

Predicción de una medida de riesgo de impago por segmentos de crédito

Miguel Cabello, Camila Rodriguez y Diego Yamunaque

Banco Central de Reserva del Perú

18 de octubre de 2024

- Es necesario monitorear el incumplimiento (riesgo de crédito). El BCRP lo hace por la supervisión macroprudencial.
- Por conveniencia, el indicador debe reflejar una transición del impago al pago.
- Los modelos a usar deben incluir series de tiempo y permitir el uso de escenarios para predecir el riesgo.
- La estimación no será micro. Se procederá a agregar los datos.

Definiciones

Sea S_{it} el saldo de crédito del cliente i en el periodo t . $y_{it}^d = 1$ si saldo en t tiene atraso de más de d días o si presenta créditos castigados, 0 en otro caso.

- 1 Probabilidad (incondicional) de impago:

$$\alpha_{it} = \mathbb{P}(y_{it}^{90} = 1)$$

- 2 Probabilidad de transición hacia impago:

$$\pi_{it} = \mathbb{P}(y_{it}^{90} = 1 | y_{it-12}^0 = 0)$$

Heterogénea por individuos y periodos. No es observable.

Definiciones

- Probabilidad de transición hacia impago promedio (agregada):

$$\pi_t = \mathbb{E}_t[\pi_{it}] = \int_{[0,1]} \pi_{it} f_t(\pi_{it}) d\pi_{it}$$

Estimador a la probabilidad de transición

- La probabilidad de transición promedio π_t se aproxima como:

$$\hat{\pi}_t = \sum_{i=1}^N \hat{\pi}_{it} \omega_{it},$$

con ω_{it} un peso específico para la unidad i en el periodo t , y $\hat{\pi}_{it}$ una medida de aproximación de la probabilidad de transición individual.

Estimadores a la probabilidad de transición: Alternativas

- Opción de conteo:

$$\hat{\pi}_t = \sum_{i=1}^N 1(y_{it}^{90} = 1, y_{it-12}^0 = 0) \times \frac{1(y_{it-12}^0 = 0)}{\sum_{i=1}^N 1(y_{it-12}^0 = 0)}.$$

- Ponderador por tamaño relativo de cartera *actual*:

$$\hat{\pi}_t = \sum_{i=1}^N 1(y_{it}^{90} = 1, y_{it-12}^0 = 0) \times \frac{S_{it} 1(y_{it-12}^0 = 0)}{\sum_{i=1}^N S_{it} 1(y_{it-12}^0 = 0)}$$

- Ponderador por tamaño relativo de cartera *inicial*:

$$\hat{\pi}_t = \sum_{i=1}^N 1(y_{it}^{90} = 1, y_{it-12}^0 = 0) \times \frac{S_{it-12} 1(y_{it-12}^0 = 0)}{\sum_{i=1}^N S_{it-12} 1(y_{it-12}^0 = 0)}$$

- El término $\frac{S_{it}1(y_{it-12}^0=0)}{\sum_{i=1}^N S_{it}1(y_{it-12}^0=0)}$ no es **necesariamente** un estimador **consistente** de la densidad condicional.
- Por conveniencia se emplea el estimador con ponderador relativo a cartera.
- Se puede interpretar como proporción de cartera que corre riesgo de deteriorarse.
- El ratio de Morosidad = $\sum_{i=1}^N 1(y_{it}^{90} = 1) \cdot \frac{S_{it}}{\sum_{i=1}^N S_{it}}$ sólo muestra proporción de cartera deteriorada a día de hoy. No captura transición.

- **Objetivo: Predecir** fuera de muestra la medida de riesgo de impago o deterioro de la cartera de crédito por segmentos de crédito.
- Nuestra medida se interpreta como: **proporción de la cartera agregada (en $t - 12$) de un segmento de crédito que pasa de una situación sin atraso en $t - 12$ a una condición de impago de 90 días o más en t , o presenta créditos castigados en t .**
- SBS lo llama **ratio de incumplimiento** y la emplea como una aproximación de la probabilidad de impago o default.
- Se le puede llamar *probabilidad de transición* (hacia el impago).
- Filtro adicional: solo se considera el impago cuando este represente más del 20% de la cartera del individuo (firma) respecto a su deuda en la entidad financiera.

Transición al impago: Ejemplo

Supuesto: saldo en $t - 12$ no se encuentra en default.

Cuadro: Cálculo de Transición al impago

UA	S_{t-12}	S_t	S. atrasado	y (p. t-12)	y (p. t)	S. riesgo (p. t-12)	S. riesgo (p. t)
1	100	50	0	0	0	0	0
2	100	60	30	1	1	100	60
3	50	80	15	1	0	50	0
4	0	40	0	—	—	—	—
5	100	5	5	0	1	0	5
6	50	0	0	0	0	0	0
Total	400	235	50	2	2	150	65

Transición al impago: Ejemplo - II

- Ratio de morosidad: $\frac{50}{235} = 21,3\%$
- Medida de conteo (incluyendo missing final): $\frac{2}{5} = 40\%$
- Medida de conteo (no missings): $\frac{2}{4} = 50\%$
- Medida ponderada por la cartera en t : $\frac{65}{235} = 27,7\%$
- Medida ponderada por la cartera en $t - 12$ (incluyendo missing final): $\frac{150}{400} = 37,5\%$

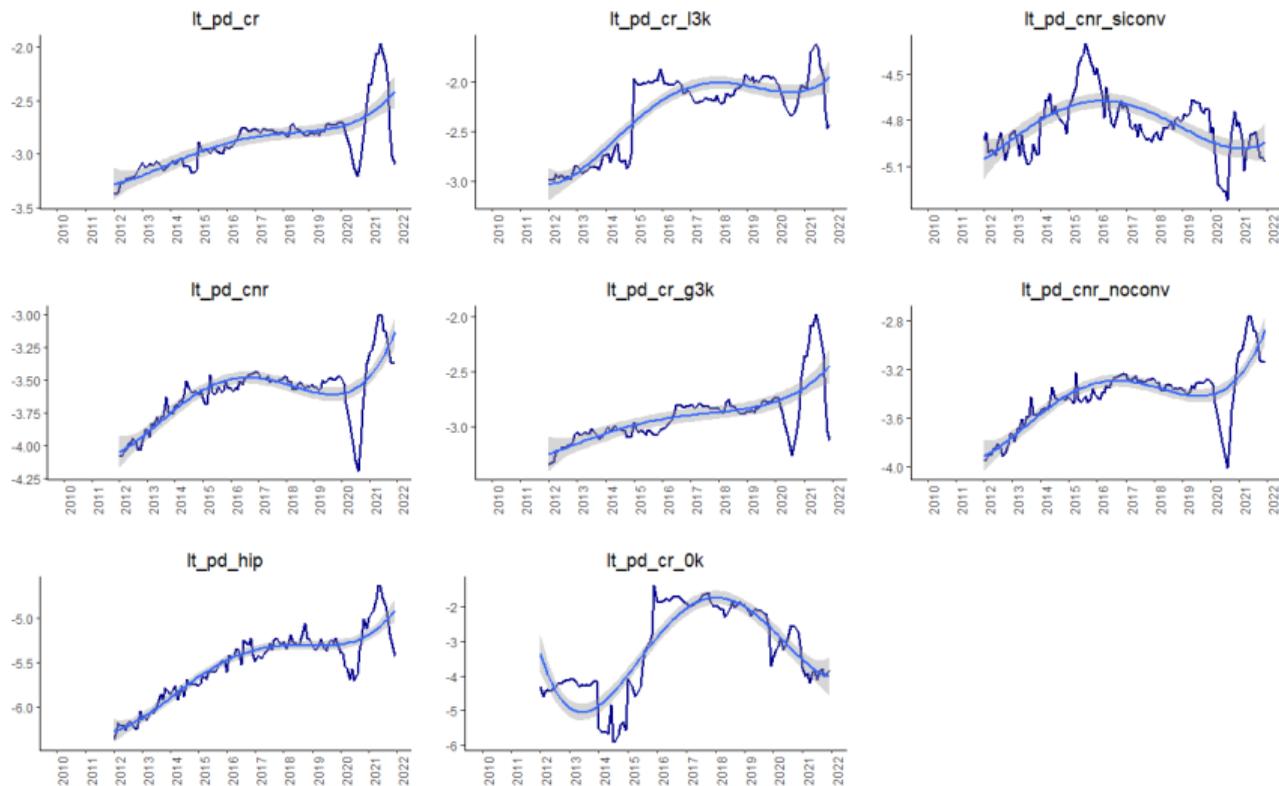
- Muestra total: 2012m01 - 2023m06
 - Estimación: 2012m01 - 2021m12
 - Ejercicios de Predicción: 2022m01 - 2023m06
 - Predicción fuera de muestra: 2023m07 - 2023m12
- Modelos de predicción:
 - Univariados: ARIMA
 - Mutlivariado: VAR en niveles y primeras diferencias // regularizados.

- Medida de transición está en $[0, 1]$. Para evitar predicciones fuera de rango, transformación monotónica (Transformación logística):

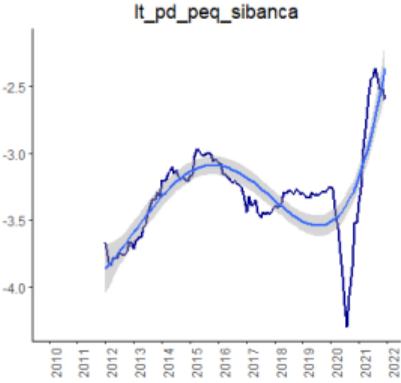
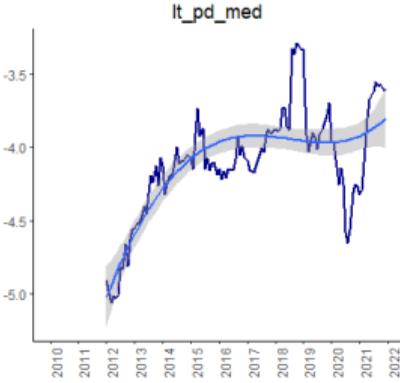
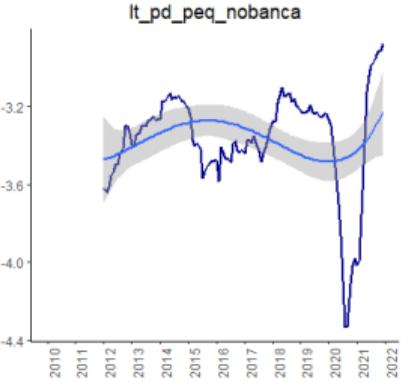
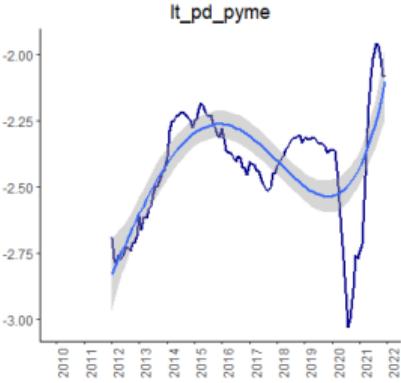
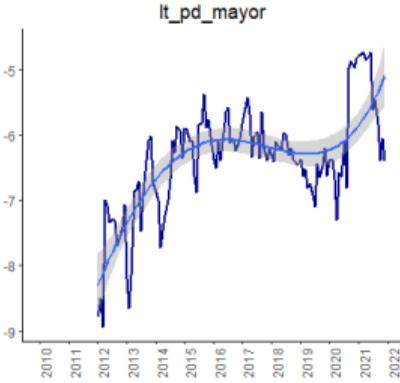
$$y_{it} = \log \left(\frac{RI_{it}}{1 - RI_{it}} \right)$$

- Ajustamos modelo lineal para y_{it}
- Predicciones se construyen así: $\hat{T}I_{t+h|t} = \frac{\exp(\hat{y}_{t+h|t})}{1 + \exp(\hat{y}_{t+h|t})}$
- Segmentos
 - Familias: Consumo Revolvente y No Revolvente e Hipotecario.
 - Empresas: Mayorista (corporativo y grandes), Medianas y Pyme (Pequeña y Microempresa).

Evolución de la Transición hacia Impago - Segmento Familias



Evolución de la Transición hacia Impago - Segmento Empresas



- Se consideraron 4 tipos de modelos ARIMA
 - $ARIMA(p, d, 0)$ (Modelo 1)
 - $ARIMA(p, d, 0)$ con tendencia lineal (Modelo 2)
 - $ARIMA(p, d, 0)$ con tendencia cuadrática (Modelo 3)
 - $ARIMA(p, d, 0)$ con tendencia cúbica (Modelo 4)

Modelos Multivariados: VAR

- Variables extras:
 - PIB y Consumo reales;
 - Ingreso Real Promedio y Tasa de desempleo;
 - Expectativas: economía a 12 meses, sector a 3 meses, demanda a 1 mes;
 - Precios de cobre, petróleo;
 - Tipo de cambio, EMBI Spread y colocaciones.
- Modelo $VAR(p)$ en niveles (si hay cointegración):

$$\mathbf{X}_t = \mathbf{C}\mathbf{D}_t + \sum_{k=1}^p \Phi_k \mathbf{X}_{t-k} + \mathbf{u}_t$$

- Modelo $VAR(p)$ en diferencias:

$$\Delta \mathbf{X}_t = \mathbf{C} + \sum_{k=1}^p \Phi_k \Delta \mathbf{X}_{t-k} + \mathbf{u}_t$$

Modelos VAR regularizados

- Siguiendo a [Kock y Callot \(2012\)](#), variable i del vector x_t en modelo VAR se puede expresar como:

$$x_{t,i} = X_t' \phi_i^* + u_{t,i}, \quad i = 1, 2, \dots, n;$$

con $X_t = (x_{t-1}' \cdots x_{t-p}')'$ y ϕ_i^* es un vector $np \times 1$.

- Con una muestra de tamaño T , el sistema anterior es

$$x_i = X \phi_i^* + u_i, \quad i = 1, 2, \dots, n;$$

con $x_i = (x_{T,i} \ x_{T-1,i} \ \cdots \ x_{p+1,i})$ y $X = (X_T \ X_{T-1} \ \cdots \ X_{p+1})'$.

- Matricialmente, el sistema anterior toma la forma:

$$x = Z \phi^* + u,$$

donde $x = [x_1' \ x_2' \ \cdots \ x_n']'$, $Z = I_n \otimes X$ y $\phi^* = (\phi_1^{*'} \ \phi_2^{*'} \ \cdots \ \phi_n^{*'})'$.

- Estimador LASSO, $\hat{\phi}_{LASSO}$:

$$\mathcal{L}_T(\beta) = \|x - Z\beta\|^2 + \zeta_T \sum_{i=1}^{n^2 p} |\beta_i|$$

Combinación de Predicciones

- Sea $f_{t,i}^h$ predicción h -pasos adelante del modelo i para variable y_t . $\mathbf{f}_t = (f_{t,1}^h \quad f_{t,2}^h \quad \cdots \quad f_{t,M}^h)$
- Se construye $\hat{y}_{t,COMB}^h = \mathbf{w}'\mathbf{f}_t$, con \mathbf{w} un vector de pesos.
- Opciones evaluadas:
 - $w_i = \frac{1}{M}$.
 - \mathbf{w} que minimice MSE de predicción.
 - $w_i = \frac{MSE_i^{-1}}{\sum_{i=1}^M MMSE_i^{-1}}$.

RIs de Familias: Consumo revolvente, no revolvente e hipotecario

Cuadro: Variables endógenas en modelo VAR

	Niveles	Diferencias
1	$y_{Revolvente}$	$\Delta y_{Revolvente}$
2	$y_{noRevolvente}$	$\Delta y_{noRevolvente}$
3	$y_{Hipotecario}$	$\Delta y_{Hipotecario}$
4		$\Delta \log(PIB)$
5		$\Delta \log(Consumo)$
6		$\Delta \log(IngProm)$
7		$Desempleo$
8		$ExpEcon12M$
9		$\Delta Colocaciones_{Consumo}$
10		$\Delta Colocaciones_{Hipotecario}$

Cuadro: Variables endógenas en modelo VAR

Niveles	Diferencias
1	$\Delta y_{Mayorista}$
2	$\Delta \log(PIB)$
3	$ExpEcon12M$
4	$ExpSector3M$
5	$ExpDemanda1M$
6	$\Delta \log(P_{Cobre})$
7	$\Delta \log(P_{Oil})$
8	$\Delta \log(TCambio)$
9	$EMBI$
10	$\Delta Colocaciones_{Mayorista}$

Cuadro: Variables endógenas en modelo VAR

Niveles	Diferencias
1	$\Delta y_{Mediana}$
2	Δy_{PYME}
3	$\Delta_{12M} \log(PIB)$
4	$ExpEcon12M$
5	$ExpSector3M$
6	$ExpDemanda1M$
7	$Desempleo$
8	$\Delta \log(P_{Cobre})$
9	$\Delta \log(P_{Oil})$
10	$\Delta \log(TCambio)$
11	$EMBI$
12	$\Delta Colocaciones_{Medianas}$
13	$\Delta Colocaciones_{PYME}$

Medidas de Predicción: Consumo Revolvente

Cuadro: Evaluación de modelos de Predicción - Consumo Revolvente

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,147	0,112	0,921	0,874
Modelo 2	0,510	0,477	5,765	5,673
Modelo 3	0,266	0,228	0,954	0,851
Modelo 4	0,556	0,522	6,180	6,077
<i>VAR Cointegrado</i>				
Intercepto	0,207	0,174	0,416	0,338
Tendencia	0,204	0,172	0,368	0,299
<i>Multivariados en diferencias</i>				
VAR-single	0,285	0,254	1,921	1,769
Lasso VAR-single	0,239	0,206	1,606	1,496
VAR-joint	0,587	0,475	5,637	3,239
Lasso VAR-joint	0,174	0,141	1,029	0,972
<i>BVAR en niveles</i>				
VAR Bayesiano	0,215	0,157	0,673	0,572

Medidas de Predicción: Consumo No Revolvente

Cuadro: Evaluación de modelos de Predicción - Consumo No Revolvente

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,092	0,076	0,423	0,398
Modelo 2	0,094	0,067	0,378	0,311
Modelo 3	0,116	0,101	0,385	0,359
Modelo 4	0,250	0,241	1,679	1,652
<i>VAR Cointegrado</i>				
Intercepto	0,086	0,072	0,268	0,221
Tendencia	0,204	0,181	0,512	0,481
<i>Multivariados en diferencias</i>				
VAR-single	0,215	0,186	0,834	0,806
Lasso VAR-single	0,126	0,111	0,568	0,543
VAR-joint	0,405	0,302	18,314	8,193
Lasso VAR-joint	0,279	0,215	1,068	1,012
<i>BVAR en niveles</i>				
VAR Bayesiano	0,118	0,103	0,478	0,450

Medidas de Predicción: Hipotecario

Cuadro: Evaluación de modelos de Predicción - Hipotecario

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,027	0,022	0,042	0,033
Modelo 2	0,035	0,029	0,129	0,125
Modelo 3	0,156	0,152	0,143	0,139
Modelo 4	0,029	0,023	0,056	0,051
<i>VAR Cointegrado</i>				
Intercepto	0,026	0,021	0,056	0,049
Tendencia	0,033	0,028	0,079	0,063
<i>Multivariados en diferencias</i>				
VAR-single	0,036	0,030	0,083	0,072
Lasso VAR-single	0,027	0,022	0,040	0,033
VAR-joint	1,014	0,426	49,039	29,405
Lasso VAR-joint	0,027	0,021	0,039	0,033
<i>BVAR en niveles</i>				
VAR Bayesiano	0,029	0,023	0,054	0,046

Cuadro: Evaluación de modelos de Predicción - Mayorista

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,179	0,130	0,368	0,331
Modelo 2	0,177	0,125	0,255	0,229
Modelo 3	0,219	0,167	0,431	0,392
Modelo 4	0,224	0,175	0,623	0,455
<i>Multivariados en diferencias</i>				
VAR-single	1,187	0,842	0,896	0,634
Lasso VAR-single	0,194	0,142	0,376	0,338
VAR-joint	0,387	0,305	0,697	0,555
Lasso VAR-joint	0,195	0,142	0,376	0,338
<i>BVAR en niveles</i>				
VAR Bayesiano	0,187	0,129	0,381	0,343

Cuadro: Evaluación de modelos de Predicción - Medianas

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,346	0,229	0,966	0,793
Modelo 2	0,347	0,237	0,979	0,894
Modelo 3	0,484	0,385	1,720	1,704
Modelo 4	0,350	0,233	0,872	0,791
<i>Multivariados en diferencias</i>				
VAR-single	1,430	1,248	2,727	2,631
Lasso VAR-single	0,344	0,225	0,858	0,699
VAR-joint	1,450	1,236	3,089	2,925
Lasso VAR-joint	0,344	0,226	0,846	0,689
<i>BVAR en niveles</i>				
VAR Bayesiano	0,389	0,278	1,460	1,304

Cuadro: Evaluación de modelos de Predicción - PYME

Modelo	Horizonte $h = 1$		Horizonte $h = 6$	
	RMSE	MAE	RMSE	MAE
<i>Univariados</i>				
Modelo 1	0,242	0,191	0,892	0,796
Modelo 2	0,239	0,189	0,665	0,561
Modelo 3	2,275	2,176	2,145	2,044
Modelo 4	0,341	0,284	2,466	2,429
<i>Multivariados en diferencias</i>				
VAR-single	1,098	0,944	5,575	5,384
Lasso VAR-single	0,318	0,262	1,824	1,642
VAR-joint	0,859	0,771	5,421	5,295
Lasso VAR-joint	0,224	0,178	1,069	0,975
<i>BVAR en niveles</i>				
VAR Bayesiano	0,336	0,270	1,526	1,430

Figura: Predicción a fin de Año - Consumo Revolvente

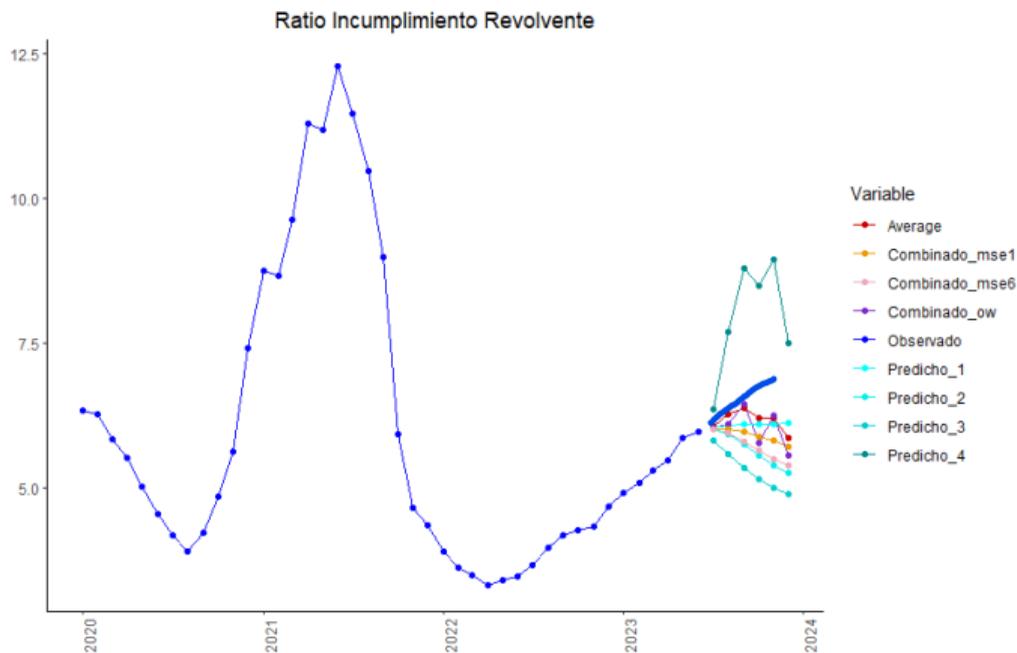


Figura: Predicción a fin de Año - Consumo NoRevolvente

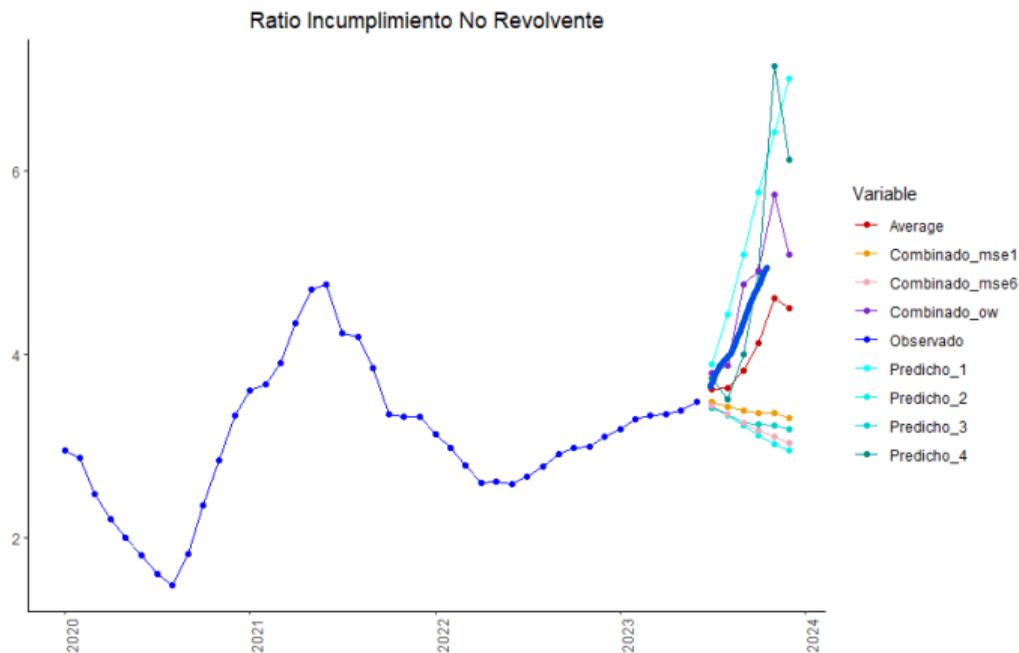


Figura: Predicción a fin de Año - Hipotecario

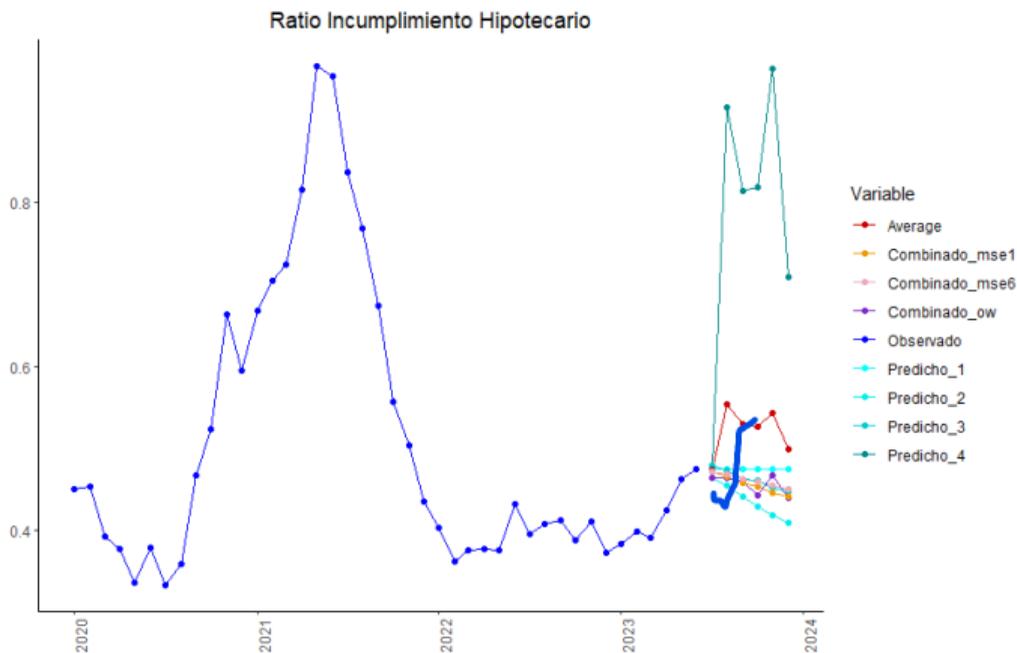
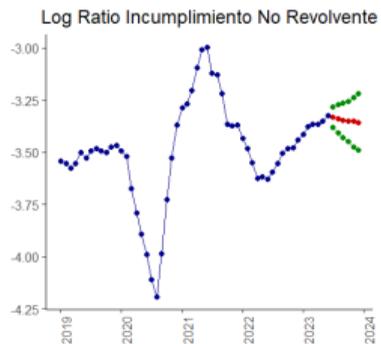
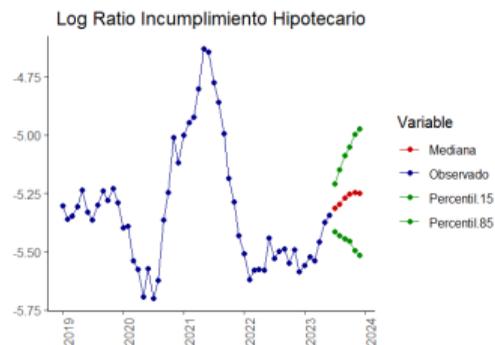
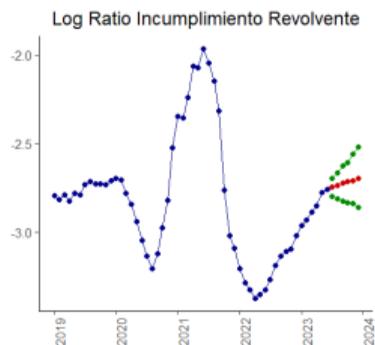


Figura: Predicción 6 meses adelante - log RIs



Comparación con datos Realizados

Figura: Predicción fuera de muestra versus Valores Realizados

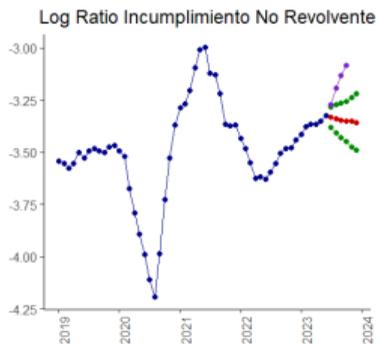
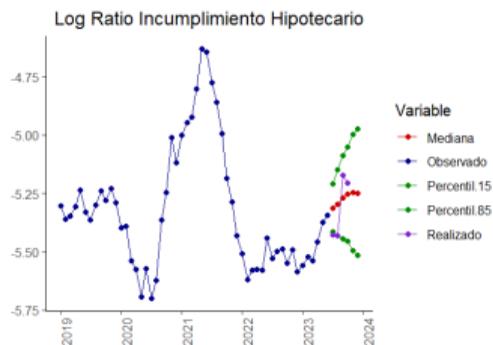
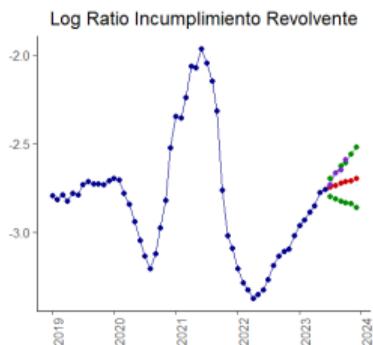
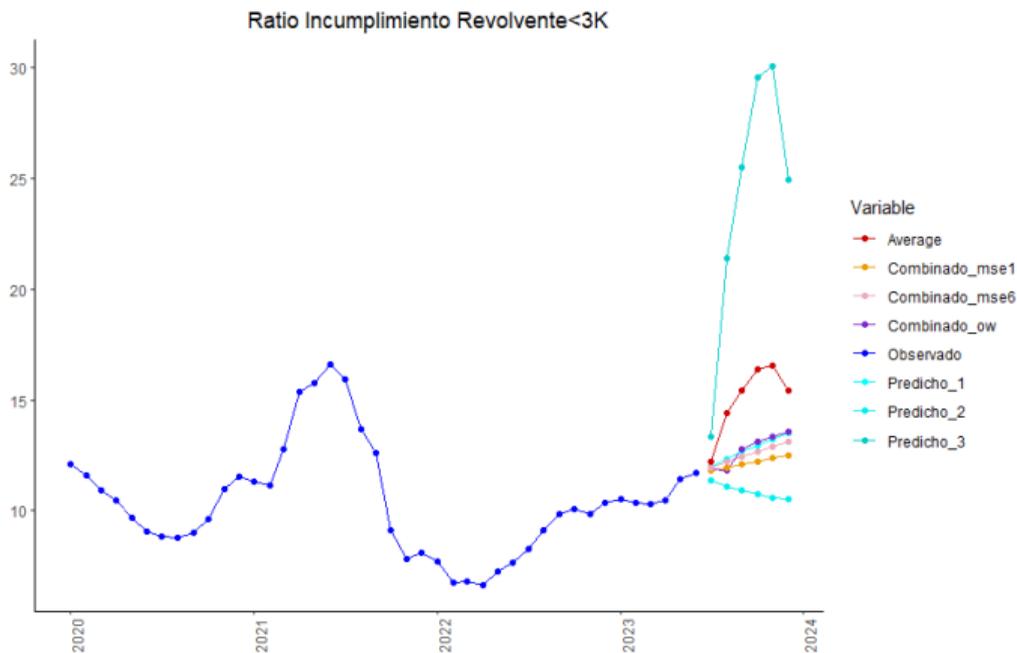


Figura: Predicción a fin de Año - Consumo Revolvente < 3K



Predicción Fin de Año: Revolvente - Línea mayor a 3K

Figura: Predicción a fin de Año - Consumo Revolvente > 3K

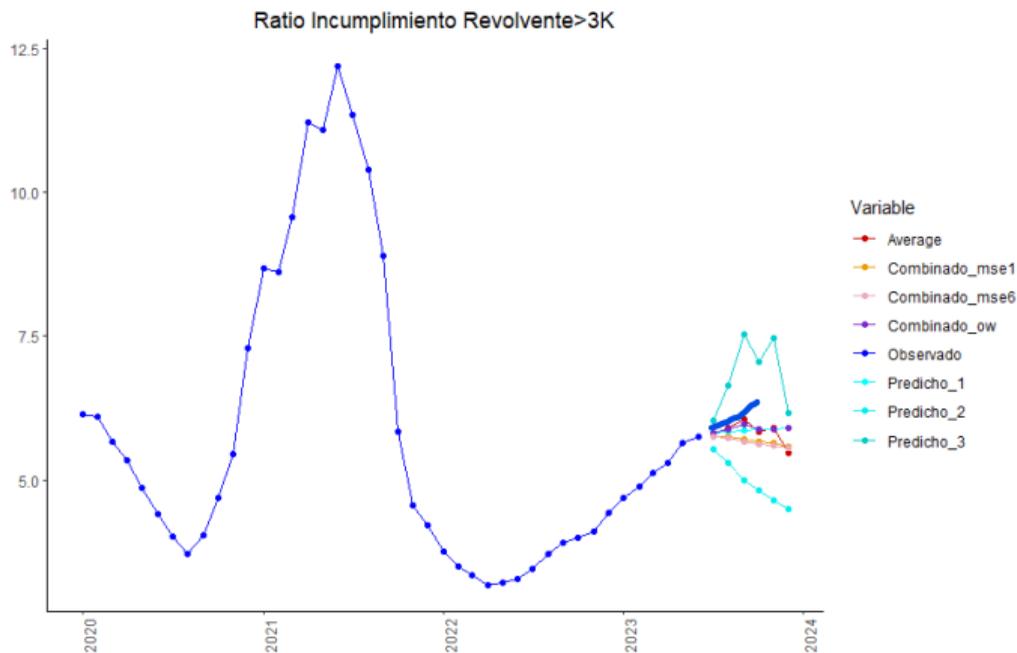
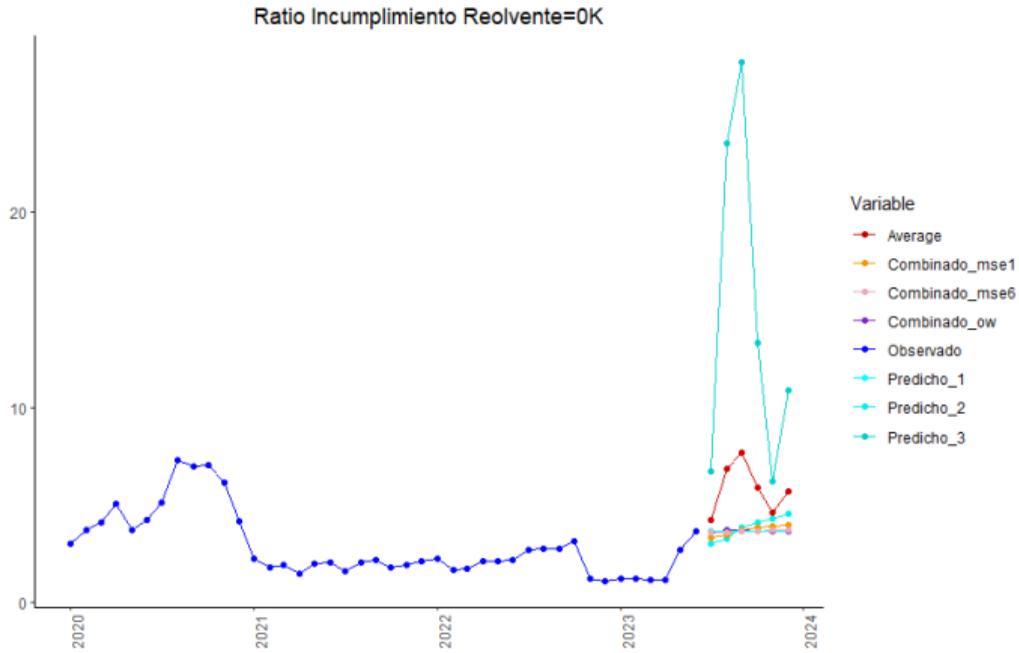
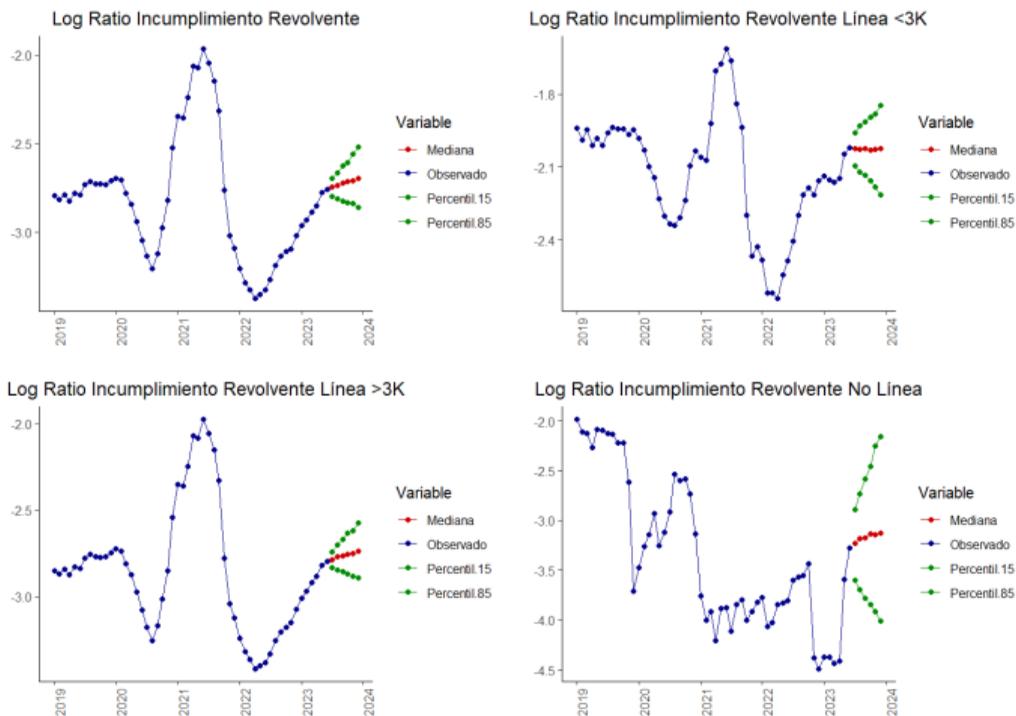


Figura: Predicción a fin de Año - Consumo Revolvente 0K



Predicción Fin de Año: Revolvente - sub-segmentos

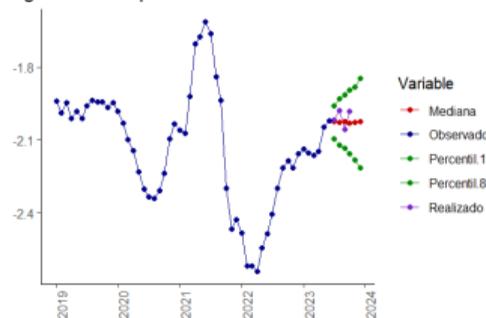
Figura: Predicción 6 meses adelante - log RIs



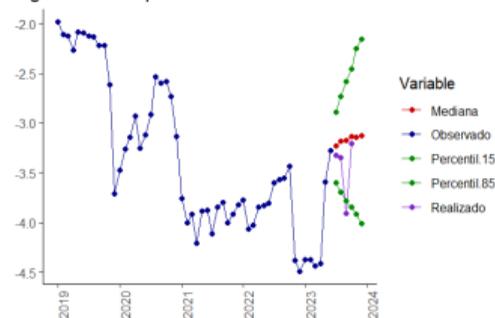
Comparación con datos Realizados

Figura: Predicción fuera de muestra versus Valores Realizados

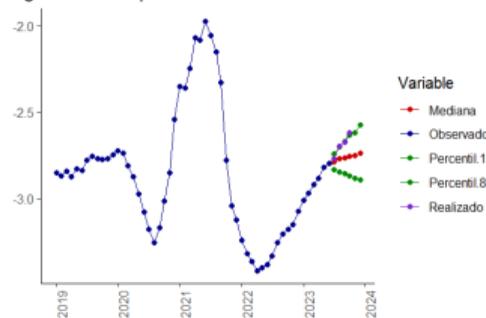
Log Ratio Incumplimiento Revolvente - Línea <3K



Log Ratio Incumplimiento Revolvente - No Línea



Log Ratio Incumplimiento Revolvente - Línea >3K



Predicción Fin de Año: No Revolvente - Sin Convenio

Figura: Predicción a fin de Año - Consumo No Revolvente sin convenio

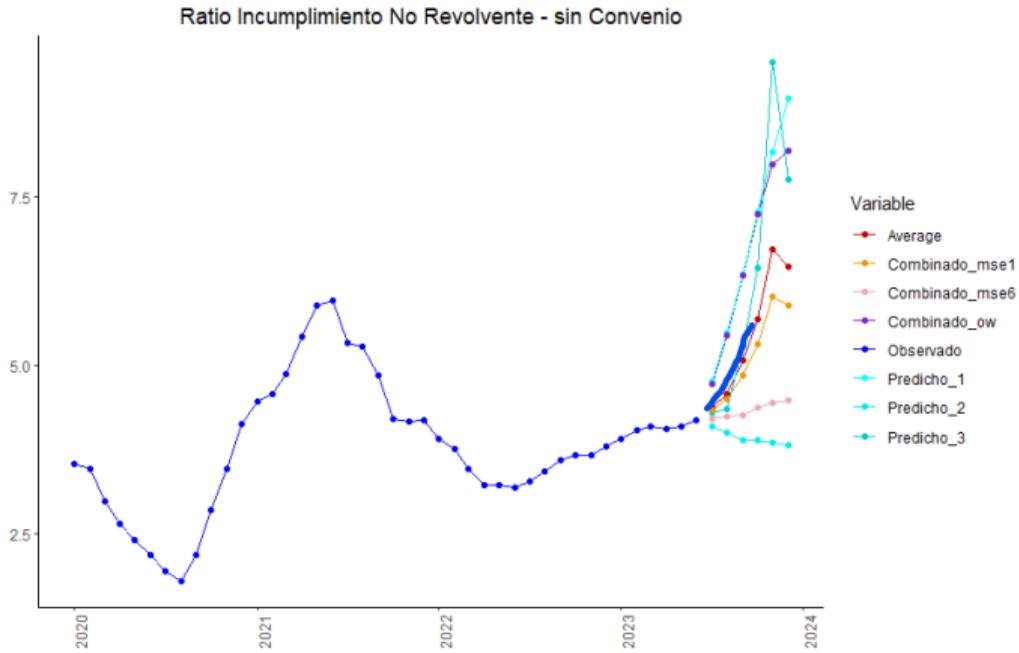
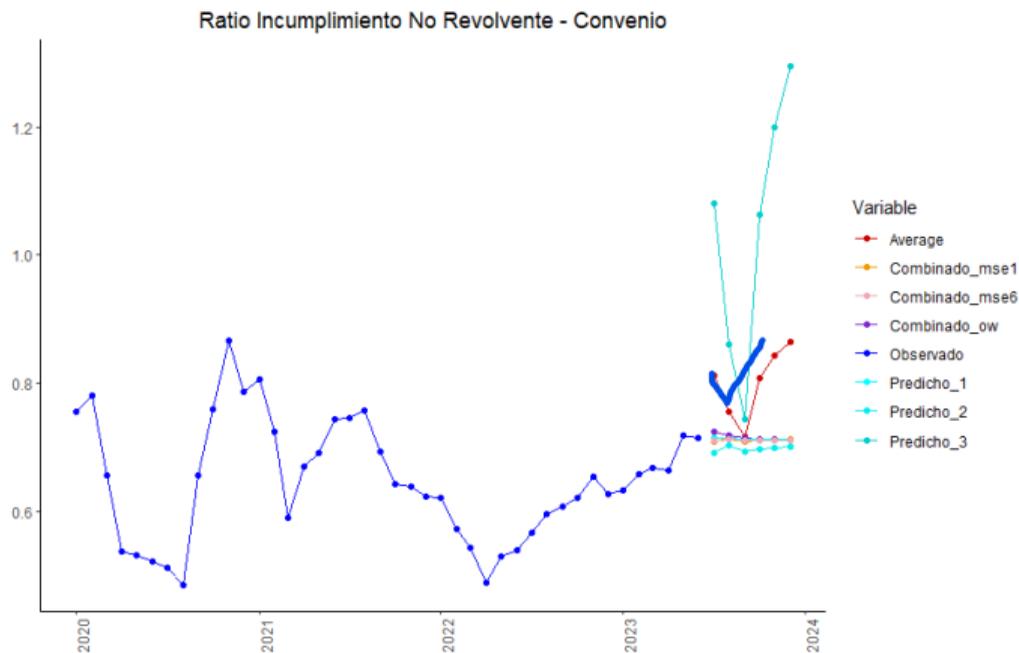
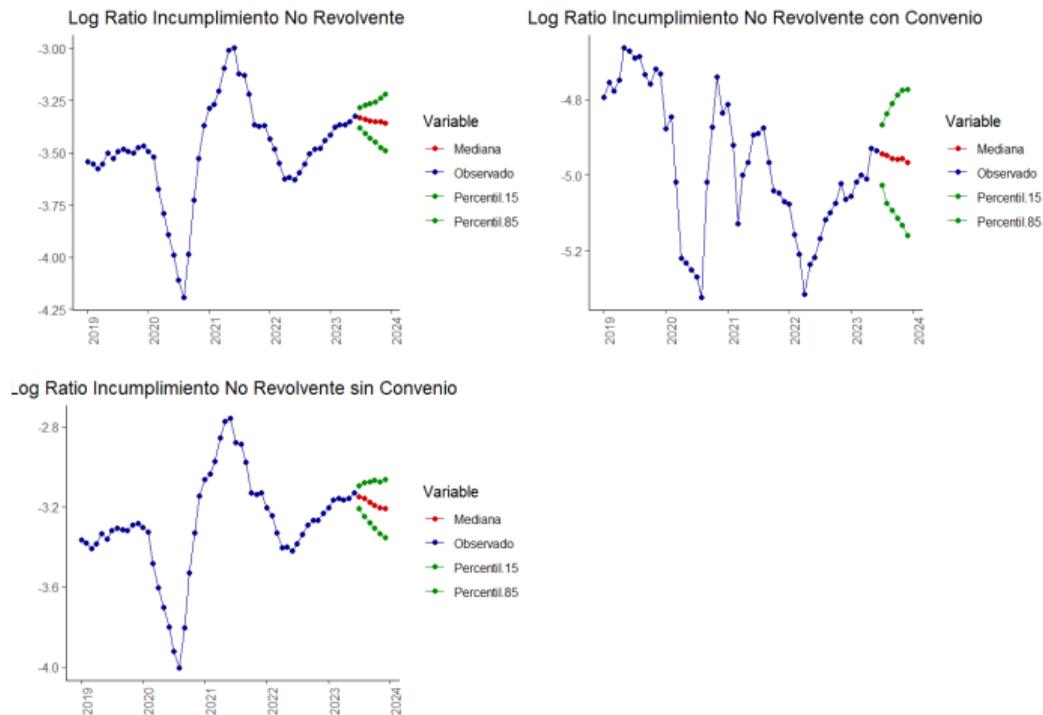


Figura: Predicción a fin de Año - Consumo No Revolvente con convenio



Predicción Fin de Año: No Revolvente - sub-segmentos

Figura: Predicción 6 meses adelante - log RIs



Comparación con datos Realizados

Figura: Predicción fuera de muestra versus Valores Realizados

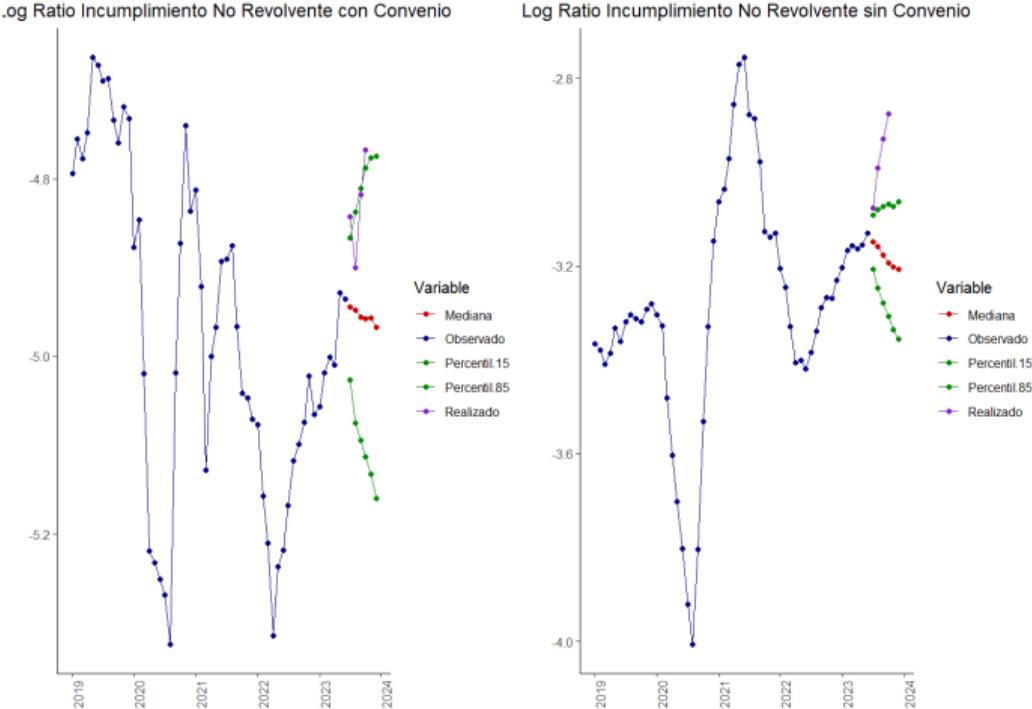


Figura: Predicción a fin de Año - Mayorista

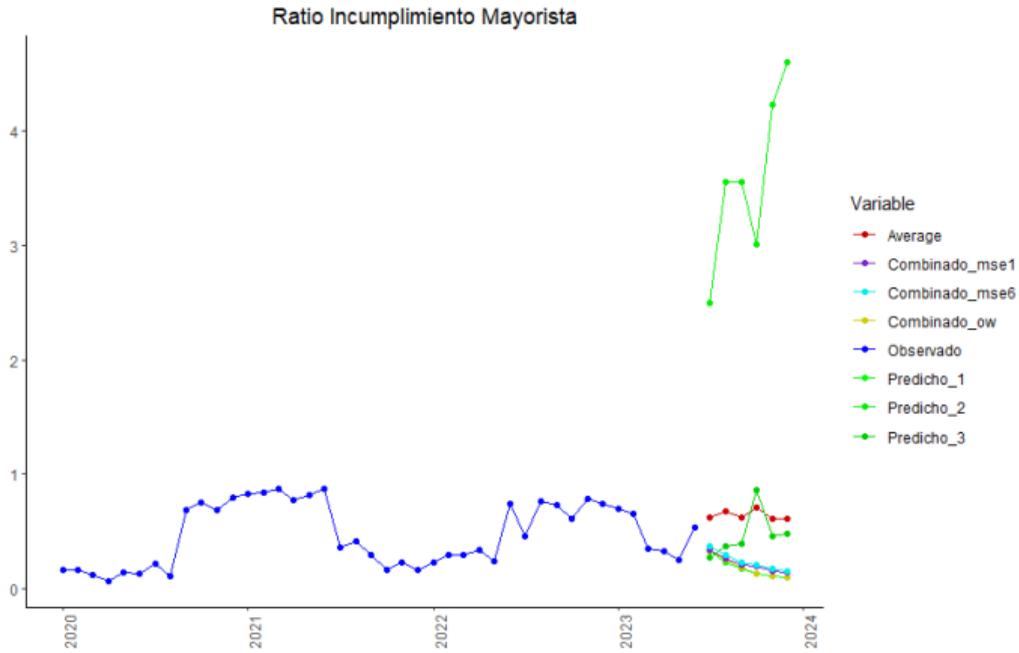


Figura: Predicción a fin de Año - Mediana

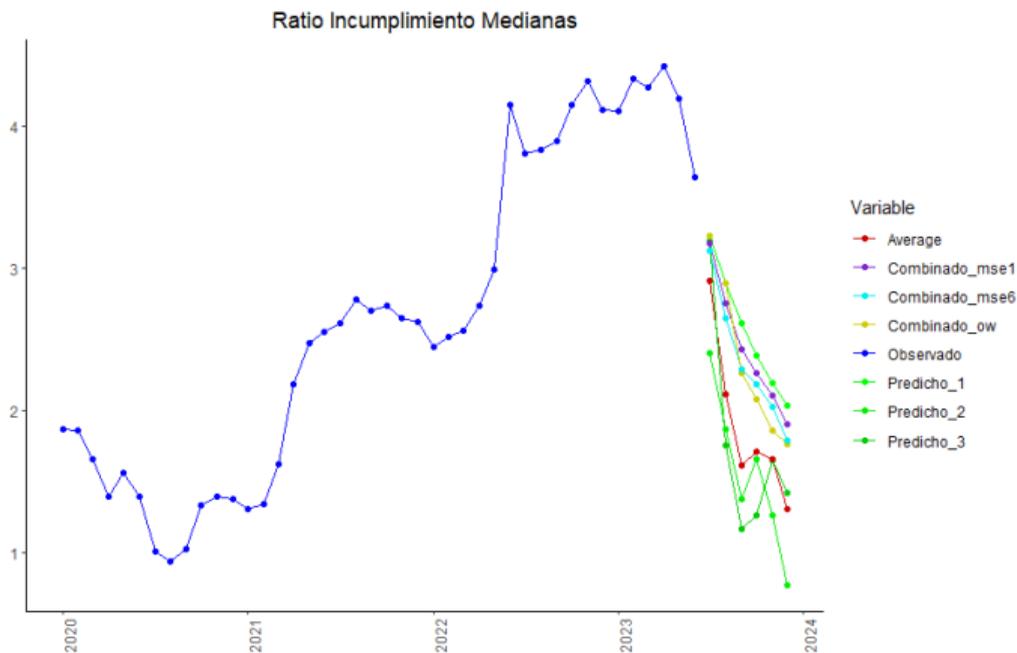


Figura: Predicción a fin de Año - PYME

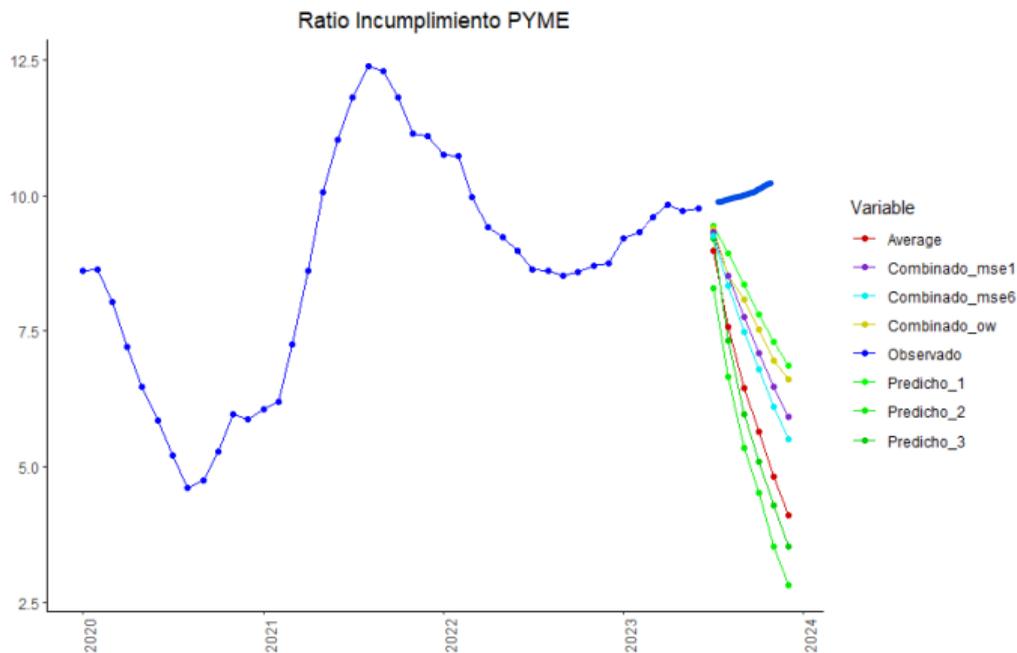


Figura: Predicción 6 meses adelante - log RIs

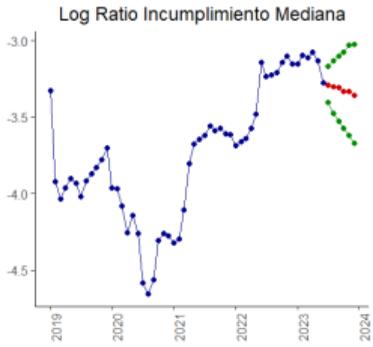
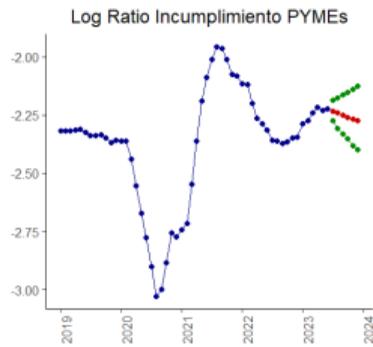
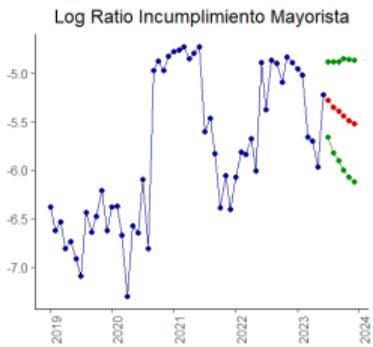


Figura: Predicción fuera de muestra versus Valores Realizados

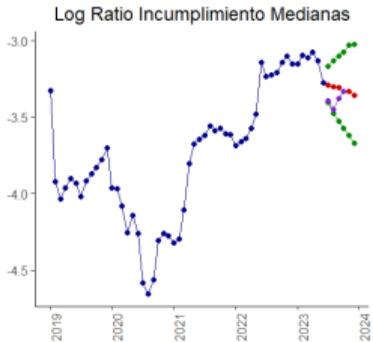
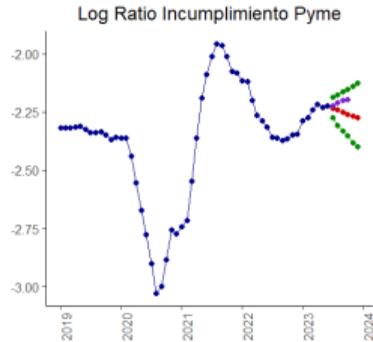
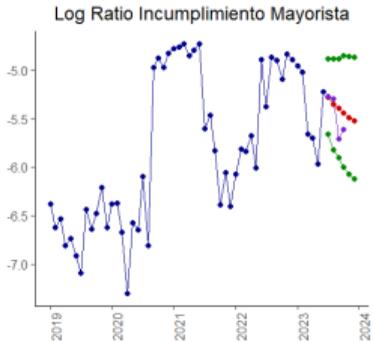
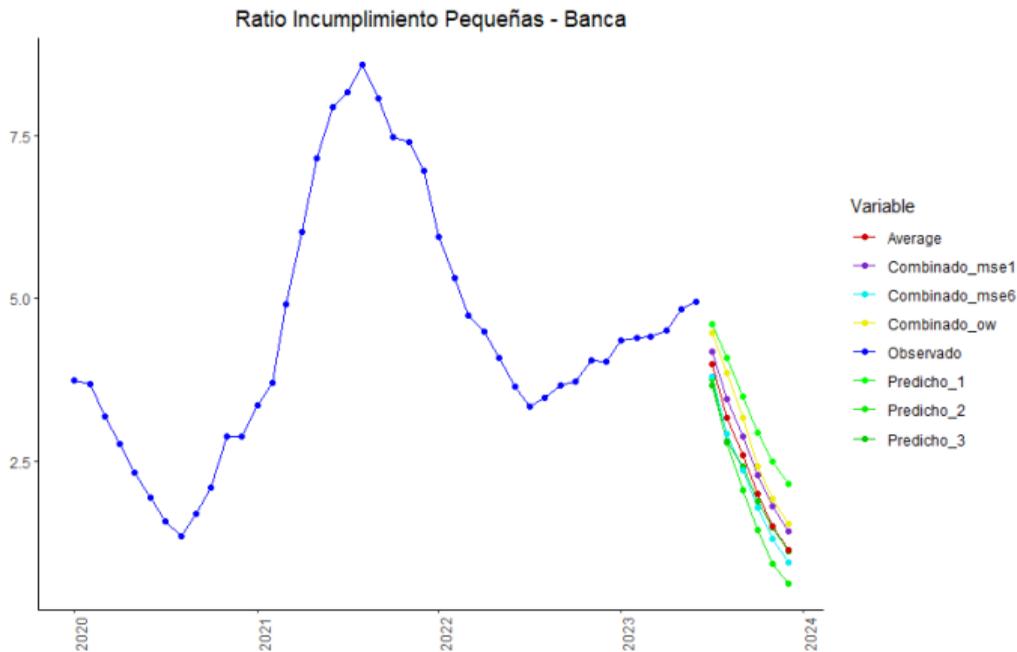


Figura: Predicción a fin de Año - Pequeñas (con entidades bancarias)



Predicción Fin de Año: Pequeñas con no Bancos

Figura: Predicción a fin de Año - Pequeñas (con entidades no bancarias)

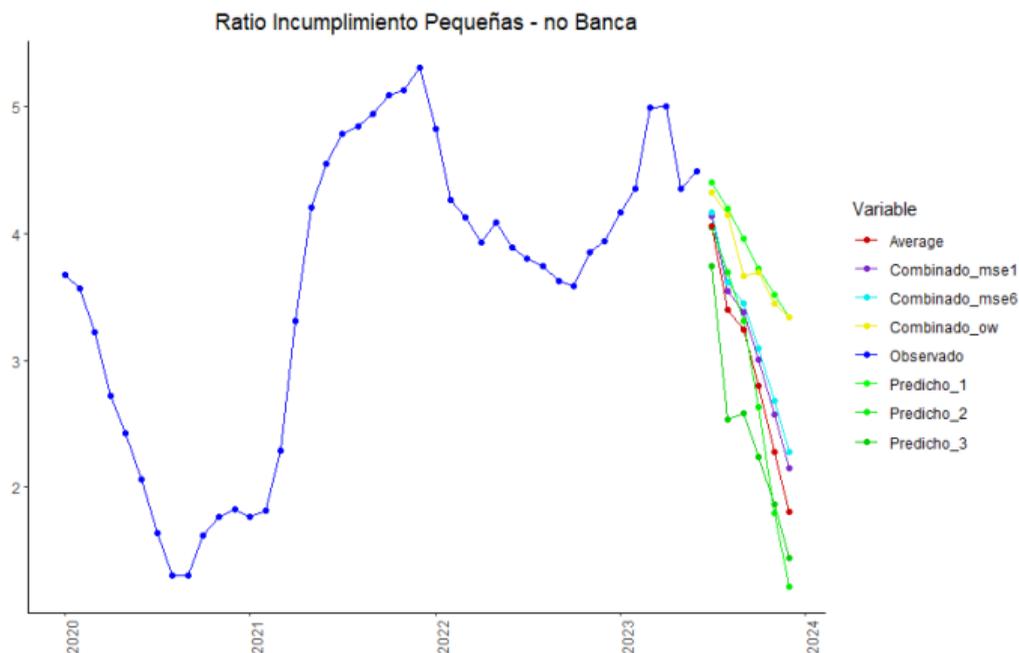
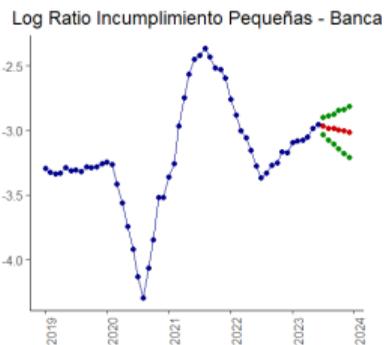
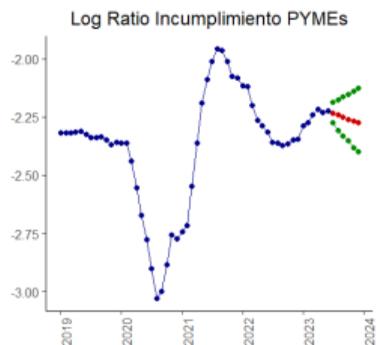


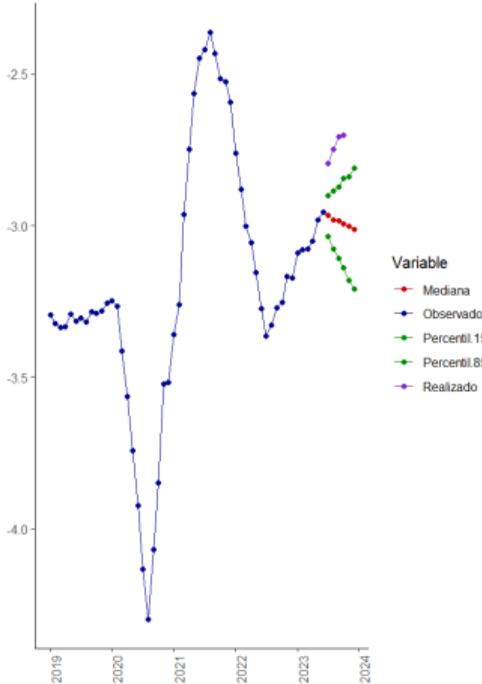
Figura: Predicción 6 meses adelante - log RIs



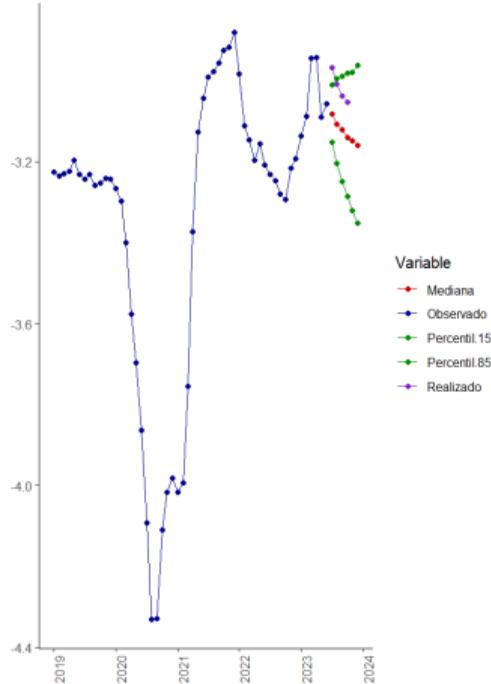
Comparación con datos Realizados

Figura: Predicción fuera de muestra versus Valores Realizados

Log Ratio Incumplimiento Pequeñas - Bancos



Log Ratio Incumplimiento Pequeñas - No Bancos



Conclusión y mejoras

- Los modelos multivariados mejoran en ciertos casos la precisión de los modelos univaridos, especialmente si tienen regularización.
- Los modelos con parámetros cambiantes en el tiempo muestran un mejor desempeño cuando predicen fuera de muestra que los modelos con parámetros invariantes.
- Estos modelos permiten calcular predicciones condicionadas.
- Es posible incorporar más variables de política en los modelos.
- ¿EAD y LGD?

Predicción de una medida de riesgo de impago por segmentos de crédito

Miguel Cabello, Camila Rodriguez y Diego Yamunaque

Banco Central de Reserva del Perú

18 de octubre de 2024