0.826	**0.905**	0.806	**0.905**
Sub-Major Group (2)	Precision	0.793	**0.889**	0.788	0.888
	Recall	0.787	**0.885**	0.767	0.884
	F1-Score	0.782	**0.883**	0.755	**0.883**
Minor Group (3)	Precision	0.765	**0.875**	0.755	**0.875**
	Recall	0.753	**0.868**	0.736	0.867
	F1-Score	0.744	**0.865**	0.711	**0.865**
Unit Group (4)	Precision	0.728	**0.860**	0.671	0.859
	Recall	0.731	**0.851**	0.709	0.850
	F1-Score	0.712	**0.845**	0.660	0.844
		Unweighted			
Major Group (1)	Precision	0.815	0.870	0.809	**0.873**
	Recall	0.811	0.877	0.770	**0.879**
	F1-Score	0.809	0.871	0.781	**0.874**
Sub-Major Group (2)	Precision	0.778	0.854	0.776	**0.858**
	Recall	0.750	0.844	0.701	**0.849**
	F1-Score	0.750	0.840	0.707	**0.846**
Minor Group (3)	Precision	0.727	0.831	0.723	**0.833**
	Recall	0.717	0.825	0.675	**0.831**
	F1-Score	0.702	0.815	0.663	**0.819**
Unit Group (4)	Precision	0.648	**0.794**	0.621	0.793
	Recall	0.646	**0.791**	0.597	**0.791**
	F1-Score	0.624	**0.775**	0.576	0.774

Notes: Model evaluation using repeated stratified k-fold cross-validation (5 folds, 5 repeats). Models predict four-digit Unit Group codes; coarser levels aggregate each prediction up the hierarchy. FT = FastText; BETO = Spanish pre-trained BERT; ALBETO = Spanish ALBERT (A Lite BERT); DistilBETO = Spanish DistilBERT. Bold indicates the best score per row; ties are bolded jointly.

A.4 Disaggregated CASEN Comparison

Figure 10: Job contracts versus CASEN distribution, by occupational group



Notes: Each panel compares the monthly share of active job contracts classified into a CIUO level-1 group (solid line) with that group's share of employment in the CASEN 2022 survey (dashed line), between February 2022 and June 2025. Panels are ordered by the CASEN share.

B Economic Sectors

Table 9: Economic sector classification

	Economic Sector	Group
1	Agriculture	Control
2	Fruit Farming/Horticulture	Control
3	Livestock Farming	Control
4	Forestry	Control
5	Fishing	Treatment
6	Copper Mining	Control
7	Oil and Natural Gas Extraction	Control
8	Other Mining	Control
9	Fishing Industry	Control
10	Other Food Industries	Control
11	Wine Production	Control
12	Production of Other Beverages and Tobacco Products	Treatment
13	Textile Industries	Control
14	Wood and Furniture Manufacturing	Control
15	Cellulose, Paper and Printing Products	Treatment
16	Fuel Production	Control
17	Chemical Industries	Treatment
18	Rubber and Plastics Production	Treatment
19	Non-metallic Mineral Manufacturing	Treatment
20	Basic Metal Industry	Control
21	Manufacturing of Metal Products, Machinery and Equipment	Control
22	Other Manufacturing Industries	Treatment
23	Electricity Generation	Treatment
24	Supply of Electricity, Gas, Water and Waste Management	Control
25	Building Construction	Control
26	Engineering Construction and Specialized Construction Activities	Treatment
27	Automotive Trade	Treatment
28	Wholesale Trade	Treatment
29	Retail Trade	Treatment
30	Restaurants and Hotels	Treatment
31	Transportation and Storage	Treatment
32	Communications	Treatment
33	Information Services	Treatment
34	Financial Activities	Treatment
35	Real Estate Activities	Treatment
37	Business Services Activities	Treatment
38	Public Education	Treatment
39	Private Education	Treatment
40	Public Health	Control
41	Private Health	Control
42	Other Services	Treatment
43	Public Administration	Control

C Robustness Results

Table 10: Effects on WFH job posting

	Dependent variable: Number of WFH postings: $\log(1+\eta_{ft})$		
	Control Group	Treatment Group	Full Sample
	(1)	(2)	(3)
Post	0.004***	0.017***	
	(0.001)	(0.001)	
Treat x Post			0.012***
			(0.001)
Firm Effects	✓	✓	✓
Time Effects			✓
AVG $\log(1+\eta_{ft})$ (Pre)	0.0002	0.0033	
% Change			3.7056
Observations	61530	239400	300930
N. of groups	1465	5700	7165
R^2	0.003	0.007	0.001

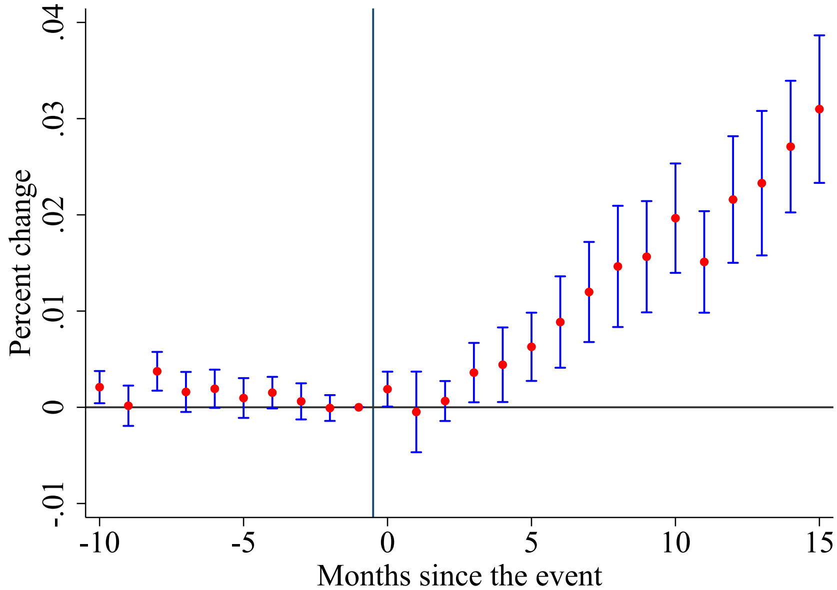
Notes: Summarizes the results of the regression defined in Equation 3, shown for each treatment group defined by $Treat_{fs}$. The dependent variable is the log of one plus the number of WFH job postings in a firm in a month, $\log(1+\eta_{ft})$. Standard errors are clustered at the firm level and are in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table 11: Effects on WFH job posting by CIUO major group classification

	Dependent variable: Number of WFH postings: $\log(1+\eta_{ft})$							
	k=1	k=2	k=3	k=4	k=5	k=7	k=8	k=9
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat x Post	0.003***	0.013***	0.007***	0.010***	0.005***	0.000	-0.000*	0.000
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)
Firm Effects	✓	✓	✓	✓	✓	✓	✓	✓
Time Effects	✓	✓	✓	✓	✓	✓	✓	✓
AVG $\log(1+\eta_{ft})$ (Pre)	0.0004	0.0024	0.0032	0.0013	0.0012	0.0003	0.0002	0.0001
% Change	7.5898	5.6199	2.1916	7.7423	4.4422	0.6977	-1.9198	1.1265
Observations	50442	156996	137718	90300	77322	100212	47712	55944
N. of groups	1201	3738	3279	2150	1841	2386	1136	1332
R^2	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000

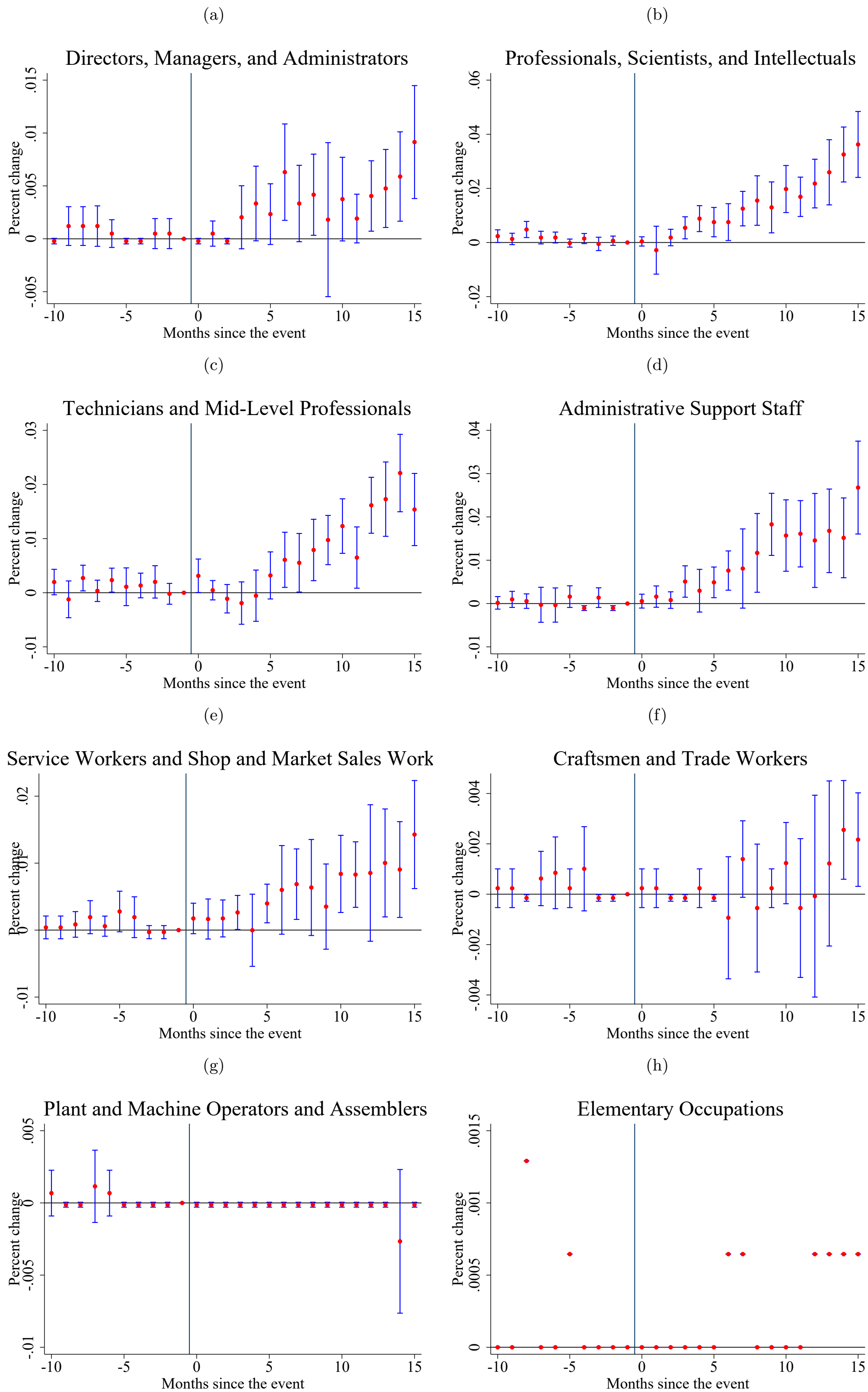
Notes: Summarizes the results of the regression defined in Equation 3 by CIUO code k, where each major group corresponds to the following label: (1) "Directors, Managers, and Administrators". (2) "Professionals, Scientists, and Intellectuals". (3) "Technicians and Mid-level Professionals". (4) "Administrative Assistants". (5) "Service Workers and Shop and Market Sales Workers". (6) "Farmers and Skilled Agricultural, Forestry, and Fishery Workers". (7) "Craftsmen and Trade Workers". (8) "Plant and Machine Operators and Assemblers". (9) "Elementary Occupations". The dependent variable is the log of one plus the number of WFH job postings in a firm in a month, $\log(1+\eta_{ft})$. Standard errors are clustered at the firm level and are in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Figure 11: Event Study on the COVID-19 impact on WFH job posting



Notes: The dependent variable is the log of one plus the number of WFH job postings in a firm in a month, $\log(1+\eta_{ft})$. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

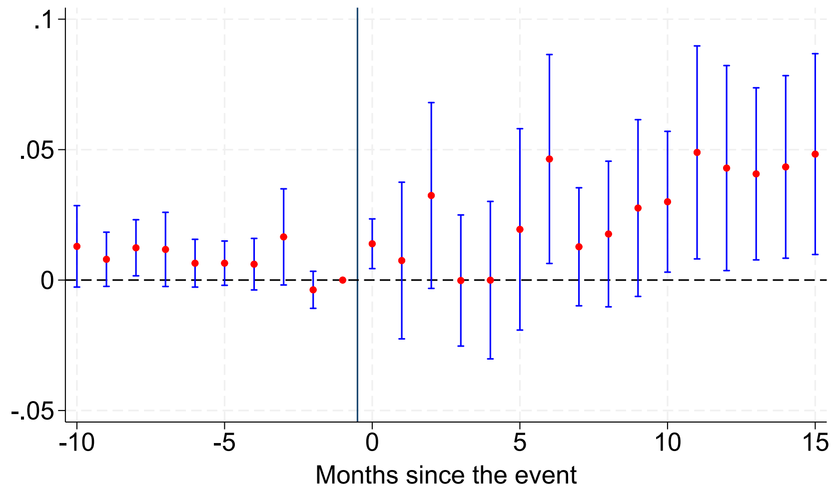
Figure 12: Event Studies on the COVID-19 impact on WFH job posting by occupation



Notes: The dependent variable is the log of one plus the number of WFH job postings in a firm in a month, $\log(1 + \eta_{ft})$. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

C.1 Firm Treatment

Figure 13: Event Study on the COVID-19 impact on WFH job posting



Notes: The dependent variable is the ratio of WFH job postings to all job postings in a firm in a month. This figure shows the estimated coefficients β_g, $g \in G$, from the regression defined in Equation 4. The vertical bars denote the 95% confidence interval.

Documentos de Trabajo **Banco Central de Chile**	**Working Papers** **Central Bank of Chile**
NÚMEROS ANTERIORES	PAST ISSUES
La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc. Existe la posibilidad de solicitar una copia impresa con un costo de Ch$500 si es dentro de Chile y US$12 si es fuera de Chile. Las solicitudes se pueden hacer por fax: +56 2 26702231 o a través del correo electrónico: bcch@bcentral.cl.	Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper. Printed versions can be ordered individually for US$12 per copy (for order inside Chile the charge is Ch$500.) Orders can be placed by fax: +56 2 26702231 or by email: bcch@bcentral.cl.

DTBC - 1089
A Deep Learning Model for Classifying Job Positions
Diego Donoso, Roberto Gillmore, Jesica Olivares, Dagoberto Quevedo

DTBC - 1088
From Invoices to Economic Monitoring Using Machine Learning Emiliano Luttini, Matías Pizarro, Dagoberto Quevedo, Marco Rojas

DTBC - 1087
Inflation Illiteracy: What Does Microdata Reveal About Chilean Consumers?
Carlos Medel

DTBC - 1086
Structural Synchronicity: A Granular Analysis of China Shocks and Global Inflation
Benjamín Álvarez, Gabriel Zelpo

DTBC – 857 (Updated)
Immigration in Emerging Countries: A Macroeconomic Perspective Agustín H. Arias, Juan Guerra-Salas

DTBC – 1085
Anatomy of the investment network in a commodity-dependent economy
Roberto Gillmore, Rodrigo Heresi, David Kohn, Dagoberto Quevedo, Nicolás Rivera

DTBC – 1084
Inversión Agregada en Chile: Factores Globales, Domésticos, y Mineros
Arturo Claro, Camilo Levenier, Carlos A. Medel

DTBC – 1083
Regulating Vertical Markets through Delegation
Andrea Canales, Nicolás Figueroa, Hugo E. Silva

DTBC – 1082
Sovereign Wealth Funds and Optimal Foreign Reserves
Miguel Acosta-Henao, Humberto Martínez, Carlos Rondón-Moreno

DTBC – 1081
Conditional Bayesian Quantile Regressions for Forecasting the GDP Growth Distribution in a Small Open Economy
Jorge Fornero, Carlos Molina

DTBC – 1080
Fundamental Drivers of Financial Conditions
Elías Albagli, Guillermo Carlomagno, Javier Ledezma, María Teresa Reszczynski

DTBC – 1079
Hospital Choice, C-sections, and long-term maternal health
Ramiro de Elejalde, Eugenio Giolito

DTBC – 1078
Zero Energy Day: How Nationwide Blackouts Affect the Economy
Luis Gonzales, Koichiro Ito, Mar Reguant

DTBC – 1015 (Updated)
Fiscal Consolidations in Commodity-Exporting Countries: A DSGE Perspective
Manuel González-Astudillo , Juan Guerra-Salas , Avi Lipton

DTBC – 1077
Un sistema de proyección de demanda por efectivo en Chile: Actualización y propuesta
Nicolás Leiva, Carlos A. Medel

DTBC – 1076
Inflation Heterogeneity and Differential Effects of Monetary and Oil Price Shocks
Felipe Martínez

DTBC – 1075
Consumption Insurance over the Life Cycle
Enzo Cerletti, Tomás Cortés

DOCUMENTOS DE TRABAJO Septiembre 2026