

Aplicación de los modelos de espacio de estados para modelar los datos de tasas de desempleo en Lima Metropolitana

Application of state space models to model unemployment rate data in Metropolitan Lima

Anna Sikov^{1*} 

¹*Departamento de Ing. Económica y Estadística, Universidad Nacional de Ingeniería, Lima, Perú*
E-mail: asikov@uni.edu.pe

Recibido (Received): 23/06/2020 Aceptado (Accepted): 07/10/2020 Publicado (Published): 15/12/2020

RESUMEN

La dinámica de la tasa de desempleo a lo largo del tiempo refleja el desempeño del mercado laboral, uno de los determinantes de bienestar más importantes en cualquier economía. La tasa de desempleo es uno de los principales indicadores utilizados para analizar la prosperidad y la situación económica de un país; permite, además, analizar la economía desde dos perspectivas: la primera, desde la oferta y, la segunda, desde la demanda. En esta investigación, vamos a analizar la dinámica de desempleo en Lima Metropolitana.

Para este fin, utilizaremos los datos de la Encuesta Permanente del Empleo (EPE) desde el año 2002 hasta el 2018 para calcular las tasas de desempleo mensuales. Estamos interesados en estudiar la dinámica de desarrollo de la tasa de desempleo en Lima Metropolitana, además en comparar las dinámicas de desarrollo de dicha tasa en distintos conos, hombres y mujeres, y también las dinámicas en distintos grupos de edades. El objetivo principal de nuestra investigación es estudiar la dinámica de las tasas de desempleo en distintos grupos sociales para identificar posibles problemas de desigualdad de oportunidades en los mismos grupos, y explicar las raíces económicas y sociales que generan esta desigualdad. Para estimar las tasas de desempleo mensuales y sus respectivas varianzas, utilizamos el estimador Horvitz-Tompson. Para analizar la dinámica de las series, ajustamos el modelo de estructura básica, que es un caso especial de modelos de espacio de estados.

Palabras clave: *tasa de desempleo, modelo de espacio de estados, modelo de estructura básica, tendencia, Filtro de Kalman, muestreo, estimador Horvitz-Tompson.*

ABSTRACT

The dynamics of the unemployment rate over time reflects the performance of the labor market, one of the most important determinants of well-being in any economy. The unemployment rate is one of the main indicators used to analyze the prosperity and economic situation of a country, and also allows analyzing the economy from two perspectives: the first from the supply side and from the demand side. In this research we are going to analyze the dynamics of unemployment in Metropolitan Lima.

For this purpose, we will use the data from the Permanent Employment Survey - EPE from 2002 to 2018 to calculate the monthly unemployment rates. We are interested in studying the development dynamics of the unemployment rate in Metropolitan Lima, and also in comparing the development dynamics of the said rate in different cones, men and women, and also the

dynamics in different age groups. The main objective of our research is to study the dynamics of unemployment rates in different social groups to identify possible problems of inequality of opportunities in the same groups, and to explain the economic and social roots that generate this inequality. To estimate the monthly unemployment rates and their respective variances we use the Horvitz-Tompson estimator. To analyze the dynamics of the series we adjust the basic structure model, which is a special case of state space models.

Keywords: *Unemployment rate, state space model, basic structure model, trend, Kalman filter, sampling, Horvitz-Tompson estimator*

1. INTRODUCCIÓN

El desempleo es uno de los principales indicadores utilizados para analizar la prosperidad y la situación económica de un país. El desempleo es un fenómeno económico puesto que surge y se desarrolla en el funcionamiento del sistema económico. Por el otro lado, el desempleo también es un fenómeno social generado por estructuras sociales de exclusión y discriminación a ciertas poblaciones, que se manifiesta principalmente en las formas de discriminación por género, edad o etnia. Por ejemplo, estiman que la tasa de desempleo es generalmente mayor entre las mujeres y entre los jóvenes. En un informe del Centro de Investigación y Desarrollo (INEI/CIDE) del año 2001, los autores identifican los siguientes grupos sociales afectados por el desempleo en el Perú: los jóvenes, las mujeres y los pobres, y explican cuáles son las raíces de este problema desde el punto de vista económico. En la Figura 1, están presentadas las tasas de desempleo para hombres y para mujeres en Lima Metropolitana.

Se puede observar claramente que existe una brecha en las tasas de desempleo entre los dos grupos, aunque esta disminuye con el tiempo.

El análisis de la tasa de desempleo es primordial, puesto que su dinámica a lo largo del tiempo permite analizar la economía desde dos perspectivas: la primera, desde la oferta, toda vez que el aumento de la tasa de desempleo disminuye la producción de bienes y servicios para una economía; y, por el lado de la demanda, un aumento de esta tasa disminuye el poder adquisitivo de los consumidores, por lo cual ambos efectos generan impacto directo en la estabilidad económica.

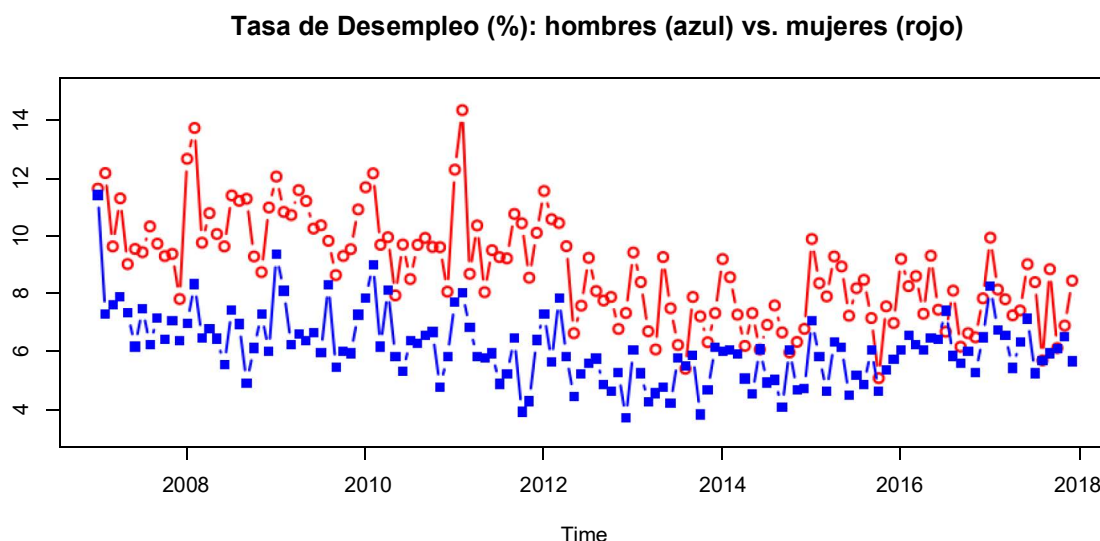
Además, es importante estudiar la dinámica de las tasas de desempleo en distintos grupos sociales para identificar posibles problemas de desigualdad de oportunidades en los mismos grupos y proponer nuevas políticas públicas en el sector laboral para combatir estos problemas.

En la región de América Latina, el análisis de los datos de desempleo es generalmente realizado por los bancos centrales, que principalmente utilizan métodos descriptivos de los

datos, teniendo como objetivos principales generar indicadores sobre empleo e ingreso, así como generar indicadores de evolución del empleo. Las conclusiones ayudan a afianzar las políticas públicas en el sector laboral.

Bellani, D., Garca, P. y Pasten, E. (2002) estudiaron la relación entre el desempleo y las vacantes laborales a largo plazo, considerando las fricciones ocurridas por el lado de la oferta y la demanda de trabajo (la curva de Beveridge). Los autores en mención, asimismo, analizaron los datos trimestrales del número de desempleados y del número de vacantes en la economía, aplicando un modelo de regresión lineal. Se concluye que la dinámica de la tasa de desocupación se desarrolla según los ciclos económicos ocurridos en el periodo de estudio. Restrepo (2008) analizó

Figura 1. Tasas de desempleo para hombres (línea azul) y mujeres (línea roja). Fuente INEI.



los datos de la tasa de desempleo mensuales y trimestrales obtenidos del Banco Central de Chile y el Instituto Nacional de Estadísticas para estimar la tasa de desempleo no aceleradora de la inflación (NAIRU), definida como aquella tasa de desempleo congruente con un entorno de inflación estable. En otras palabras, si el desempleo se ubica por encima de la NAIRU, existe gran probabilidad de que la inflación decaiga a largo plazo. El objetivo principal de este trabajo es utilizar la estimación de la NAIRU, cuantificar la brecha de desempleo. Lasso, F., Zárate, H. (2019) describen la naturaleza de las variaciones de la tasa de desempleo. Los autores utilizaron los datos mensuales de la Gran Encuesta Integrada de Hogares provistos por el Departamento Administrativo Nacional de Estadística que fueron colectados en las 13 principales ciudades de Colombia. En este trabajo, fue aplicado el modelo de autor regresión vectorial (VAR) a la serie de las tasas de desempleo. Arango, L., Ros, A. (2015) estudiaron la duración del desempleo en Colombia. Buscaron interpretar el hecho de que la tasa de desempleo de las mujeres es más duradero comparado con la de los hombres (4% mayor). Con ese fin, se utilizan datos mensuales de la Gran Encuesta

Integrada de Hogares enfocándose en las características de los individuos en el mercado laboral. Los autores en mención demuestran que la mayor tasa de desempleo de las mujeres está vinculada con los ingresos de la pareja y que el ingreso más alto de la pareja aminora los esfuerzos de búsqueda de trabajo. Adicionalmente, se resalta que la duración del desempleo en hogares con niños es mayor en las mujeres. Juszik, D. y Mateos, N. (2017) estudian los determinantes del desempleo en Argentina en el periodo 2003 a 2015, desde el punto de vista de la dinámica de las variables institucionales (el salario mínimo y seguro de capacitación), empleo y macroeconómicas (la intensidad del capital y términos de intercambio). En este estudio, la tasa de desempleo fue modelada utilizando los modelos autorregresivos. Los autores concluyen que la tasa de desempleo puede ser explicada, tanto por variables institucionales, como por variables macroeconómicas. Es por ello que los aumentos del salario mínimo se relacionan con disminución del desempleo. De las variables macroeconómicas, se destaca la formación bruta de capital físico como porcentaje del PBI, ya que incrementos de esta generan caídas en la tasa de desempleo en el largo plazo. Además, se pone hincapié en el contexto internacional, puesto que ante un favorable contexto internacional se notó reducciones de la tasa de desempleo, sobre todo en países en vías de desarrollo, cuyas principales exportaciones son materias primas.

El objetivo principal de esta investigación es analizar cómo ha evolucionado el desempleo entre 2002-2018 en Lima Metropolitana. Específicamente, nos interesa el análisis comparativo de la evolución de la tasa de desempleo a lo largo del tiempo según el sexo y la ubicación geográfica (cinco conos de Lima: Lima Norte, Lima Centro, Lima Este, Lima Sur y el Callao) para identificar las brechas laborales entre grupos poblacionales diferentes. Nosotros utilizaremos el modelo Estructural Básico que representa un caso particular de los modelos de espacio de estados planteados por Harvey (1989). Estos modelos son muy flexibles y, por ende, comunes en análisis de series de tiempo económicos. La ventaja principal del uso de estos modelos es que los mismos permiten estimar los componentes no observables de la serie: el nivel, la tendencia y los efectos de estacionalidad para cada punto de tiempo, utilizando el filtro de Kalman.

Estimación del nivel, la tendencia y los efectos de estacionalidad permiten estudiar el desarrollo de estos componentes a lo largo del tiempo y, por ello, explicar la dinámica de la serie (por ejemplo, explicar si los cambios ocurren por el cambio de la tendencia o pueden ser explicados por la variabilidad de datos). Otra ventaja del modelamiento de los datos es que estos nos brindan la posibilidad de pronosticar los valores futuros con cierto nivel de significancia.

La idea de ajustar los modelos de espacio de estados a los datos de desempleo no es nueva. Dichos modelos, por ejemplo, ya fueron utilizados para analizar los datos de desempleo de Chile y Colombia. Análisis de los datos de desempleo de Chile utilizando modelos de espacio de estados y filtro de Kalman fue presentado por Contreras, S., Pino O. y Pizzinga, A. (2006). El objetivo de esta investigación fue estudiar el desarrollo de los efectos de los

componentes de tendencia y estacionalidad de las series de desempleo a lo largo del tiempo, y explicar la dinámica de la serie desde un punto de vista económico. Los autores utilizan los datos brindados por el Instituto Nacional de Estadística (INE) de Chile sobre la tasa de desocupación de cuatro regiones: Nuble, Concepción, Arauco y Bo-Bo, correspondientes al periodo de enero 1996 hasta setiembre 2004.

2. LA TASA DE DESEMPLEO

La tasa de desempleo representa la proporción de personas desempleadas que están buscando activamente empleo entre la Población Económicamente Activa (PEA), que es la suma de número de empleados y número de desempleados.

Sigue entonces que para la medición de la tasa de desempleo se requiere definir los conceptos empleado y "desempleado". Los desempleados son personas que se encuentran en edad de trabajar (más de 14 años de edad en el caso del Perú) y tienen las siguientes características:

- No trabajaban durante el periodo de referencia, es decir, no estaban en un empleo remunerado o por cuenta propia (autoempleo).
- Estaban disponibles para un empleo remunerado o por cuenta propia en el periodo de referencia.
- Estaban buscando trabajo (se incluyen personas que no estaban buscando trabajo, pero que en un futuro cercano (no más de tres meses) tendrán una participación en el mercado laboral, por ejemplo, personas que estaban haciendo arreglos para empezar un trabajo).

De igual manera, los empleados son todas aquellas personas que se encuentran en edad de trabajar y quienes, durante un periodo específico de tiempo, estaban con empleo remunerado o como trabajadores independientes.

3. LOS DATOS

En Perú, la entidad gubernamental responsable de producir las estadísticas laborales sociales es el Instituto Nacional de Estadística e Informática (INEI). Con la finalidad de estudiar las características del mercado laboral, el INEI recolecta información a través de dos encuestas: una a nivel nacional, que es la Encuesta Nacional de Hogares (ENAHO) y otra a nivel de Lima Metropolitana y el Callao, que es la Encuesta Permanente del Empleo (EPE), cuyo principal objetivo es medir el empleo y desempleo. Los datos de ENAHO son utilizados para publicar las estadísticas laborales a escala nacional y con frecuencia anual, mientras los datos de la EPE sirven para elaborar y publicar las estadísticas del mercado laboral en Lima Metropolitana y el Callao, con una frecuencia mensual.

En este trabajo, utilizaremos los datos proporcionados por la EPE. El método de muestreo utilizado por la EPE es probabilístico bietápico. En la primera etapa, una muestra de conglomerados fue seleccionada implementando el método de muestreo PPT (Sistemático con Probabilidad Proporcional al Tamaño), donde el tamaño de conglomerado es definido por el número de viviendas en dicho conglomerado. Los conglomerados son las unidades primarias de muestreo para la EPE y se definen como conjunto de bloques de viviendas contiguas.

En la segunda etapa, una muestra de 4 viviendas es seleccionada en cada conglomerado, para que así cada vivienda tenga la misma probabilidad de ser incluida en la muestra.

En cada vivienda seleccionada a la muestra, fueron investigados todos los integrantes (los datos colectados incluyen las características sociodemográficas, así como información necesaria para definir si la persona es empleada o desempleada).

Además, para cada miembro de dicha vivienda fue calculado (por el INEI) el factor de expansión. El factor de expansión que corresponde a una persona seleccionada en la muestra se interpreta como la cantidad de personas en la población, que representa la misma persona.

Las muestras de los años 2002- 2004 incluyen 600 conglomerados en total, que fueron divididos en 12 submuestras de 50 conglomerados, de las cuales mensualmente se visitó 4 (200 conglomerados) para investigar 4 viviendas de cada uno. De esta manera, en un trimestre, las viviendas fueron seleccionadas de todos los 600 conglomerados. De forma similar, las muestras de los años 2005-2018 incluyen 1 200 conglomerados divididos en 6 submuestras de 200, de las cuales se visitó 2 submuestras mensualmente, visitando las viviendas de todos los 1 200 conglomerados en un trimestre. Desde el 2007 a la fecha, la muestra trimestral es de 1 200 conglomerados, de los cuales son visitados mensualmente 400 conglomerados.

De la EPE se obtendrá la estimación de la tasa de desempleo mensual con las varianzas correspondientes en Lima Metropolitana. Los datos se refieren al periodo 2002-2018. Como fue mencionado, las series de las tasas de desempleo del dicho periodo serán calculadas para hombres y mujeres, de edades distintas en distintos conos de Lima Metropolitana y el Callao. Las dichas tasas y varianzas serán obtenidas utilizando el método de Horvitz-Tompson (Horvitz-Tompson, 1952).

4. METODOLOGÍA

En esta sección, presentamos la metodología para analizar los datos de desempleo. El análisis consiste en dos partes. Primero, a partir de los datos individuales de la EPE, estimamos las tasas de desempleo mensuales.

Estas tasas son calculadas utilizando el estimador Horvitz-Thompson (Horvitz and Tompson, 1952). Además, calculamos las varianzas de sus respectivos estimadores. Así obtenemos la serie mensual de tasas de desempleo y la serie de las varianzas correspondientes. En la segunda parte, aplicamos el modelo de estructura básica (BSM), definida a continuación. Dicho modelo permite estimar el nivel, la tendencia y los efectos de estacionalidad de la serie.

4.1 ESTIMACIÓN DE LAS TASAS DE DESEMPLEO

Como ya fue mencionado, para estimar las tasas de desempleo, utilizamos los datos individuales de las personas seleccionadas en la muestra. Para cada persona investigada, observamos la siguiente información: sexo, edad, el cono donde está su domicilio, la información si la misma persona pertenece a la población económicamente activa, si estaba empleada en el tiempo de investigación, según las definiciones dadas en el capítulo 3, y, además, el factor de expansión.

Sea S_t la muestra seleccionada de una población U_t en el tiempo t , donde $t = 1, \dots, T$ y $T = 168$, donde $t = 1$ corresponde a enero de 2005, y corresponde a diciembre de 2018. Designamos por n_t el tamaño de la muestra S_t , o sea el número total de personas investigadas en el tiempo t . Designamos por D_{ti} una variable indicadora, tal que $D_{ti} = 1$ si la persona i , seleccionada en la muestra que corresponde al tiempo t , está desempleada y 0 caso contrario, y por A_{ti} una variable indicadora que es igual a 1 si la persona i pertenecía a la población económicamente activa en el tiempo t y 0 caso contrario.

El verdadero valor de la tasa de desempleo general en el tiempo t , R_t se obtiene de la siguiente manera:

$$R_t = \frac{\sum_{i \in U_t} D_{ti}}{\sum_{i \in U_t} A_{ti}} = \frac{D_t}{A_t} \quad (1)$$

donde el numerador representa el verdadero número total de las personas desempleadas en la población investigada en el tiempo t , y el denominador es el verdadero tamaño de la población económicamente activa en el tiempo t .

En realidad, nunca observamos los valores de A_{ti} y D_{ti} para toda la población investigada, entonces dicha tasa no puede ser calculada de forma exacta. Sin embargo, la misma puede ser estimada utilizando la información de la encuesta.

Sea r_t la estimación de la tasa de desempleo general en el tiempo t . Entonces,

$$r_t = \frac{\sum_{i=1}^{n_t} w_{ti} D_{ti}}{\sum_{i=1}^{n_t} w_{ti} A_{ti}} = \frac{\widehat{D}_t}{\widehat{A}_t} \quad (2)$$

donde w_{ti} es el factor de expansión que corresponde a la persona $i \in S_t$, que fue investigada en el tiempo t .

El numerador de la ecuación (2) es el estimador Horvitz-Tompson (Horvitz and Tompson, 1952) para el número total de desempleados en la población en el tiempo t , y el denominador de la misma ecuación es un estimador Horvitz-Tompson para el tamaño de población económicamente activa en el tiempo t . Entonces, la estimación de la tasa de desempleo se obtiene como una razón de dos estimadores, donde cada uno se obtiene ponderando el valor de la observación en cada persona por su factor de expansión y, luego, sumando todas las personas de la muestra. Calculando las estimaciones r_t para cada tiempo t se obtiene la serie las tasas de desempleo para el periodo de investigación.

En la Figura 2 presentamos las estimaciones de la tasa de desempleo en Lima Metropolitana, obtenidas aplicando la fórmula (2).

4.2 MODELOS DE ESPACIO DE ESTADOS

Los modelos de espacio de estados son una clase amplia de modelos de series de tiempo, que son muy populares en estudios de econometría. La mayoría de los modelos (por ejemplo, los modelos tipo SARIMA) pueden ser presentados de forma de un modelo de espacio de estados. La ventaja de estos modelos es que los mismos permiten utilizar una técnica llamada de Filtro de Kalman para estimar los componentes del mismo modelo y, además, predecir los valores futuros.

Primero, vamos a definir los modelos de espacio de estados y, más adelante, explicaremos la idea del Filtro de Kalman. Para estudiar las dinámicas de las series de desempleo, utilizaremos el modelo de Estructura Básica, que está definido en la Sección 5.2.3.

4.2.1 MODELOS DE ESPACIO DE ESTADOS

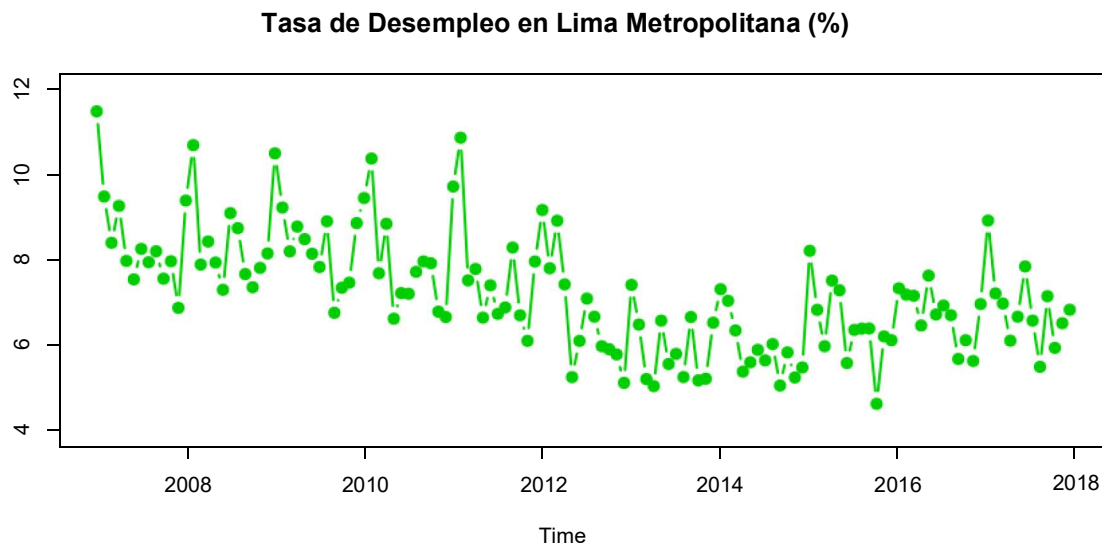
El modelo de espacio de estados se define mediante dos ecuaciones: la ecuación de observación y la ecuación de estados.

La ecuación de observación define la relación entre el valor de la variable de interés en el tiempo t , y_t , y el vector de los coeficientes desconocidos, β_t :

$$y_t = Z_t \beta_t + \varepsilon_t, \varepsilon_t \sim N(0, \Sigma_t) \quad (3)$$

donde y_t es un vector de una dimensión D , β_t es un vector de estados de una dimensión p , Z_t es una matriz, cuya dimensión es $D \times p$, ε_t es un vector de errores aleatorios y Σ_t es la matriz de varianzas de ε_t . Se supone también que los errores ε_t son independientes a lo largo del tiempo, es decir, $E(\varepsilon_{t_1}, \varepsilon_{t_2}) = 0$ si $t_1 \neq t_2$. En la mayoría de las aplicaciones Z_t y Σ_t no dependen del tiempo t . Cabe mencionar que si β_t y Σ_t no dependen del tiempo; la ecuación de observación define un modelo de regresión clásica.

Figura 2. Estimación de la tasa de desempleo en Lima Metropolitana



La ecuación de estados define el modelo de desarrollo estocástico del vector β_t :

$$(\beta_t = T\beta_{t-1}) + \eta_t \quad (4)$$

donde T es la matriz de transición, cuya dimensión es $p \times p$, η_t es un vector de errores aleatorios, tal que $\eta_t \sim N(0, H)$ y $E(\eta_{t_1}, \eta_{t_2}^t) = 0$ si $t_1 \neq t_2$. Además, $E(\varepsilon_{t_1}, \eta_{t_2}^t) = 0, \forall t_1, t_2$.

Para la descripción más detallada del modelo de espacio de estados, ver el libro de Harvey (1989).

4.2.2 EL FILTRO DE KALMAN

El Filtro de Kalman es un algoritmo recursivo que está formado por dos etapas, dados los valores iniciales β_0 y $P_0 = E(\widehat{\beta}_0 - \beta_0)(\widehat{\beta}_0 - \beta_0)^t$

1. En el tiempo $t-1$, dado el estimador para β_{t-1} , $\widehat{\beta}_{t-1}$, se predice el valor de β_t :

$$\widehat{\beta}_{t|t-1} = T\widehat{\beta}_{t-1}$$

2. En el tiempo t , la predicción de t es actualizada, utilizando el valor de la serie, observado en el tiempo t , es decir, y_t :

$$\widehat{\beta}_t = \widehat{\beta}_{t|t-1} + P_{t|t-1}^{-1} Z_t^T F_t^{-1} (Y_t - Z_t \widehat{\beta}_{t|t-1})$$

donde $P_{t|t-1} = TP_{t-1}T^t + H$ y $F_t = Z_t P_{t|t-1} Z_t^T + \Sigma_t$

Entonces, aplicando el algoritmo del Filtro de Kalman se obtienen los predictores del vector de estados para cada tiempo t .

Implementación del algoritmo Filtro de Kalman requiere definición de los valores iniciales para el vector β_0 y la matriz P_0 . En la mayoría de aplicaciones utilizan un vector de ceros para inicializar $\widehat{\beta}_0$, es decir $\beta_{0,j} = 0, j = 1, \dots, p$, y una matriz $P_{0,jk} = MI_{(j=k)}$, donde M es un valor grande comparando con los valores de la serie y_t .

4.2.3 MODELO DE ESTRUCTURA BÁSICA

Sea y_t la estimación (obtenida de la EPE) de la tasa de desempleo mensual para Lima Metropolitana en el tiempo t , y Y_t su correspondiente valor poblacional (el verdadero valor). Observamos la serie de las estimaciones y_t . Obviamente, las estimaciones de las tasas de desempleo, basadas en los datos de la encuesta, no coinciden con los verdaderos valores; sin embargo, podemos suponer que $y_t = Y_t + e_t$, donde e_t es un error muestral que sigue un proceso de ruido blanco, $e_t \sim N(0, \sigma^2)$, y que los errores e_t son independientes a lo largo del tiempo. Ahora, supongamos, que la serie Y_t puede ser descompuesta de la siguiente forma:

$$Y_t = L_t + S_t + I_p, I_t \sim N(0, \sigma_t^2),$$

donde

$$L_t = L_{t-1} + R_{t-1} + \eta_t^L, \eta_t^L \sim N(0, \sigma_L^2)$$

$$R_t = R_{t-1} + \eta_t^R, \eta_t^R \sim N(0, \sigma_R^2)$$

$$S_t = \sum_{j=1}^6 S_{jt}$$

$$S_{jt} = S_{j,t-1} \cos \omega_j + S_{j,t-1}^* \sin \omega_j + \eta_{jt}^S, \eta_{jt}^S \sim N(\sigma_S^2)$$

$$S_{j,t}^* = -S_{j,t-1} \sin \omega_j + S_{j,t-1}^* \cos \omega_j + \eta_{jt}^S, \eta_{jt}^S \sim N(\sigma_S^2)$$

$$\omega_j = \frac{2\pi j}{12}$$

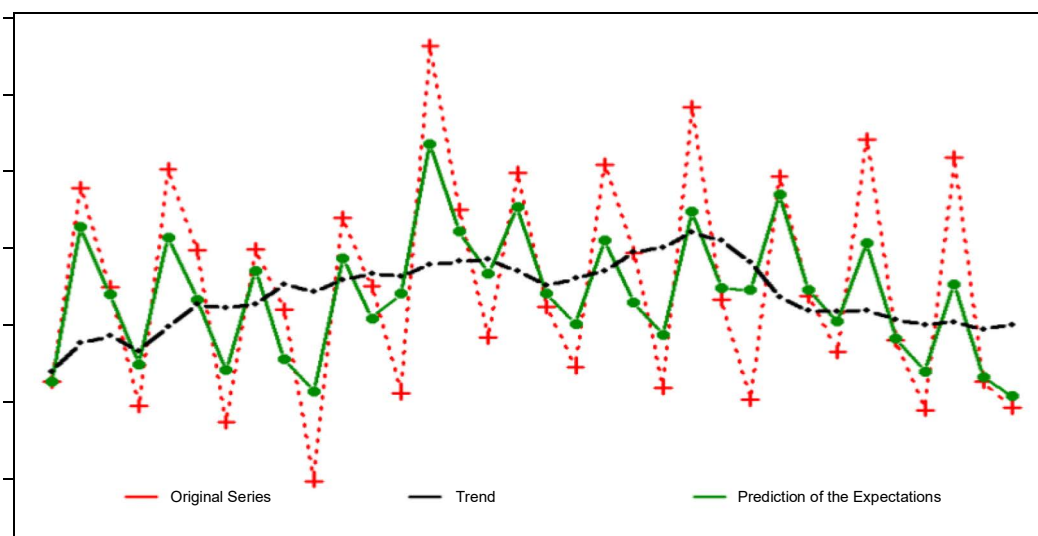
donde los componentes L_t , R_t , S_t y I_t definen el nivel de la tendencia lineal en el tiempo t , la pendiente de tendencia, el efecto estacional y el término irregular de la serie, respectivamente, operando en el tiempo t . El modelo para la tendencia se aproxima a una tendencia local lineal, mientras que el modelo para el efecto estacional usa la descomposición cíclica del componente estacional, dividido en seis subcomponentes, correspondientes a las 6 frecuencias (armónicas) de una serie estacional mensual. Los ruidos agregados permiten que los efectos estacionales evolucionen estocásticamente en el tiempo. El modelo definido arriba se llamó el Modelo Estructural Básico (*Basic Structural Model*) y es muy común en análisis de series de tiempo económicos. La Figura 3 muestra un ejemplo de una serie original y su descomposición en las componentes L_t , R_t , S_t y I_t y donde los puntos rojos presentan los valores originales de la serie, la curva negra representa el nivel de la tendencia de la serie, las diferencias entre dos puntos consecutivos de la curva del nivel son los pendientes locales de la serie, las diferencias entre la curva negra y los puntos verdes presentan los efectos de estacionalidad (positivo o negativo) y, normalmente, las diferencias entre los puntos rojos y puntos verdes son los términos irregulares (los errores) de la serie.

Entonces, dado $y_t = Y_t + e_t$ la serie de las estimaciones y_t puede ser modelada utilizando el modelo BSM,

$$y_t = L_t + S_t + e_t, \hat{e}_t \sim N(0, \sigma_e^2)$$

donde el error \hat{e}_t contiene el término irregular I_t de la serie Y_t y el error muestral e_t .

Cabe mencionar que para aplicar el Filtro de Kalman, tenemos que primero estimar los parámetros desconocidos del modelo definido: σ_e^2 , σ_L^2 , σ_R^2 , σ_S^2 . Para estimar dichos parámetros, utilizamos el método de máxima verosimilitud (vea el libro de Harvey, 1989).

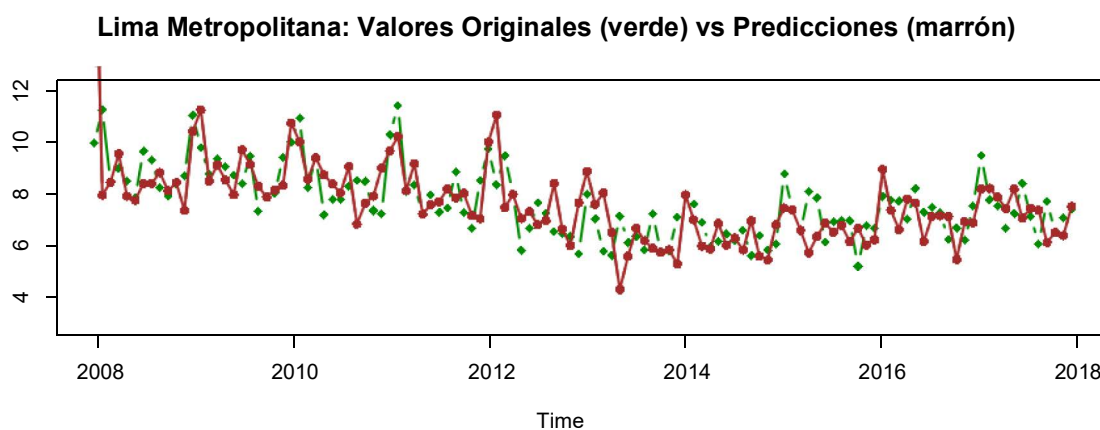
Figura 3. Descomposición de una serie en componentes según el modelo BSM.

5. APLICACIÓN DEL MODELO BSM A LOS DATOS DE DESEMPLEO

En esta sección, presentamos y analizamos los resultados de aplicación del algoritmo Filtro de Kalman a los datos de desempleo en varios conos de Lima Metropolitana.

5.1. ANÁLISIS GENERAL DE LAS DINÁMICAS DE LAS SERIES DE LAS TASAS DE DESEMPLEO EN LIMA METROPOLITANA

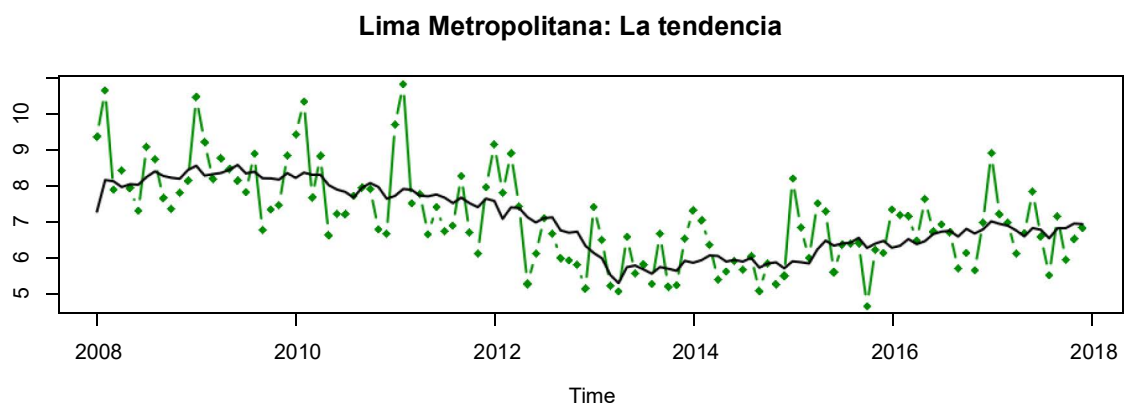
En esta sección, presentamos los resultados de aplicación del algoritmo Filtro de Kalman a los datos de las tasas de desempleo en Lima Metropolitana y el Callao. Para justificar el método utilizado, primero vamos a comparar los valores originales de las tasas de desempleo con sus respectivas predicciones, hechas un mes adelante.

Figura 4. Comparación de los valores de la serie original de las tasas de desempleo con los valores predichos bajo el modelo BSM.

Los resultados (Figura 4) indican que, pese a algunas diferencias en dichos valores, pueden ser atribuidas a las varianzas de los errores. Las predicciones obtenidas son generalmente buenas.

En el siguiente gráfico, presentamos la dinámica de la serie de las tasas de desempleo.

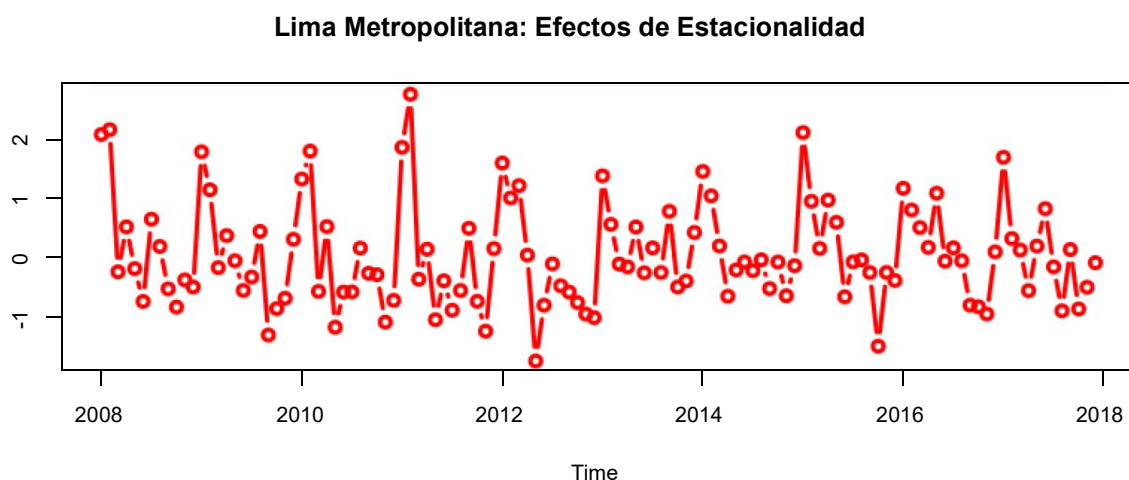
Figura 5. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia, estimada bajo el modelo BSM.



El gráfico demuestra que hay una caída de la tendencia entre los años 2008 y 2013, especialmente en el año 2012, y después, una subida persistente a partir del año 2015.

En la Figura 6, se puede observar que la serie de las tasas de desempleo es una serie que tiene un patrón de estacionalidad fuerte. Los efectos de estacionalidad son más altos en enero y febrero, y más bajos en octubre y noviembre. Para los otros meses, el efecto de estacionalidad es alrededor de cero. Es decir, en enero y febrero, las tasas de desempleo son más altas y, en octubre y noviembre, son más bajas.

Figura 6. Las estimativas de los efectos de estacionalidad a lo largo del tiempo.



5.2. ANÁLISIS COMPARATIVO DE LAS DINÁMICAS DE LAS SERIES DE LAS TASAS DE DESEMPLEO ENTRE HOMBRES Y MUJERES EN LIMA METROPOLITANA

Como en el caso anterior, primero, vamos a comparar los valores originales de las series de las tasas de desempleo con las predicciones correspondientes (Figuras 7 y 8).

Figura 7. Comparación de los valores de la serie original de las tasas de desempleo con los valores predichos bajo el modelo BSM (los datos para hombres).

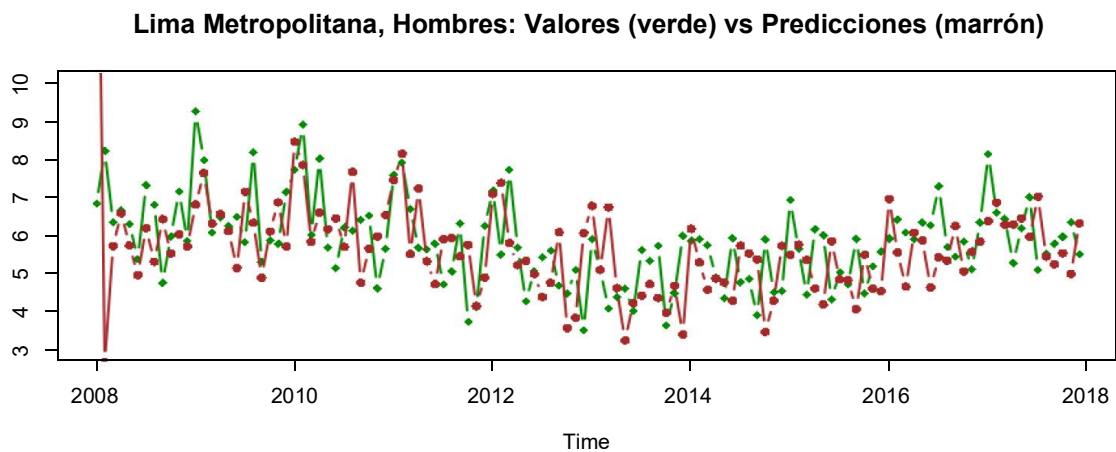
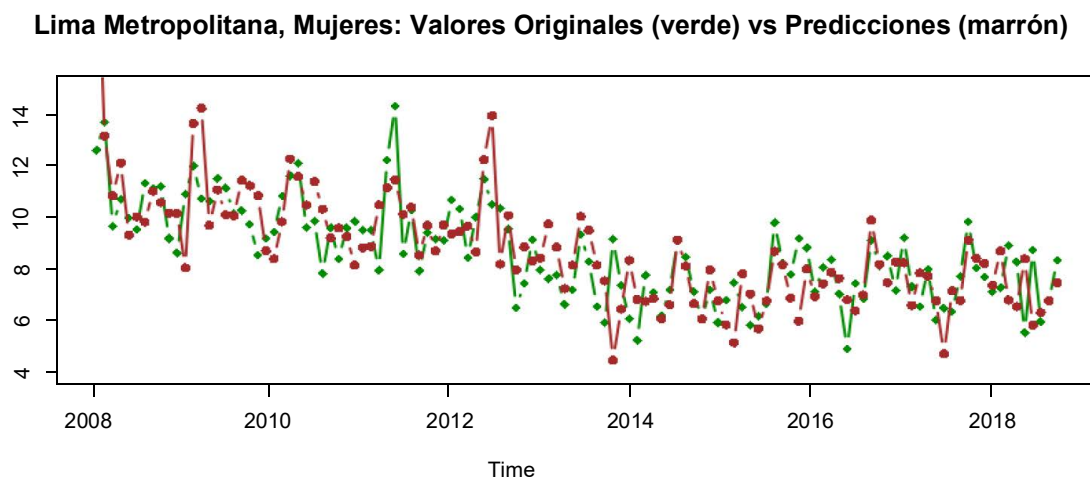


Figura 8. Comparación de los valores de la serie original de las tasas de desempleo con los valores predichos bajo el modelo BSM (los datos para mujeres).



Los gráficos demuestran que los valores originales y las predicciones son suficientemente próximos, justificando el uso del método propuesto. Los dos siguientes gráficos muestran las series originales de las tasas de desempleo y las curvas de la tendencia para hombres (Figura 9) y para mujeres (Figura 10).

Figura 9. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para hombres).

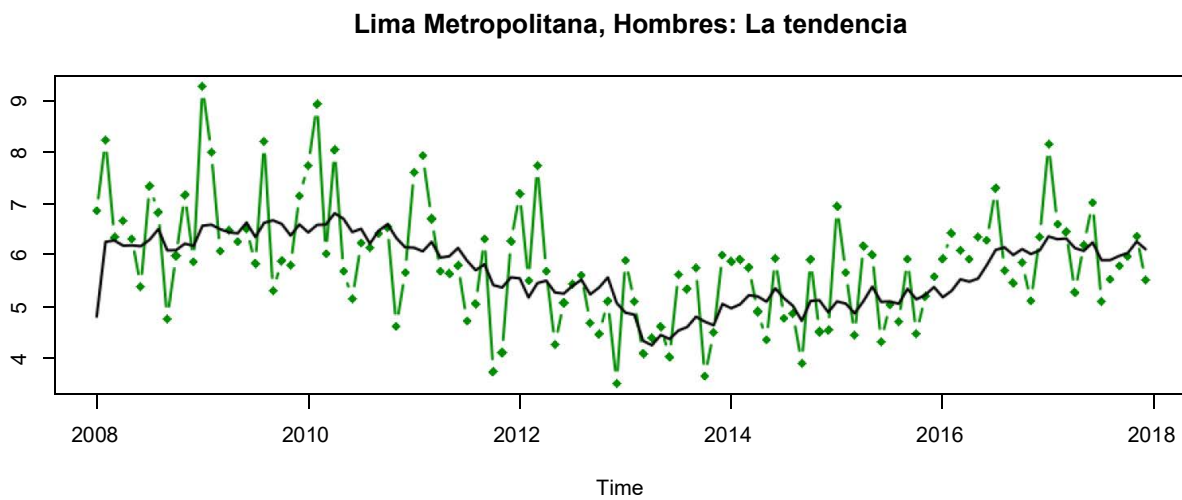
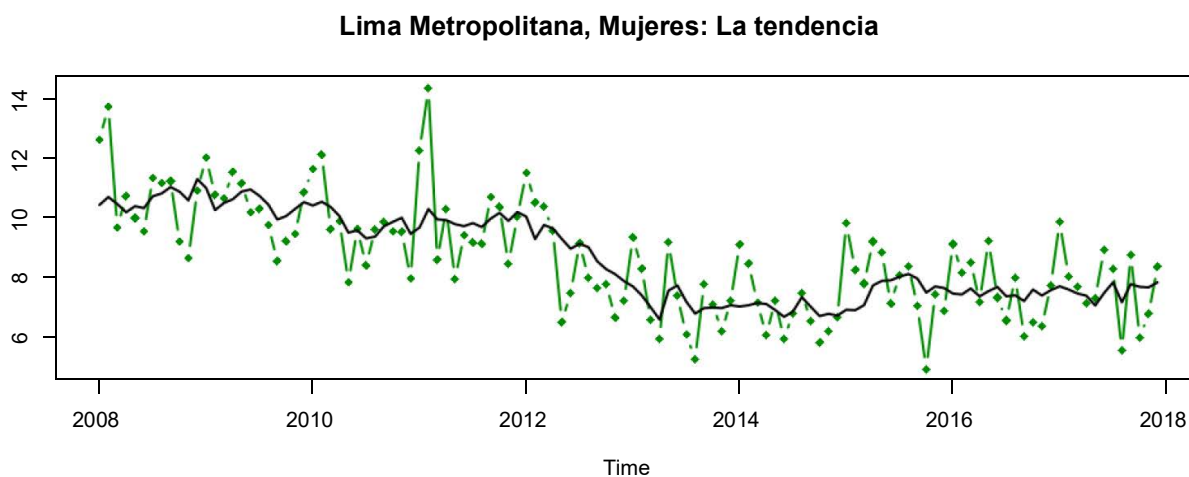


Figura 10. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para mujeres).

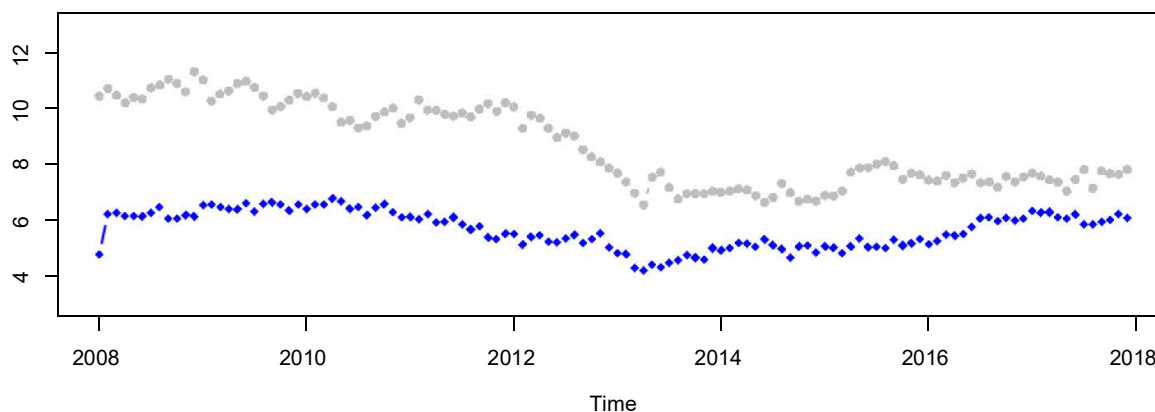


En el caso de hombres, la tendencia presenta una caída entre los años 2011 y, después, a partir del año 2014 una subida hacia los valores observados en los años 2008-2011. En el caso de mujeres, también se puede observar una caída entre los años 2012 y 2014, pero después del año 2014 la tendencia se mantiene constante.

En el siguiente gráfico, comparamos las tendencias de las tasas de desempleo entre hombres y mujeres.

Figura 11. Comparación de las tendencias de las tasas de desempleo entre hombres y mujeres.

Lima Metropolitana, Hombres (azul) vs Mujeres (plomo): La tendencia

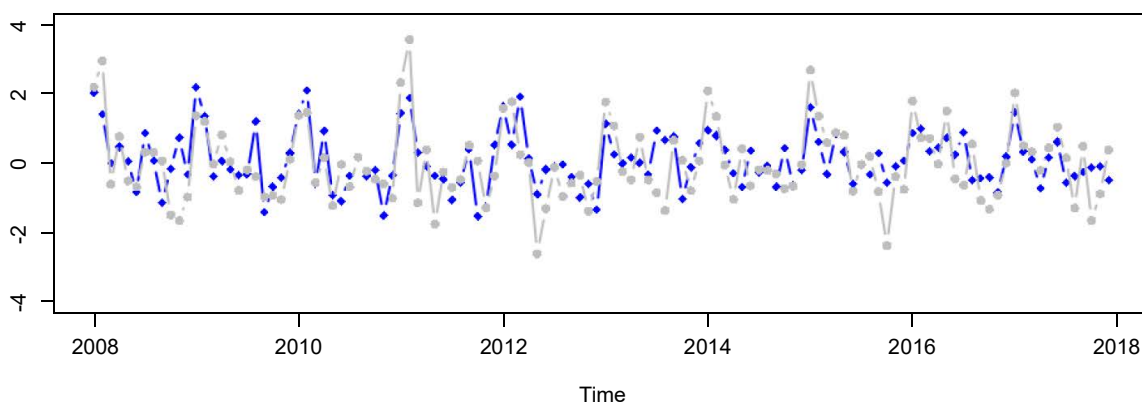


Se puede observar una brecha de alrededor de 4% en el año 2008 que va disminuyendo a lo largo del período de investigación y, en el año 2017, llega al nivel de menos que 2%.

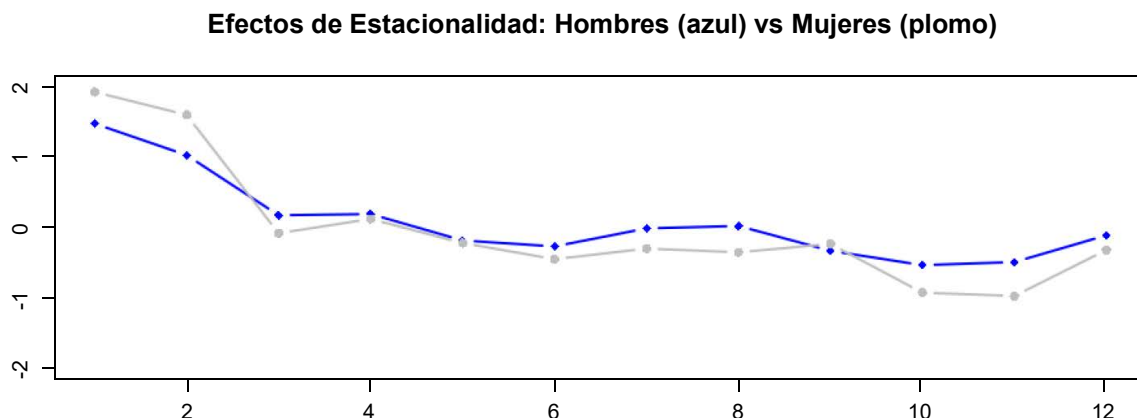
Las Figuras 12 y 13 muestran los efectos de estacionalidad de las series de las tasas de desempleo para hombres y para mujeres.

Figura 12. Comparación de los efectos de estacionalidad de las tasas de desempleo entre hombres y mujeres a lo largo del tiempo.

Lima Metropolitana, Hombres (azul) vs Mujeres (plomo): Efectos de Estacionalidad



Los gráficos demuestran comportamiento parecido de los efectos de estacionalidad en los dos casos, donde los efectos de estacionalidad mayores se observan en enero y febrero, y menores en octubre y noviembre.

Figura 13. Comparación de los efectos de distintos meses en las tasas de desempleo.

5.3. ANÁLISIS COMPARATIVO DE LAS DINÁMICAS DE LAS SERIES DE LAS TASAS DE DESEMPLEO EN CONOS DISTINTOS DE LIMA METROPOLITANA Y EL CALLAO

En esta sección, el Cono 1 se refiere a Lima Norte, Cono 2 es Lima Centro, Cono 3 se refiere a Lima Oeste, Cono 4 es Lima Sur y Cono 5 es el Callao.

En las Figuras 14 y 15, presentamos los resultados de comparación de los valores originales de las series con las predicciones obtenidas bajo el modelo BSM para dos conos diferentes: Lima Norte y Lima Sur. Los resultados indican que el uso del modelo BSM produce buenas predicciones. Para los otros conos, los gráficos son parecidos.

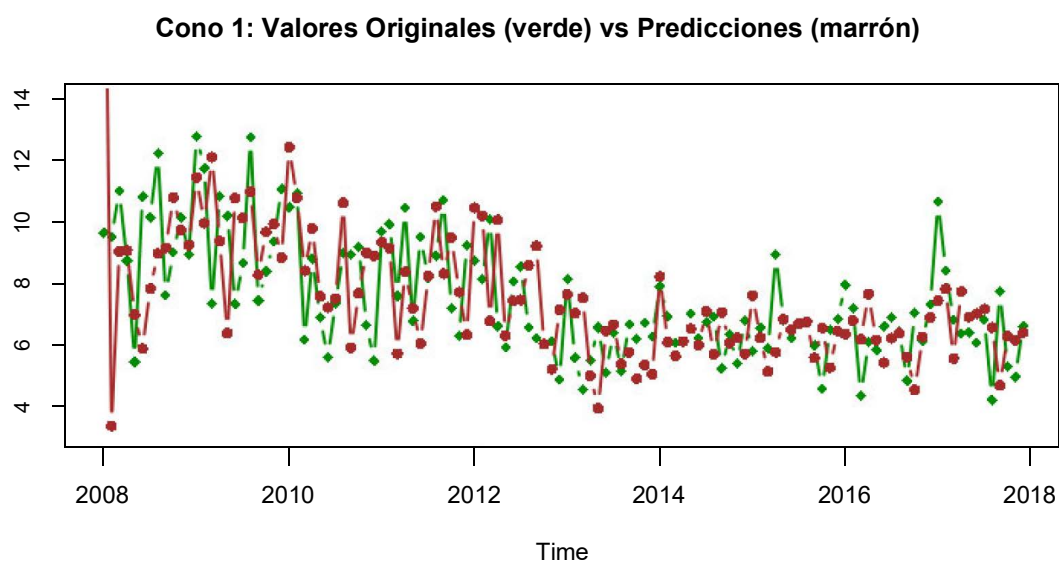
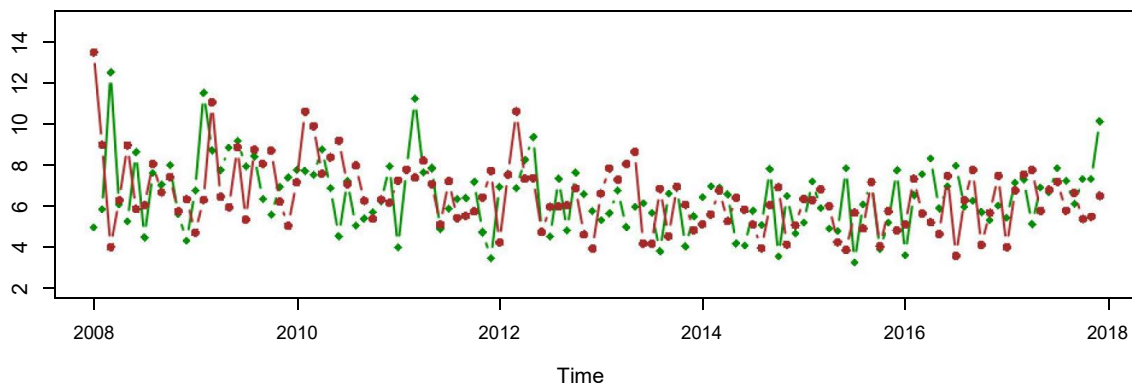
Figura 14. Comparación de los valores de la serie original de las tasas de desempleo con los valores predichos bajo el modelo BSM (los datos para Lima Norte).

Figura 15. Comparación de los valores de la serie original de las tasas de desempleo con los valores predichos bajo el modelo BSM (los datos para Lima Sur).

Cono 4: Valores Originales (verde) vs Predicciones (marrón)



Enseguida, presentamos para cada cono los gráficos con los valores originales de las series vs. las curvas de las tendencias (Figuras 16-20).

Figura 16. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para el cono de Lima Norte).

Cono 1: La tendencia

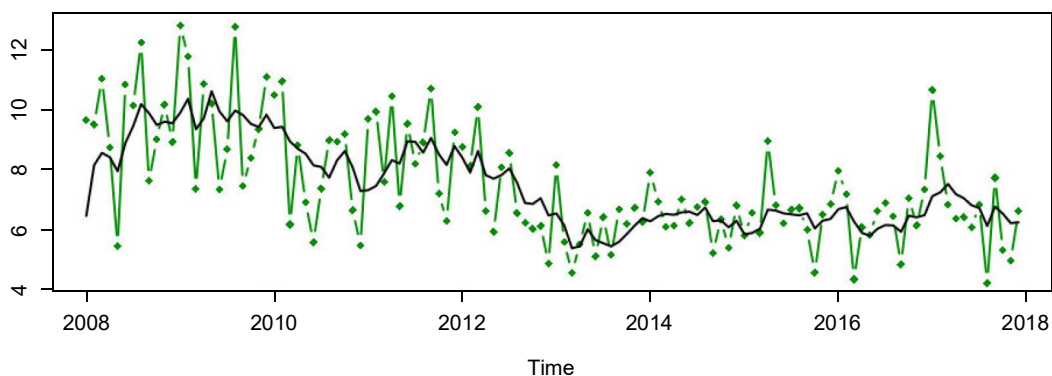


Figura 17. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para el cono de Lima Centro).

Cono 2: La tendencia

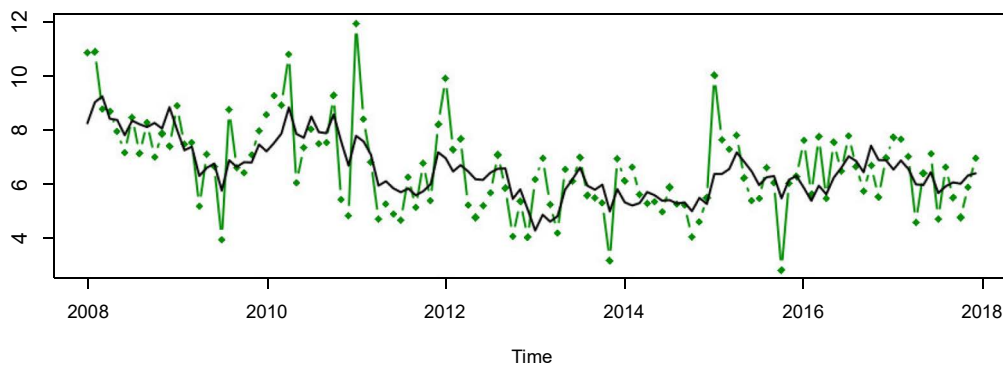


Figura 18. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para el cono de Lima Oeste).

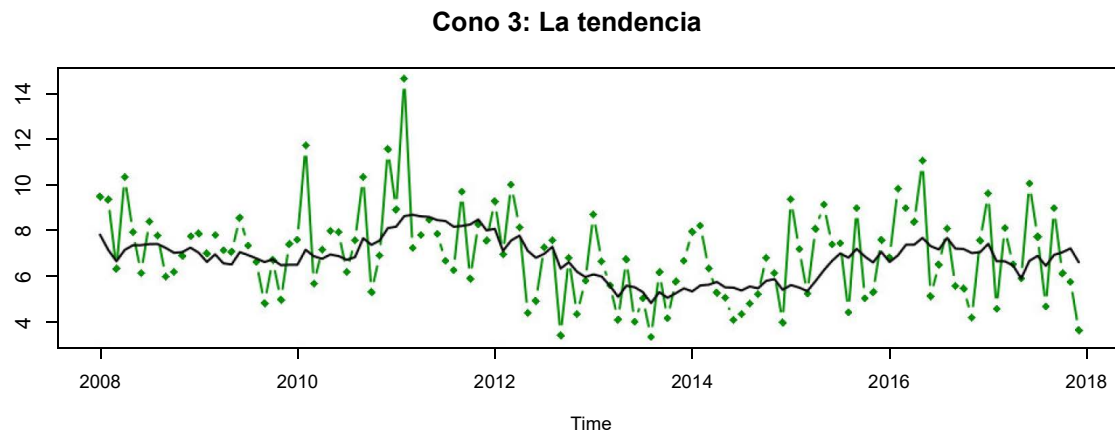


Figura 19. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para el cono de Lima Sur).

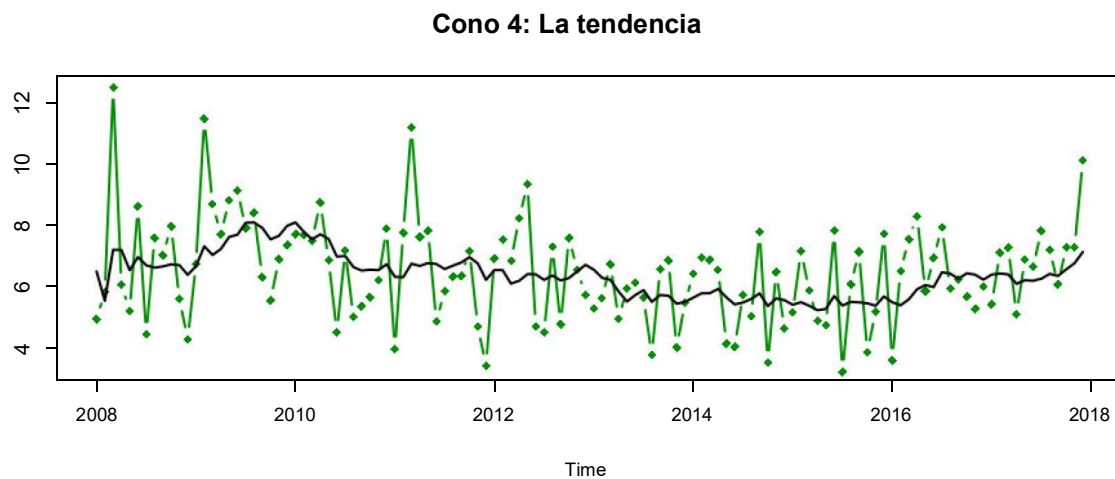
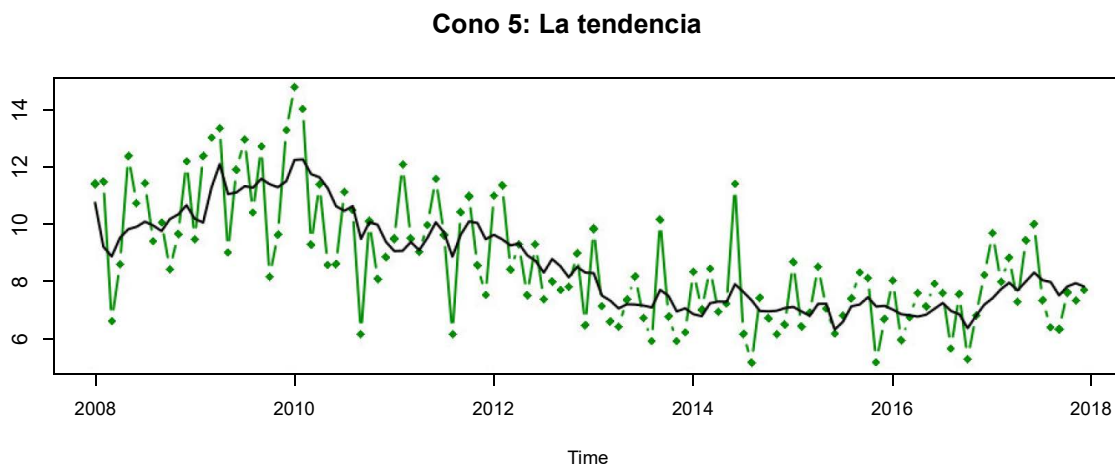
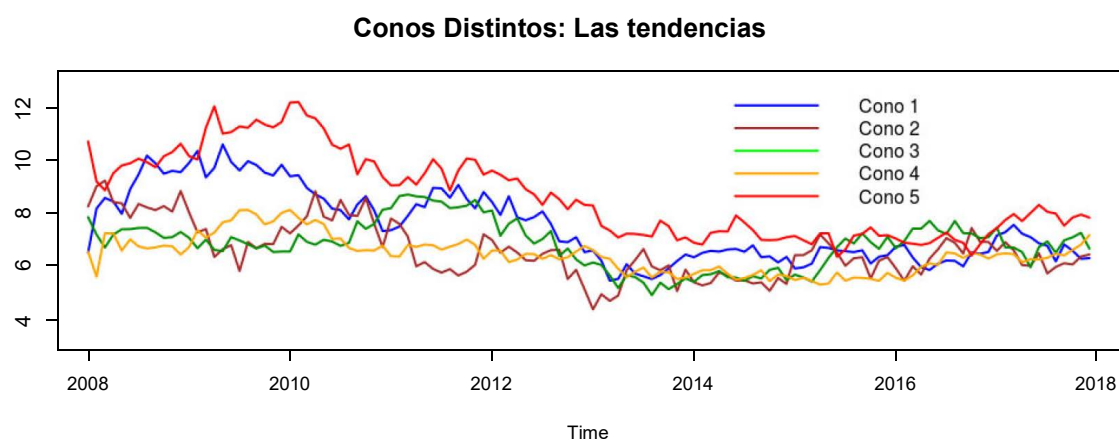


Figura 20. Los valores de la serie original de las tasas de desempleo vs. la curva de la tendencia estimada bajo el modelo BSM (los datos para el cono del Callao).



La Figura 21 demuestra comparación de las tendencias entre conos distintos. Se puede observar brechas en las tendencias de las tasas de desempleo en conos distintos en los años 2008-2012 que están disminuyendo a lo largo del período de investigación. Se observan tasas de desempleo mayores en el Callao, comparando con los otros conos. En el cono Lima Norte las tasas de desempleo son generalmente más altas que en Lima Centro, Lima Este y Lima Sur. Los conos Lima Centro y Lima Sur presentan niveles más estables de dichas tasas.

Figura 21. Las curvas de las tendencias para conos distintos.



Las Figuras 22-26 presentan los efectos de estacionalidad para los conos distintos para Lima Metropolitana y el Callao.

Figura 22. Las estimativas de los efectos de estacionalidad a lo largo del tiempo- Lima Norte.

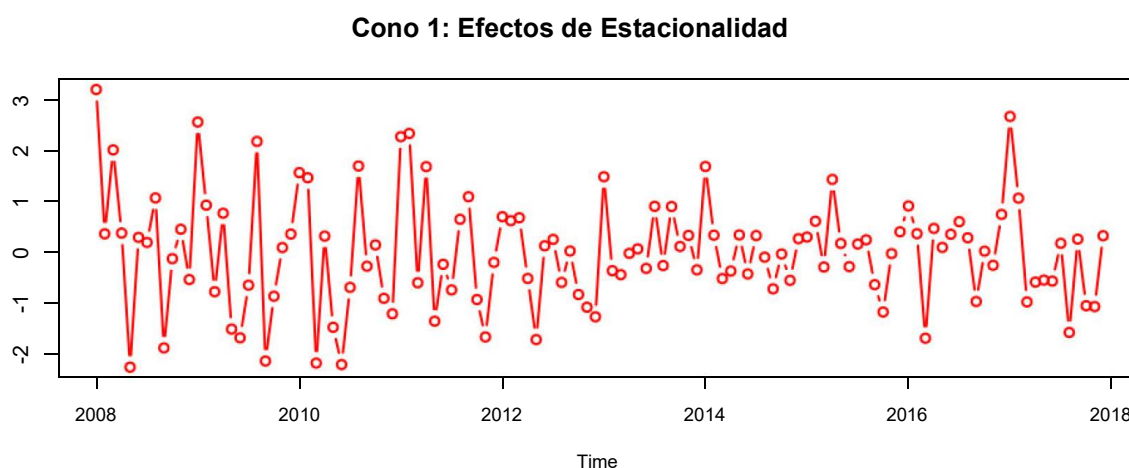


Figura 23. Las estimativas de los efectos de estacionalidad a lo largo del tiempo- Lima Centro.

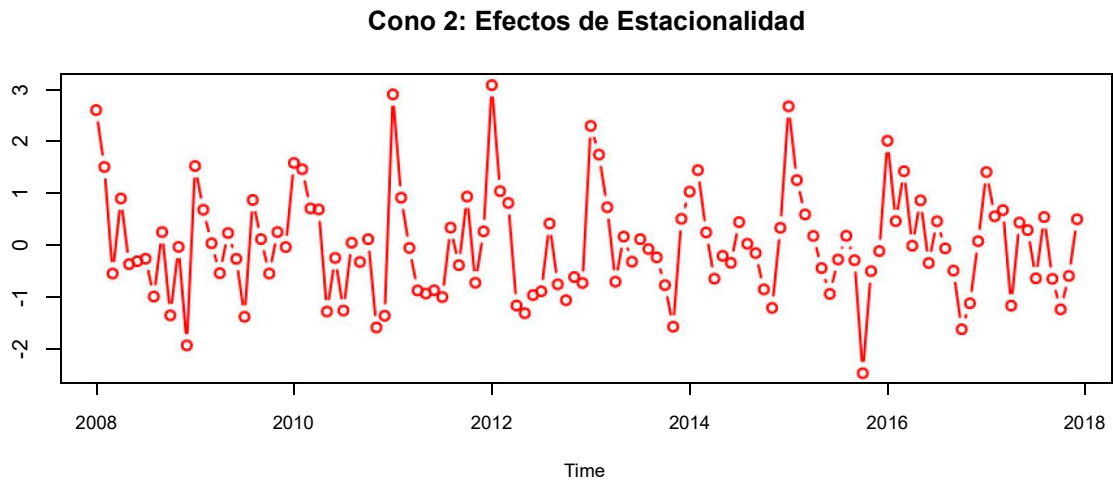


Figura 24. Las estimativas de los efectos de estacionalidad a lo largo del tiempo- Lima Oeste.

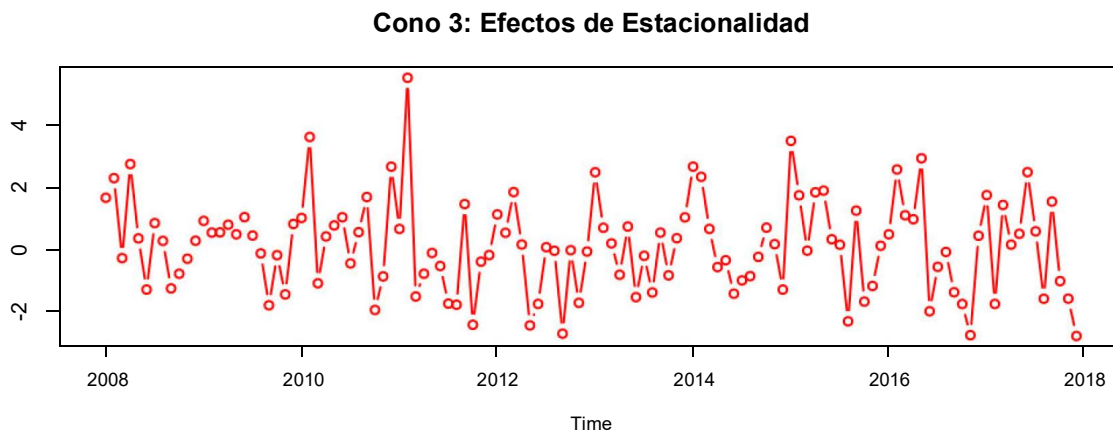


Figura 25. Las estimativas de los efectos de estacionalidad a lo largo del tiempo- Lima Sur.

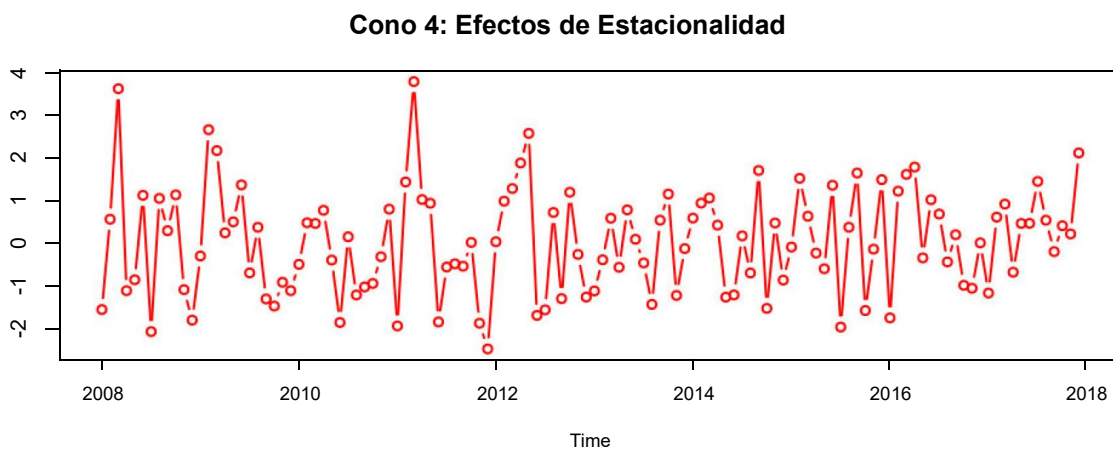
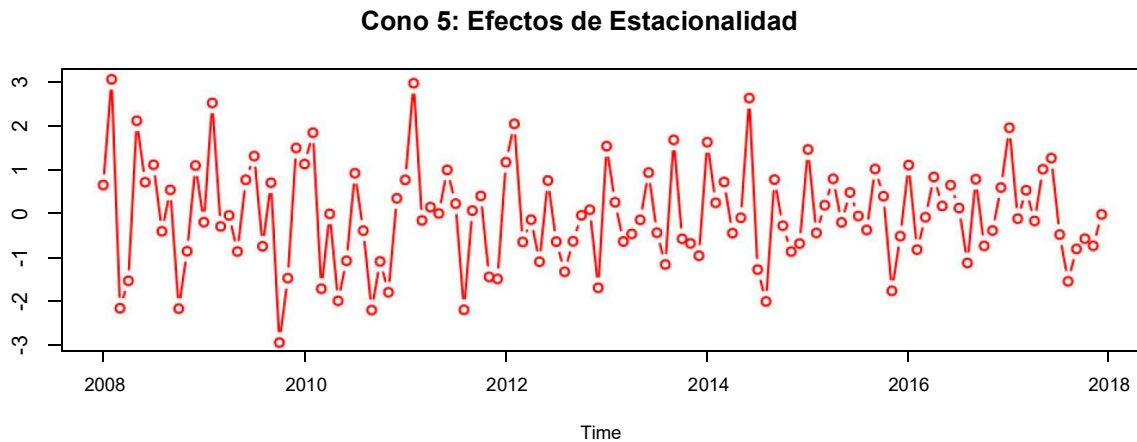


Figura 26. Las estimativas de los efectos de estacionalidad a lo largo del tiempo- El Callao.



Los gráficos demuestran que, en todos los conos, las series de las tasas de desempleo presentan efectos de estacionalidad, y que los patrones en todos los conos son parecidos, donde los efectos más altos se observan en enero y febrero y, más bajos, en octubre y noviembre. Esto es consistente con los resultados en las secciones anteriores.

CONCLUSIONES

- En el presente trabajo, ajustamos el modelo BSM a los datos de las tasas de desempleo, obtenidas de la encuesta de EPE. La aplicación del Filtro de Kalman permite la estimación de las componentes no observables de las series, y, consecuentemente, el análisis de las dinámicas de dichas tasas. Los resultados demuestran que hay brechas en las curvas de las tendencias de las dinámicas de las tasas de desempleo entre hombres y mujeres, y entre conos diferentes, que están disminuyendo a lo largo del periodo de investigación. Para investigar estas brechas, es necesario analizar más desagregados en grupos, definidos por el género, el cono y, posiblemente, la edad y el sector económico. Por otro lado, este análisis exigirá la aplicación de las técnicas del análisis en áreas pequeñas, puesto que puede ocurrir que en algunos grupos el número de observaciones será insuficiente para obtener estimativas confiables de las tasas de desempleo. Esto requiere más análisis, y puede ser una línea de investigación futura.
- Los resultados de investigación demuestran también que las series de las tasas de desempleo tienen efectos fuertes de estacionalidad, y que el patrón de estacionalidad es parecido en todos los grupos investigados.

AGRADECIMIENTOS

El presente trabajo fue financiado por el Instituto de Investigación de la Facultad de Ingeniería Económica y Estadística de la Universidad Nacional de Ingeniería (IECOS-UNI). El punto de vista presentado en este trabajo corresponde a los autores y no representa ninguna posición oficial de la institución a la que los autores están aliados.

La autora agradece a la asistente de investigación de este proyecto Loribeth Cristina Prudencio Rodríguez.

REFERENCIAS

- Arango, L. & Posada, C. (2013). El desempleo en Colombia. *Banco de la República*. Recuperado de <https://www.banrep.gov.co/docum/ftp/borra265pdf.pdf>
- Arango, L. & Ros, A. (2015). Duración del desempleo en Colombia: género, intensidad de búsqueda y anuncios de vacantes. Banco de la Reserva- Colombia. Borradores de economía. N866. Recuperado de <https://publications.iadb.org/es/publicacion/13851/duracion-del-desempleo-en-colombia-genero-intensidad-de-busqueda-y-anuncios-de>
- Ball, L., De Roux, N. & Hofstetter, M. (2011). *Unemployment in Latin America and the Caribbean*. International Monetary Fund. Recuperado de <https://www.elibrary.imf.org/view/journals/001/2011/252/article-A001-en.xml>
- Belapatiño, V., Céspedes, N. & Gutiérrez, A. (2014). La duración del desempleo en Lima Metropolitana. *Revista Estudios Económicos*. Banco Central de Reserva del Perú. Recuperado de <https://www.bcrp.gob.pe/docs/Publicaciones/Revista-Estudios-Economicos/27/ree-27-cespedes-gutierrez.pdf>
- Bellani, D., Garca, P. & Pasten E. (2002). *Curva de Beveridge, vacantes y desempleo: Chile 1996.I-2002.II*. Banco Central de Chile. Documentos de trabajo (191). Recuperado de https://www.bcentral.cl/documents/33528/133326/bcch_archivo_167954_es.pdf/e43da653-8889-7955-9896-959728f9f653?t=1573282942020
- Contreras, S., Pino O., & Pizzinga A.(2006). *Aplicación de la metodología espacio estado en el análisis de las series de desempleo: Caso región del Bío-Bío*. Theoría 25(1). Chile: Universidad del Bío Bío: Recuperado de <https://www.redalyc.org/pdf/299/29915107.pdf>
- Contreras-Reyes, J., & Idrovo, B. (2011). En busca de un modelo benchmark univariado para predecir la tasa de desempleo de Chile. *Cuadernos de Economía*, 30(55), 105-125. Retrieved May 10, 2021, Recuperado de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-47722011000200006&lng=en&tlng=es

- Durbin, J., & Quenneville, B. (1997). Benchmarking by State Space Models. *International Statistical Review / Revue Internationale De Statistique*, 65(1), 23-48. doi:10.2307/1403431.
- Harvey, A. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781107049994
- Judzik, D. & Mateos, N. (2017). *Sobre los determinantes de la tasa de desempleo en Argentina (2003-2015)*. Pontificia Universidad Católica Argentina, Facultad de Ciencias Humanas. Recuperado de <https://repositorio.uca.edu.ar/handle/123456789/9975>
- Lasso, F. (2013). La dinámica del desempleo urbano en Colombia. Recuperado de <https://www.banrep.gov.co/es/borrador-667>
- Lasso, F. & Zárate, H. (2019). *Pronosticar la tasa de desempleo colombiana utilizando los flujos de la fuerza laboral*. Banco de la Reserva-Colombia: Borradores de Economía. N1073. Recuperado de <http://doi.org/10.32468/be.1073>
- Marchionni, M., Gasparini, L., & Edo, M. (2019). Brechas de género en América Latina. Un estado de situación. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1401>
- Restrepo, J. (2008). *Estimaciones de la NAIRU para Chile*. Banco Central de Chile. Recuperado de <https://www.cemla.org/PDF/ic/2008-ic/IC-16.pdf>
- Centro de Investigación y Desarrollo (INEI/CIDE). (2001). *¿Qué sabemos sobre el desempleo en el Perú?* Recuperado de https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0489/Libro.pdf
- Organización Internacional del Trabajo (2013). *Modelo de Proyección de Empleo para el Perú*. Lima: OIT/ Oficina de la OIT para los países Andinos, 1a ed. Recuperado de https://www.ilo.org/wcmsp5/groups/public/---americas/---ro-lima/---sro-lima/documents/publication/wcms_236122.pdf

Los artículos publicados por IECOS pueden ser compartidos a través de la licencia Creative Commons: CC BY 4.0 Perú. Permisos lejos de este alcance pueden ser consultados a través del correo revistas@uni.edu.pe.

