

Índice de Anuncios de Empleo con Base en Anuncios de Periódicos en Línea

LEÓN FERNÁNDEZ BUJANDA

BANCO DE MÉXICO

JUNIO 2021

Aclaratoria: Las opiniones y conclusiones aquí presentadas son responsabilidad exclusiva del autor y no reflejan necesariamente las del Banco de México o su Junta de Gobierno.

Introducción

Motivación:

- Ausencia de indicadores que den seguimiento a la evolución de la demanda laboral en México
- Utilización de herramientas de analítica de datos

Propósito: introducir dos indicadores experimentales que den seguimiento a la evolución de las vacantes de empleo en la economía mexicana.

Índice de Anuncios Impresos de Empleos (IAIE) :

- Conteo de anuncios de búsqueda de empleo en la sección de clasificados de la versión impresa de un importante diario de circulación nacional en México.
- Frecuencia: Trimestral
- Período: 2005 hasta 2021.
- Cinco ejemplares por trimestre, elegidos al azar de una estratificación por días de la semana (domingo, lunes, martes, miércoles y viernes).

Introducción

Índice de Anuncios Electrónicos de Empleo (IAEE):

- Conteo de anuncios de búsqueda de empleo en la sección de empleo del portal electrónico de anuncios clasificados del diarios utilizado para el IAIE.
- Frecuencia: Semanal.
- Período: Desde Marzo de 2019
- Los anuncios se recogen a través de un *webscrapper* que se activa los lunes de cada semana.
- Se recopila información cualitativa de cada anuncio.
- Se utiliza una técnica de procesamiento de lenguaje natural con la herramienta Python para clasificar los anuncios en ocupaciones utilizando la descripción del anuncio.
 - Sistema Nacional de Clasificación de Ocupaciones 2018 (SINCO): división y niveles de grupo (1 y 2 dígitos)

Introducción

Resultados principales:

- Disminución secular en el uso de medios impresos para anunciar los trabajos disponibles.
 - Mayor intensidad del uso de Internet para promover vacantes en portales digitales especializados.
- Se observan fluctuaciones cíclicas a pesar de la tendencia decreciente.
 - El pico histórico se produce antes de la recesión de 2009 y 2010.
- Covid-19 resultó en un colapso en las vacantes para México y, a pesar de su recuperación, estas aún no han alcanzado los niveles pre-pandémicos.
 - 400 anuncios semanales antes de la pandemia, baja a 50 anuncios al principio de ésta, y 200 más recientemente.

Estructura

- Revisión de la Literatura
- IAIE
- IAEE
- Conclusiones

Revisión de la Literatura

El indicador guía para la elaboración de IAIE e IAEE: *Help-Wanted Index* del *Conference Board*.

Ejemplos más recientes:

- Índices de Vacantes de Chile (Belani et al., 2002),
- y el de Colombia (Alvarez y Hofstetter, 2013, 2014).

Los índices de vacantes construidos a partir de anuncios de empleo no pretenden medir el nivel de vacantes *per se*.

- Los índices de vacantes a partir de anuncios de prensa siguen muy de cerca las fluctuaciones de las medidas obtenidas directamente a partir de información de establecimientos (Abraham 1987).
- Existe una relación positiva entre los índices de anuncios de empleo de otros países y las tasas de contrataciones (Amoah 2000).

Revisión de la Literatura

Problemas de representatividad: la probabilidad de anuncios de empleos publicados varía por ocupación.

- Ponderar los anuncios de empleo por la participación del empleo total de cada ocupación a partir de información de las encuestas de hogares.

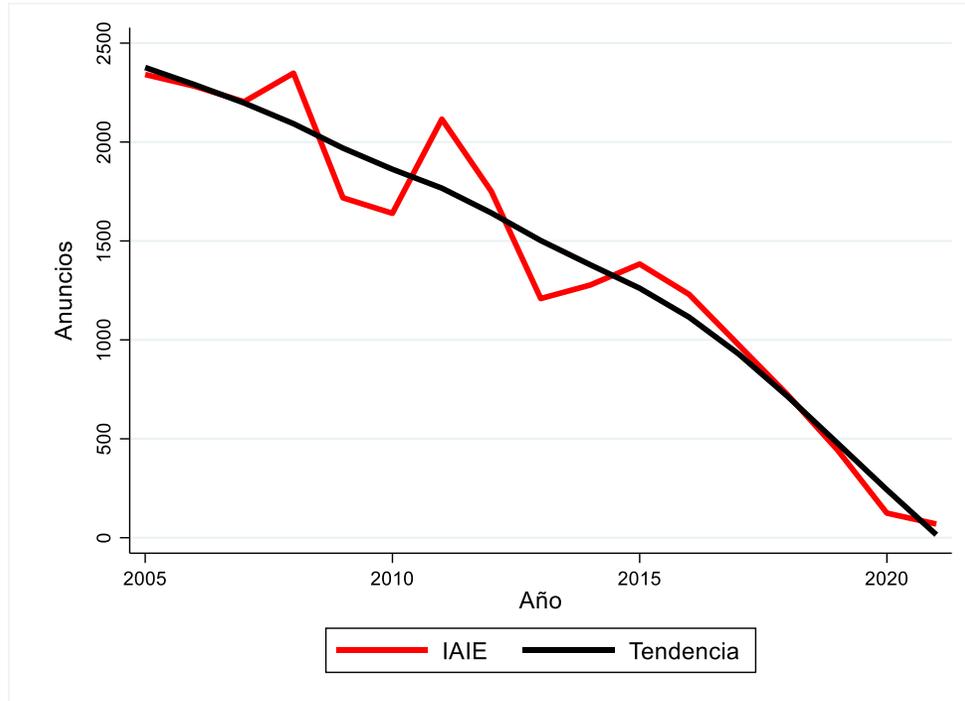
Cambios tecnológicos ocurridos en los últimos veinte años: las empresas han sustituido los anuncios de empleo en las versiones impresas de los periódicos por anuncios en las versiones en líneas de las mismas y en portales de empleo.

- Webscrapping (Campos-Vazquez et al. 2020).

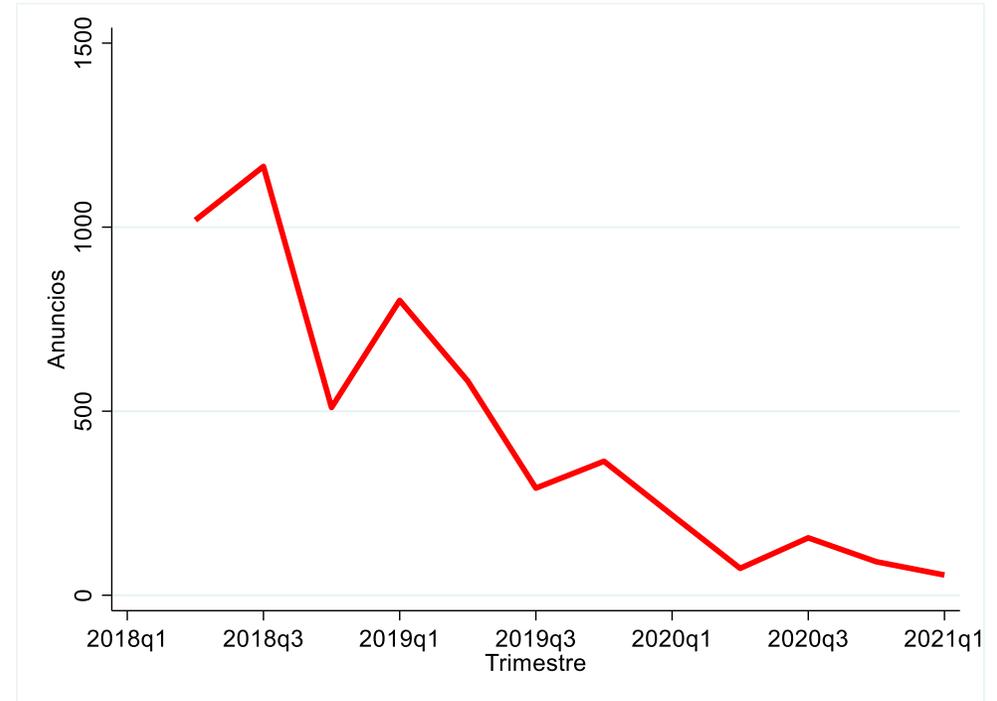
Además de las referencias directamente relacionadas con la construcción de índices de vacantes, esta propuesta aspira a contribuir a la larga literatura sobre curva de Beveridge.

Índice de Anuncios Impresos de Empleos

ÍNDICE DE ANUNCIOS IMPRESOS DE EMPLEO PROMEDIO ANUAL



ÍNDICE DE ANUNCIOS IMPRESOS DE EMPLEO PROMEDIO TRIMESTRAL



Fuente: Fuente: Elaborado por Banco de México con datos avisos clasificados impresos de un diario nacional.

Índice de Anuncios Electrónicos de Empleos

El IAEE se construye a partir de una base de datos que incluye los textos de los anuncios de empleo de un sitio web de avisos clasificados de un importante diario de circulación nacional.

Información se recolecta a través de un *webscrapper*, el cual se activa todos los lunes.

La base de datos para la realización el cálculo de este índice contiene la siguiente información:

- Fecha
- Subcategoría
- Descripción corta
- Descripción del empleo

El IAEE se construye con base en el número de anuncios (cada anuncio es una vacante).

Índice de Anuncios Electrónicos de Empleos

Con los campos descripción corta y descripción completa se pueden clasificar los anuncios en ocupaciones (grupo principal) a partir del Sistema Nacional de Clasificación de Ocupaciones (SINCO).

Se construyó un diccionario de palabras clave que pueden estar relacionados con los puestos de trabajo descritos por el SINCO para cada grupo principal.

- E.g. El grupo principal 83 hace referencia a conductores de transporte, por lo que las palabras para este grupo tienen la siguiente forma: ['taxista', 'uber', 'chofer', 'trailer', 'trailer', ...]
- Se incluye la profesión (e.g. trailer) y términos para abarcar la mayor cantidad de palabras que puedan describir la vacante solicitada (e.g. trailer, camion tres toneladas, etc.)

Índice de Anuncios Electrónicos de Empleos

Veamos un ejemplo. Consideremos el siguiente anuncio:

“Empleos Maestros Centro (Área 4), 6040, Cuauhtémoc, Ciudad de México, México. Solicitamos profesora de inglés para primaria con teacher. Presentarse en Laguna del Carmen 167, Colonia Anáhuac, cerca metro Colegio Militar. Anuncio – 0000369588”

La limpieza del texto produciría el siguiente output:

['teacher', 'carmen', 'colegio', 'primaria', 'solicitamos', 'cerca', 'profesora', 'presentarse', 'anahuac', 'empleos', 'ingles', 'colonia', 'militar', 'maestros']

El anuncio claramente solicita una profesora de inglés. Al utilizar la función descrita, se obtendría la siguiente lista de palabras:

['ingles', 'maestros', 'profesora']

Índice de Anuncios Electrónicos de Empleos

A partir de esta clasificación preliminar se encuentra en desarrollo un modelo supervisado de *machine learning*.

Se compararon diversos modelos para la clasificación de texto utilizando la clasificación preliminar para el periodo 2019M03 a 2021M05.

El mejor modelo fue definido como aquel que tuviera el mayor *accuracy*. El modelo ganador fue un *Linear Support Vector Classification* (LSVC) con un *accuracy* de 93.55%.

- Un modelo SVM construye un hiperplano o conjunto de hiperplanos que se usan para la clasificación. Intuitivamente, una separación adecuada se obtiene con el hiperplano que tenga la mayor distancia respecto al dato más cercano de cualquier clase. En el caso de un LSVC se utiliza un kernel lineal.
- El modelo estimado considera un intercepto, una función de pérdida *squared hinge*, y una estrategia de clasificación *one-vs-rest*.

Índice de Anuncios Electrónicos de Empleos

CLASIFICACIÓN PRELIMINAR
NÚMERO DE ANUNCIOS



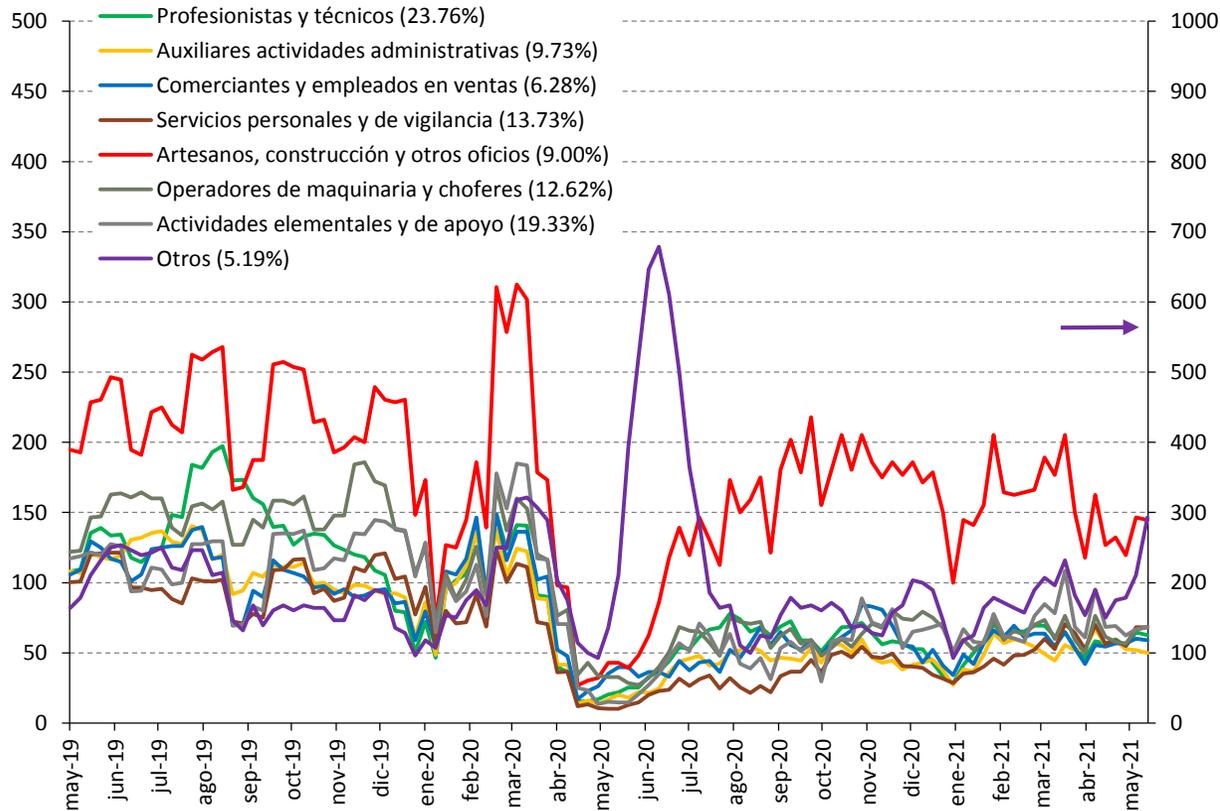
CLASIFICACIÓN LSVC
NÚMERO DE ANUNCIOS



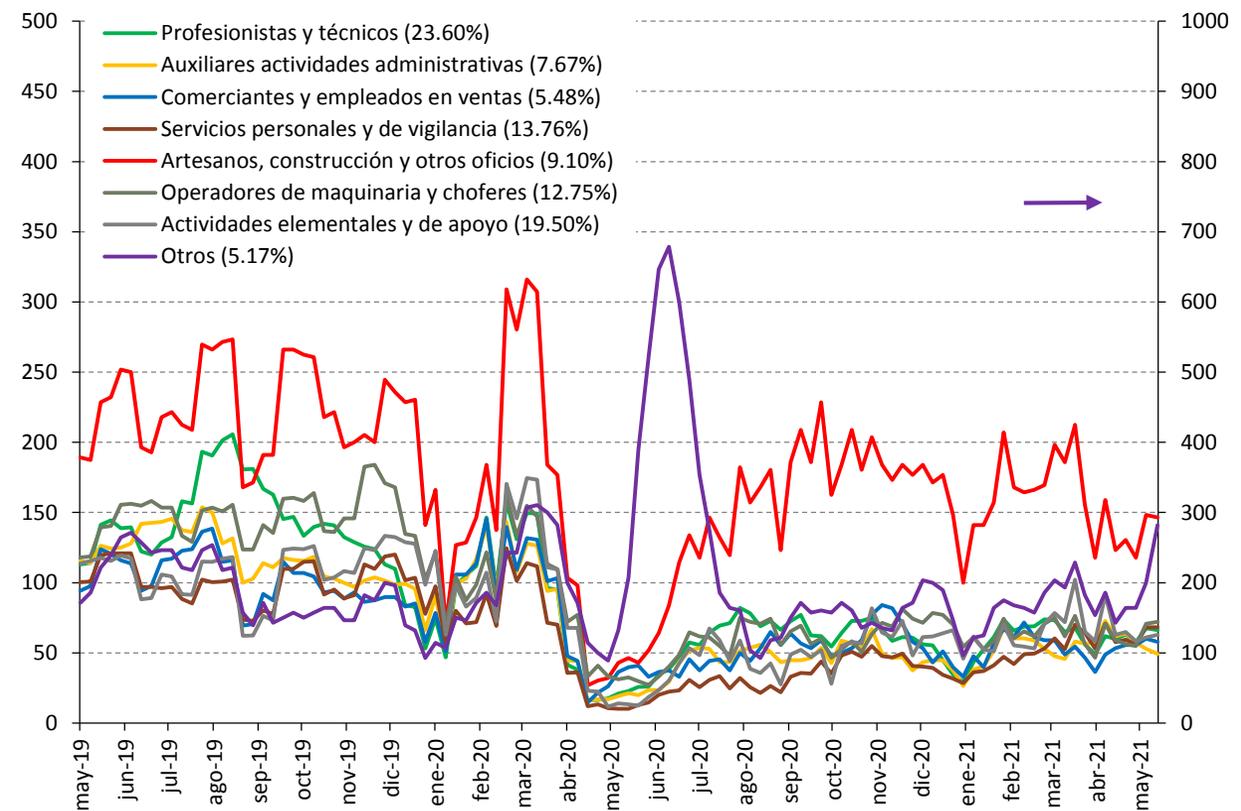
Fuente: Elaborado por Banco de México con datos del portal de avisos clasificados de un diario nacional.

Índice de Anuncios Electrónicos de Empleos

CLASIFICACIÓN PRELIMINAR POR OCUPACIÓN NÚMERO DE ANUNCIOS



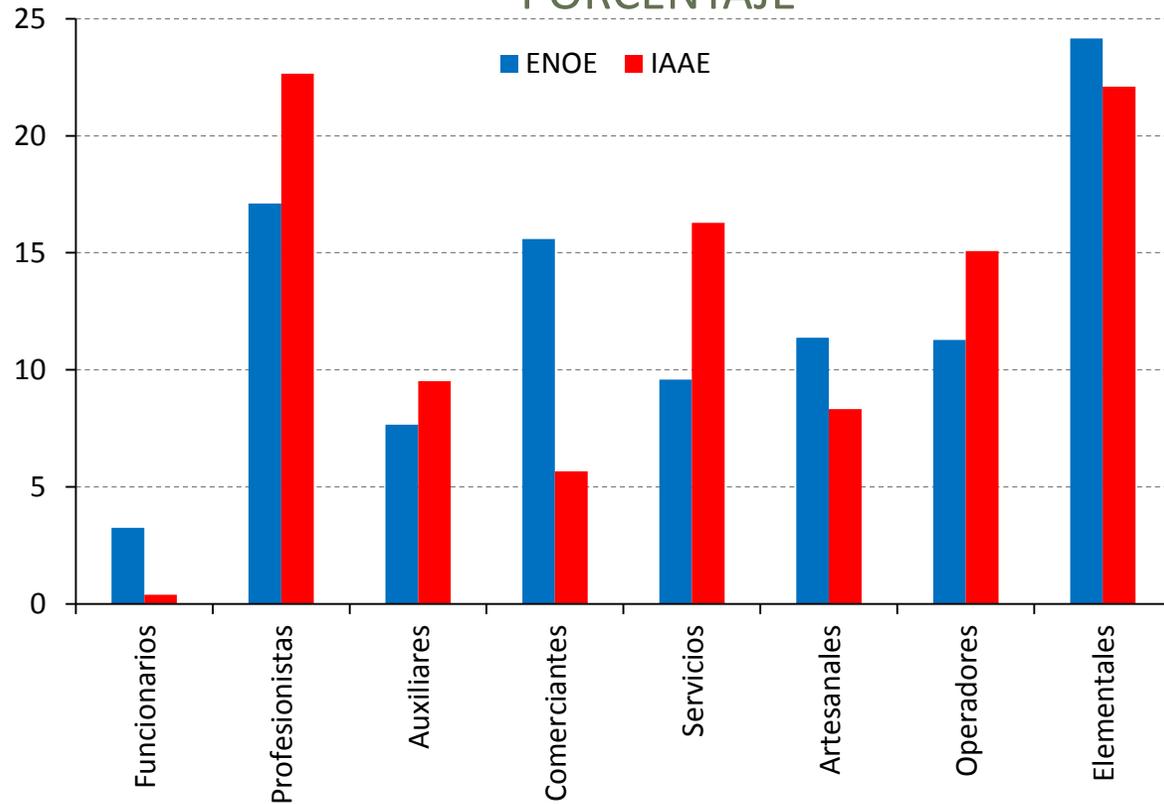
CLASIFICACIÓN LSVIC NÚMERO DE ANUNCIOS



Fuente: Elaborado por Banco de México con datos del portal de avisos clasificados de un diario nacional.

Índice de Anuncios Electrónicos de Empleos

DISTRIBUCIÓN DE OCUPACIONES
PORCENTAJE



IAEE PONDERADO Y SIN PONDERAR
NÚMERO DE ANUNCIOS



Comentarios Finales

Se introducen dos indicadores cuyo objetivo es hacerle seguimiento a la evolución de las vacantes en el mercado laboral.

Resultados

- Tendencia decreciente, lo cual se podría explicar por cambios estructurales en la manera en que se realizan las búsquedas para llenar las vacantes.
- Comportamiento cíclico de la economía mexicana durante el período de estudio.
- Reducción en los esfuerzos de búsqueda al final y al principio del año.
- Desplome desde abril de 2020 asociado a la pandemia de COVID-19.

Próximos pasos

- OCR a los anuncios impresos.
- Mejora de los modelos de clasificación automáticos.
- Recopilar data faltante de los anuncios impresos.