

Regresión Avanzada Proyecto Final

Alejandra Lelo de Larrea Ibarra 00012433 Laura López Santibañez Jácome 000144088

Dante Ruiz Martínez 000183340

10 de diciembre del 2018

Resumen

Agregar un resumen al final.

Índice

| | |
|---|----------|
| 1. Introducción | 2 |
| 1.1. Problema | 2 |
| 1.2. Objetivo | 2 |
| 1.3. Hipótesis | 2 |
| 2. Datos | 2 |
| 2.1. Descripción de los datos | 2 |
| 2.2. Análisis exploratorio de los datos | 2 |
| 3. Estimación de Modelos | 6 |
| 3.1. Datos Originales | 6 |
| 3.1.1. Modelo de Regresión Clásico | 6 |
| 3.1.2. GLM Estático | 7 |
| 3.1.3. GLM Dinámico | 9 |

1. Introducción

1.1. Problema

1.2. Objetivo

1.3. Hipótesis

2. Datos

2.1. Descripción de los datos

Tabla 1: Resumen de los Datos

| Código | Variable | Unidades | Fuente | Ticker |
|------------------|---|------------------------------|-----------|----------------|
| WTI | West Texas Intermediate | Dólares por barril | Bloomberg | CL1 Comdty |
| JPM Dollar Index | JPM Dollar Index | Unidades | Bloomberg | FXJPEMCI Index |
| VIX | Chicago Board Options Exchange SPX Volatility Index | Unidades | Bloomberg | VIX Index |
| Prod. OPEP | Producción Total de Petróleo de la OPEP | Millones de barriles por día | Bloomberg | OPCRTOTL Index |
| Dem. Petróleo | Demanda Total de Petróleo de la OPEP | Millones de barriles por día | Bloomberg | OPCBRTOT Index |
| TBILL-10YR | Tasa de Largo Plazo de Estados Unidos | Por ciento | FRED | DGS10 |
| TBILL-1YR | Tasa de Corto Plazo de Estados Unidos | Por ciento | FRED | DGS1 |

Se tienen observaciones mensuales del West Texas Intermediate (WTI), del JPM Dollar Index, del Chicago Board Options Exchange SPX Volatility Index (VIX), de la producción y demanda de petróleo por parte de la OPEP, así como las tasas de corto y largo plazo en Estados Unidos. La tabla 1 resume las fuentes de información de los datos a utilizar en este trabajo. La muestra contiene 225 observaciones que corresponden al periodo de enero del 2000 a septiembre del 2018. Para estimar los modelos se utiliza el periodo comprendido entre enero del 2000 a junio del 2018; mientras que el horizonte de pronóstico va de julio a septiembre del 2018. La siguiente sección muestra un análisis exploratorio de los datos.

2.2. Análisis exploratorio de los datos

La variable WTI representa el precio de los contratos de futuros del petróleo crudo (para contratos de 1000 barriles). Como se puede apreciar en la figura 1, desde el inicio de la muestra y hasta antes de la crisis financiera del 2008, el precio del WTI presentaba una tendencia creciente. Posteriormente, durante la crisis presentó una caída significativa pasando de 140 dólares por barril a menos de 50 dólares por barril. Después de dicha crisis, el precio del WTI se recuperó alcanzando en algunos periodos precios superiores a los 100 dólares por barril. Sin embargo, en el primer trimestre del 2014 y hasta finales del 2015 (periodo de interés, marcado con la barra gris) volvió a presentar una caída drástica llegando a aproximadamente 36 dólares por barril. Dicha caída se ha revertido en los últimos, pero no parece regresar a los niveles previos al 2014 e incluso podría tratarse de un cambio estructural.

La figura 2 muestra la serie de tiempo de cada una de las variables explicativas que se van a utilizar en la estimación de los modelos; por su parte, la tabla 2 muestra algunas estadísticas descriptivas de las mismas. El JPM Dollar Index, figura 2 a) representa el índice del tipo de cambio del dólar americano respecto a las monedas de algunas economías emergentes. En este sentido, es una medida de la fortaleza del dólar. Cabe

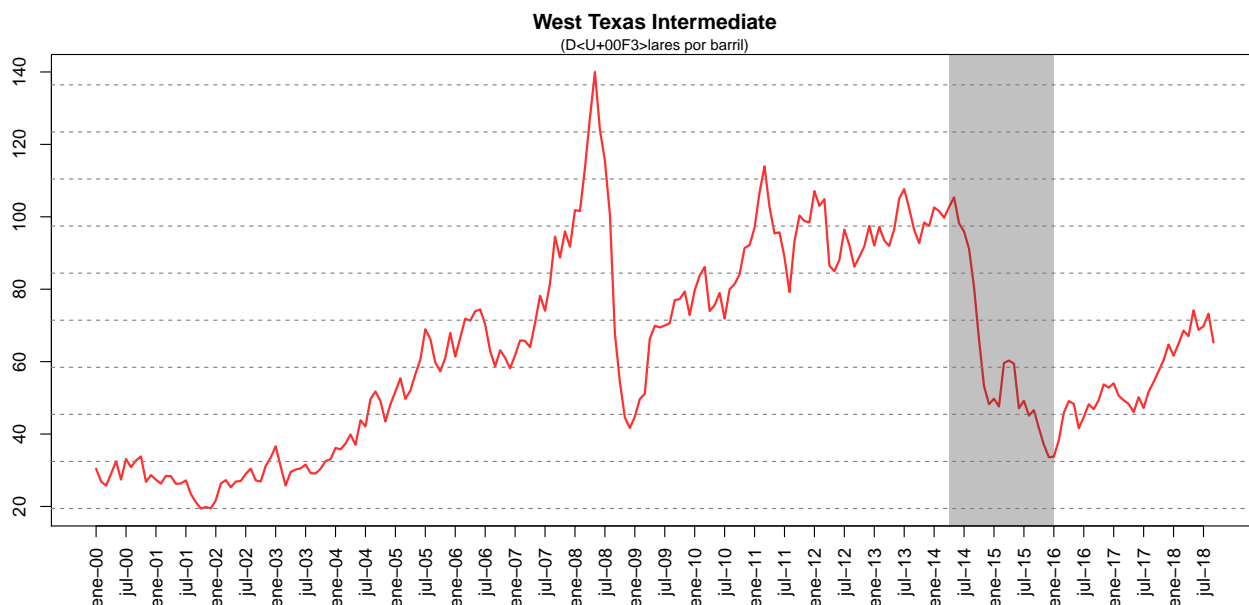


Figura 1: Serie de Tiempo del precio del West Texas Intermediate

Tabla 2: Estadísticas Descriptivas de las Variables de Estudio

| | WTI | JPM Dollar Ind. | VIX | Prod. OPEP | Dem. Petr<U+00F3>leo | TBILL-10YR | TBILL-1YR |
|------------|--------|-----------------|-------|------------|----------------------|------------|-----------|
| Mediana | 60.57 | 92.05 | 17.47 | 29.95 | 86.20 | 3.56 | 1.24 |
| Media | 62.69 | 89.76 | 19.67 | 29.87 | 87.41 | 3.51 | 1.86 |
| Moda | 101.58 | 107.07 | 13.29 | 28.08 | 84.70 | 2.30 | 0.12 |
| Varianza | 726.88 | 163.02 | 64.80 | 4.38 | 52.46 | 1.53 | 3.40 |
| Desv.Est. | 26.96 | 12.77 | 8.05 | 2.09 | 7.24 | 1.24 | 1.84 |
| Coef. Var. | 0.43 | 0.14 | 0.41 | 0.07 | 0.08 | 0.35 | 0.99 |
| Min | 19.44 | 60.91 | 9.51 | 24.10 | 73.80 | 1.50 | 0.10 |
| Max | 140.00 | 112.43 | 59.89 | 34.14 | 101.30 | 6.66 | 6.33 |
| Rango | 120.56 | 51.52 | 50.38 | 10.04 | 27.50 | 5.16 | 6.23 |

destacar que existen otros indicadores similares (como el DXY), pero que ponderan por el tipo de cambio de las economías avanzadas que, en su mayoría, no son productoras de petróleo por lo que el impacto en el precio del WTI sería menor. Desde 2002 y previo a la crisis del 2008, el dólar presentó una fortaleza creciente frente a las economías emergentes; mientras que durante la crisis del 2008 se debilitó de manera importante. A pesar de haber tenido una rápida recuperación en el periodo post-crisis, desde el 2011 el JPM Dollar Index presentó su mayor caída (en el periodo muestral) manteniéndose relativamente estable desde inicios del 2016, pero volviendo a debilitarse en este último año.

La figura 2 b) muestra la serie de tiempo del VIX, el cual refleja una estimación de la volatilidad futura del mercado basada en el promedio ponderado de las volatilidades implícitas para una amplia gama de precios de ejercicio. La serie del VIX presenta valores altos al inicio de la muestra y, como era de esperarse, presenta un gran pico en la crisis financiera del 2008 pues es el momento de mayor incertidumbre en los mercados a lo largo de periodo muestral. Cabe destacar que durante el periodo de interés la estimación de la volatilidad futura del mercado se mantuvo relativamente estable.

La producción total de petróleo por parte de los países miembros de la OPEP se puede encontrar en la figura 2 c). De manera general, se puede apreciar una tendencia creciente a lo largo de todo el periodo muestra, aunque con algunos episodios de baja en la producción del crudo. Es interesante destacar que, durante la crisis financiera del 2008, la producción de petróleo por parte de la OPEP tuvo una fuerte caída; mientras que en el periodo de interés, la producción de petróleo se incremento a un ritmo elevado, llegando a máximos históricos a finales del 2016. De hecho, en la literatura, el incremento en la oferta del petróleo se reconoce

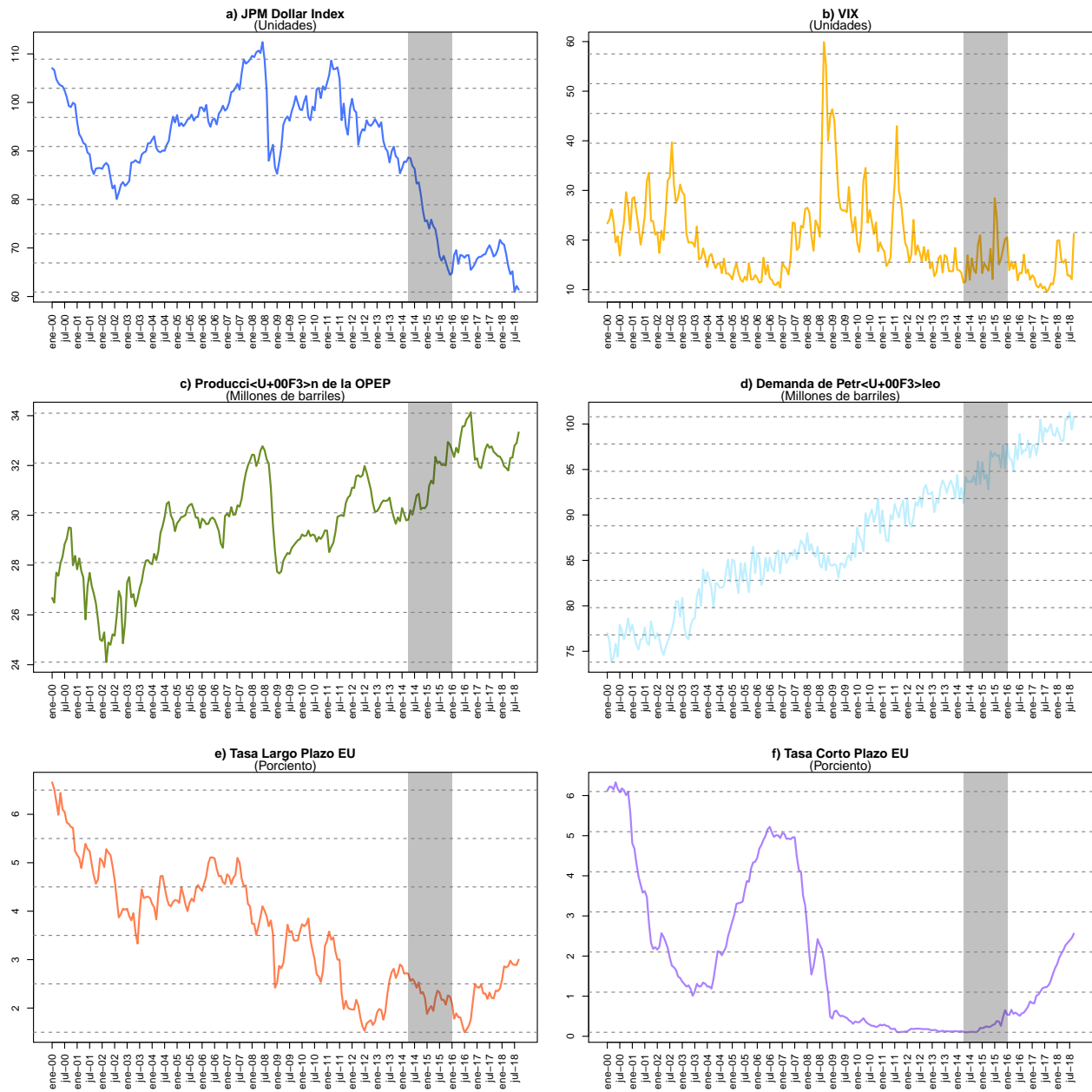


Figura 2: Serie de Tiempo de los Regresores

como una de las principales causas de la caída de los precios de crudo.

La figura 2 d) muestra la serie de tiempo de la demanda de petróleo tanto por parte de países miembros de la OPEP como de los no miembros de la misma. Análogo al caso de la producción de petróleo, la demanda de este commodity se ha incrementado a lo largo de toda la muestra. Además, esta serie parece tener un componente estacional, incrementándose la demanda por petróleo en los meses de invierno y disminuyendo durante el verano. Durante el periodo de interés destaca un incremento de la demanda en el segundo semestre del 2015; sin embargo este no es el incremento más pronunciado de toda la muestra. Posiblemente, esta variable ayude a capturar tendencia en la estimación de los modelos.

La gráfica de la serie de tiempo de la tasa de largo plazo de Estados Unidos (Bono del Tesoro a 10 años) se puede encontrar en la figura 2 e). En general, la tasa de largo plazo ha disminuido a lo largo de la muestra, representando que se esperaban tasas de corto plazo menores a las observadas. A diferencia del resto de los indicadores, la tasa de largo plazo no muestra un comportamiento marcadamente distinto durante la crisis del 2008 (existen otros periodos con caídas similares a las del 2008 e incluso mayores). Durante el periodo de interés, esta tasa no muestra un comportamiento extraordinario. Sin embargo, en los últimos dos años se ha presentado un incremento de la tasa de largo plazo; lo cual, está en línea con la expectativa de normalización de la política monetaria de Estados Unidos.

Por último, la figura 2 f) muestra la evolución del Bono del Tesoro a un año de Estados Unidos; es decir, de la tasa de corto plazo. Se puede observar que entre el año 2000 y 2004, la tasa de corto plazo disminuyó significativamente pasando de niveles superiores al 6 %, a niveles cercanos al 1 %. Entre el 2004 y mediados del 2006, se dio un fuerte incremento en esta tasa alcanzando niveles cercanos al 5 % y manteniéndose estable en este nivel por un par de años. Ante la crisis del 2008, la tasa de corto plazo comenzó con un ciclo de expansión de la política monetaria para fomentar el consumo y la inversión. Estados Unidos mantuvo su tasa de corto plazo cerca del cero desde el 2009 y hasta finales del 2015 con la finalidad de reactivar la economía. Finalmente, ante las señales de recuperación económica, y como se aprecia en la gráfica, a partir del 2016 se inicio un ciclo de alzas de tasas de interés para normalizar la política monetaria de dicho país, lo que se refleja en una tasa de corto plazo mayor hacia finales de la muestra.

Tabla 3: Matriz de correlaciones de las variables de estudio

| | WTI | JPM Dollar Ind. | VIX | Prod. OPEP | Dem. OPEP | TBILL-10YR | TBILL-1YR |
|-----------------|-------|-----------------|-------|------------|-----------|------------|-----------|
| WTI | 1.00 | 0.38 | -0.14 | 0.46 | 0.44 | -0.42 | -0.35 |
| JPM Dollar Ind. | 0.38 | 1.00 | 0.19 | -0.34 | -0.56 | 0.50 | 0.34 |
| VIX | -0.14 | 0.19 | 1.00 | -0.34 | -0.37 | 0.16 | -0.04 |
| Prod. OPEP | 0.46 | -0.34 | -0.34 | 1.00 | 0.82 | -0.62 | -0.23 |
| Dem. OPEP | 0.44 | -0.56 | -0.37 | 0.82 | 1.00 | -0.84 | -0.56 |
| TBILL-10YR | -0.42 | 0.50 | 0.16 | -0.62 | -0.84 | 1.00 | 0.84 |
| TBILL-1YR | -0.35 | 0.34 | -0.04 | -0.23 | -0.56 | 0.84 | 1.00 |

La tabla 3 muestra la matriz de correlaciones de todas las variables en la muestra. En particular estamos interesados en conocer la correlación de cada uno de los regresores con el precio del WTI. Ninguna variable tiene una correlación superior a 0.5 con la variable de interés, sin embargo varias de ellas podrían considerarse altas. El índice del dólar, el nivel de producción de la OPEP y la demanda del Petróleo tiene una correlación alta y positiva con el WTI (0.3823844, 0.4564846, y 0.4431912 respectivamente); por lo que esperaríamos que en los distintos modelos a estimar, estas variables fueran significativas y con coeficiente positivo. Tanto la tasa de largo como la de corto plazo tienen una correlación alta, pero negativa, con el precio del petróleo (-0.4230761 y -0.3541499), por lo tanto. Por último el VIX tiene una correlación negativa pero muy pequeña con el WTI -0.1392221 por lo que podría no ser una variable significativa a la hora de estimar los modelos.

3. Estimación de Modelos

Como se mencionó con anterioridad, uno de los objetivos de este proyecto es analizar la crisis de los precios del petróleo entre el 2014 y 2016. En esta sección se procede a estimar distintos modelos de regresión para analizar si las variables de la muestra tienen impacto en el precio del WTI y, en particular, conocer la forma en que cada una de las variables explicativas incidió en ésta durante el periodo de interés.

La estimación de los modelos se hace en tres bloques: utilizando los datos originales, estandarizando los datos y transformando los datos (secciones 3.1, ?? y ??). Para cada uno de estos bloques se estiman al menos un modelo de regresión lineal normal clásico (con enfoque frecuentista), un modelo lineal generalizado estático, un modelo dinámico y un modelo dinámico con suavizamiento. A lo largo de cada una de las secciones se hace referencia a las razones por las cuales se decide estimar uno u otro modelo, los valores DIC y pseudo- R^2 obtenidos, así como los problemas que se presentaron en cada paso. Por último, al final de la sección se realiza una comparación de los modelos. Cabe mencionar que se presenta la estimación del modelo de regresión lineal normal clásico únicamente con la finalidad de comparar los modelos bayesianos que hemos visto en clase; ya que, en muchas ocasiones, una regresión lineal frecuentista es la primera opción que viene al a mente al estimar modelos de regresión.

3.1. Datos Originales

En esta sección se utilizan los datos originales; es decir, sin ningún tipo de escalamiento o transformación para estimar los distintos modelos.

3.1.1. Modelo de Regresión Clásico

Como primer modelo, y con fines de comparación, se estima una regresión lineal múltiple (con enfoque frecuentista) ya que, en la mayoría de los casos, ésta sería la primer opción a utilizar. El modelo a estimar se define como

$$WTI_t = \beta_0 + \beta_1 JPM_t + \beta_2 VIX_t + \beta_3 ProdOPEP_t + \beta_4 DemPet_t + \beta_5 TBILL10_t + \beta_6 TBILL1_t + \epsilon_t \quad (1)$$

Tabla 4: Coeficientes Estimados Para el Modelo (1)

| | Estimaci | Desv. Est. | Valor t | Valor-p |
|-----------------|----------|------------|---------|---------|
| Intercepto | -394.79 | 30.20 | -13.07 | 0.00 |
| JPM Dollar Ind. | 1.94 | 0.08 | 24.74 | 0.00 |
| VIX Ind. | -0.02 | 0.11 | -0.18 | 0.86 |
| Prod. OPEP | 1.57 | 0.81 | 1.95 | 0.05 |
| Dem. OPEP | 2.77 | 0.31 | 8.79 | 0.00 |
| TBILL-10YR | 0.61 | 2.14 | 0.29 | 0.78 |
| TBILL-1YR | -3.65 | 1.06 | -3.43 | 0.00 |

Como se puede apreciar en la tabla 4, todos los regresores son significativos al 95 % de confianza salvo el VIX y la tasa de largo plazo. En cuanto a los regresores que sí son significativos, el signo de los coeficientes coincide con lo esperado a partir de la matriz de correlaciones. En este sentido, ante un incremento de una unidad en el JPM Dollar Index, el precio del WTI se incrementaría en 1.94 dólares por barril; un incremento de un millón de barriles producidos llevaría a un incremento del precio del WTI de 1.57 dólares por barril, un incremento de un millón de barriles en la demanda del petróleo estaría asociado a un incremento en el precio del WTI de 2.77 dólares por barril y un incremento de un punto porcentual en la tasa de tasa de corto plazo disminuiría el precio del WTI en -3.65 dólares por barril. Este modelo tiene un coeficiente de $R^2 = 0.81$; es decir, el modelo explica un 81.3 % de la varianza de los datos. Este valor servirá para comparar con el valor de la pseudo R^2 de los modelos bayesianos.

3.1.2. GLM Estático

El segundo modelo a estimar es un modelo lineal generalizado normal estático; de hecho, este modelo coincide con un Modelo de Regresión Normal Lineal. El modelo está definido como:

$$\begin{aligned} WTI_t &\sim N(\mu_t, \tau_i) \\ \mu_t &= \beta_0 + \beta_1 JPM_t + \beta_2 VIX_t + \beta_3 ProdOPEP_t + \beta_4 DemPet_t + \beta_5 TBILL10_t + \beta_6 TBILL1_t \\ \tau_i &= \tau \end{aligned} \tag{2}$$

Como distribución inicial de los coeficientes β_j se utiliza una normal no informativa; esto es $\beta_j \sim N(0, 0.001)$ con $j = 0, 1, \dots, 6$. La estimación del modelo se realizó utilizando JAGS en R. Se corrieron 20000 simulaciones para 2 cadenas con un periodo de calentamiento de 2000 observaciones y sin adelgazamiento de la cadena ($n.thin=1$). En cuanto al ajuste del modelo, éste tuvo un valor $DIC = 1807.701$ y una pseudo- $R^2 = 0.79$, por lo que el modelo está capturando aproximadamente el 79.2% de la varianza de los datos.

Tabla 5: Coeficientes Estimados para el Modelo (2)

| | Media | Mediana | Moda | 2.5 % | 97.5 % | Prob. |
|-----------------|---------|---------|---------|---------|---------|-------|
| Intercepto | -182.94 | -183.24 | -185.96 | -231.03 | -133.32 | 0.00 |
| JPM Dollar Ind. | 1.84 | 1.84 | 1.81 | 1.67 | 2.00 | 0.00 |
| VIX Ind | -0.29 | -0.29 | -0.19 | -0.52 | -0.06 | 0.01 |
| Prod. OPEP | 0.52 | 0.53 | 0.10 | -1.18 | 2.20 | 0.27 |
| Dem. OPEP | 1.26 | 1.26 | 1.39 | 0.63 | 1.87 | 0.00 |
| T-Bill 10YR | -10.76 | -10.75 | -9.10 | -14.77 | -6.89 | 0.00 |
| T-Bill 1YR | -0.82 | -0.83 | -1.28 | -2.93 | 1.31 | 0.23 |

La tabla 5 muestra los coeficientes estimados para este modelo. Como estimadores puntuales se tienen la media (bajo pérdida cuadrática), la mediana (bajo pérdida absoluta) y la moda (bajo pérdida vecindad); además se incluye la estimación por intervalos. Como se puede apreciar, los estimadores puntuales de la media y la mediana son muy parecidos para todos los coeficientes; mientras que el estimador puntual de la moda es el que más difiere en todos los casos. Todos los regresores son significativos — y sus respectivos intervalos de probabilidad no contienen al cero — salvo los correspondientes al nivel de producción de la OPEP y a la tasa de corto plazo. Estos resultados contrastan con los del modelo de regresión lineal clásico.

De esta manera, para los coeficientes significativos se tiene que un incremento de una unidad en el índice del dólar (todo lo demás constante) estará relacionado con un incremento de 1.84 dólares por barril en el precio del WTI; un incremento de una unidad en el índice de volatilidad conlleva a una disminución de 0.29 dólares por barril en el precio del petróleo, un incremento de un millón de barriles en la demanda del petróleo incrementará su precio en 1.26 dólares por barril y un incremento de un punto porcentual en la tasa de largo plazo estará relacionado con una disminución de 10.76 dólares por barril en el precio del WTI.

La figura 3 muestra el valor observado, el valor pronosticado y los intervalos de predicción para cada uno de los regresores vs el precio del WTI. Se puede notar que la variabilidad en todos los regresores es muy alta y que muchas observaciones quedan fuera de las bandas de predicción.

Por último, la figura 4 muestra la serie de tiempo observada para el precio del WTI, los valores pronosticados y el pronóstico para el tercer trimestre del 2018 junto con sus intervalos de predicción. El modelo captura de forma correcta la evolución del precio del WTI en la mayoría de los periodos; sin embargo, en algunos momentos la predicción sobrestima o subestima el verdadero valor como es el caso de la crisis del 2008 y, sobre todo, del periodo de interés. Si bien el pronóstico fuera de muestra tiene una tendencia que coincide con lo observado, el valor observado se sale de la banda de predicción.

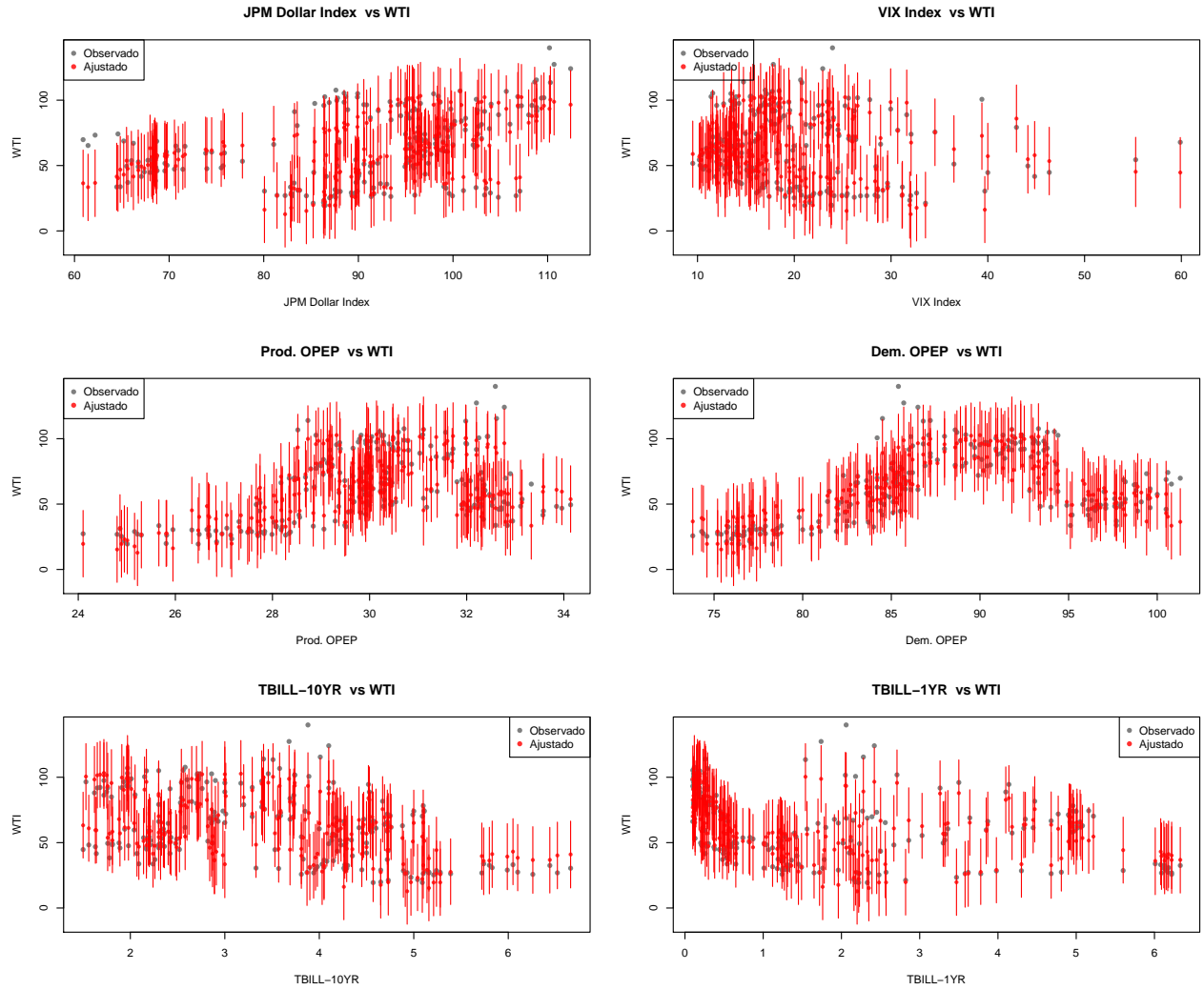


Figura 3: Regresores vs WTI: Modelo (2)

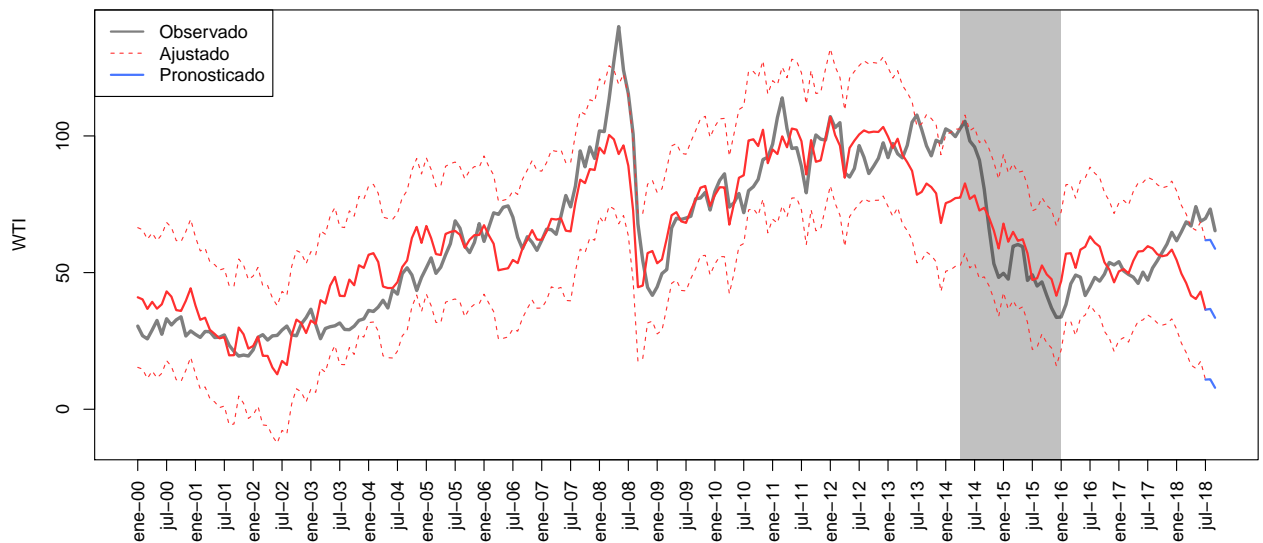


Figura 4: Ajuste y Prediccin: Modelo (2)

3.1.3. GLM Dinámico

Dado que el modelo estático parece no replicar correctamente la dinámica del WTI en algunos momentos del tiempo, como tercer modelo se estima un modelo lineal generalizado normal dinámico definido como:

$$\begin{aligned} WTI_t &= \beta_{0,t} + \beta_{1,t}JPM_t + \beta_{2,t}VIX_t + \beta_{3,t}ProdOPEP_t + \beta_{4,t}DemPet_t + \beta_{5,t}TBILL10_t + \beta_{6,t}TBILL1_t + \epsilon_t \\ \beta_{j,t} &= \beta_{j,t-1} + \omega_t \end{aligned} \quad (3)$$

donde $\epsilon_t \sim N(0, V^{-1})$ y $\omega_t \sim N(0, W^{-1})$. Como distribución inicial de los coeficientes β_j se utiliza una normal no informativa; esto es $\beta_{j,0} \sim N(0, 0.001)$ con $j = 0, 1, \dots, 6$. Por su parte, como distribución inicial para ϵ_0 y ω_0 se utilizan distribuciones gamma no informativas; es decir $\epsilon_0 \sim \text{gamma}(0.1, 0.1)$ y $\omega_0 \sim \text{gamma}(0.1, 0.1)$. La estimación del modelo se realizó utilizando JAGS en R. Se corrieron 20000 simulaciones para 2 cadenas con un periodo de calentamiento de 2000 observaciones y sin adelgazamiento de la cadena (`n.thin=1`). En cuanto al ajuste del modelo, éste tuvo un valor $DIC = 22587.06$ y una pseudo- $R^2 = 1$, por lo que el modelo está capturando aproximadamente el 100 % de la varianza de los datos.

La figura 5 muestra el valor observado, el valor pronosticado y los intervalos de predicción para cada uno de los regresores vs el precio del WTI. Se puede notar que la variabilidad en todos los regresores para la mayoría de las observaciones ahora es mínima y que prácticamente todas las predicciones coinciden con los valores observados. Esto podría estar hablando de problemas de sobre ajuste en el modelo o de problemas numéricos en JAGS.

La figura 6 muestra la serie de tiempo observada para el precio del WTI, los valores pronosticados y el pronóstico para el tercer trimestre del 2018 junto con sus intervalos de predicción. Efectivamente, el modelo está sobreajustando a los datos. Este replica de manera exacta la evolución del precio del WTI a lo largo de la muestra y el ancho del intervalo de predicción es mínimo. En contraste, la predicción para el tercer trimestre del 2018 es muy mala y los intervalos de predicción muy anchos ya que el modelo no es capaz de generalizar los patrones importantes.

Por último, la figura ?? muestra la serie de tiempo de los coeficientes estimados en el modelo. Cabe destacar que en este modelos, ninguno de los coeficientes salen significativos en ningún momento del tiempo. Esto se puede apreciar en el hecho de que la serie de tiempo de la mayoría de los coeficientes es plana, salvo con unos pequeños picos en la crisis del 2008 para los primeros cuatro regresores.

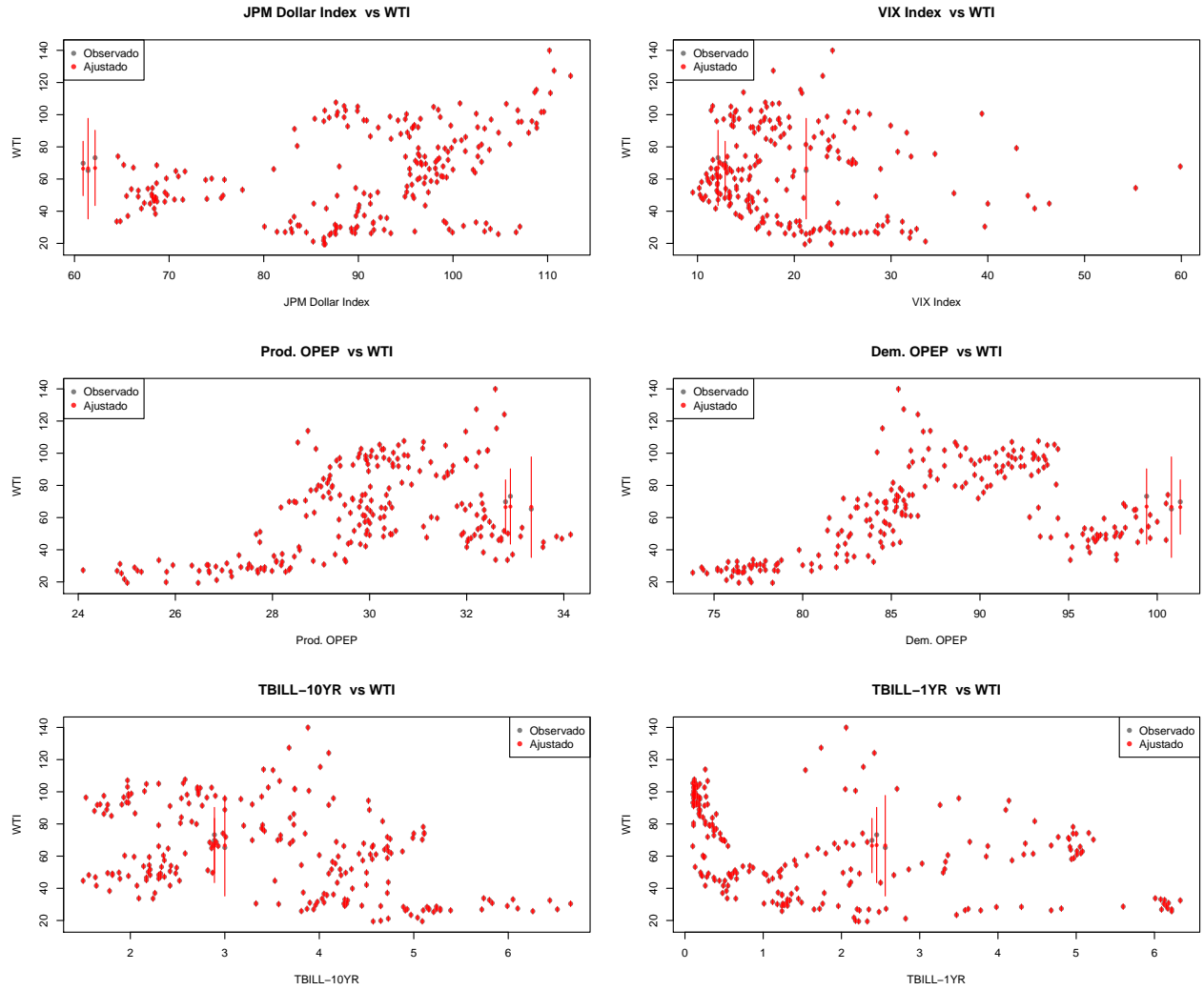


Figura 5: Regresores vs WTI: Modelo (3)

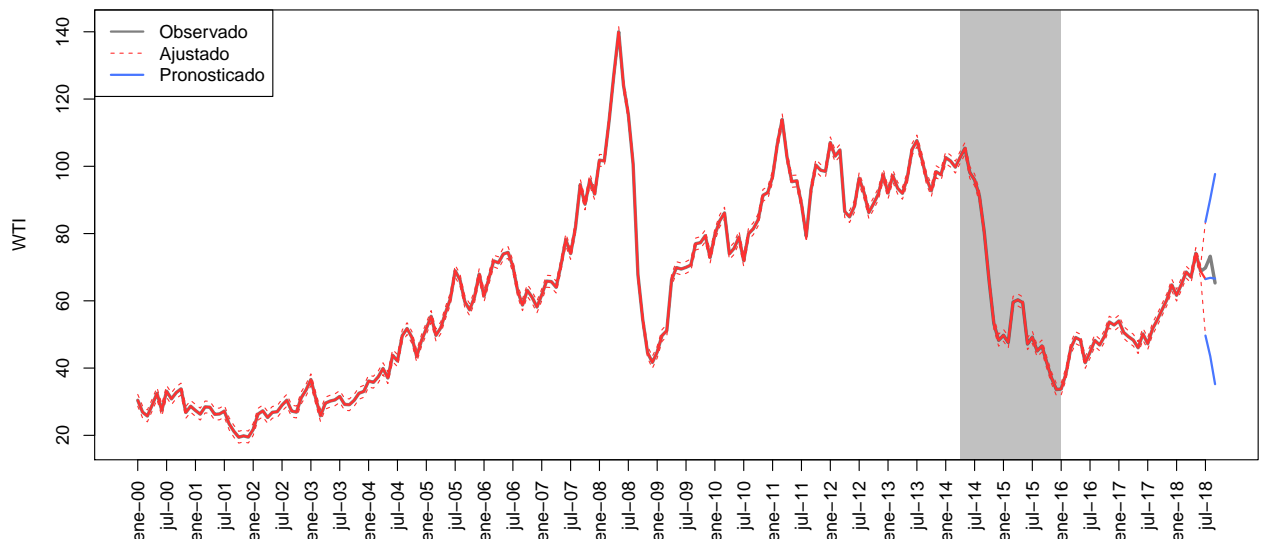


Figura 6: Ajuste y Prediccin: Modelo (3)

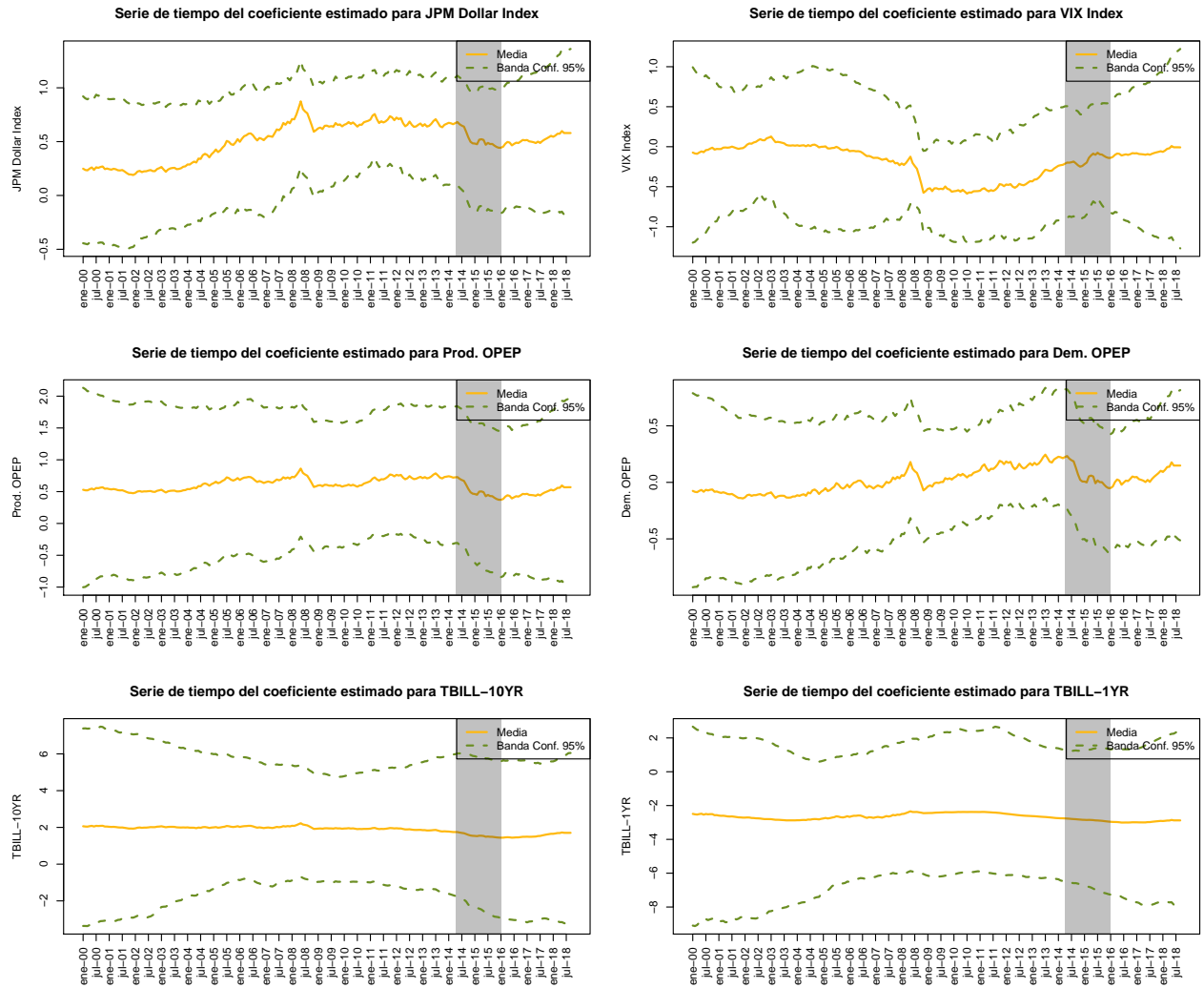


Figura 7: Coeficientes Estimados: Modelo (3)