



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL**

**SEGUIMIENTO A MODELOS DE RIESGO DE CRÉDITO PARA
MICROEMPRESARIOS**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL
INDUSTRIAL**

ROMINNA ANDREA JIMÉNEZ OLIVERA

**PROFESOR GUÍA:
CRISTIÁN BRAVO ROMÁN**

**MIEMBROS DE LA COMISIÓN:
PATRICIO VALENZUELA AROS
SEBASTIÁN MALDONADO ALARCÓN**

Este trabajo ha sido parcialmente financiado por el Centro de Finanzas del
Departamento de Ingeniería Industrial de la Universidad de Chile

**SANTIAGO DE CHILE
MARZO 2014**

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniera Civil Industrial
POR: Rominna Andrea Jiménez Olivera
FECHA: 5/3/2014
PROFESOR GUÍA: Cristián Bravo Román

SEGUIMIENTO A MODELOS DE RIESGO DE CRÉDITO PARA MICROEMPRESARIOS

En Chile, los microempresarios son cada vez más. En el año 2004 el 81% de las empresas existentes correspondía a microempresarios. Es por esto que las oportunidades de crédito también están en aumento. Para que las instituciones puedan definir a quien asignar o no crédito, generalmente utilizan *Credit Scoring*. El problema que se presenta ante esta metodología, es que los modelos pierden poder discriminante en el tiempo, debido a cambios en la población y en la distribución de las variables que se evalúan.

Esta memoria diseña un proceso de negocio, basado en test estadísticos, que permite determinar el momento óptimo para ajustar los modelos de riesgo de crédito a consumidores. La metodología que se sigue para obtenerlo, consiste en definir un test supervisado, test K-S, y un test no supervisado, Fieller Stability Measure, para estimar una medida de pérdida y el momento óptimo de la re-calibración. Con esto, se construye el proceso de negocio y se evalúa en datos de una institución financiera real.

Los datos con los que se cuenta para realizar estas estimaciones consideran a carteras de microempresarios de sectores no rurales. El periodo de estudio va desde Enero del 2010 a Septiembre del 2012, y la muestra total incluye un total de 83.137 registros. Con estos datos se evalúa el comportamiento de la curva K-S versus la pérdida y se obtiene con un error del 22%, al medir el valor del aumento de la pérdida dada la baja porcentual del estadístico KS.

Aplicando el método Fieller Stability Measure se verifica cuales meses las variables no cumplen con la condición de mantenerse dentro de los límites aceptables. Con esto se concluye que la alerta de acción (para re-calibrar el modelo) ocurre cuando durante tres meses seguidos el intervalo de la variable traspasa los límites del modelo. Así mismo se establece como alerta de precaución cuando el intervalo de la variable se sale de los límites por un mes. Se comprueba que para el caso de alerta de acción la pérdida que se recupera justifica el costo de re-calibración. Al contrario del caso de la alerta de precaución en donde no es rentable, e incluso aumenta la pérdida, al recalibrar el modelo.

A Pipe, Patty, Mis Papás y Lore.

Agradecimientos

Primero quiero agradecer a Felipe Vildoso, por su constante amor, por estar siempre que lo necesito, y por incluso preocuparse más de mí que por él mismo. Por ser el mejor del mundo conmigo y por haberme escogido. Pipe, sin ti nunca jamás habría logrado llegar a este momento, me demostraste desde el día anterior al examen de mecánica que podemos lograr cualquier cosa juntos, te amo.

A Marina Olivera y Luis Jiménez, por darme la vida, un hogar y el amor infinito de todos estos años. Sin su preocupación, cariño y dedicación jamás habría soñado siquiera con entrar a la U. Los quiero infinito, son los mejores.

A Patricia Jiménez, mi hermana y amiga, gracias por todo Patty, en especial por esos bailes, haces de mi vida algo mucho más más entretenido. Eres mi hermana favorita, tú ya sabes todo, te quiero mucho.

A Loreto Jiménez, por su apoyo incondicional y todo su cariño. Loretito, eres la mejor, te quiero montones, gracias por dejarme contar contigo y estar siempre ahí.

A mis tíos, primos y abuelos, por su apoyo y cariño. En especial a mis tías abuelas y a mi abuelita, siempre preocupadas por cómo iba y por ayudar en todo lo que pudieran.

A mis amigos de la universidad. Tania, Stefano, Rurik, Piero, Pasti, Pancho Barrientos, Nolguin, Nico, Jhanis, Ilo, Hojmi, Fito, Felipe, Eve, Andrés Cordova, Alfredo, Alexis, gracias a todos por haber hecho de la universidad un lugar grato en el que me sentí acogida inmediatamente y hasta el final. Quiero mencionar especialmente a Tania Ocampo, Stefano Cimma y Daniela Pavez, que de alguna u otra manera estuvieron presentes desde el primer día de clases hasta el día de hoy, los quiero niños.

Al equipo docente que me acompañó durante el E y el F, especialmente a Cristián Bravo y Patricio Valenzuela, por su disposición para ayudar y su buena voluntad para siempre responder mis dudas.

Gracias a todos por estar presentes en mi vida, sin ustedes no podría haber llegado hasta aquí.

Tabla de Contenido

Capítulo 1: Introducción	1
1.1 Descripción del Proyecto.....	2
1.2 Objetivo General	2
1.3 Objetivos Específicos.....	3
1.4 Metodología.....	3
1.5 Alcances	3
1.6 Resultados Esperados.....	4
1.7 Plan de Trabajo	4
Capítulo 2: Marco Teórico	6
2.1 Proceso de Extracción del Conocimiento (KDD).....	6
2.2 Credit Scoring	6
2.3 Concept Drift	7
2.4 Test Supervisados y no Supervisados.....	8
2.4.1 Test de Kolmogorov–Smirnov.....	8
2.4.2 Stability Index (SI)	9
2.4.3 Fieller Stability Measure.....	10
Capítulo 3: Tratamiento de los Datos.....	12
3.1 Selección y Preprocesamiento	12
3.2 Transformación.....	13
3.3 Evaluación	14
Capítulo 4: Metodología para Anticipar la Re-calibración	15
4.1 Planteamiento del Problema.....	15
4.2 Test Supervisado	16
4.2.1 Evaluación	16
4.2.2 Estimación de Medidas	17
4.3 Test No Supervisado	19
4.3.1 Estimación de la Alerta.....	20
4.3.2 Costo Económico de la Re-calibración.....	22
Capítulo 5: Resultados	23

5.1 Fieller Stability Measure.....	23
5.2 Stability Index.....	25
5.3 Análisis de Resultados	26
Capítulo 6: Conclusiones	29
6.1 Acerca de la Metodología para Anticipar la Re-calibración.....	29
6.2 Acerca de los Objetivos.....	29
6.3 Recomendaciones a la Institución.....	30
6.4 Trabajo Futuro	30
Bibliografía	32
Anexos	33
Anexo A	33
Relación entre el aumento porcentual en el estadístico K-S y la disminución porcentual de la pérdida económica.....	33
Anexo B.....	34
Fieller Stability Measure para todas las variables sin realizar ninguna re-calibración del modelo.....	34
Fieller Stability Measure para las variables que quedaron al realizar la re-calibración del modelo para Noviembre 2010 en adelante.....	42
Anexo C.....	50
Datos de la Regresión Logística de la Re-calibración de Octubre 2011.....	50
Anexo D.....	53
Estadístico K-S y Pérdidas Económicas por cartera: Sin ninguna re-calibración.....	53
Re-calibración Noviembre 2010.....	54
Re-calibración Octubre 2011.....	55
Re-calibración Febrero 2012.....	56
Anexo E.....	58
Stability Index.....	58

Índice de Figuras

Figura 1: K-S en el tiempo.....	18
Figura 2: K-S Vs Pérdida Económica	18
Figura 3: Disminución Pérdida Real Vs Disminución Pérdida Predicha	19

Figura 4: Variable Demográfica 1	20
Figura 5: Variable Deuda 1A	20
Figura 6: Variable Deuda 1D	24
Figura 7: Variable Operación 2C	24

Índice de Tablas

Tabla 1: Carta Gantt	5
Tabla 2: Datos Extraídos que Cursan Operaciones	12
Tabla 3: Estadísticos Descriptivos de las Variables.....	14
Tabla 4: Crecimiento K-S Vs Disminución Pérdida Económica	17
Tabla 5: Pérdidas con y sin Re-calibración Octubre 2011	23
Tabla 6: Pérdidas con y sin Recalibración Febrero 2012	25
Tabla 7: Pérdidas con y sin Re-calibración Junio 2012	25

Tabla 1: Carta Gantt	5
Tabla 2: Datos Extraídos que Cursan Operaciones	12
Tabla 3: Estadísticos Descriptivos de las Variables.....	14
Tabla 4: Crecimiento K-S Vs Disminución Pérdida Económica	17
Tabla 5: Pérdidas con y sin Re-calibración Octubre 2011	23
Tabla 6: Pérdidas con y sin Recalibración Febrero 2012	25
Tabla 7: Pérdidas con y sin Re-calibración Junio 2012	25

Capítulo 1

Introducción

En el presente informe se aborda un problema presente en las instituciones que otorgan créditos. Este es, no saber en qué momento los modelos que poseen para tomar decisiones dejan de ser predictivos y se vuelven obsoletos.

El *credit score* es una herramienta que se utiliza para entregar una puntuación a postulantes de créditos, y más allá de entregar la respuesta a si el cliente es bueno o malo (si pagará o no), entrega la probabilidad, de que dado un conjunto de variables que lo describen, el cliente cumpla con su obligación [10].

La problemática presentada nace principalmente de la pérdida de la utilidad y de la estabilidad de la población estudiada para la creación del modelo. Esto ocurre con el pasar del tiempo, ya que los modelos utilizados se tratan como estacionarios cuando en realidad son dinámicos, por lo que las poblaciones varían, y ahora los atributos que hacían de sus participantes, buenos o malos clientes, puede que probablemente no entreguen mayor información o incluso información errónea, por lo que frente a un cambio en la población se hace necesario reajustar estos factores.

Actualmente existen estudios que tratan de captar este fenómeno y dar a conocer sus consecuencias. Por ejemplo en [9] se dan a conocer distintos métodos de detección de cambio aplicando distintos tipos de test supervisados. Por otro lado en [3] se proponen técnicas estadísticas con el fin de determinar cuándo los cambios en las características de la población pueden afectar el desempeño de los modelos. Y en [1] se muestra como incluso pequeñas diferencias en las predicciones pueden significar grandes pérdidas o ganancias económicas, situación muy importante a considerar al ver cuánto es lo que se pierde o se podría ganar gracias a realizar un ajuste adecuado de los modelos a tiempo.

El aporte de este trabajo será configurar una alerta que pueda predecir el ajuste de los modelos antes de que se incurra en pérdidas.

1.1 Descripción del Proyecto

El proyecto pretende resolver la problemática que se produce a raíz de no saber cuándo se está trabajando con un modelo de predicción de riesgo de crédito obsoleto. Para esto se diseña un proceso de negocio por medio del cual es posible identificar el momento óptimo en el que debe realizarse el ajuste a los modelos, es decir, que presente una alerta ante posibles cambios para detectar de manera oportuna la necesidad de ajuste. Esto es realizado tomando en cuenta dos variables, el decaimiento presente en la curva K-S y los cambios en la estabilidad poblacional.

En estos momentos este es un tema importante en una prestigiosa institución financiera nacional, en donde la gerencia de riesgo de microempresas tiene que tolerar un año de pérdidas, antes de saber que el modelo está fallando, ya que es recién en ese momento cuando los modelos reflejan su bajo poder predictivo. Esto se da con este desfase ya que para la evaluación que se realiza para conocer la situación de cada uno de los microempresarios se requieren doce meses de información para conocer su comportamiento. Esta cantidad de información es requerida para que el estadístico K-S represente los resultados de la mejor manera posible, ya que si no se cuenta con una cantidad de meses previos del comportamiento del cliente, los resultados que se entregarán serán anómalos. Por ejemplo, K-S muy altos, que se deben a que en períodos de tiempo más reducidos sólo van a fallar los clientes que realmente son muy malos, en cambio luego de considerar un tiempo prudente se va a mostrar la situación real de las carteras. De esta forma, luego de doce meses de comportamiento se sabe si el cliente cayó en default y se le clasifica como mal o buen cliente.

La idea es que este proyecto les permita evitar esas pérdidas, dándoles a conocer anticipadamente cuándo es momento de hacer cambios en el modelo, de este modo, se podrá conocer con anterioridad el comportamiento de las variables de las distintas carteras de microempresarios para saber si es necesario o no recalibrar el modelo.

1.2 Objetivo General

Diseñar un proceso de negocio que permita determinar el momento óptimo para ajustar los modelos de riesgo de créditos a microempresarios de Banco Estado Microempresas aplicando test estadísticos.

1.3 Objetivos Específicos

- Definir una metodología para medir pérdidas por baja en la predicción de los modelos.
- Definir por cada punto de K-S el monto de la pérdida esperada.
- Definir en base a los test supervisados y no supervisados conocidos, el mejor momento para re-calibrar.
- Obtener el proceso de negocio que permita determinar de forma eficiente el momento óptimo para el cambio en el modelo.
- Evaluar el funcionamiento del proceso en los modelos de riesgo de crédito a microempresarios de Banco Estado Microempresas.

1.4 Metodología

- Estudiar la literatura existente en modelos de predicción de riesgo de créditos.
- Construir modelos mediante el proceso KDD para su posterior uso en las estimaciones de las medidas.
- Estimar una medida de pérdida en base a la información con la que se cuenta.
- Definir test supervisados útiles.
- Estimar el valor de la pérdida económica y el momento óptimo de la re-calibración para los test supervisados.
- Definir test no supervisados útiles.
- Estimar el valor de la pérdida económica y el momento óptimo de la re-calibración para los test no supervisados.
- Construir proceso de negocio.
- Evaluar el proceso de negocio en datos reales del banco.

1.5 Alcances

El proyecto aquí presentado se basará principalmente en el estudio de los test supervisados y no supervisados para la obtención de un proceso de negocio que indique el momento óptimo en que debe ser ajustado un modelo de predicción de riesgo de crédito a microempresarios. Dada la naturaleza del proyecto y su propósito, está realizado especialmente para un banco nacional. De hecho se harán pruebas con datos reales del banco para comprobar la utilidad efectiva del método.

La memoria se abordará desde dos puntos de vista, del decaimiento de la curva K-S y de la estabilidad que presenta la población. A pesar de que está enfocada en la predicción del ajuste de un modelo de riesgo de crédito para

microempresarios, se analizará la posibilidad de, eventualmente, aplicar el proceso en cualquier tipo de situación en donde se necesite estudiar y predecir el comportamiento de clientes, por lo tanto no se limitaría sólo a la banca. Sin embargo, los detalles específicos de esta generalización no se encuentran en los alcances de esta memoria.

1.6 Resultados Esperados

Se espera, gracias a las observaciones en la curva K-S y en los cambios poblacionales, poder configurar una alerta que indique cuándo es preciso realizar un ajuste del modelo predictivo. Esto, mediante la obtención de dos tipos de alerta, una de precaución y otra de acción. Cuando exista alerta de precaución habrá que poner atención a los posibles cambios que puedan ocurrir, de manera de tener monitoreado el modelo. Cuando exista alerta de acción el ajuste en el modelo será necesario.

1.7 Plan de Trabajo

Los campos en la Carta Gantt corresponden a los pasos de la planificación del trabajo para realizar esta memoria. A continuación el detalle de cada uno de ellos:

- Entrega Informe IN6908: entrega del presente informe al profesor guía para que este entregue su recomendación sobre el trabajo, aprobado o reprobado.
- Revisión de Literatura: revisión y estudio de la literatura existente referente a la predicción de riesgo de *credit scoring*.
- Introducción: actualización y mejora del presente informe para ser incorporado como primer capítulo del informe final de IN6909.
- Datos: revisión, estudio, selección y transformación de los datos que se evaluarán. También hay que estimar una medida de pérdida con la información que se tenga.
- Marco Teórico: diseño, y mejora del marco teórico que hace referencia a todos los conceptos empleados en esta memoria.
- Evaluar Test Supervisados: definir qué test supervisados serán útiles y estimar una medida de pérdida para ellos y el momento óptimo para realizar la re-calibración del modelo de predicción.
- Evaluar Test No Supervisados: definir qué test no supervisados serán útiles y estimar una medida de pérdida para ellos y el momento óptimo para realizar la re-calibración del modelo de predicción.

- Resultados: analizar los resultados entregados de las experiencias anteriores y escoger la metodología que mejor se adapte a la resolución del problema.
- Conclusiones: concluir respecto al trabajo realizado entregando el proceso de negocio que permita saber el momento preciso en que se deben recalibrar los modelos y configurar una alarma.
- Correcciones a informe: actualizar el informe de acuerdo al feedback entregado por profesores, realizando las mejoras pertinentes.
- Entrega Informe IN6909.

Tabla 1: Carta Gantt

Fechas		22- jul	29- jul	05- ago	12- ago	19- ago	26- ago	02- sep	09- sep	16- sep	23- sep	30- sep	07- oct	14- oct	21- oct	28- oct	04- nov	11- nov	18- nov	25- nov	02- dic	09- dic	16- dic	23- dic	
Meses		Julio		Agosto				Septiembre				Octubre				Noviembre				Diciembre					
Semanas de Clases		15	16	17	-	1	2	3	4	5	-	6	7	8	9	10	11	12	13	14	15	16	17	18	
I N 6 9 0 8	Entrega Informe IN6908																								
	Revisión de Literatura																								
I N 6 9 0 9	Introducción																								
	Datos																								
	Marco Teórico																								
	Evaluar Test Supervisados																								
	Evaluar Test No Supervisados																								
	Resultados																								
	Conclusiones																								
	Entrega Informe Final																								

Capítulo 2

Marco Teórico

Este capítulo contempla los conceptos necesarios para comprender las herramientas utilizadas en la memoria.

2.1 Proceso de Extracción del Conocimiento (KDD)

En este trabajo es necesario definir qué variables son útiles y cuáles no. El proceso que se utiliza para llevar a cabo la extracción de la información es KDD. Este proceso abarca cinco etapas: la selección, el preprocesamiento, la transformación, el minado de datos y su evaluación [8].

Una vez sean seleccionados los datos se pasan por un filtro, un preprocesamiento, en donde los datos serán inspeccionados buscando posibles errores o *outliers*, de este modo se obtendrán los datos limpios. Cuando se cuente con los datos procesados se procede a la transformación, en donde se llevarán a cabo cambios en los datos, por ejemplo el convertir algunas variables en dicotómicas, esto, con el fin de facilitar la lectura e interpretación de los datos. Al llegar a este punto se contará con la data consolidada y mediante el minado de estos datos se podrán obtener patrones, de donde por medio de su interpretación y posterior evaluación se llegará a obtener el conocimiento que se requiere.

El proceso KDD busca principalmente poder manejar una gran cantidad de datos más fácilmente facilitando el acceso a la información relevante que desea obtenerse. Que para este caso en particular consiste en el comportamiento de los microempresarios postulantes a créditos.

2.2 Credit Scoring

La herramienta utilizada por el banco en estos momentos para modelar el riesgo que tiene cada cliente es *credit scoring*. El cual consiste en una metodología mediante la cual se le asigna un puntaje (score) a cada uno de los postulantes, puntaje que depende de una serie de variables evaluadas por expertos, entre las que se encuentran datos relevantes para estimar el comportamiento futuro que tendrá el postulante, por ejemplo datos socioeconómicos, o de su comportamiento financiero en tiempo pasado.

El modelo estándar usado para esto es una función logit:

$$\text{Logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}$$

Donde p_i , β_k y x_{ki} son la probabilidad posterior, los coeficientes de la regresión y las variables respectivamente [10]. Aplicando función exponencial a ambos lados de la ecuación se obtiene:

$$p(y_i = 1|x_i) = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^V \beta_j x_{ij})}}$$

De esta forma se está "obligando" a p_i a tomar un valor entre 0 y 1, que es el valor esperado para una probabilidad, aplicado a este caso en particular, a la probabilidad de caer en *default*, es decir de ser un cliente considerado como malo que no cumple con su obligación de pago.

Cada vez que se re-calibre el modelo, es a través de una regresión logística, en donde mediante el estadístico del test de Wald [7] es el criterio de decisión que se utiliza para decidir si la variable sigue siendo significativa o no. En caso de que este sea menor a dos para alguna de las variables, no se sigue utilizando en el modelo.

Estos modelos requieren de un seguimiento constante ya que, en general, luego de dos años ya no poseen el mismo poder predictivo original por lo que se vuelven obsoletos. Esto quiere decir que el modelo ya no está entregando información fidedigna respecto a lo que se quiere estimar, situación que se debe a cambios que ocurren en la población y hacen que el modelo ya no discrimine de manera correcta a los clientes buenos de los malos. Esto puede deberse a cambios de distribución en las variables o bien a la pérdida de su capacidad discriminante. Para el primer caso se utilizará el Stability Index [5] y para el segundo el test de Kolmogorov-Smirnov [6].

2.3 Concept Drift

El *concept drift* se define como el problema que se presenta en análisis predictivo cuando los datos no son estables y cambian con el tiempo, haciendo que el modelo que se construyó con los datos originales sea inconsistente con los datos nuevos. Esto hace que una actualización regular del modelo sea necesaria. [11] En esta memoria se utiliza el *concept drift* como método para detectar cambios en las variables a analizar.

2.4 Test Supervisados y no Supervisados

Los tests que se estudian en esta memoria para su posterior evaluación son de dos tipos, test supervisados y test no supervisados.

En este contexto se entenderá test no supervisado como aquel que no requiere la salida real de la variable objetivo. Por lo tanto test supervisado es en donde se requiere conocer el valor de la función evaluada para poder detectar cambios en el modelo que se está evaluando. Dado que, para los test supervisados no se necesita conocer este dato para obtener conclusiones, se puede evaluar el cambio de variables o de los parámetros en la regresión sin la necesidad de saber, para este caso, si el cliente era efectivamente bueno o malo.

En muchas aplicaciones es deseable poder detectar cambios en los modelos antes de obtener salidas reales, es decir, no tener que esperar a que el modelo deje de ser válido para saber que es necesario recalibrarlo. Como es justamente el caso que se quiere evitar en esta memoria y es por eso que resulta fundamental evaluar este tipo de test para comprobar su efectividad en riesgo de crédito.

Debido a esto para la metodología que se implementará para generar la alerta se utilizarán los dos tipos de test, no supervisados si es que no se tiene información del comportamiento del cliente y supervisado si es que se cuenta con al menos un año de antigüedad en la institución.

2.4.1 Test de Kolmogorov–Smirnov

El test K-S es un test no paramétrico, del tipo supervisado, que es usado para comparar dos muestras, o una muestra con una distribución de probabilidad. El estadístico que entrega el test es la distancia máxima que existe entre dos distribuciones acumuladas [6].

La función de distribución F_n para n observaciones X_i independientes e idénticamente distribuidas se define:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}$$

Donde $I_{X_i \leq x}$ es la indicatriz, igual a 1 si es que $X_i \leq x$ y a 0 si no.

Las hipótesis que se comparan son:

- La hipótesis nula: las distribuciones de ambas poblaciones son iguales
- La hipótesis alternativa: las distribuciones de ambas poblaciones son distintas.

Como se mencionó el estadístico muestra la máxima distancia que existe entre ambas distribuciones, y viene dado por:

$$D = \max[F_B(x) - F_M(x)]$$

Donde F_B es la distribución acumulada de score de los clientes malos y F_M es la distribución acumulada de los clientes buenos de valores de la segunda muestra.

El estadístico de K-S que nade a partir de esta función sigue una distribución normal tipificada, por lo que si su valor es mayor que el correspondiente a "z" al nivel de confianza deseado es posible anunciar que hay evidencia estadística suficiente para rechazar la hipótesis que dice que las poblaciones son iguales, y por lo tanto concluir que la variable en juego es útil al modelo para distinguir entre clientes buenos y malos [9]. Mientras mayor sea D indica una diferencia más importante entre las dos distintas clases que se evalúan, que es lo que se busca para comprobar que los clientes "malos" tienen una distribución distinta a la de los clientes "buenos".

2.4.2 Stability Index (SI)

SI es un test del tipo no supervisado y sirve para medir la similitud entre poblaciones distintas [5]. Para este caso se considera la distribución de las poblaciones mediante *scorecard*:

$$SI = \sum_{i=1}^m (R_i - O_i) \ln \frac{R_i}{O_i}$$

Donde m es el número de clases de la variable, R_i representa el porcentaje de observaciones en el conjunto de referencia (de entrenamiento) de la clase i, y O_i el porcentaje de observaciones de la clase i en el conjunto observado (el de prueba).

Valores más altos de SI indican cambios sustanciales en la población. Si el SI es menor o igual a 0,1, no hay cambio significativo. Si el índice es mayor a

0,1, pero menor o igual a 0,25 ha habido un cambio menor. Si el SI es mayor a 0,25 ha habido un cambio significativo y es preciso recalibrar el modelo.

2.4.3 Fieller Stability Measure

Este método es no supervisado y se basa en dos cantidades, el Intervalo de Confianza de Fieller y el Maximum Deviation Measure.

El intervalo de confianza de Fieller proporciona un rango para la variable independiente que se quiere medir. Esta corresponde a la proporción de las medias de la población nueva y de la original, Fieller Stability Measure:

$$FSM = \mu_n / \mu_o$$

Donde μ_n es la media de la población nueva y μ_o la media de la población original. El intervalo de confianza para esta variable corresponde a:

$$(DSM_l, DSM_u) = \frac{1}{1-g} \left[\frac{\bar{x}_n}{\bar{x}_o} \pm \frac{t_{r,\alpha}}{\bar{x}_o} \left(S_n^2(1-g) + \frac{\bar{x}_n^2}{\bar{x}_o^2} S_o^2 \right)^{1/2} \right]$$

$$g = \frac{t_{r,\alpha}^2 S_n^2}{\bar{x}_n^2}$$

En donde \bar{x}_n y \bar{x}_o corresponden a la media y S_n y S_o a las desviaciones estándar de la variable del modelo nuevo y del original respectivamente. El factor $t_{r,\alpha}$ es utilizado para normalizar y corresponde a una t de student con r grados de libertad a un nivel de significancia de α .

El intervalo de Fieller es usado ampliamente en medicina. Se utiliza para medir la efectividad según el costo de tratamientos médicos, comparando el costo total al tratar a dos poblaciones distintas, con y sin el tratamiento en cuestión. [2]

Fieller Stability Measure establece el intervalo en el cual se mueve la variable. Sin embargo para poder evaluarlas se necesitan límites para estos intervalos de manera de establecer cuando éstas presentan cambios que se salen de la norma. Este método consiste en limitar los valores que puede tomar este intervalo mediante la máxima desviación de los coeficientes betas respectivos. Cada variable v tiene un intervalo dado por:

$$\beta_v \in [\widehat{\beta}_v - \alpha\sigma_{\widehat{\beta}_v}, \widehat{\beta}_v + \alpha\sigma_{\widehat{\beta}_v}]$$

En donde α a una significancia del 95% toma el valor de 1,96 y $\sigma_{\widehat{\beta}_v}$ es la desviación estándar de $\widehat{\beta}_v$. Y normalizando este resultado se puede obtener un intervalo que limite DSM_l, DSM_u :

$$MD_l = \frac{\widehat{\beta} - 2\sigma_{\widehat{\beta}}}{\widehat{\beta}}, \quad MD_u = \frac{\widehat{\beta} + 2\sigma_{\widehat{\beta}}}{\widehat{\beta}}$$

El método funciona verificando si es que el intervalo de confianza para la variable (DSM_l, DSM_u) se sale o no de los límites establecidos por (MD_l, MD_u) [4]. Si es que lo hace quiere decir que un cambio es detectado y la variable debe ser estudiada para la re-calibración del modelo.

Capítulo 3

Tratamiento de los Datos

La base de datos con la que se trabaja se limita a los microempresarios de los sectores de transporte, comercio, servicios profesionales y manufactura, dejando de lado la agricultura y la pesca. Se cuenta con información desde Enero del 2010 hasta Septiembre del 2012.

3.1 Selección y Preprocesamiento

Para realizar la extracción de los datos se accede directamente a la base de datos de la institución, y a través de una consulta en MySQL¹ se extraen todos los datos que cumplan con la condición necesaria para ser tratados, es decir, que tengan operaciones cursadas. De aproximadamente 500.000 movimientos registrados se obtienen 83.137 que cumplen con estas condiciones, de donde el 5,5% corresponde a clientes que caen en default. Los únicos datos nulos que posee la muestra corresponden a valores faltantes en la marca de clase que indica si el cliente cayó en default o no. Son 1088, que equivale a un 1,3% del total, por lo tanto, dado el bajo porcentaje que representan de la muestra total son eliminados. El detalle de lo que queda finalmente está en la Tabla 2.

Tabla 2: Datos Extraídos que Cursan Operaciones

Mes	Clientes	Mes	Clientes	Mes	Clientes
ene-10	2938	dic-10	4291	nov-11	2305
feb-10	2700	ene-11	3138	dic-11	2349
mar-10	2659	feb-11	0	ene-12	1801
abr-10	2685	mar-11	3612	feb-12	1481
may-10	2813	abr-11	3091	mar-12	1926
jun-10	2834	may-11	3249	abr-12	1348
jul-10	3152	jun-11	3122	may-12	1300
ago-10	3363	jul-11	2885	jun-12	1597
sep-10	3289	ago-11	3056	jul-12	1415
oct-10	2973	sep-11	2567	ago-12	1946
nov-10	3602	oct-11	2048	sep-12	1600

¹ Gestor de Base de Datos que utiliza el lenguaje de consultas SQL.

Capta la atención el hecho de que no exista ningún registro en Febrero 2011. Esto probablemente sucede debido al criterio de exclusión utilizado, ya que en el mes de Febrero todas las operaciones que se cursaron eran del tipo excluyente, es decir los clientes antiguos no contaban con movimientos. Esto se debe a que a Febrero 2011 pertenecen los registros para los cuales no se pudo calcular el criterio (cliente bueno o malo) debido a la falta de datos y que tuvo que eliminarse luego (los 1088 registros mencionados en el párrafo anterior). Así mismo puede verse que el número de registros decrecen a medida que el tiempo avanza, esto también se debe a que el criterio de exclusión sólo aplica a los movimientos de los clientes antiguos.

3.2 Transformación

Dado que el objetivo de esta memoria es evaluar el comportamiento del modelo actual, se eliminan todo el resto de variables que no están en él, así como también el ID de cada individuo y sus datos demográficos. La muestra final consiste en once variables que conforman el modelo que la institución utiliza actualmente. Cinco variables de deuda, dos variables demográficas y cuatro relacionadas con la operación.

Dado que no se cuenta con los datos originales con los que se realizó el modelo, se deben convertir en dummies todas las variables. Esto, ya que a pesar de que no estén los datos con los que se entrenó el modelo original, si se tienen los coeficientes beta de éste, pero de las once variables como dummies. Por lo tanto para que sean comparables, estas variables también deben convertirse en dummies.

Estas once variables se convirtieron en variables dummies, quedando en total 27 variables. En la tabla 3 puede verse que las letras A, B, C y D indican la pertenencia al mismo grupo, es decir forman parte de la variable original. Por ejemplo en el caso de Variable Operación 1, los tramos considerados hacen que existan tres variables dummies, Variable Operación 1A, Variable Operación 1B y Variable Operación 1C.

Se observa el detalle de sus desviaciones estándar, medias y fuentes de procedencias en la Tabla 3. Es importante que las variables se encuentren por tramos ya que es así como se comporta la población, por ejemplo, el comportamiento de una persona de 28 años no difiere mayormente del de una de 29.

Tabla 3: Estadísticos Descriptivos de las Variables

Variables Dummies	Desviación Estándar	Media	Fuente
Variable Operación 1A	0,485	0,621	Artificial
Variable Operación 1B	0,365	0,158	Artificial
Variable Operación 1C	0,249	0,067	Artificial
Variable Deuda 1A	0,362	0,155	Artificial
Variable Deuda 1B	0,466	0,320	Artificial
Variable Deuda 1C	0,403	0,204	Artificial
Variable Deuda 1D	0,353	0,146	Artificial
Variable Demográfica 1	0,459	0,302	Interna
Variable Deuda 2	0,462	0,310	Sistema Financiero
Variable Deuda 3A	0,438	0,259	Sistema Financiero
Variable Deuda 3B	0,414	0,220	Sistema Financiero
Variable Deuda 3C	0,381	0,176	Sistema Financiero
Variable Deuda 3D	0,358	0,151	Sistema Financiero
Variable Deuda 4A	0,271	0,080	Sistema Financiero
Variable Deuda 4B	0,314	0,111	Sistema Financiero
Variable Deuda 4C	0,500	0,492	Sistema Financiero
Variable Operación 2A	0,487	0,388	Artificial
Variable Operación 2B	0,296	0,097	Artificial
Variable Operación 2C	0,159	0,026	Artificial
Variable Operación 3A	0,494	0,424	Artificial
Variable Operación 3B	0,401	0,201	Artificial
Variable Demográfica 2A	0,449	0,280	Interna
Variable Demográfica 2B	0,310	0,108	Interna
Variable Demográfica 2C	0,116	0,014	Interna
Variable Demográfica 2D	0,066	0,004	Interna
Variable Operación 4	0,422	0,232	Interna
Variable Deuda 5	0,324	0,119	Sistema Financiero

3.3 Evaluación

Dada la poca cantidad de meses con los que se cuenta, las carteras a estudiar serán de frecuencia mensual, para así tener más meses con los que se pueda comparar. El conjunto de entrenamiento utilizado para el test Stability Index está compuesto desde Enero a Septiembre del 2010. Bajo el supuesto que esta muestra es mucho más similar, por la cercanía en tiempo al modelo original, que el resto de los meses. Luego, desde Octubre del 2010 a Septiembre 2012 cada mes representa una cartera. Las regresiones logísticas realizadas son hechas a través del software Rapid Miner.

Capítulo 4

Metodología para Anticipar la Re-calibración

Con el fin de determinar los pasos para conocer de antemano el momento de re-calibración se definen el problema, los tipos de cambios a medir y cómo reconocer el momento para realizar el ajuste del modelo.

4.1 Planteamiento del Problema

Los modelos de Credit Scoring se tratan como estacionarios cuando en realidad son dinámicos. Esto hace que a medida que avanza el tiempo las estimaciones que entrega el modelo pierdan poder predictivo. Esto se debe a cambios que se presentan en la población, que pueden provocar pérdida de la capacidad discriminante de las variables que componen el modelo o cambios en su distribución.

Una de las formas que actualmente se utiliza para medir la calidad del modelo es el test K-S, que cuantifica la divergencia existente entre la distribución de los clientes buenos y los malos. Es decir la diferencia, para un mismo puntaje de corte, de la proporción acumulada de los clientes buenos aceptados y la de los malos aceptados. Idealmente el valor entregado por el estadístico K-S para el puntaje de corte que maximiza esta diferencia se asemeja al puntaje de corte del modelo. A medida que avanza el tiempo el estadístico K-S se empieza a degradar, lo que implica que para el puntaje de corte óptimo, que maximiza la diferencia entre clientes buenos y malos, la diferencia es cada vez más pequeña volviendo más difícil discernir entre aquellos clientes que sí cumplirán su obligación y los que no. Es por esto que se hace necesario establecer una alerta que en base al comportamiento temprano del K-S indique cuando el modelo debe ser recalibrado.

Por otro lado, el cambio que ocurre puede medirse desde dos puntos de vista, la existencia de *concept drift* en alguna de las variables y la pérdida económica que se presenta por aceptar clientes malos y rechazar clientes buenos. Para saber si es que existe *Concept Drift* se utiliza Feller Stability Measure en cada una de las variables, y si para alguna de ellas su intervalo de confianza se sale de los límites establecidos por el modelo indica la existencia de cambio en esa variable específica.

Existen dos costos involucrados al equivocarse gracias al modelo. Son los costos de aceptar un cliente malo y el de rechazar a un cliente bueno. La cantidad de clientes malos que fueron aceptados es fácil de obtener ya que se cuenta con la información de quienes no cumplieron, en cambio la de los clientes buenos que fueron rechazados no se posee de manera directa, por lo que la pérdida que se considera cuenta sólo a los clientes malos que fueron aceptados.

4.2 Test Supervisado

Como se mencionó anteriormente, el test K-S sirve para saber qué tan bien discrimina el modelo a los clientes buenos de los malos. Por lo que se busca establecer una relación entre pérdida/aumento porcentual del estadístico K-S y el aumento/disminución de la pérdida económica de la cartera. Para poder medir esta última, la pérdida sólo considera el costo por aceptar un cliente malo, ya que se desconoce el comportamiento futuro de quienes fueron rechazados inicialmente impidiendo saber si es que hubiesen resultado ser buenos clientes o no.

4.2.1 Evaluación

No se cuenta con el monto solicitado por cada microempresario por lo que se utilizan los datos de la solicitud promedio que se maneja. Bajo el supuesto de que el crédito promedio solicitado es de \$3.000.000 y que la exposición ante el no pago de éste es del 85%² la pérdida que se tiene por cartera asciende a:

$$Pérdida = \$3.000.000 \cdot N^{\circ}malos(s) \cdot 0,85$$

En donde "s" representa el puntaje de corte utilizado. En este caso corresponde al puntaje de corte que maximiza el K-S. Para generar la relación entre ambas el crecimiento del K-S se obtiene mediante:

$$100\% - \frac{KS_t}{KS_{t-1}}$$

En donde KS_t corresponde al K-S del mes t y KS_{t-1} al K-S del mes t-1. La disminución de la pérdida económica por punto porcentual de KS se calcula:

² Tanto el monto solicitado como el porcentaje de pérdida se basan en promedios de los datos que posee la institución.

$$\frac{Pérdida_t/Total Clientes_t}{Pérdida_{t-1}/Total Clientes_{t-1}} - 100\%$$

Donde $Pérdida_t$ representa la pérdida económica en el mes t y $Total Clientes_t$ a la cantidad de clientes tanto buenos como malos en el mes t . $Pérdida_{t-1}$ y $Total Clientes_{t-1}$ corresponden respectivamente a la pérdida económica y al total de clientes en el mes $t-1$. En la Tabla 4 puede observarse esta relación.

Tabla 4: Crecimiento K-S Vs Disminución Pérdida Económica

Mes	Crecimiento K-S	Disminución de la pérdida por punto de K-S
oct-10	14%	72%
nov-10	-7%	-31%
dic-10	-30%	-18%
ene-11	-1%	-17%
mar-11	0%	5%
abr-11	16%	84%
may-11	4%	-53%
jun-11	-36%	-30%
jul-11	14%	159%
ago-11	-10%	-34%
sep-11	15%	33%
oct-11	8%	-9%
nov-11	-26%	-4%
dic-11	-4%	-9%
ene-12	2%	-8%
feb-12	7%	22%
mar-12	-1%	15%
abr-12	-38%	-56%
may-12	18%	-4%
jun-12	-6%	50%
jul-12	4%	-11%
ago-12	7%	19%
sep-12	27%	65%

4.2.2 Estimación de Medidas

En el gráfico 1 puede observarse que el K-S tiende a mantenerse en los mismos valores, no se muestra una disminución sostenida como se esperaría. Lo que indica que las carteras son relativamente estables. Esto hace que no sea posible predecir el comportamiento del K-S ya que no

presenta ninguna tendencia, pero si permite que se pueda establecer una relación entre el K-S y la pérdida.

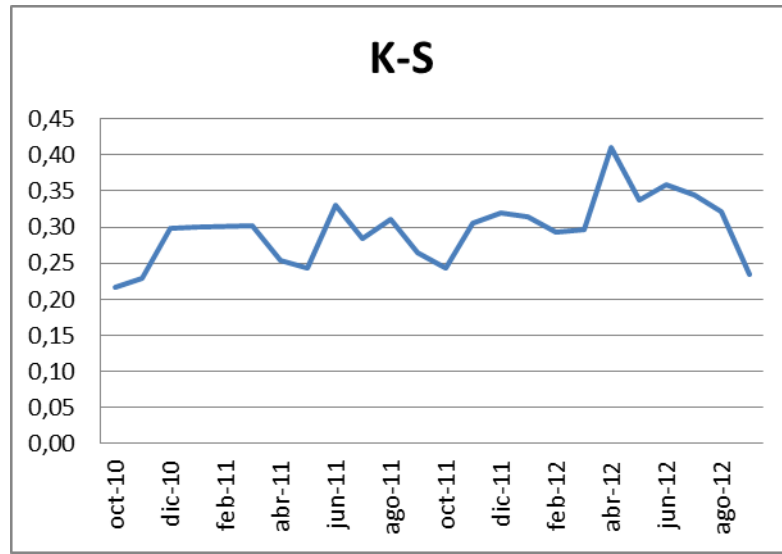


Figura 1: K-S en el tiempo

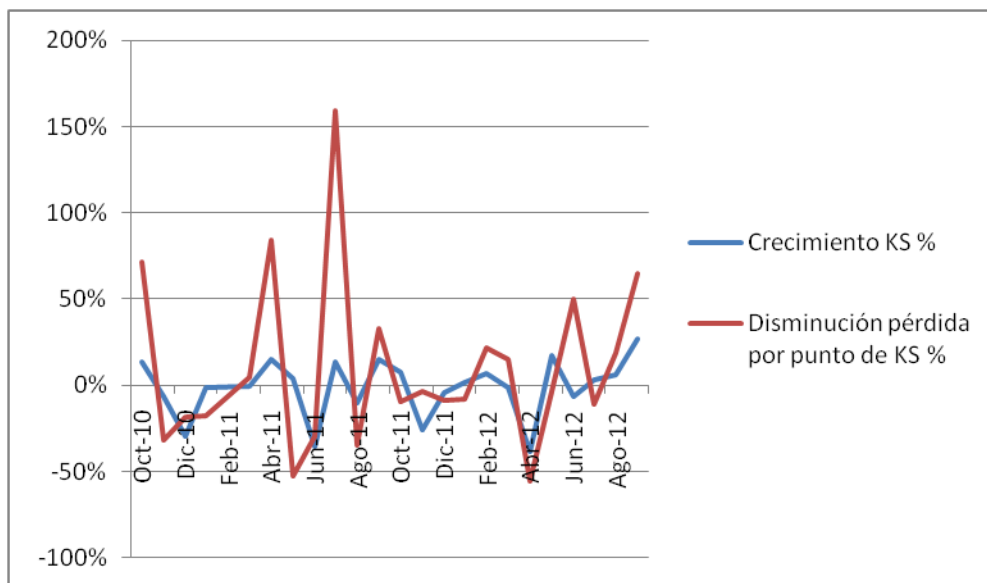


Figura 2: K-S Vs Pérdida Económica

Al representar esta relación puede verse en el gráfico 2 que aunque el cambio porcentual de K-S subestima la pérdida, de igual manera captura su tendencia. Es por lo que se calculó el coeficiente de correlación entre la pérdida de todas las carteras y los distintos K-S que presentaban y además se realizó una regresión lineal con el fin de poder establecer esta dependencia. De donde se obtiene:

$$\text{Disminución Pérdida} = 0,122041433 + \text{Crecimiento KS} \cdot 1,66987054$$

Esta regresión indica que por cada punto porcentual que disminuye el estadístico K-S de un periodo a otro, la pérdida económica aumenta un 1,67%.

Aplicando esta regresión a los datos conocidos, el error que ésta presenta en los 24 meses de información que se posee equivale a un promedio de \$67.133 por cliente y a \$212.938.848 del total.

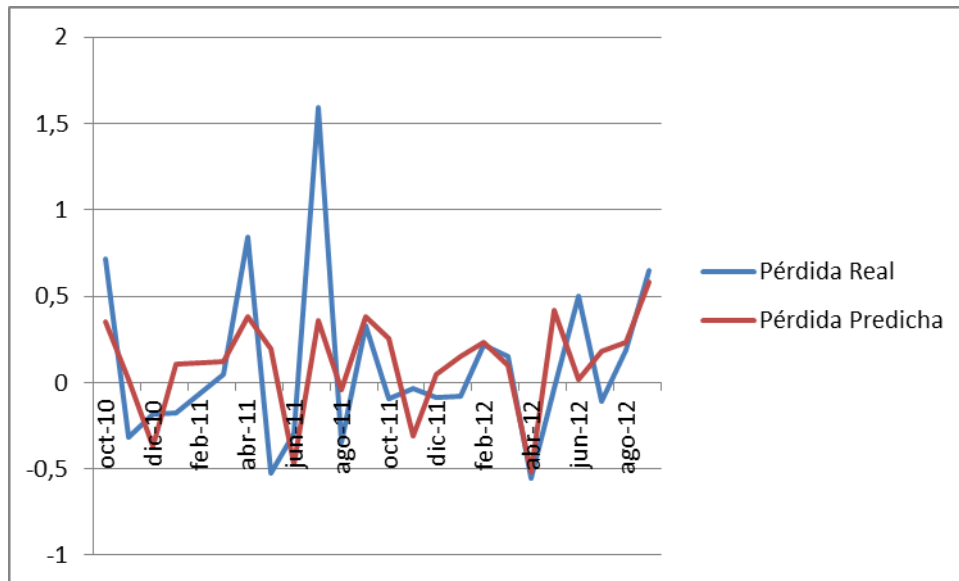


Figura 3: Disminución Pérdida Real Vs Disminución Pérdida Predicha

La pérdida real asciende a \$953.700.000, versus los \$740.761.125 que pueden verse en el gráfico 3, o sea está subestimando la pérdida un 22%.

4.3 Test No Supervisado

Así como se busca generar una alerta para los clientes de los que se posea información anterior de su comportamiento, también se necesita saber con anterioridad cómo será el comportamiento de quienes no poseen historia. Para poder hacerlo hace falta un test que no requiera conocer la salida real de cada uno de los clientes de la cartera, es decir, que no requiera conocer si el cliente es bueno o malo. El test Fieller Stability Measure cumple con esta condición.

4.3.1 Estimación de la Alerta

Lo primero corresponde a calcular el intervalo de Fieller para todas las variables. Como este contempla la desviación de la muestra original y de la muestra observada, nuevamente la muestra original será la de los datos correspondientes a Enero del 2010 hasta Septiembre del mismo año, recordando que no se poseen los datos originales. En base a ello, los intervalos obtenidos para cada variable pueden verse en el Anexo B. A continuación una variable estable versus una que se sale de los límites:

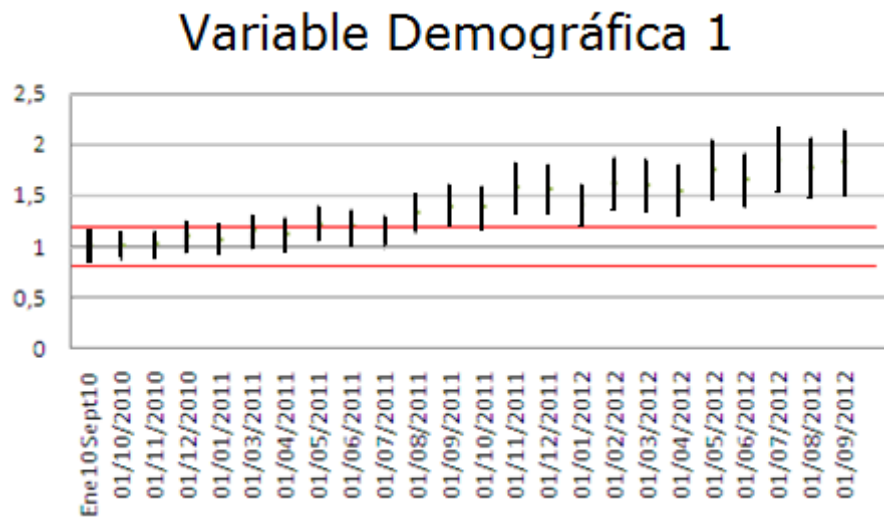


Figura 4: Variable Demográfica 1

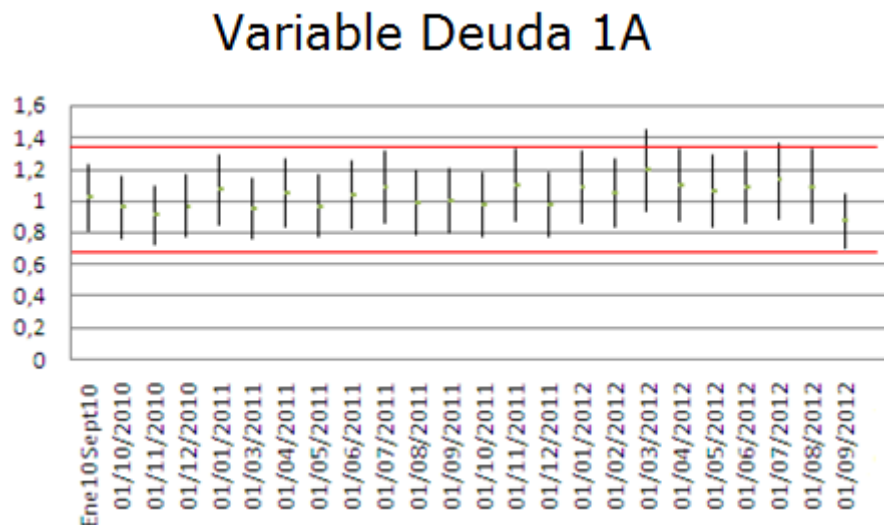


Figura 5: Variable Deuda 1A

La mayoría de las variables tienen el comportamiento de la Variable Deuda 1^a que puede verse en el gráfico 5. Son cinco las variables que se salen de

los límites del modelo en algún instante, al igual que la Variable Demográfica 1. Puede verse en el gráfico 4 que el primer mes que el intervalo de Fieller de la variable cruza los límites establecidos por el modelo es en Agosto 2011. Y que además en todas las variables en donde ocurre, cada vez que sus intervalos salen de los límites del modelo una o dos veces vuelven a tomar estabilidad. Por lo que la alerta de acción que se testea es que la variable se escape de los intervalos durante tres meses seguidos, y la alerta de precaución es salirse uno o dos meses.

Para comprobar la hipótesis establecida en el punto anterior se realiza una re-calibración en el mes de Octubre 2011, con la cartera del mismo mes, ya que este es el tercer mes que para la Variable Demográfica 1 su intervalo se encuentra fuera de los límites del modelo. Al recalibrar el modelo, de las 27 variables dummy que se tenían, quedan once (quedan tres grandes variables de las once originales), ya que todas las demás ya no son significativas. Entre ellas se va la Variable Demográfica 1 que justamente era la que indicó realizar la re-calibración. El criterio utilizado para dejar variables en el modelo es que el estadístico del test de Wald [7] sea mayor a dos para alguna de las variables dummie que conforma los grandes grupos (las cinco de deuda, dos demográficas y cuatro de operación).

Al llevar a cabo las re-calibraciones considerando el hecho de que la muestra en juego presentaba un gran desbalance con respecto a clientes buenos y malos (95% y 5%) se ponderó el peso de los malos para así hacer el modelo más balanceado. Lo importante de la asignación de pesos a las variables radica en que dado el desequilibrio que existe en la muestra entre clientes buenos y malos, la regresión logística podría decidir hacer un modelo en el cual prediga que todos son buenos clientes y tenga un error de sólo un 5%.

Luego, para verificar que efectivamente la re-calibración no debe realizarse antes de la alerta de acción que corresponde a tres meses seguidos saliendo de los límites del modelo, se recalibra el modelo donde se presentan alertas de precaución, es decir, cuando los intervalos cruzan los límites una o dos veces, con el fin de observar si es que efectivamente compensaba el costo de la re-calibración.

Como el modelo no se ha recalibrado nunca, en general cualquier calibración que se realice, aunque la alerta corresponda sólo a un mes, hará que disminuyan las pérdidas, en menor proporción que en el caso de cuando la alerta corresponde a tres meses, pero de todas maneras indica que se debiese recalibrar. Como se busca que esto no pase sino que las re-calibraciones se lleven a cabo solamente cuando realmente valga la pena no

se utiliza el modelo original para comparar, sino un modelo entrenado con los datos desde Enero a Noviembre 2010 que contempla una situación más estable para los meses que quedan, haciéndolos realmente comparables.

4.3.2 Costo Económico de la Re-calibración

Existen dos costos asociados a la re-calibración del modelo. Estos son, el reajustar el peso de alguna o algunas variables, y, tener que agregar o quitar variables. Hay que considerar que este ajuste puede realizarse *in-house* o por alguna institución externa, en donde los costos ascienden a por lo menos 1.300.000 dólares, es decir alrededor de \$650.000.000. Para este caso en particular se consideran los costos asociados a una re-calibración realizada in-house.

El costo por reajustar alguna de las variables del modelo alude al trabajo de tres meses aproximadamente. Esto considera tres ingenieros trabajando full time con al menos tres años de experiencia, esto implica un sueldo bruto de \$1.900.000³ cada uno, es decir \$17.100.000 en total. En caso de tener que agregar o quitar variables la operación se vuelve más extensa, y con la misma fuerza laboral se tarda aproximadamente seis meses, con lo que el costo ascendería a \$34.200.000.

³ Fuente: El salario de un Ingeniero Civil Industrial de la Universidad de Chile al cuarto año de egreso varía entre \$1.800.000 a \$2.000.000. Fuente: Futuro Laboral.

Capítulo 5

Resultados

5.1 Fieller Stability Measure

Puede verse en la tabla 5 que con este nuevo modelo la pérdida que se tenía desde noviembre 2011 a septiembre 2012 se reduce de \$953.700.000 a \$897.600.000. Así mismo el K-S que se tenía en noviembre 2011 antes de la re-calibración era 0,25 y al recalibrar aumentó a 0,26.

Tabla 5: Pérdidas con y sin Re-calibración Octubre 2011

Carteras	Sin Re-calibración	Re-calibración Octubre 2011
nov-11	112.200.000	170.850.000,00
dic-11	104.550.000	114.750.000,00
ene-12	76.500.000,00	56.100.000,00
feb-12	76.500.000,00	56.100.000,00
mar-12	114.750.000,00	63.750.000,00
abr-12	35.700.000,00	38.250.000,00
may-12	85.425.000,00	40.800.000,00
jun-12	61.200.000,00	71.400.000,00
jul-12	48.450.000,00	56.100.000,00
ago-12	131.325.000,00	117.300.000,00
sep-12	107.100.000,00	89.250.000,00
Total	953.700.000,00	874.650.000,00

Al utilizar los datos hasta Octubre 2010 para realizar la re-calibración las variables que quedan son las de operación 1, deuda 1, demográfica 1, deuda 3, deuda 4, operación 2 y operación 3. Siete de las once originales. Con este nuevo modelo las variables se comportan con relativa estabilidad excepto variable deuda 1D y variable operación 2C, en donde justamente se da el tipo de alerta que se quiere evaluar.

Luego para comprobar la efectividad de la alerta de precaución, es decir cuando los intervalos se salen de los límites solamente un mes, se utilizan las variables después de ser calibradas en noviembre 2010 como se explicó anteriormente.

Variable Deuda 1D

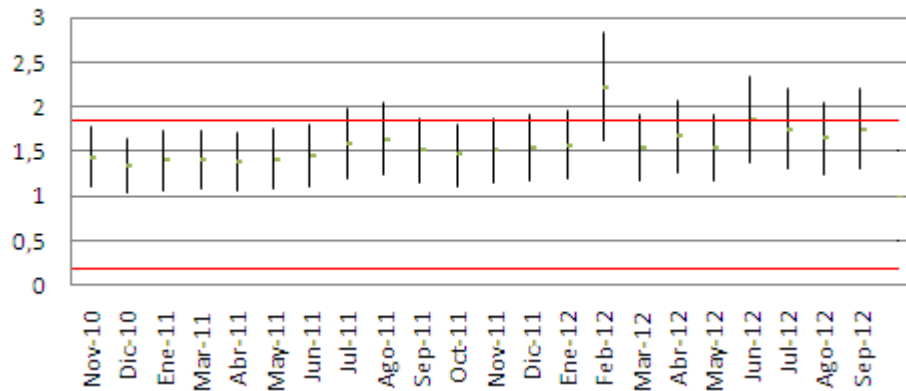


Figura 6: Variable Deuda 1D

Puede verse en la figura 6 que en el mes de Febrero hay una alerta de precaución ya que el intervalo se sale de los límites, pero sólo durante un mes. Es importante recordar que en el día a día esta información futura no se tendrá, lo único que se sabrá es que en el mes de Febrero el intervalo se salió de los límites, pero no se sabe si se va a seguir manteniendo esta tendencia. Es por ello que se probará qué ocurre si es que se calibra ante solamente un mes como alerta.

Variable Operación 2C

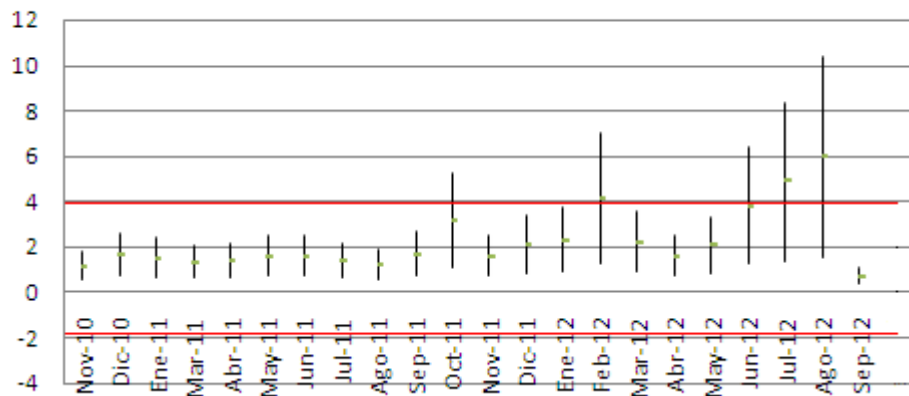


Figura 7: Variable Operación 2C

Luego en el segundo caso, en la figura 7, quiere verse qué pasa si se calibra el modelo ante una salida de los límites de la media del intervalo. Lo que ocurre en los meses de Julio, Agosto y Septiembre. Para ver qué pasa en cada uno de ambos casos se recalibra el modelo en Febrero 2012 con los datos del mismo mes y luego se ve que ocurre desde Marzo 2012 en adelante tanto en la situación con re-calibración como en la sin re-

calibración. Lo mismo para el segundo caso, donde se recalibra con los datos de Julio para ver qué ocurre con la pérdida de Agosto y Septiembre 2012 en ambos casos.

Al recalibrar el modelo con los datos hasta Febrero 2012 quedan sólo las cuatro Variables de deuda. La pérdida que resulta desde Marzo 2012 a Septiembre 2012 es de \$757.350.000, en cambio sin realizar esta re-calibración \$405.000.000. La pérdida aumentó al realizar la re-calibración cuando sólo se esperó un mes para llevarla a cabo. El detalle se observa en la tabla 6.

Tabla 6: Pérdidas con y sin Re-calibración Febrero 2012

Carteras	Sin Re-calibración	Re-calibración Febrero 2012
mar-12	81.600.000	163.200.000
abr-12	17.850.000	15.300.000
may-12	58.650.000	84.150.000
jun-12	86.700.000	155.550.000
jul-12	40.800.000	96.900.000
ago-12	91.800.000	135.150.000
sep-12	28.050.000	107.100.000
Total	405.000.000	757.350.000

En el caso de la re-calibración del modelo con los datos hasta Junio 2012 ocurrió lo mismo. Al realizar la re-calibración la pérdida aumentó de \$160.650.000 a \$323.850.000 como puede observarse en la tabla 7.

Tabla 7: Pérdidas con y sin Re-calibración Junio 2012

Carteras	Sin Re-calibración	Re-calibración Junio 2012
jul-12	40.800.000	71.400.000,00
ago-12	91.800.000	130.050.000,00
sep-12	28.050.000	122.400.000,00
Total	160.650.000	323.850.000

5.2 Stability Index

Para tener una comparación con otro test no supervisado conocido se calculó el Stability Index. Es importante mencionar que dado que no se tienen los datos con los que se realizó el modelo original, el Stability Index tampoco pudo calcularse con estos datos. Para obtener su valor se utilizaron los datos

de entrenamiento del conjunto de Enero 2010 a Septiembre 2012, por lo que los resultados que pueda indicar para este apartado no son exactos.

Para ambos casos, en las variables específicas en las que se presentaba la alerta, el Stability Index indicó cambio total ya que se presenta un valor mayor a 0,25. En el caso de la variable "Variable Deuda 1D" en Febrero es 1,7 y para la variable "Variable Operación 2C" en Junio de 2,84. De donde pueden concluirse dos posibles causas para esta situación. El hecho de que no se cuente con los datos originales puede hacer que con los datos utilizados se esté sobreestimando el valor del Stability Index y esto es lo que provoca que aunque Fieller Stability Measure indique sólo una alerta, el Stability Index proponga cambio total. La otra razón que podría estar provocando estos resultados es que el Stability Index es mucho más sensible y detecta cambios enseguida aunque no sean significativos en el futuro, por lo que no es un buen indicador para este tipo de problemas.

5.3 Análisis de Resultados

Es importante recordar que en el día a día no se conoce el comportamiento futuro de las carteras, por lo que cobra gran relevancia poder distinguir entre una alerta en la que se deba recalibrar y otra en la que sólo se debe monitorear para observar que pasa a continuación y recién en ese momento decidir si la re-calibración debe llevarse a cabo o no. Esta diferencia puede verse en los tres casos presentados, en Octubre 2011, Febrero 2012 y Junio 2012.

En el caso de la primera re-calibración la disminución de pérdida obtenida por ajustar el modelo en Octubre 2011, versus dejarlo sin ningún cambio, es \$56.100.000, lo que dado los costos que implica realizar una re-calibración conviene realizar. Así mismo en el caso de la re-calibración en los meses en donde el intervalo de Fieller se salía de los límites sólo una vez, Febrero 2012 y Junio 2012, la re-calibración incluso aumentó la pérdida en ambos casos, por lo que la re-calibración no debiese realizarse en los casos en que sólo exista una alerta de precaución. De todas maneras con estos resultados se comprueba que calibrar en estos casos donde no existe una alerta de acción termina empeorando la situación, lo que resulta totalmente lógico, ya que si eventualmente el intervalo iba a volver a estar dentro de los límites del modelo se comprueba que era sólo un mes *outlier* y que no hacía falta recalibrar todo el modelo sólo por esa anomalía específica.

Aunque tanto el intervalo de Fieller como el Stability Index son ambos no supervisados, se puede ver empíricamente que no arrojan los mismos

resultados. Cuando en Fieller sólo se recomienda recalibrar en la minoría de los meses, que es donde las variables comienzan a mostrar un comportamiento fuera de la norma, en Stability Index se detecta cambio en 24 de las 27 variables.

Como puede verse en la tabla 8 el Stability Index sólo se comporta adecuadamente en cuatro de las veintisiete variables, donde indica que ha habido un cambio mínimo o que no existe cambio.

Tabla 8: Detección del Cambio por parte de los índices para cada variable

Variables	Stability Index	Fieller Stability Measure
Variable Operación 1A	No	No
Variable Operación 1B	Si	No
Variable Operación 1C	Si	No
Variable Deuda 1A	Si	No
Variable Deuda 1B	Si	No
Variable Deuda 1C	Si	No
Variable Deuda 1D	Si	Si
Variable Demográfica 1	Si	Si
Variable Deuda 2	Si	No
Variable Operación 2A	No	No
Variable Operación 2B	Si	No
Variable Operación 2C	Si	Si
Variable Demográfica 2A	Si	No
Variable Demográfica 2B	Si	No
Variable Demográfica 2C	Si	No
Variable Demográfica 2D	Si	No
Variable Operación 3A	No	No
Variable Operación 3B	Si	No
Variable Deuda 3A	Si	No
Variable Deuda 3B	Si	No
Variable Deuda 3C	Si	No
Variable Deuda 3D	Si	No
Variable Operación 4	Si	No
Variable Deuda 4A	Si	No
Variable Deuda 4B	Si	No
Variable Deuda 4C	No	No
Variable Deuda 5	Si	No

La tabla 8 muestra para que variables el Stability Index y Fieller Stability Measure detectan cambio. Para el caso del Stability Index "No" se indica en la tabla cuando al menos para 14 carteras el índice se mantiene dentro de los límites aceptables para la variable, es decir, es menor a 0,25, "Si" en

caso contrario. En el caso de Fieller Stability Measure "SI" se muestra en la tabla cuando el intervalo de confianza de Fieller traspasa los límites establecidos por el modelo.

Para conocer mayor detalle de cada uno de los valores de SI para todas las variables en todos los períodos revisar Anexo E.

Capítulo 6

Conclusiones

Las conclusiones de la presente memoria abarcan la metodología utilizada para anticipar la re-calibración, el cumplimiento de los objetivos y el trabajo futuro que puede nacer a partir de aquí.

6.1 Acerca de la Metodología para Anticipar la Re-calibración

La alerta de acción que se establece, recalibrar luego de tres meses en que alguna variable tuviese su intervalo fuera de los límites, fue acertada, ya que se demostró que se disminuyó la pérdida y que además la disminución era suficiente como para asumir posibles costos de re-calibración. La alerta de precaución también fue efectiva ya que se demostró que al recalibrar el modelo ante solamente un mes en que el intervalo se salía de los límites establecidos, la pérdida no tan sólo no disminuía sino que aumentaba, verificando el error que significaría recalibrar el modelo con la información de un mes que se escapaba de la normalidad.

Dependiendo de la aversión al riesgo que se tenga, puede ocuparse también el resultado entregado por la relación que se establece entre el test K-S y la pérdida que se genera. Se menciona la aversión al riesgo ya que el error que presenta este modelo es de un 22%. De todas maneras si es que se pretende considerar esta alerta además de las establecidas por Fieller Stability Measure, es importante tener en cuenta que esta predicción realizada por el K-S esté siempre siendo seguida en el tiempo y actualizándose ya que lo más probable es que el error que se cometa vaya creciendo a medida que el tiempo avanza.

6.2 Acerca de los Objetivos

Los objetivos de la memoria se cumplen. Se define una metodología para medir pérdidas por baja en la predicción de los modelos, ésta hace referencia a la pérdida que se sufre por cada punto de K-S que se pierde. Relación que fue encontrada pero con un error no despreciable, por lo que finalmente queda a criterio del lector si utilizarlo o no al momento de llevar a cabo la metodología.

Se definió también en base al test supervisado (test K-S) y al no supervisado (Fieller Stability Measure) el mejor momento para calibrar. El primero mencionado anteriormente y el segundo definido ante la presencia de una alerta de acción, que se define como el tercer mes en el que el intervalo de Fieller de una variable se ha salido de los límites del modelo.

Finalmente se evaluó el funcionamiento del proceso en el modelo de riesgo de crédito a microempresarios en una institución financiera real.

6.3 Recomendaciones a la Institución

Se recomienda en primer lugar repetir el experimento con los últimos meses de los que se posee información. Así se puede verificar si los datos siguen comportándose de manera similar a los tratados en esta memoria o hay que establecer algún nuevo tipo de alerta.

Es importante considerar que en esta memoria las carteras escogidas fueron de un mes, a causa de la escasa cantidad de datos disponibles. Por lo que resultaría muy importante realizar un análisis de sensibilidad respecto al tamaño de las carteras, pudiendo quedar trimestrales o quizás semestrales.

Una vez conocida la alerta de acción y precaución (en caso de que haya cambiado) se debe establecer el cálculo automático de los intervalos de Fieller y los límites respectivos para cada una de las variables. De esta manera al finalizar cada período se podrá saber de inmediato como es el comportamiento de las variables y se podrán efectuar los cambios necesarios a tiempo.

6.4 Trabajo Futuro

En un futuro se puede considerar trabajar con esta misma metodología, pero con los datos de los montos de crédito reales, de manera de poder hacer un seguimiento total de las carteras y ver cómo actúa la curva de pérdidas en relación al monto solicitado.

De la misma forma que en el caso anterior queda propuesto incluir en el cálculo de la pérdida esperada no sólo el valor de aceptar a un cliente malo, sino también el costo de oportunidad que se tiene al rechazar a clientes que eventualmente hubiesen resultado ser buenos.

En este trabajo se hicieron pruebas de re-calibración del modelo, no se intentó incluir variables nuevas ya que escapa del alcance de esta memoria.

Por lo tanto es posible incorporar que variaciones deberían hacerse o considerar en caso de querer hacer un trabajo más acabado al respecto reconstruyendo el modelo completo y no sólo ajustando el peso de las variables que ya se poseen.

Como se mencionó anteriormente en los alcances de esta memoria, dada la naturaleza de la metodología planteada, el proceso podría usarse en cualquier situación en donde se necesite predecir el comportamiento de algún grupo de individuos o de algún evento. Se recomienda una generalización del proceso que pueda aplicarse a cualquier situación similar. En esta memoria no se profundizó en ello, a causa de los alcances de ésta.

Bibliografía

- [1] A. BLÖCHLINGER y M. LEIPPOLD, «Economic Benefit of Powerful Credit Scoring,» *Journal of Banking & Finance* 30, 2005.
- [2] C. BRAVO y S. MALDONADO, «Detecting Significant Changes Using a Model-Dependent Dataset Shift Test,» 2013.
- [3] C. BRAVO, S. MALDONADO y R. WEBER, «Experiencias Prácticas en la Medición de Riesgo Crediticio de Microempresarios Utilizando Models de Credit Scoring,» *Ingeniería de Sistemas*, vol. XXIV, 2010.
- [4] C. BRAVO, S. MALDONADO y R. WEBER, «Granting and managing loans for micro-entrepreneurs: New developments and practical experiences.,» *European Journal of Operational Research* (2013), <http://dx.doi.org/10.1016/j.ejor.2012.10.040>.
- [5] G. CASTERMANS, D. MARTENS, T. VAN GESTEL, B. HAMERS y B. BAESENS, «An Overview and Framework for PD Backtesting and Benchmarking,» *Journal of the Operational Research Society*, 2010.
- [6] D. CIESLAK y N. CHAWLA, «Detecting Fractures in Classifier Performance,» *Seventh IEEE International Conference on Data Mining*, 2007.
- [7] C. GOURIÉROUX, A. HOLLY y A. MONF, «Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters,» *Econometrica*, vol. 50, 1982.
- [8] J. HAN y K. M., «Data Mining: Concepts and Techniques, » USA: Morgan Kaufmann Publishers, 2001.
- [9] F. J. MASSEY JR, «The Kolmogorov-Smirnov Test for Goodness of Fit,» *Journal of the American Statistical Association*, vol. 46, 1951.
- [10] N. SIDDIQI, «Credit Risk Scorecard, » New Jersey: Wiley, 2006.
- [11] A. TSYMBAL, «The problem of concept drift: definitions and related work,» *Trinity College Dublin, Ireland*, 2004.

Anexos

Anexo A

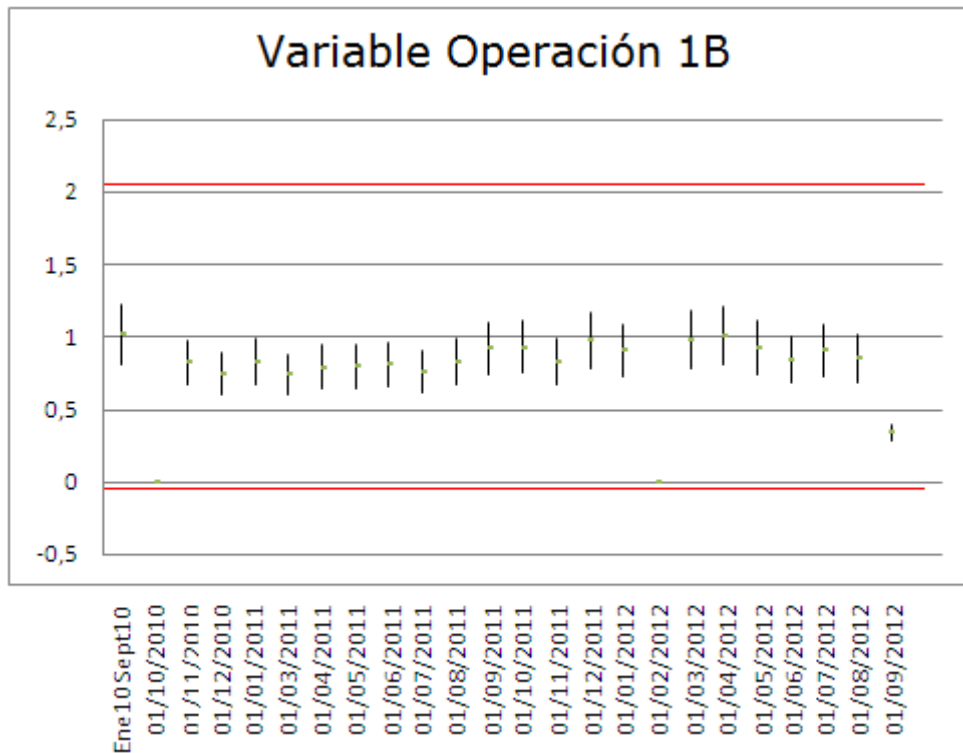
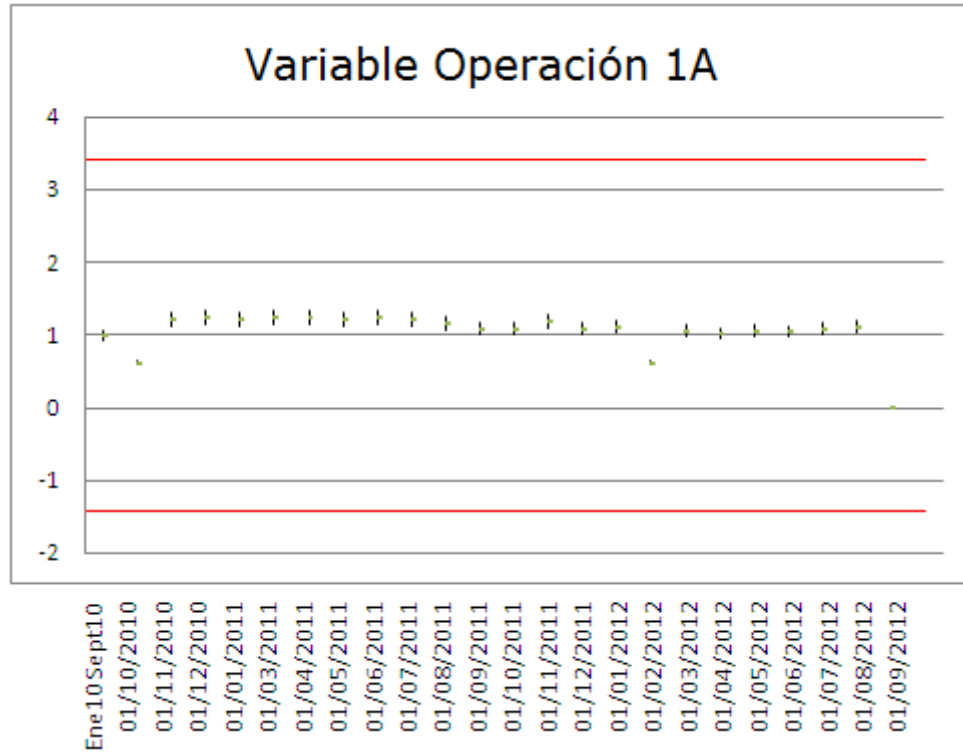
Relación entre el aumento porcentual en el estadístico K-S y la disminución porcentual de la pérdida económica.

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	0,122041433	0,085957387	1,419789932	0,170344374	0,056716739	0,300799604
KS	1,669870535	0,506739744	3,295321822	0,003447274	0,616047553	2,723693517

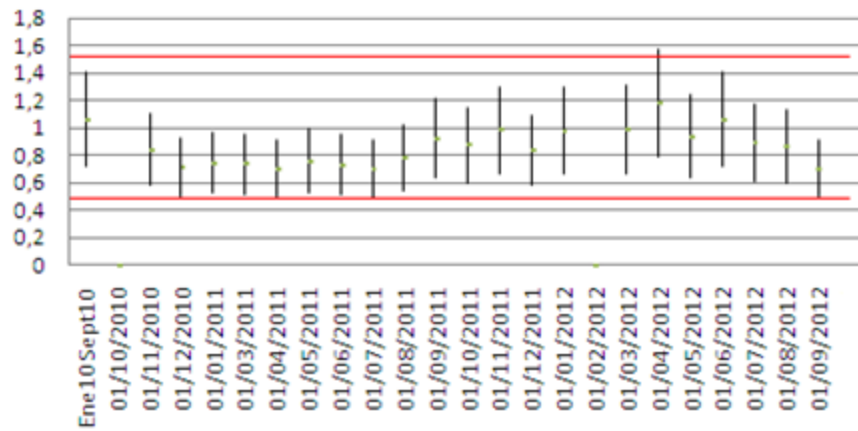
<i>Estadísticas de la regresión</i>	
Coeficiente de correlación múltiple	0,583822421
Coeficiente de determinación R ²	0,34084862
R ² ajustado	0,309460459
Error típico	0,411457342
Observaciones	23

Anexo B

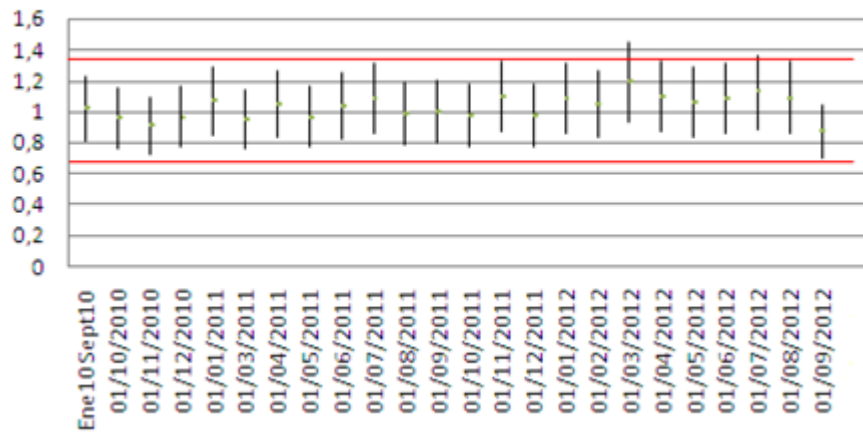
Fieller Stability Measure para todas las variables sin realizar ninguna recalibración del modelo



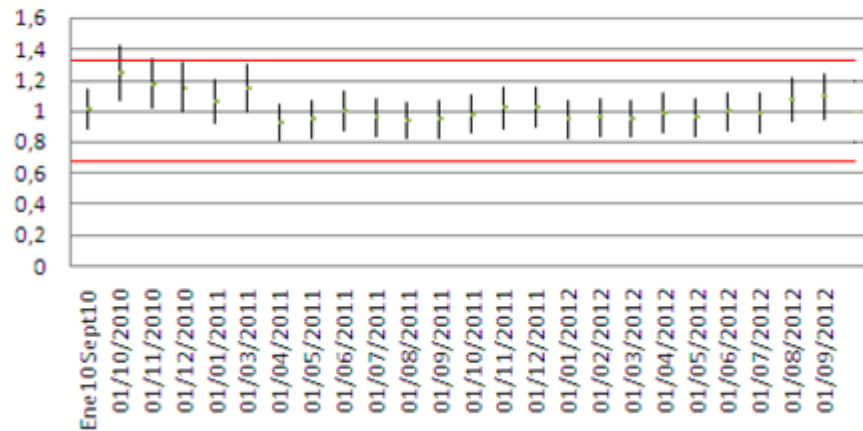
Variable Operación 1C



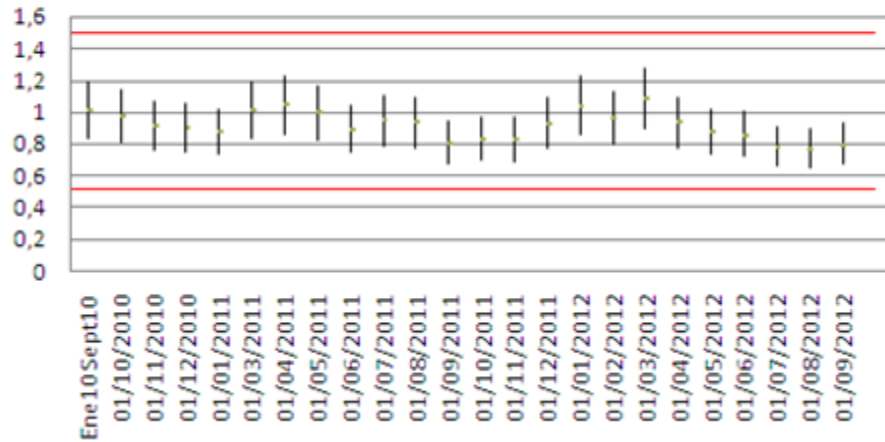
Variable Deuda 1A



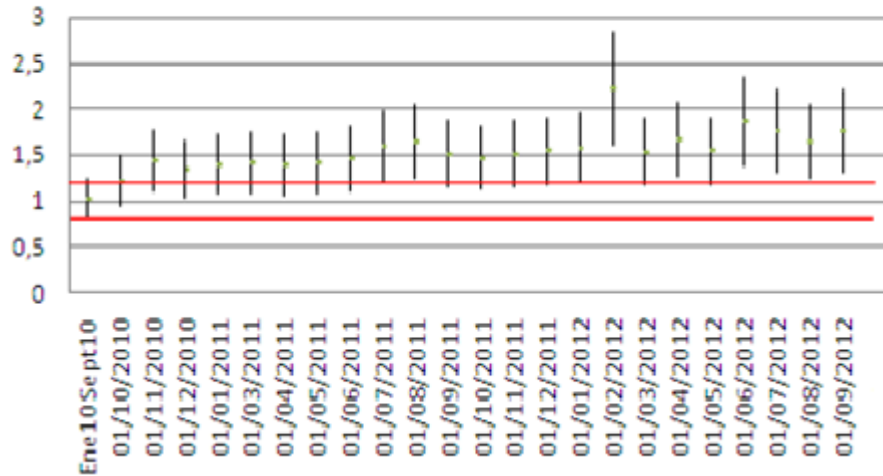
Variable Deuda 1B



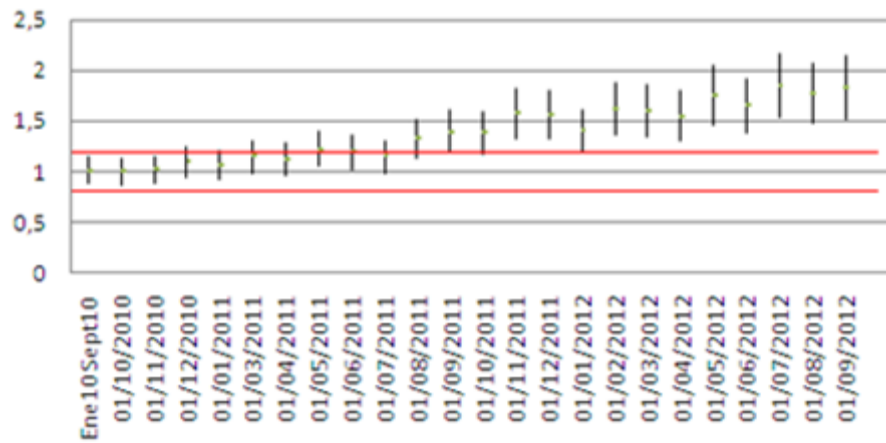
Variable Deuda 1C



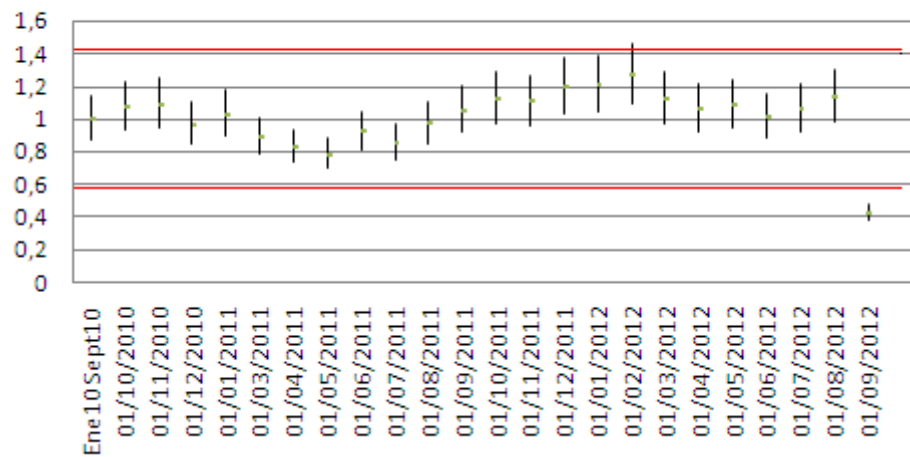
Variable Deuda 1D



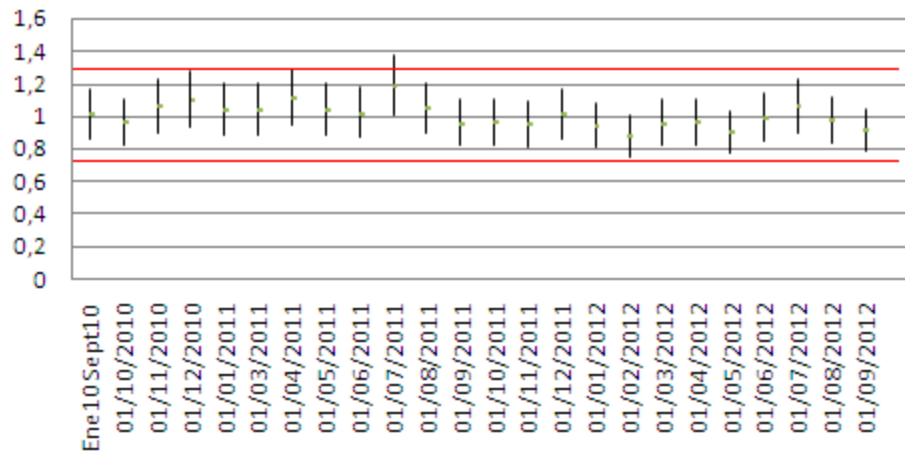
Variable Demográfica 1



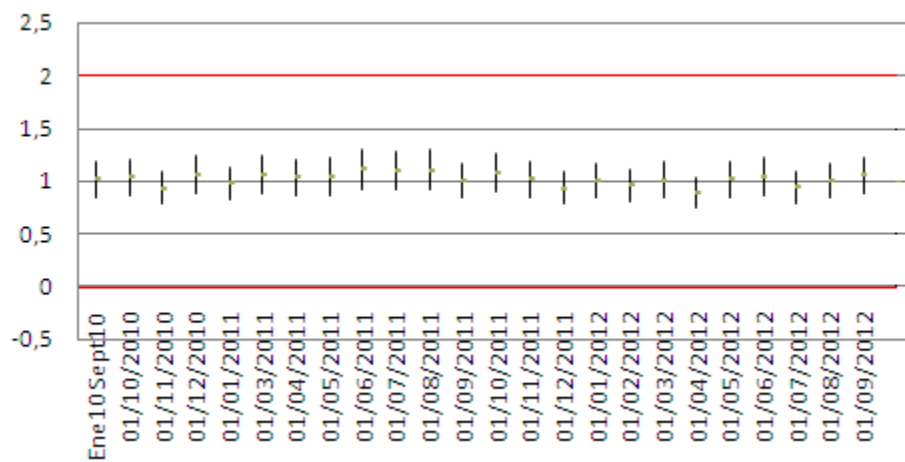
Variable Deuda 2



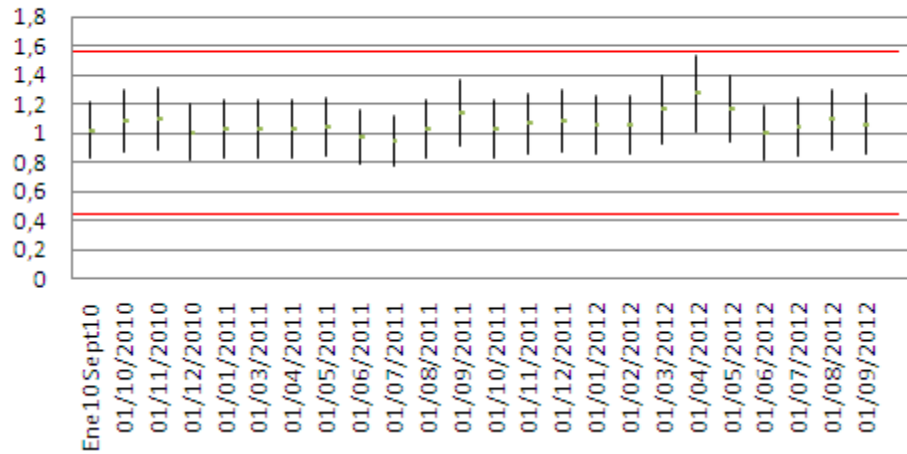
Variable Deuda 3A



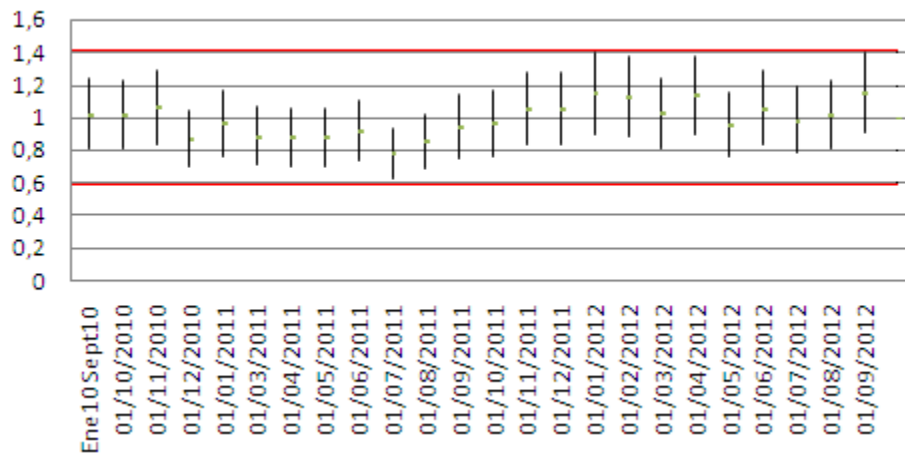
Variable Deuda 3B



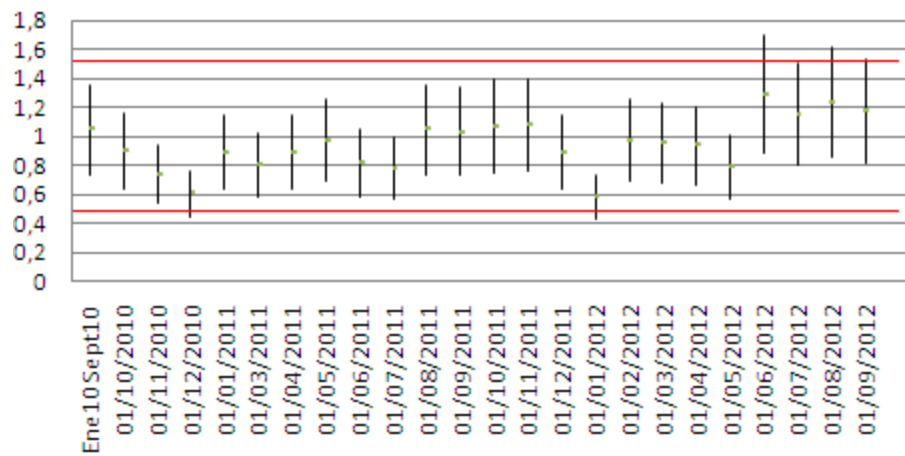
Variable Deuda 3C



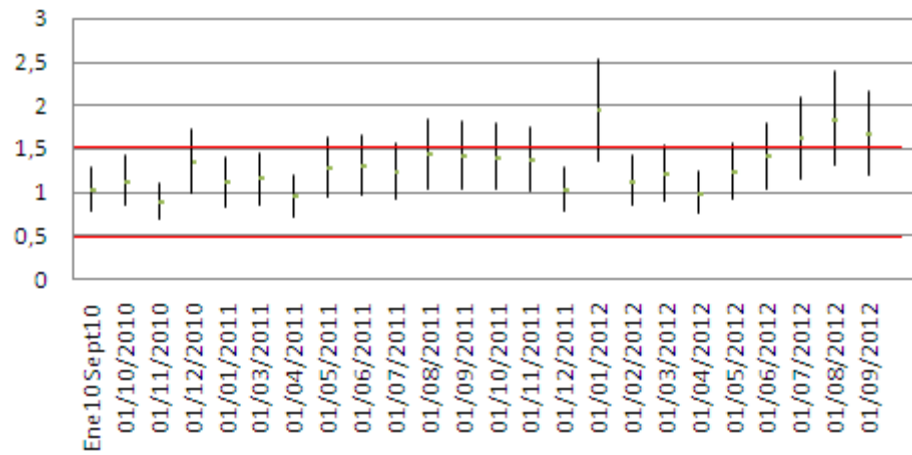
Variable Deuda 3D



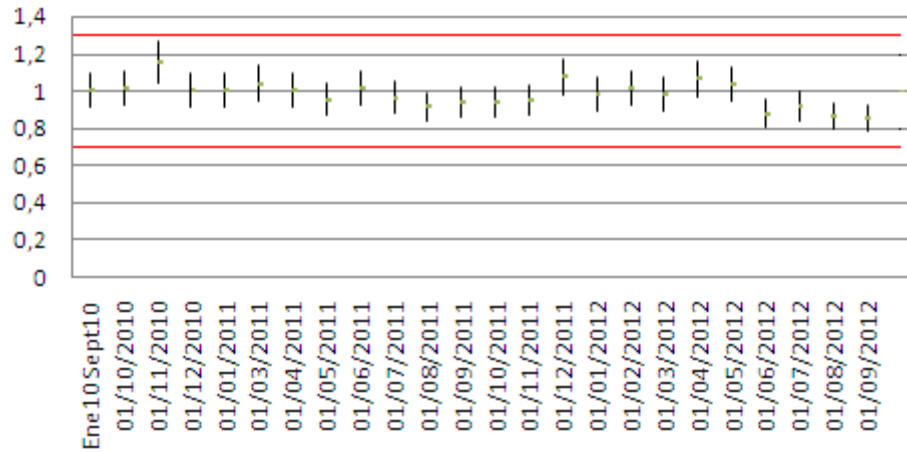
Variable Deuda 4A



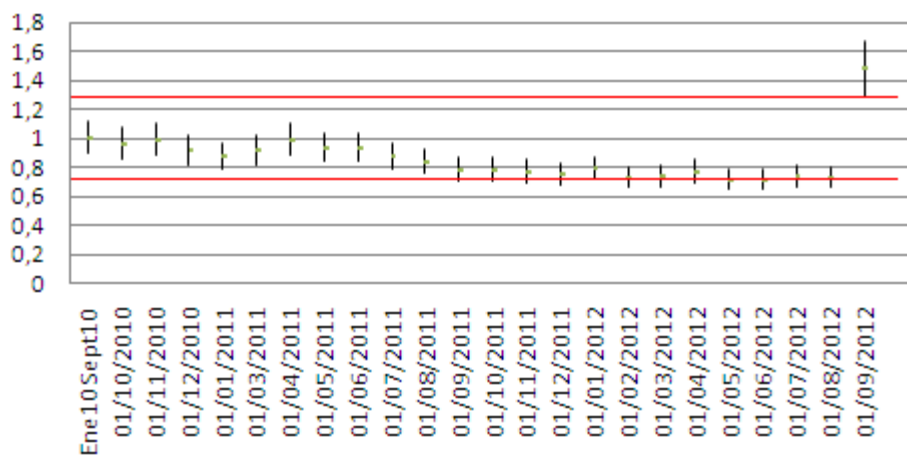
Variable Deuda 4B



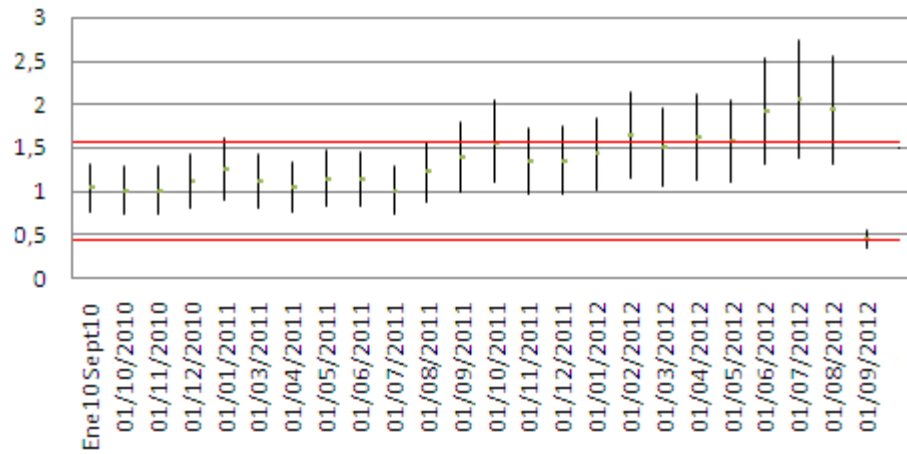
Variable Deuda 4C



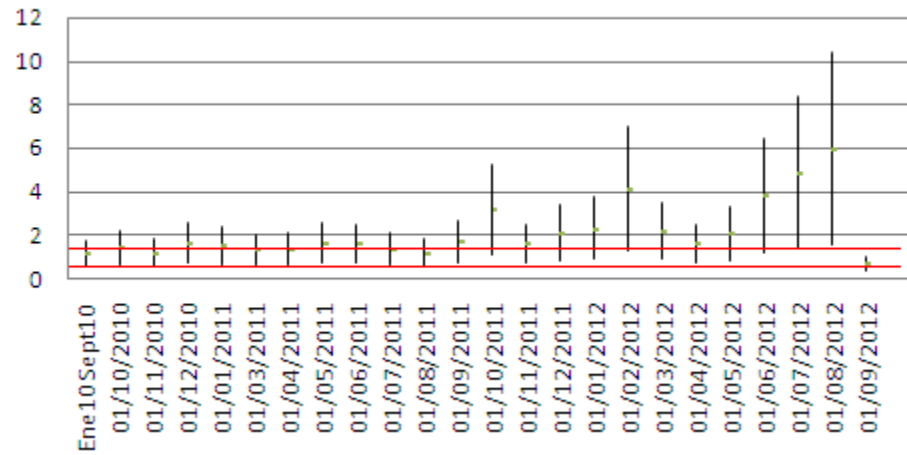
Variable Operación 2A



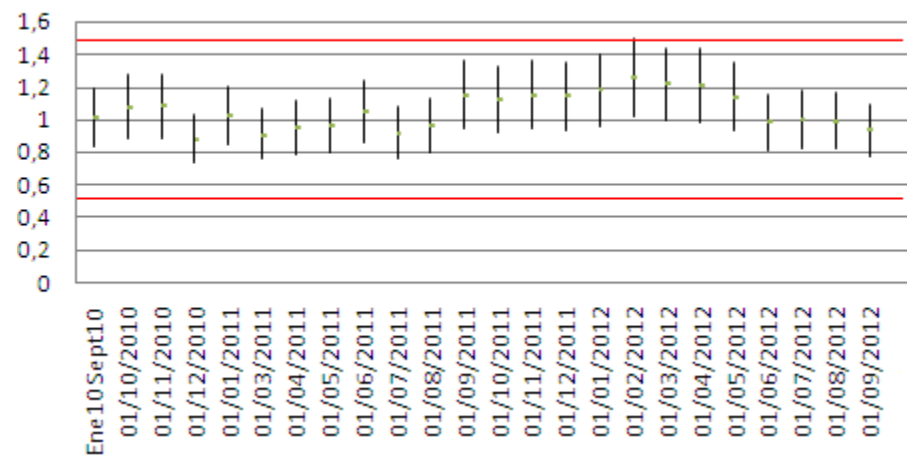
Variable Operación 2B



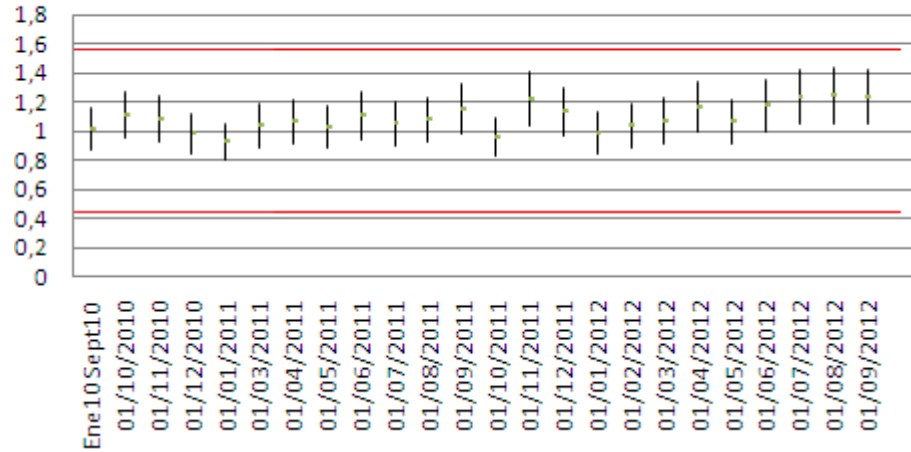
Variable Operación 2C



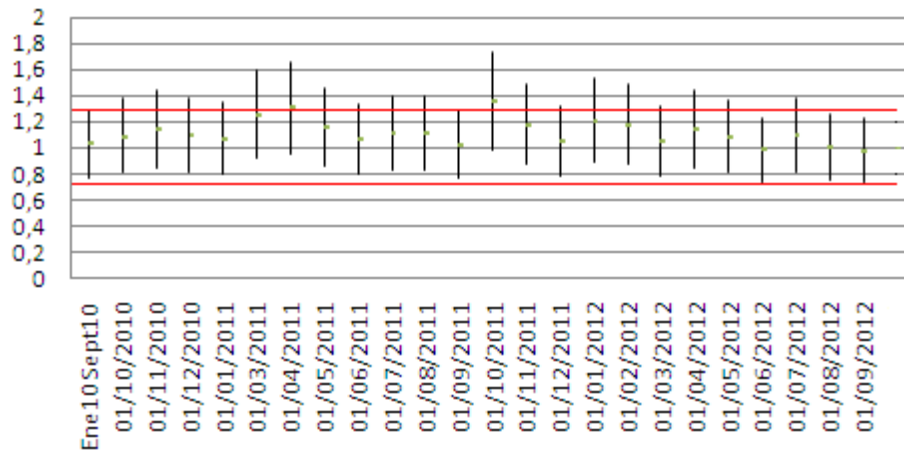
Variable Operación 3B



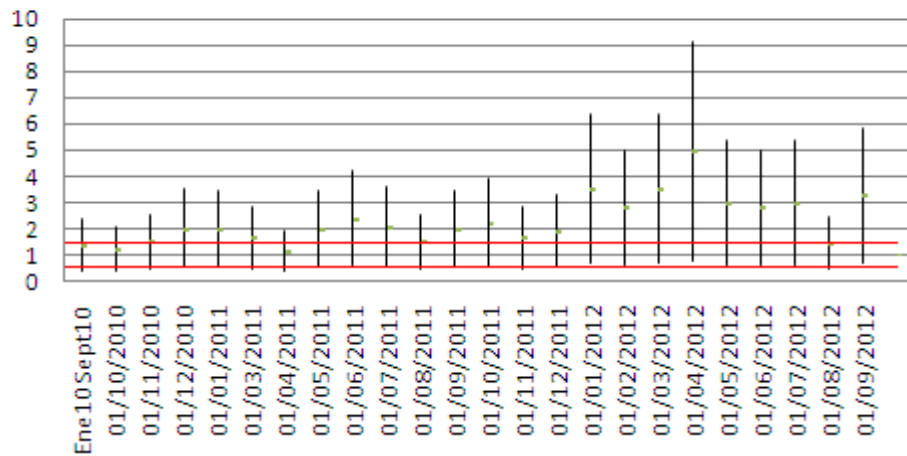
Variable Demográfica 2A



Variable Demográfica 2B



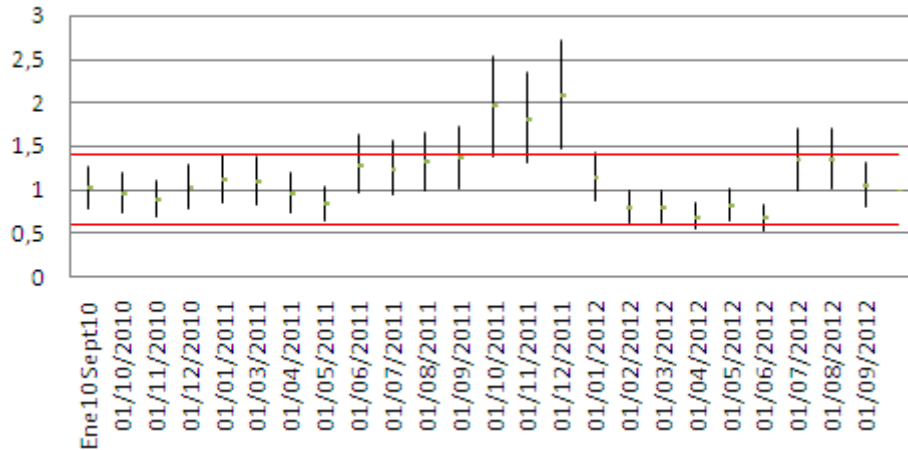
Variable Demográfica 2C



Variable Operación 4

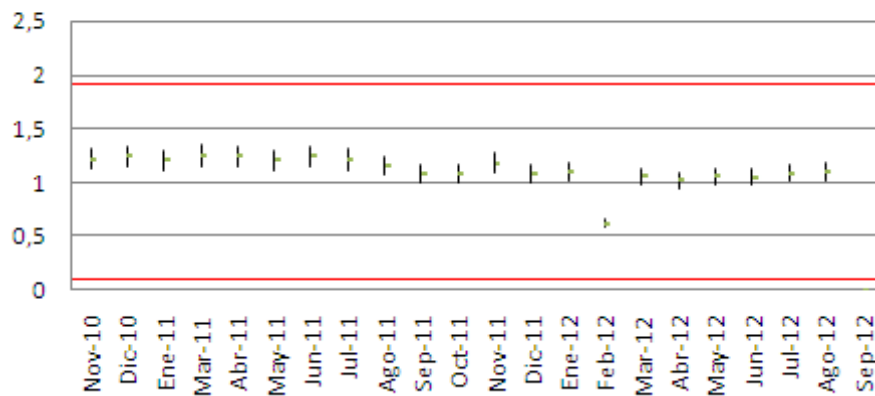


Variable Operación 5

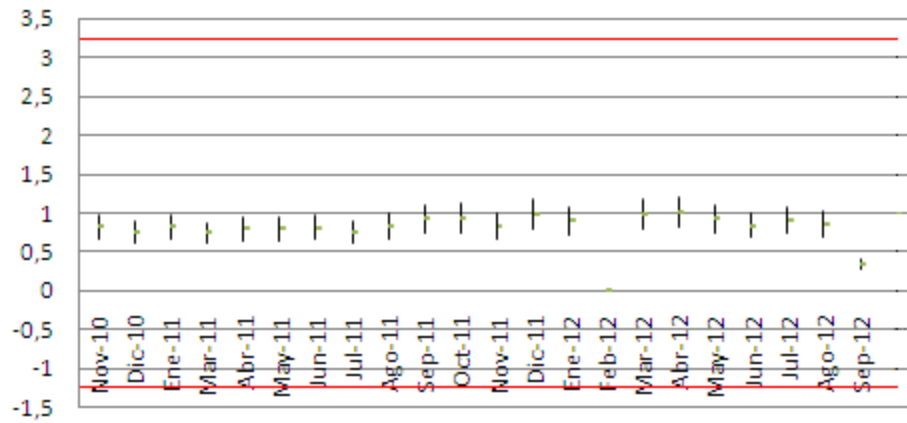


Fieller Stability Measure para las variables que quedaron al realizar la recalibración del modelo para Noviembre 2010 en adelante

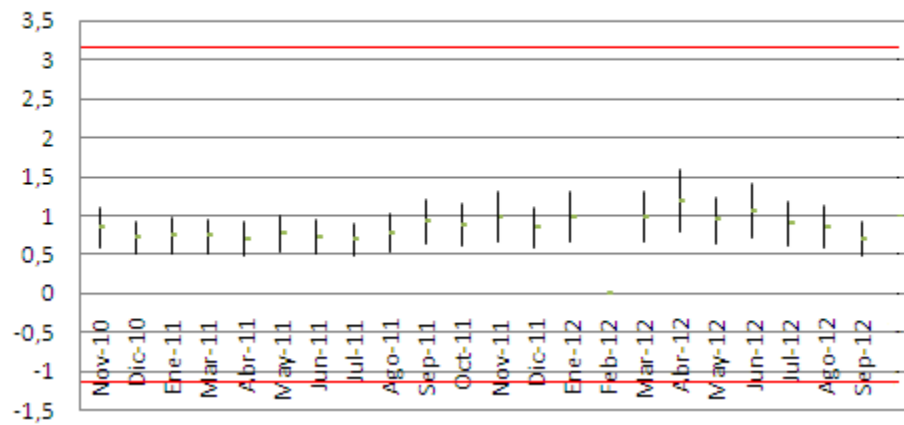
Variable Operación 1A



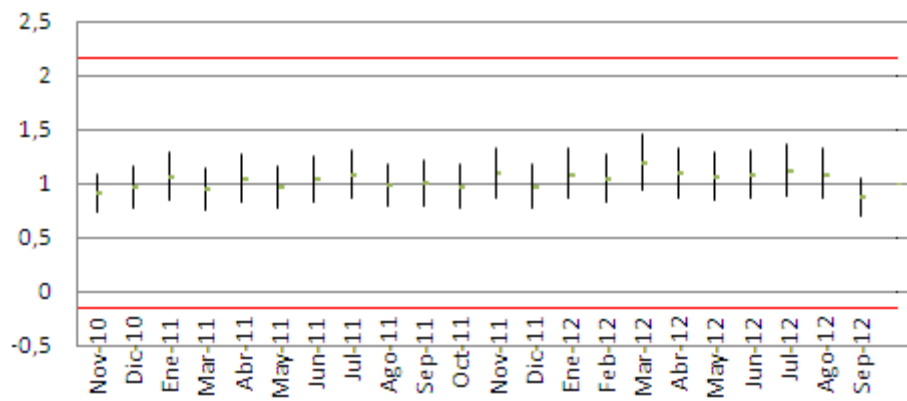
Variable Operación 1B



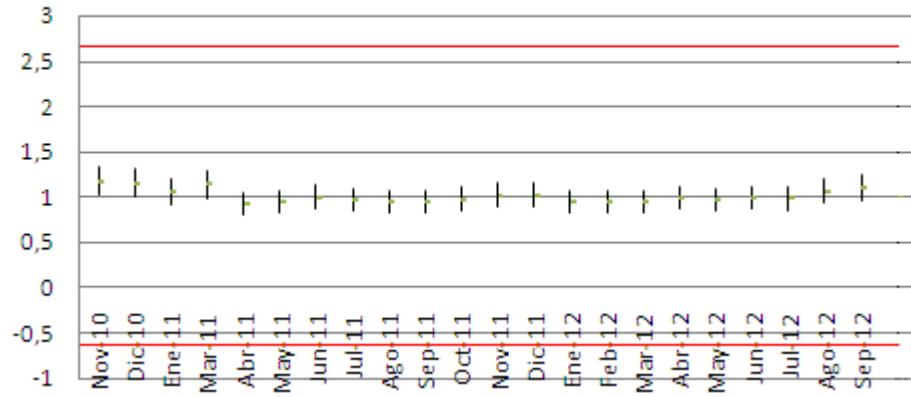
Variable Operación 1C



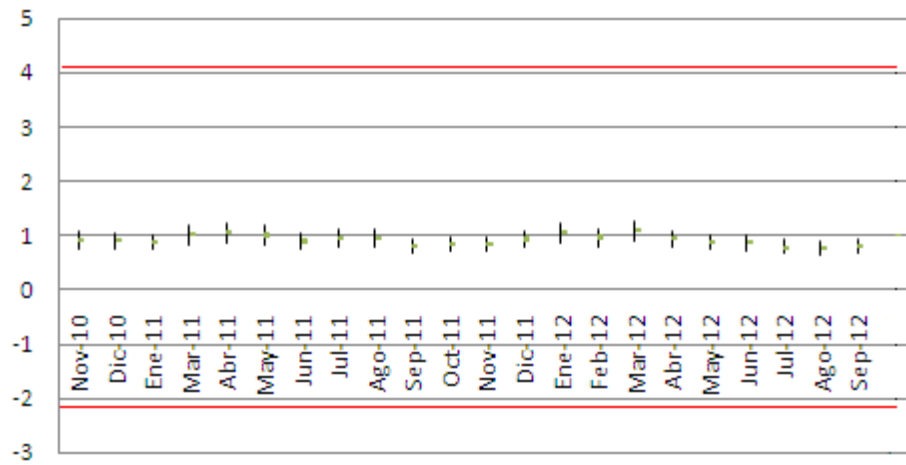
Variable Deuda 1A



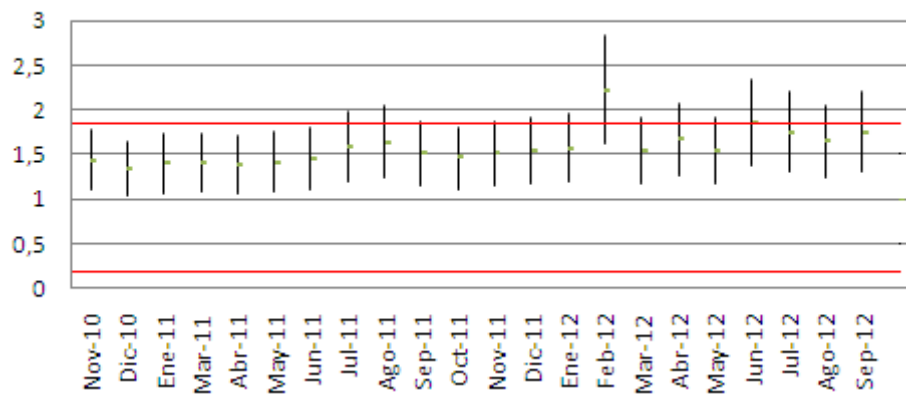
Variable Deuda 1B



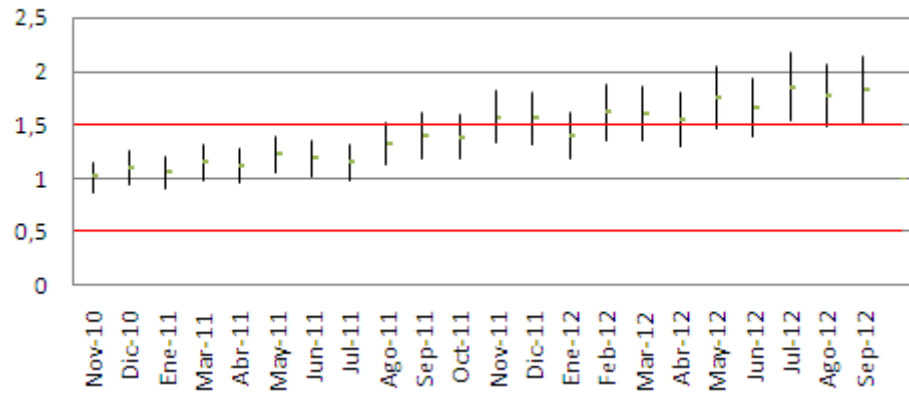
Variable Deuda 1C



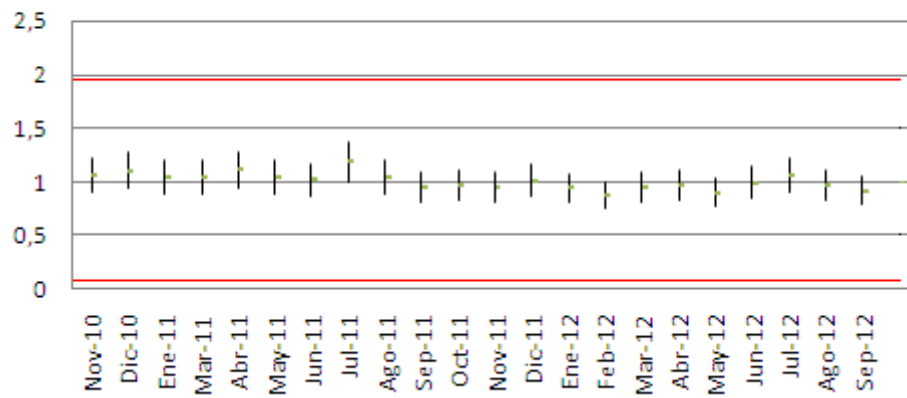
Variable Deuda 1D



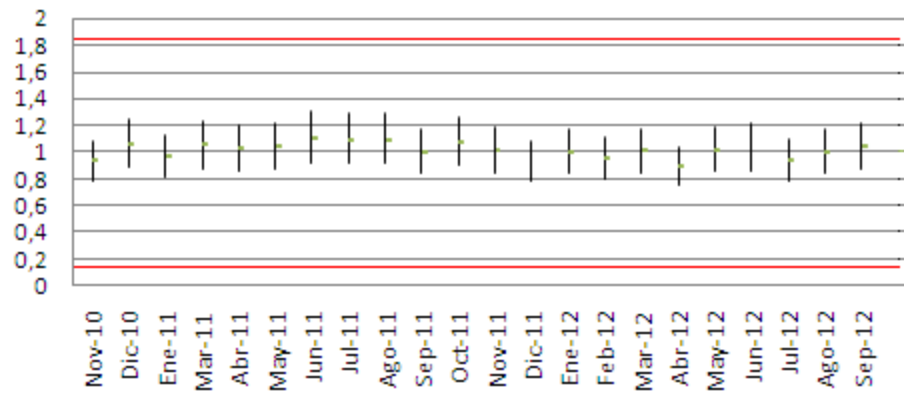
Variable Demográfica 1



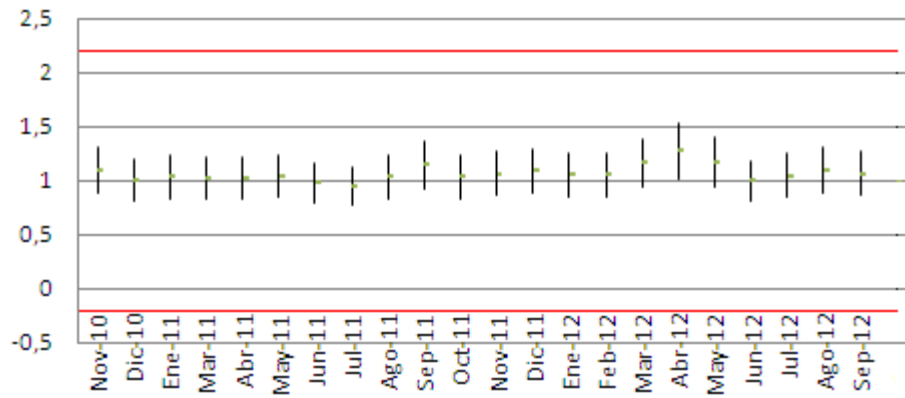
Variable Deuda 3A



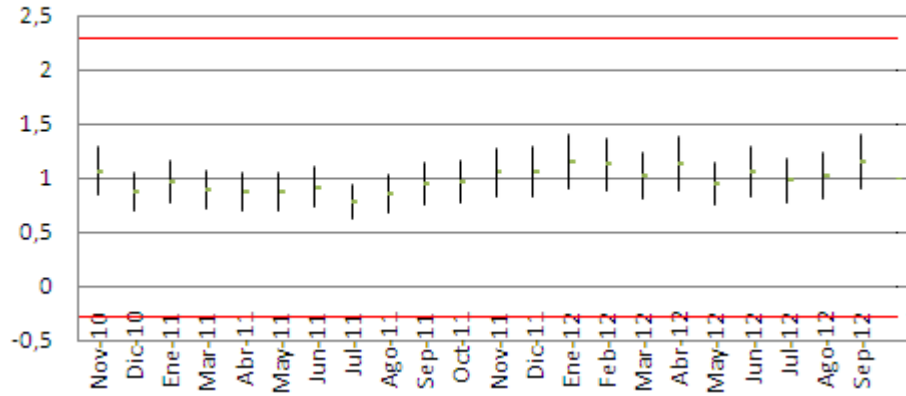
Variable Deuda 3B



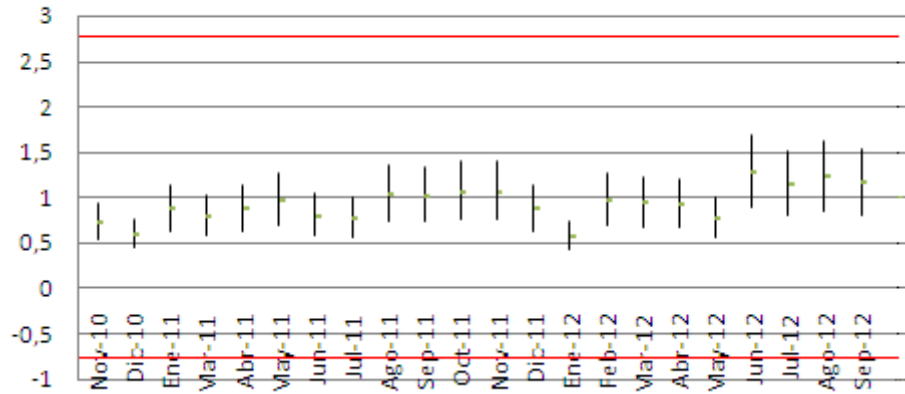
Variable Deuda 3C



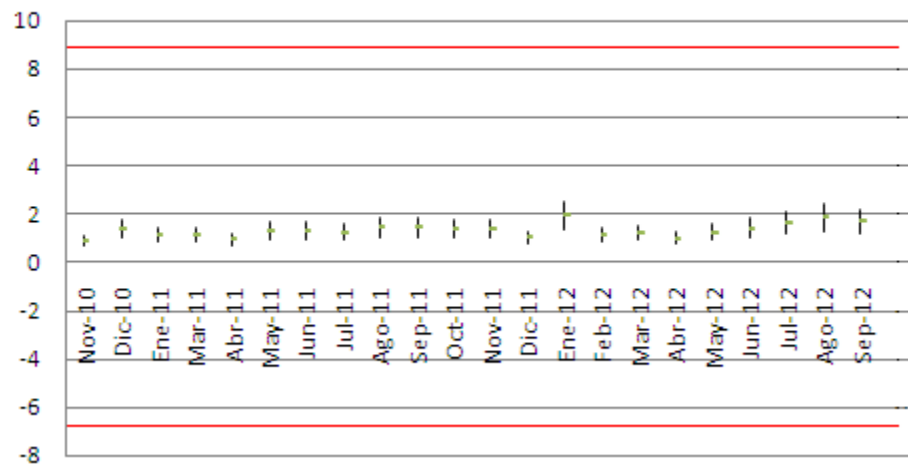
Variable Deuda 3D



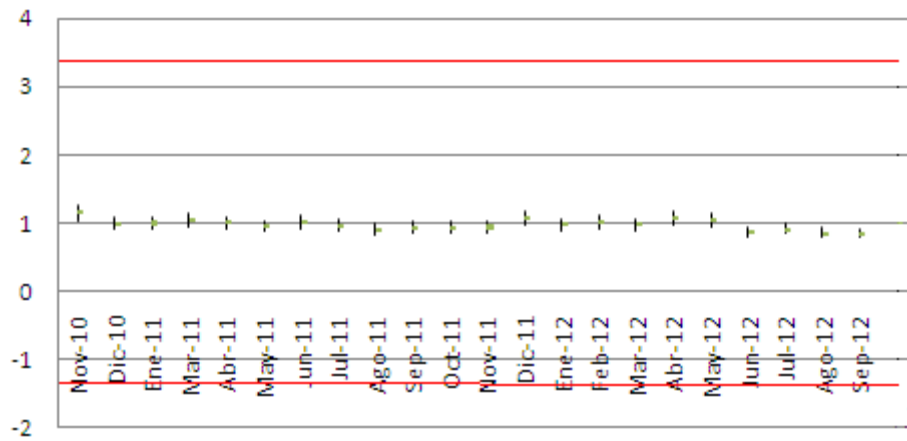
Variable Deuda 4A



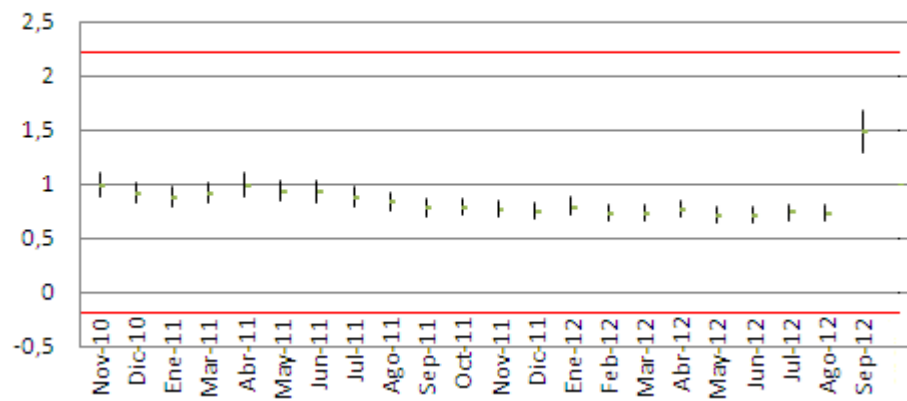
Variable Deuda 4B



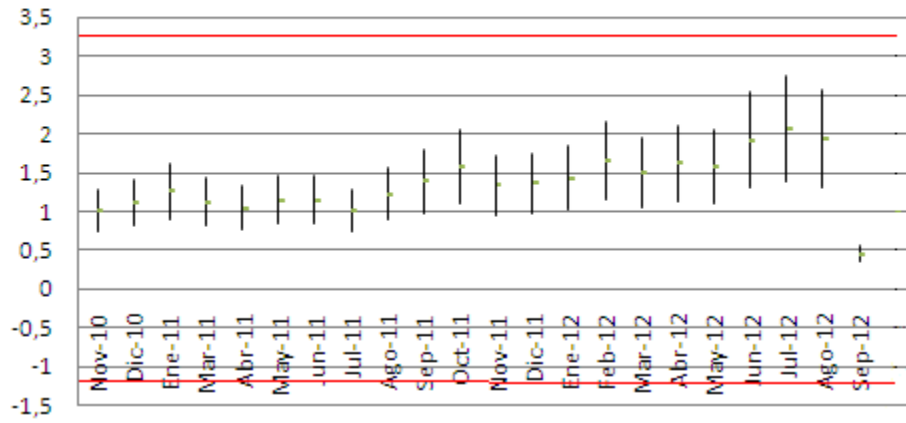
Variable Deuda 4C



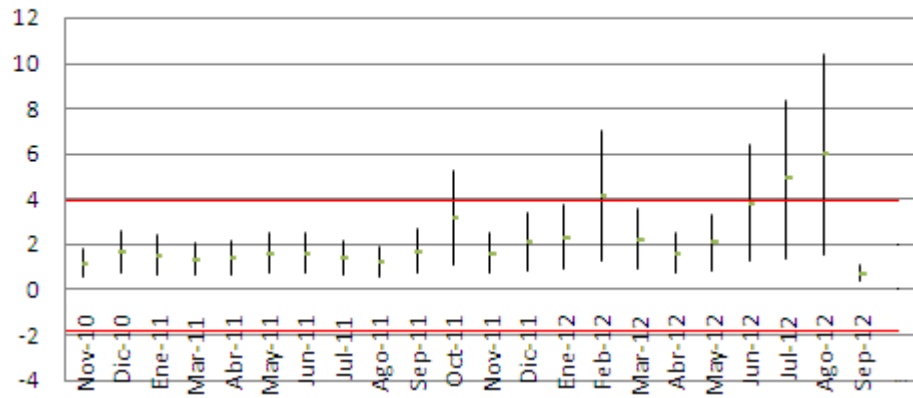
Variable Operación 2A



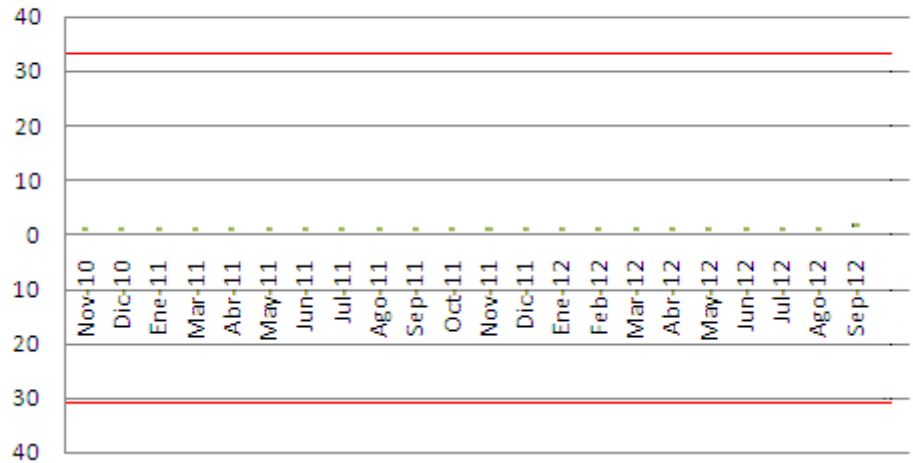
Variable Operación 2B



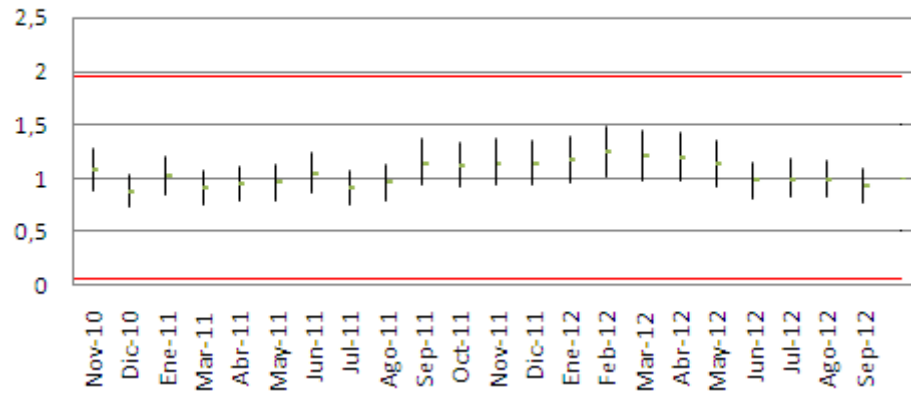
Variable Operación 2C



Variable Operación 3A



Variable Operación 3B



Anexo C

Datos de la Regresión Logística de la Re-calibración de Octubre 2011.

LogisticRegression

Bias (offset): 0.620 (SE: 0.648, Wald: 0.915)

Coefficients:

beta(Variable Deuda 1A) = -0.427 (SE: 0.658, Wald: 0.421)

beta(Variable Deuda 1B) = -0.921 (SE: 0.576, Wald: 2.558)

beta(Variable Deuda 1C) = -0.126 (SE: 0.568, Wald: 0.049)

beta(Variable Deuda 1D) = 1.014 (SE: 0.648, Wald: 2.445)

beta(Variable Deuda 3A) = -1.389 (SE: 0.596, Wald: 5.433)

beta(Variable Deuda 3B) = -0.611 (SE: 0.558, Wald: 1.199)

beta(Variable Deuda 3C) = -0.488 (SE: 0.589, Wald: 0.684)

beta(Variable Deuda 3D) = -0.361 (SE: 0.605, Wald: 0.356)

beta(Variable Deuda 4A) = 0.696 (SE: 0.708, Wald: 0.964)

beta(Variable Deuda 4B) = 1.575 (SE: 0.634, Wald: 6.173)

beta(Variable Deuda 4C) = -0.452 (SE: 0.488, Wald: 0.858)

Datos de la Regresión Logística de la Re-calibración de Noviembre 2010.

LogisticRegression

Bias (offset): -0.651 (SE: 0.494, Wald: 1.738)

Coefficients:

beta(Variable Operación 1A) = -0.665 (SE: 0.301, Wald: 4.873)

beta(Variable Operación 1B) = -0.362 (SE: 0.404, Wald: 0.802)

beta(Variable Operación 1C) = -0.513 (SE: 0.553, Wald: 0.860)

beta(Variable Deuda 1A) = -0.714 (SE: 0.410, Wald: 3.029)

beta(Variable Deuda 1B) = -0.401 (SE: 0.327, Wald: 1.503)
 beta(Variable Deuda 1C) = -0.210 (SE: 0.327, Wald: 0.414)
 beta(Variable Deuda 1D) = 0.911 (SE: 0.373, Wald: 5.984)
 beta(Variable Demográfica 1) = 0.930 (SE: 0.230, Wald: 16.276)
 beta(Variable Deuda 3A) = 0.877 (SE: 0.405, Wald: 4.696)
 beta(Variable Deuda 3B) = 0.909 (SE: 0.382, Wald: 5.670)
 beta(Variable Deuda 3C) = 0.664 (SE: 0.397, Wald: 2.804)
 beta(Variable Deuda 3D) = 0.679 (SE: 0.436, Wald: 2.421)
 beta(Variable Deuda 4A) = 0.423 (SE: 0.368, Wald: 1.324)
 beta(Variable Deuda 4B) = 0.094 (SE: 0.366, Wald: 0.066)
 beta(Variable Deuda 4C) = -0.252 (SE: 0.294, Wald: 0.732)
 beta(Variable Operación 2A) = -0.458 (SE: 0.273, Wald: 2.819)
 beta(Variable Operación 2B) = 0.317 (SE: 0.352, Wald: 0.808)
 beta(Variable Operación 2C) = 0.499 (SE: 0.705, Wald: 0.501)
 beta(Variable Operación 3A) = -0.017 (SE: 0.272, Wald: 0.004)
 beta(Variable Operación 3B) = 0.779 (SE: 0.362, Wald: 4.620)

Datos de la Regresión Logística de la Re-calibración de Marzo 2012.

LogisticRegression

Bias (offset): 0.869 (SE: 1.604, Wald: 0.294)

Coefficients:

beta(Variable Operación 1A) = -1.754 (SE: 1.195, Wald: 2.153)
 beta(Variable Operación 1B) = -0.635 (SE: 1.337, Wald: 0.226)
 beta(Variable Operación 1C) = -2.314 (SE: 2.288, Wald: 1.023)
 beta(Variable Deuda 1A) = -0.621 (SE: 1.360, Wald: 0.209)
 beta(Variable Deuda 1B) = 0.856 (SE: 1.026, Wald: 0.696)

beta(Variable Deuda 1C) = 2.361 (SE: 1.109, Wald: 4.536)
beta(Variable Deuda 1D) = -1.312 (SE: 11.285, Wald: 0.014)
beta(Variable Deuda 3A) = 0.231 (SE: 1.000, Wald: 0.054)
beta(Variable Deuda 3B) = 0.117 (SE: 1.204, Wald: 0.009)
beta(Variable Deuda 3C) = -0.009 (SE: 1.188, Wald: 0.000)
beta(Variable Deuda 3D) = 2.294 (SE: 1.287, Wald: 3.178)
beta(Variable Deuda 4A) = 0.334 (SE: 1.338, Wald: 0.062)
beta(Variable Deuda 4B) = -0.219 (SE: 1.275, Wald: 0.029)
beta(Variable Deuda 4C) = -1.957 (SE: 0.959, Wald: 4.164)

Anexo D

Estadístico K-S y Pérdidas Económicas por cartera: Sin ninguna re-calibración

Cartera	K-S	Pérdida Promedio por Cliente	Pérdida Total
Ene10 Sept10	0,250280928007422	49.982,97	1.320.900.000,00
oct-10	0,215704321474636	85.771,95	255.000.000,00
nov-10	0,229735370988820	58.759,02	211.650.000,00
dic-10	0,297624424005434	48.135,63	206.550.000,00
ene-11	0,300699663685616	39.818,36	124.950.000,00
mar-11	0,301240306223901	41.652,82	150.450.000,00
abr-11	0,254229082417269	76.722,74	237.150.000,00
may-11	0,242878516736586	36.181,37	117.300.000,00
jun-11	0,329808286951144	25.458,94	79.050.000,00
jul-11	0,283685466503696	66.048,30	188.700.000,00
ago-11	0,311556023884791	43.350,00	130.050.000,00
sep-11	0,263627595210491	57.770,27	145.350.000,00
oct-11	0,242957848416421	51.049,80	104.550.000,00
nov-11	0,305631345511109	44.508,00	112.200.000,00
dic-11	0,319202048218477	44.508,00	104.550.000,00
ene-12	0,314125814854978	42.476,40	76.500.000,00
feb-12	0,292723034011906	51.654,29	76.500.000,00
mar-12	0,296204620462046	59.579,44	114.750.000,00
abr-12	0,410216718266254	26.483,68	35.700.000,00
may-12	0,337365591397849	65.711,53	85.425.000,00
jun-12	0,358142155678982	38.321,85	61.200.000,00
jul-12	0,344938643608891	34.240,28	48.450.000,00
ago-12	0,321820471307789	40.621,79	131.325.000,00
sep-12	0,233874110020899	66.937,50	107.100.000,00

Re-calibración Noviembre 2010

Cartera	K-S	Pérdida Promedio por Cliente	Pérdida Total
Ene10 Sept10	0,250280928007422	49.982,97	1.320.900.000,00
oct-10	0,215704321474636	85.771,95	255.000.000,00
nov-10	0,400876046657955	29.733,48	107.100.000,00
dic-10	0,296111133464773	19.114,38	68.850.000,00
ene-11	0,386290088623113	33.981,12	122.400.000,00
mar-11	0,240094003612588	54.511,38	196.350.000,00
abr-11	0,359531073108732	56.635,20	204.000.000,00
may-11	0,347107004095348	37.520,82	135.150.000,00
jun-11	0,412502093968489	24.069,96	86.700.000,00
jul-11	0,400335918762360	31.857,30	114.750.000,00
ago-11	0,433644686249874	22.654,08	81.600.000,00
sep-11	0,399373238069091	27.609,66	99.450.000,00
oct-11	0,355579162732137	20.530,26	73.950.000,00
nov-11	0,479928067225186	19.822,32	71.400.000,00
dic-11	0,478017634219947	19.822,32	71.400.000,00
ene-12	0,579423324195153	14.158,80	51.000.000,00
feb-12	0,347789540264323	26.193,78	94.350.000,00
mar-12	0,483325913205707	22.654,08	81.600.000,00
abr-12	0,728073886068810	4.955,58	17.850.000,00
may-12	0,476350461133070	16.282,62	58.650.000,00
jun-12	0,412638580931264	24.069,96	86.700.000,00
jul-12	0,610206731249376	11.327,04	40.800.000,00
ago-12	0,390902655843499	25.485,84	91.800.000,00
sep-12	0,594832076974934	7.787,34	28.050.000,00

Re-calibración Octubre 2011.

Cartera	K-S	Pérdida Promedio por Cliente	Pérdida Total
Ene10			
Sept10	0,250280928007422	49.982,97	1.320.900.000,00
oct-10	0,215704321474636	85.771,95	255.000.000,00
nov-10	0,229735370988820	58.759,02	211.650.000,00
dic-10	0,297624424005434	48.135,63	206.550.000,00
ene-11	0,300699663685616	39.818,36	124.950.000,00
mar-11	0,301240306223901	41.652,82	150.450.000,00
abr-11	0,254229082417269	76.722,74	237.150.000,00
may-11	0,242878516736586	36.181,37	117.300.000,00
jun-11	0,329808286951144	25.458,94	79.050.000,00
jul-11	0,283685466503696	66.048,30	188.700.000,00
ago-11	0,311556023884791	43.350,00	130.050.000,00
sep-11	0,272009560831766	55.743,24	140.250.000,00
oct-11	0,248567921440262	9.121,62	22.950.000,00
nov-11	0,256010762407248	77.028,85	170.850.000,00
dic-11	0,234597823767869	50.796,81	114.750.000,00
ene-12	0,294369866257837	31.149,36	56.100.000,00
feb-12	0,309874692976979	37.879,81	56.100.000,00
mar-12	0,288320498716538	33.099,69	63.750.000,00
abr-12	0,342049977885891	28.375,37	38.250.000,00
may-12	0,226075268817204	31.384,62	40.800.000,00
jun-12	0,310021733880705	44.708,83	71.400.000,00
jul-12	0,348333988383772	39.646,64	56.100.000,00
ago-12	0,278177207550673	60.277,49	117.300.000,00
sep-12	0,297669228861889	55.781,25	89.250.000,00

Re-calibración Febrero 2012

Cartera	KS	Pérdida Promedio por Cliente	Pérdida Total
Ene10			
Sept10	0,250280928007422	49.982,97	1.320.900.000,00
oct-10	0,215704321474636	85.771,95	255.000.000,00
nov-10	0,400876046657955	29.733,48	107.100.000,00
dic-10	0,296111133464773	19.114,38	68.850.000,00
ene-11	0,386290088623113	33.981,12	122.400.000,00
mar-11	0,240094003612588	54.511,38	196.350.000,00
abr-11	0,359531073108732	56.635,20	204.000.000,00
may-11	0,347107004095348	37.520,82	135.150.000,00
jun-11	0,412502093968489	24.069,96	86.700.000,00
jul-11	0,400335918762360	31.857,30	114.750.000,00
ago-11	0,433644686249874	22.654,08	81.600.000,00
sep-11	0,399373238069091	27.609,66	99.450.000,00
oct-11	0,355579162732137	20.530,26	73.950.000,00
nov-11	0,479928067225186	19.822,32	71.400.000,00
dic-11	0,478017634219947	19.822,32	71.400.000,00
ene-12	0,579423324195153	14.158,80	51.000.000,00
feb-12	0,347789540264323	26.193,78	94.350.000,00
mar-12	0,290922003858466	45.308,16	163.200.000,00
abr-12	0,735496736765772	4.247,64	15.300.000,00
may-12	0,359373235460192	23.362,02	84.150.000,00
jun-12	0,419186529933481	43.184,34	155.550.000,00
jul-12	0,350344552082293	26.901,72	96.900.000,00
ago-12	0,306710547685368	37.520,82	135.150.000,00
sep-12	0,301127658270515	29.733,48	107.100.000,00

Re-calibración Julio 2012

Cartera	KS	Pérdida Promedio por Cliente	Pérdida Total
Ene10			
Sept10	0,250280928007422	49.982,97	1.320.900.000,00
oct-10	0,215704321474636	85.771,95	255.000.000,00
nov-10	0,400876046657955	29.733,48	107.100.000,00
dic-10	0,296111133464773	19.114,38	68.850.000,00
ene-11	0,386290088623113	33.981,12	122.400.000,00
mar-11	0,240094003612588	54.511,38	196.350.000,00
abr-11	0,359531073108732	56.635,20	204.000.000,00
may-11	0,347107004095348	37.520,82	135.150.000,00
jun-11	0,412502093968489	24.069,96	86.700.000,00
jul-11	0,400335918762360	31.857,30	114.750.000,00
ago-11	0,433644686249874	22.654,08	81.600.000,00
sep-11	0,399373238069091	27.609,66	99.450.000,00
oct-11	0,355579162732137	20.530,26	73.950.000,00
nov-11	0,479928067225186	19.822,32	71.400.000,00
dic-11	0,478017634219947	19.822,32	71.400.000,00
ene-12	0,579423324195153	14.158,80	51.000.000,00
feb-12	0,347789540264323	26.193,78	94.350.000,00
mar-12	0,483325913205707	22.654,08	81.600.000,00
abr-12	0,728073886068810	4.955,58	17.850.000,00
may-12	0,476350461133070	16.282,62	58.650.000,00
jun-12	0,412638580931264	24.069,96	86.700.000,00
jul-12	0,485235860048604	19.822,32	71.400.000,00
ago-12	0,341040090458636	36.104,94	130.050.000,00
sep-12	0,258097383097383	33.981,12	122.400.000,00

Anexo E

Stability Index

SI	Variable Operación 1A	Variable Operación 1B	Variable Operación 1C
oct-10	-	-	-
nov-10	0,048121992	0,810591217	1,561955913
dic-10	0,056717363	0,725679009	1,425350201
ene-11	0,045930313	0,815085007	1,463947437
mar-11	0,060253826	0,720806325	1,454263419
abr-11	0,058251705	0,776729642	1,420041702
may-11	0,044771172	0,784601953	1,47989177
jun-11	0,058570827	0,799663158	1,446450714
jul-11	0,046896507	0,737803346	1,411204951
ago-11	0,027609152	0,822963392	1,501255122
sep-11	0,008553741	0,913321996	1,632667048
oct-11	0,043235731	1,018234812	1,679325553
nov-11	0,492687199	0,351000292	1,234260195
dic-11	0,009056241	0,962470194	1,557061612
ene-12	0,01404891	0,896241233	1,67999704
feb-12	-	-	-
mar-12	0,004837398	0,970562495	1,685618122
abr-12	0,0004837	0,991707737	1,820756947
may-12	0,004287725	3,503941575	0,967561873
jun-12	0,003199681	0,83104485	1,741439503
jul-12	0,009621212	0,896960506	1,605798251
ago-12	0,013235135	0,840459244	1,577472134
sep-12	-	0,098419866	1,413816922

SI	Variable Deuda 1A	Variable Deuda 1B	Variable Deuda 1C	Variable Deuda 1D
oct-10	0,964674492	0,55884486	0,73872605	1,219689368
nov-10	0,921807341	0,507432482	0,683259785	1,357599308
dic-10	0,973143406	0,491463177	0,664415658	1,303786087
ene-11	1,059467288	0,418786016	0,642624268	1,337294645
mar-11	0,962474172	0,484518705	0,773077135	1,344361239
abr-11	1,042248643	0,306484413	0,797050443	1,334533771
may-11	0,973794091	0,322925646	0,757406524	1,347711239
jun-11	1,034934118	0,37238207	0,658961905	1,37222882
jul-11	1,072681048	0,338822265	0,711437662	1,444707682
ago-11	0,989316553	0,320088085	0,702255373	1,467441474
sep-11	1,005453514	0,326107324	0,571876533	1,40180679
oct-11	1,079563126	0,443398964	0,698910755	1,465541982
nov-11	0,604000029	0,027125667	0,153524774	0,933309128
dic-11	0,981824424	0,388916242	0,693700082	1,417911656
ene-12	1,077740387	0,32346746	0,795377965	1,432319759
feb-12	1,043178263	0,33312783	0,725425199	1,70383903
mar-12	1,157388945	0,330514412	0,831141286	1,414674518
abr-12	1,082724376	0,361362384	0,701977001	1,480400242
may-12	0,297043056	3,974940858	3,202566461	1,990810556
jun-12	1,075161667	0,364735435	0,626378079	1,569009582
jul-12	1,107672396	0,357418115	0,546723495	1,520920687
ago-12	1,07803421	0,427877337	0,530590673	1,468976433
sep-12	0,887370381	0,447467908	0,563000787	1,521407345

SI	Variable Demográfica 1
oct-10	0,422383919
nov-10	0,4268247
dic-10	0,498065225
ene-11	0,464316797
mar-11	0,537977537
abr-11	0,516348812
may-11	0,590964517
jun-11	0,569105493
jul-11	0,537445866
ago-11	0,66239763
sep-11	0,708824009
oct-11	0,7997213
nov-11	0,343277515
dic-11	0,808085458
ene-12	0,713703227
feb-12	0,839903277
mar-12	0,833589216
abr-12	0,801113388
may-12	0,551210024
jun-12	0,861003163
jul-12	0,959672903
ago-12	0,919463631
sep-12	0,948691249

SI	Variable Deuda 2
oct-10	0,463669334
nov-10	0,475256586
dic-10	0,373114671
ene-11	0,425103407
mar-11	0,305432061
abr-11	0,251033015
may-11	0,209302923
jun-11	0,33405881
jul-11	0,270778937
ago-11	0,378740169
sep-11	0,445887947
oct-11	0,597173977
nov-11	0,080340924
dic-11	0,554057189
ene-12	0,56277715
feb-12	0,606738294
mar-12	0,501397368
abr-12	0,45144695
may-12	3,286603019
jun-12	0,411802403
jul-12	0,452771855
ago-12	0,510488412
sep-12	0,046798399

SI	Variable Deuda 3A	Variable Deuda 3B	Variable Deuda 3C	Variable Deuda 3D
oct-10	0,514056534	0,725945751	0,962852136	1,040538679
nov-10	0,600995439	0,634812733	0,973570026	1,077740387
dic-10	0,636676938	0,746426667	0,893660853	0,900725997
ene-11	0,587539816	0,670976364	0,916234945	0,991396629
mar-11	0,585601658	0,74418615	0,914274835	0,919989225
abr-11	0,644967516	0,722429859	0,912584765	0,912584765
may-11	0,585476269	0,73101245	0,924425522	0,913264031
jun-11	0,56687121	0,786569152	0,868082721	0,952181322
jul-11	0,702822419	0,778878934	0,841697904	0,810556317
ago-11	0,591515421	0,778924077	0,917859459	0,887370381
sep-11	0,513203721	0,697462847	1,005453514	0,976657503
oct-11	0,613563363	0,861224183	1,015653989	1,090633194
nov-11	0,090248159	0,245634808	0,463752876	0,587396668
dic-11	0,558595281	0,629982878	0,962470194	1,074045596
ene-12	0,498594064	0,70004069	0,936313338	1,144216447
feb-12	0,437545848	0,657445131	0,937923426	1,130334203
mar-12	0,513034538	0,7015059	1,019964555	1,047173682
abr-12	0,515781845	0,591515421	1,101065664	1,134011616
may-12	2,865933027	2,193233514	1,11275028	3,484372513
jun-12	0,542187817	0,728014776	0,888811824	1,075161667
jul-12	0,601041551	0,640301741	0,929647028	1,010556319
ago-12	0,527202541	0,698210565	0,965755865	1,041158253
sep-12	0,473709567	0,734985581	0,94204748	1,147808882

SI	Variable Deuda 4A	Variable Deuda 4B	Variable Deuda 4C
oct-10	1,465912406	1,387023534	0,072252236
nov-10	1,311490284	1,194338468	0,150188636
dic-10	1,139613669	1,534404786	0,067390044
ene-11	1,461074014	1,383881018	0,068374619
mar-11	1,373544477	1,410699459	0,087216085
abr-11	1,459943155	1,248820626	0,069356038
may-11	1,532253806	1,499918636	0,043732327
jun-11	1,388241922	1,508369165	0,072399116
jul-11	1,355913724	1,466644541	0,047810895
ago-11	1,589662487	1,582826512	0,026946377
sep-11	1,578601677	1,574567212	0,038886726
oct-11	1,694754486	1,654160836	0,089968291
nov-11	1,15498683	1,091096457	0,428210507
dic-11	1,456804425	1,320035841	0,105129652
ene-12	1,108668681	1,815068801	0,057559354
feb-12	1,53222151	1,392331195	0,072118436
mar-12	1,518035983	1,455636947	0,057465875
abr-12	1,5010602	1,283625867	0,100920381
may-12	3,129647707	2,592525626	0,001129076
jun-12	1,749042486	1,569009582	0,014091981
jul-12	1,666902725	1,674826643	0,028634602
ago-12	1,717907531	1,774101899	0,01010396
sep-12	1,677191976	1,698599684	0,007970315

SI	Variable Operación 2A	Variable Operación 2B	Variable Operación 2C
oct-10	0,190203443	1,404011006	2,574649573
nov-10	0,210356155	1,405984928	2,463986552
dic-10	0,155913761	1,481552147	2,660754508
ene-11	0,126310157	1,580004486	2,612729615
mar-11	0,156451722	1,486987234	2,530897989
abr-11	0,208352546	1,434105122	2,549745857
may-11	0,170474195	1,505717667	2,646496988
jun-11	0,16937115	1,505329042	2,638488252
jul-11	0,124311728	1,405235347	2,557461917
ago-11	0,0996291	1,559278724	2,479330342
sep-11	0,063137852	1,654310863	2,670391997
oct-11	0,126775041	1,83439183	3,058281287
nov-11	0,319700433	1,175510614	2,258106829
dic-11	0,045951006	1,639796167	2,795172981
ene-12	0,069294161	1,67999704	2,837158963
feb-12	0,032659378	1,785763374	3,102580439
mar-12	0,036041274	1,715664178	2,810892539
abr-12	0,05640338	1,773128225	2,63631852
may-12	2,271081601	4,409611484	0,000366477
jun-12	0,025286147	1,897907712	3,069259543
jul-12	0,037363639	1,951029087	3,16762166
ago-12	0,03133791	1,905406235	3,24126275
sep-12	0,537573518	0,703786496	2,134763057

SI	Variable Operación 3A	Variable Operación 3B
oct-10	0,157229307	0,841642958
nov-10	0,166732878	0,844442522
dic-10	0,214947008	0,662598218
ene-11	0,176777755	0,796113075
mar-11	0,217670645	0,691315278
abr-11	0,18156671	0,731935029
may-11	0,173068195	0,744111617
jun-11	0,205673557	0,815925973
jul-11	0,169118149	0,699969522
ago-11	0,167821208	0,744029803
sep-11	0,130490584	0,89912967
oct-11	0,214420542	0,975269173
nov-11	0,054741154	0,421016474
dic-11	0,121893371	0,893187874
ene-12	0,14802304	0,918932444
feb-12	0,13500956	0,974211498
mar-12	0,149470782	0,9453981
abr-12	0,120734694	0,939369523
may-12	0,137495535	0,255625179
jun-12	0,130334224	0,760158375
jul-12	0,152336788	0,776554171
ago-12	0,186648932	0,769914284
sep-12	0,569490152	0,716651539

SI	Variable Demográfica 2A	Variable Demográfica 2B	Variable Demográfica 2C	Variable Demográfica 2D
oct-10	0,573981615	1,38422077	2,830577328	3,822220044
nov-10	0,550439877	1,425066683	2,931890433	3,604173355
dic-10	0,467763875	1,389607407	3,064540379	3,612267063
ene-11	0,4160972	1,368104845	3,057390484	3,64053732
mar-11	0,509873699	1,499873818	2,978463303	3,731203104
abr-11	0,534394238	1,529285361	2,798802395	3,847148331
may-11	0,505513425	1,430217179	3,059818667	3,877662772
jun-11	0,567987555	1,364319434	3,134071642	3,507864745
jul-11	0,521281974	1,399302024	3,078418052	3,579855348
ago-11	0,544132114	1,400330848	2,934409003	3,524910584
sep-11	0,605844605	1,336680011	3,06383927	3,29741151
oct-11	0,536306025	1,649207583	3,180051759	3,369928972
nov-11	0,203681047	0,98310173	2,609767565	3,353555843
dic-11	0,589540608	1,351587108	3,043163274	3,645580171
ene-12	0,471017565	1,466901516	3,278325157	3,828368133
feb-12	0,509704312	1,444697221	3,197914184	3,439898055
mar-12	0,540337967	1,350006829	3,277078319	3,610606883
abr-12	0,615732753	1,420937033	3,378347564	3,309139301
may-12	4,737517481	0,277240189	1,117995594	2,213427089
jun-12	0,621720123	1,300852907	3,198777883	3,565578754
jul-12	0,662026179	1,388279956	3,22076235	3,790613984
ago-12	0,673528584	1,314044151	2,915808264	3,487094117
sep-12	0,663859221	1,292946306	3,249111295	3,653169984

SI	Variable Deuda 5
oct-10	1,200929487
nov-10	1,135208268
dic-10	1,261818384
ene-11	1,329740926
mar-11	1,313874453
abr-11	1,200758916
may-11	1,080806616
jun-11	1,443598564
jul-11	1,414203503
ago-11	1,461418376
sep-11	1,493542226
oct-11	1,852669947
nov-11	1,262167951
dic-11	1,813441173
ene-12	1,344202829
feb-12	1,043178263
mar-12	1,050244359
abr-12	0,919976469
may-12	3,521640807
jun-12	0,907817412
jul-12	1,474663002
ago-12	1,483071956
sep-12	1,274172093

SI	Variable Operación 4
oct-10	0,679281984
nov-10	0,632723112
dic-10	0,658975005
ene-11	0,665976737
mar-11	0,624839276
abr-11	0,599356804
may-11	0,61216839
jun-11	0,655222931
jul-11	0,669195038
ago-11	0,655932062
sep-11	0,661154129
oct-11	0,744025946
nov-11	0,219643331
dic-11	0,71910127
ene-12	0,748819312
feb-12	0,692125909
mar-12	0,695180295
abr-12	0,717236288
may-12	2,553781915
jun-12	0,662170884
jul-12	0,754873244
ago-12	0,624305552
sep-12	0,661424636