

Eficiencia de los modelos Poisson y Logístico en la asignación de probabilidades de incumplimiento a empresas mineras mexicanas

Salvador Cruz Aké*

Instituto Politécnico Nacional, Escuela Superior de Economía

Nora Gavira Durón

Universidad de las Américas Puebla, Escuela de Negocios y Economía

Reyna Susana García Ruíz

Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración

(Recibido 16 de marzo 2016, aceptado 03 de Agosto 2016)

Resumen

La existencia del entorno volátil y características propias del sector minero, afecta el cumplimiento de pago de créditos otorgados a empresas del sector, por lo que modelos de scoring basados en normalidad, modelos Probit o Logit, subestiman la probabilidad de impago. El objetivo del artículo es probar que los modelos Logit no capturan correctamente las probabilidades de incumplimiento de empresas del sector minero al subestimar las colas, mediante un análisis de estabilidad y confiabilidad de dichas probabilidades, lo que genera problemas de bondad de ajuste y subestimación de la probabilidad; mientras que un modelo Poisson de variable dicotómica captura los efectos de cola y estabiliza la regresión, sus estimadores y probabilidades estimadas. Los resultados indican que los modelos Logísticos no están diseñados para analizar óptimamente las variables independientes con valores extremos y que no están asociados al trimestre de operación, sino que son propios de cada empresa. Los modelos Poisson resultaron ser capaces de captar los valores extremos de la distribución; por lo que resultan más adecuados para determinar la calidad crediticia de las empresas mineras.

Clasificación JEL: C01, C13, C52, L61.

Palabras clave: Credit Scoring, Riesgo de crédito, Probabilidad de incumplimiento, Modelos Logit y de Poisson, Empresas mineras.

Efficiency of the Poisson and Logistic Models in the Allocation of Probabilities of default to Mexican Mining Companies.

Abstract

The existence of the volatile environment and characteristics of the mining sector, affects compliance with payment of loans granted to companies in the sector, making scoring models based on normal, Probit or Logit models underestimate the probability of default. The objective of this article is to prove that the Logit models do not capture properly the default probabilities of mining companies to underestimate the tails, by analysis of stability and reliability of these probabilities, creating problems for goodness of fit and underestimation

* Escuela Superior de Economía. Plan de Agua Prieta. No 66. Col Plutarco Elías Calles, Unidad Profesional Adolfo López Mateos, Casco de Santo Tomás. Del. Miguel Hidalgo. C.P. 11340, Ciudad de México, Tel. 2009-2016. Correo Electrónico:salvador.ake22@gmail.com

of the probability; while a Poisson model dummy variable captures the effects of tail and stabilizes the regression, estimators and their estimated probabilities. The results indicate that the logistic models are not designed to optimally analyze the independent variables with extreme values and are not associated with the quarter of operation, but are specific to each company. Poisson models were found to be able to capture the extreme values of the distribution; so they are best suited to determine the credit quality of mining companies.

JEL Classification : C32, E5, F10, F31, F51.

Key Words: Credit Scoring, Credit Risk, Probability of Default, Logit and Poisson Models, Mining Companies.

1. Introducción

El sector minero es tradicionalmente dinámico en México, aunque este dinamismo está sujeto a las vicisitudes de los precios de los minerales (*commodities*), y a las condiciones del entorno económico global. Muestra de estas condiciones es el hecho de que México se ubica entre los 10 principales productores de 16 diferentes minerales, Duarte¹ además de ser el primer destino en inversión en exploración minera en América Latina y el 4 en el mundo de acuerdo con el reporte publicado por (Mining, 2015). Sin embargo, en 2014, la inversión total en el sector disminuyó en 24.8% (4, 948 MDD), mientras que la Inversión Extranjera Directa, IED, en el sector decreció 59% (2, 117 MDD), todo respecto a 2013. Lo anterior generó un decremento en la capacidad de pago de las empresas mineras, con lo que no sólo se ven afectadas las instituciones que les otorgan crédito, sino también las empresas que integran la cadena productiva de los minerales y/o metales que producen (como es la construcción); lo cual se refleja en el impacto, que este sector aporta al Producto Interno Bruto (PIB) nacional por actividad económica, como se muestra en el cuadro siguiente:

Variación porcentual real anual		
Año	Minería	Construcción
2010	0.88	0.75
2011	-0.37	4.06
2012	0.87	2.45
2013	-0.15	-4.77
2014	-2.27	1.95

Fuente: Elaboración propia con datos del BOLETÍN: Evolución del Producto Interno Bruto (CEFP, 2015).

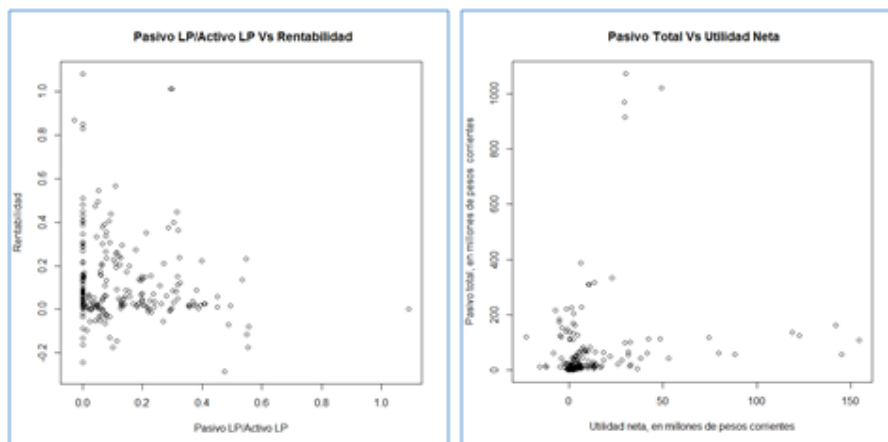
Algunos autores como (Wise & Pozo, 2002), (Tetreault, 2013) para la minería o (Rodríguez - Núñez, 2015) para el petróleo aducen esta contracción del PIB a las

¹ Plata, bismuto, fluorita, celestita, wollastonita, cadmio, molibdeno, plomo, zinc, diatomita, sal, barita, grafito, yeso, oro y cobre.

diferentes reformas fiscales² como la de 2014 , mientras que otros como (Zhang, Nieto, & Kleit, 2015) o (Haque, Topal, & Lilford, 2014) hacen notar la relevancia de la baja generalizada en el precio de los metales en la contracción del PIB. Este entorno macroeconómico global provoca una disminución de la capacidad de las empresas del ramo para hacer frente a sus obligaciones crediticias, tal y como lo mencionan (Khan, Anuar, Choo, & Tahir, 2014) o (Maslyuk & Dharmaratna, 2013).

El ramo minero resulta particularmente interesante dada la combinación de riesgos de mercado que enfrentan, a saber: el precio del bien vendido (commodity), la tasa de interés libre de riesgo (son intensivas en capital), el tipo de cambio (buena parte de la deuda está denominada en moneda extranjera) y el estar eslabonadas a industrias mayoritariamente pro cíclicas (las relacionadas con la construcción y la industria). Evidentemente, todos estos factores, así como sus interacciones y distribuciones de probabilidad repercuten en la capacidad de pago y por ende la probabilidad de incumplimiento de las empresas del ramo, creando entornos de alto apalancamiento con entradas de dinero inciertas, tal y como lo describen (Acharya, Davydenko, & Strebulaev, 2012) o (Douglas, Huang, & Vetzal, 2016). Los efectos de este entorno volátil sobre algunas variables representativas pueden apreciarse mejor en la Gráfica 1.

Gráfica 1. Distribuciones con valores extremos de variables financieras seleccionadas de empresas mineras.



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres I-2012 al II-2013.

La existencia de este entorno volátil y por ende la existencia de extremos, provoca que modelos basados en el paradigma de normalidad o por mucho

² Ésta impone cargas como el derecho especial del 7.5%, el derecho extraordinario del 0.5% a los ingresos derivados de la enajenación del oro, plata y platino y el derecho adicional sobre minería, *i.e.* 50% más de la cuota estipulada por concesiones no exploradas o no explotadas durante dos años continuos.

curtosis moderada, como lo son los modelos Probit o Logit respectivamente, subestimen la probabilidad real de incumplimiento en el pago de las obligaciones del sector minero, afectando con ello a toda la cadena crediticia del sector, estudios sobre los efectos de las colas pesadas en el riesgo crédito se pueden encontrar en el análisis realizado para evaluar la calidad de los modelos estadísticos (Dánielsson, 2008) y los problemas potenciales de sesgo e inestabilidad (Carey & Hrycay, 2001).

En este artículo, postulamos que los modelos *Probit* o *Logit* no capturan correctamente las probabilidades de incumplimiento de pago de las empresas del sector minero, al subestimar las colas, lo que genera problemas de bondad de ajuste y subestimación de dichas probabilidades, mientras que un modelo Poisson de variable dicotómica captura los efectos de cola y estabiliza la regresión, sus estimadores y probabilidades estimadas. Proponemos un modelo de *scoring* más adecuado para las empresas mineras en México, con lo que esperamos se logre disminuir el incumplimiento de pago en el sector y apoyar a las instituciones que les otorgan crédito para mejorar la calidad de su cartera crediticia.

2. Scorings de Crédito

Desde hace varios años, se han desarrollado diferentes metodologías para estimar la probabilidad de incumplimiento de pago de las empresas. Estos modelos van desde las técnicas estadísticas variadas y multivariadas como las usadas en (Cuadras, 2014), modelos de clasificación como los usados en (Chocce, González, Aguirre, Vega, & Carrazana, 2002), árboles de decisión como los postulados en (Cardona H., 2004), modelos de elección cualitativa (*Probit* y *Logit*) (Duffie & Singleton, 2003) o análisis de matrices de transición (Jarrow, Lando, & Turnbull, 1997), entre otros; en donde se evalúa la calidad crediticia de los posibles acreditados, considerando variables financieras y de calidad crediticia obtenidas por empresas especializadas, así como el posible deterioro de su calidad crediticia. El común denominador de estos modelos es su uso de razones financieras (en muchos casos se utilizan variables correlacionadas que sesgan los modelos al contribuir en el mismo concepto, como es el caso de las razones de liquidez, apalancamiento, rentabilidad, etc.) e indicadores macroeconómicos para tratar de predecir la quiebra de la empresa.

Como alternativa a estos modelos basados en razones financieras, se han desarrollado modelos internos (consideran las particularidades de cada institución otorgadora de crédito y de sus clientes). Los modelos internos más conocidos y utilizados internacionalmente son *CreditRisk* (Wilde, 1997), y *CreditMetrics* (J. P. Morgan, 1997), los cuales son una adaptación del concepto de Valor en Riesgo (VaR) desarrollado para riesgo de mercado. Es importante destacar que ambos modelos pueden ser equivalentes mediante el uso de variables latentes dicotómicas provenientes de una regresión Logit, tal y como lo señala (Gordy, 2000).

Por otra parte, es necesario mencionar algunos otros modelos internos desarrollados para predecir las quiebras de las empresas, para más detalles véase (Altman E. I., 1968), (Ohlson, 1980). Modelos internos más recientes pueden encontrarse en (Lee & Choi, 2013) quienes usan redes neurales y discriminantes sobre variables financieras para tratar de predecir la banca rota, (Delen, Kuzey, & Uyar, 2013) usan árboles de decisión sobre variables financieras para medir el

desempeño de las empresas o (Zhou, 2013) quien muestra los efectos de distintas formas de muestreo sobre variables financieras en la capacidad de los modelos para predecir las quiebras. Es importante destacar que en todos los casos se busca distinguir entre los estados de quiebra y supervivencia de la empresa, aunque en ocasiones se busca evaluar la posibilidad de un cambio en la nota crediticia.

Aunque se han propuesto una gran variedad de modelos para medir la probabilidad de incumplimiento de las empresas, algunos de ellos muy sofisticados como los propuestos por (Baesens, *et al.*, 2003), quienes analizan varios tipos de modelado contra modelos probabilísticos, obteniendo resultados similares en todos ellos o el propuesto por (Bharath & Shumway, 2008) quienes usan una aproximación basada en el modelo de (Merton, 1974) para conocer la probabilidad de incumplimiento o el modelo propuesto por (Bottazzi, Grazi, Secchi, & Tamagni, 2011) quienes usan variables macroeconómicas y financieras para predecir el incumplimiento.

Entre los modelos no probabilistas usados en la literatura para calcular dichas probabilidades, también se encuentran los propuestos por (Huang, Chen, & Wang, 2007) quienes construyen modelos de calificación de crédito basado en *Support Vector Machines* (SVM) híbridos, para evaluar la puntuación de crédito de solicitantes de tarjetas de crédito, por su parte, (Ping & Yongheng, 2011) hacen un esfuerzo similar, el que comparan con modelos de análisis discriminante lineal, regresión logística y redes neuronales.

3. Aplicación a la Industria Minera en México

En este trabajo, se busca analizar la estabilidad y con ello confiabilidad de las probabilidades de incumplimiento obtenidas mediante técnicas econométricas estándar tanto a lo largo del tiempo como a través de diferentes submuestras. Para ello se hace uso de modelos de probabilidad Logit y Poisson sobre una muestra que incluye una serie de variables financieras internas de las empresas.³ Los datos de este estudio fueron obtenidos de los estados financieros

³ ACT: Activo Circulante Total, AFT: Activo Fijo Total, AT: Activo Total, PC: Pasivo a Corto Plazo, PLP: Pasivo a Largo Plazo, PT: Pasivo Total, CS: Capital Social, CC: Capital Contable, TI: Total de Ingresos, CV: Costos de Ventas, UB: Utilidad Bruta, UO: Utilidad de Operación, UAI: Utilidad Antes de Impuestos, UN: Utilidad Neta, R_LIQUIDEZ: Ratio de Liquidez, R_SOLVENCIA: Ratio de Solvencia, NWC: Capital de Trabajo Neto, WC_A: Capital de Trabajo Sobre Activos, WC_D: Capital de Trabajo Sobre Deudas a Corto Plazo, Dias: Días de Medición Del Intervalo Tiempo, R_ENDEUDAMIENTO: Razón de Endeudamiento, R_END_INVERSI.N: Razón de Endeudamiento Sobre la Inversión a CP, R_END_INVERSIÓN_LP: Razón de Endeudamiento Sobre la Inversión a LP, CALIDAD_DEUDA: Razón de Calidad de la Deuda, COSTO_DEUDA: Coste de la Deuda, R_ESTABILIDAD: Razón de Estabilidad, AT_CC: Apalancamiento Financiero, V_AT: Rotación Total de Activos, R_ACTIVOS_FIJO: Rotación de Activos Fijos, MARGEN_UTILIDAD: Margen de Utilidad, ROA: Rendimiento Sobre Activos, ROE: Rendimiento Sobre El Capital, RENTABILIDAD: Rentabilidad Económica. Los datos de este estudio fueron obtenidos de los estados financieros de 71 empresas relacionadas con la minería en México a lo largo del periodo del II/2012 y hasta III/2013, las cuales fueron analizadas de manera conjunta, trimestral y por cuartiles agrupados según la razón de endeudamiento, la cual es la variable que resultó significativa en todas las regresiones Poisson previas (total y

de 71 empresas relacionadas con la minería en México a lo largo del periodo del II/2012 y hasta III/2013, las cuales fueron analizadas de manera conjunta, trimestral y por cuartiles agrupados según la razón de endeudamiento, la cual es la variable que resultó significativa en todas las regresiones Poisson previas (total y trimestral).

a. Aplicación del Modelo Logit

La elección de un modelo econométrico resulta relativamente clara tanto por la restricción impuesta por el hecho de que la probabilidad de incumplimiento se encuentra entre 0 y 1, como por la práctica tradicional de elegir un modelo Logit para este análisis, el cual parte del supuesto de un proceso estocástico *iid* como una variable logística de la forma:

$$Y_i = \frac{e^{\alpha + \beta_k X_{ki}}}{1 + e^{\alpha + \beta_k X_{ki}}} + \varepsilon_i. \quad (1)$$

Para mayores detalles sobre este modelo, véase (Hosmer, Lemeshow, & Sturdivant, 2013), (King & Zeng, 2001) o (Peng, Lee, & Ingersoll, 2002). En estos textos se resalta que los signos del modelo indican la dirección de movimiento de la probabilidad ante cambios en la variable independiente, aunque el valor del parámetro no coincide con la magnitud de la variación en la probabilidad. Para conocer esta sensibilidad es necesario construir el factor de sensibilidad (valuar la función de densidad en los parámetros promedio) y luego multiplicar por el valor del parámetro.

Con la finalidad de evitar (en la medida de lo posible) un sesgo por variable omitida, se inició el análisis del modelo con todo el conjunto de variables de la base de datos, para luego ir descartando la de mayor probabilidad de ser estadísticamente cero hasta que el modelo pasara todas las pruebas econométricas convencionales, las cuales fueron realizadas en el paquete R (R Core Team, 2015).⁴

En la siguiente tabla (Tabla 2) se muestran los resultados de la regresión Logística aplicada a toda la base de datos.⁵

trimestral).

⁴ Todos los parámetros estadísticamente significativos, el menor criterio de Akaike, Bondad de ajuste, normalidad e independencia de residuales. El código en R que se usó para el análisis, así como la base de datos están a disposición de los lectores interesados a vuelta de correo electrónico.

⁵ En este caso, las variables que resultaron estadísticamente significativas fueron: R_LIQUIDEZ (Cociente 3 de Liquidez), R_SOLVENCIA (Cociente de Solvencia), R_END_INVERSI.N (Cociente de Endeudamiento Sobre la Inversión a CP) y T1 (dada por el cociente de la diferencia entre Activo circulante total menos Pasivo a corto plazo entre el Activo fijo total).

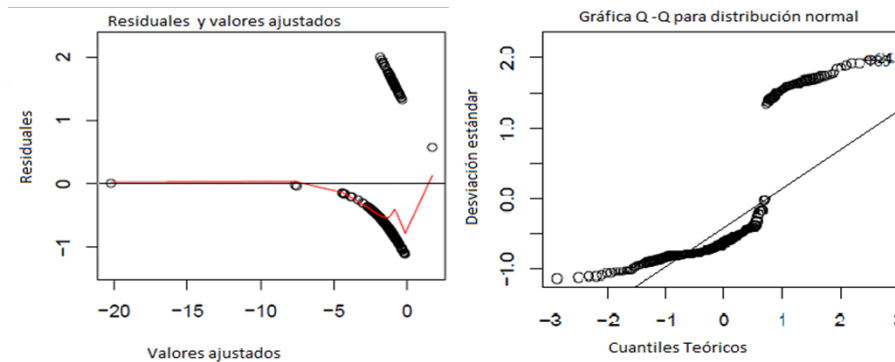
Tabla 2. Regresión logística para la muestra entera.

Variable	Variable	Pr(> z)	Estimador ajustado por la pendiente
R_LIQUIDEZ	0.0138	0.0351	0.0112
R_SOLVENCIA	-0.1029	0.0095	-0.0832
R_END_INVERSI.N	-2.5766	0.0000	-2.0840
T1	-0.2688	0.0265	-0.2174
AIC =	263.94	Pendiente =	0.8088178

Fuente: Elaboración propia con datos del BOLETÍN: Evolución del Producto Interno Bruto (CEFP, 2015).

En la Tabla 2 es llamativo el hecho de que el signo asociado a todas las variables independientes es contrario al esperado aún y cuando el parámetro es estadísticamente significativo, sin mencionar que los residuales de la regresión son claramente no normales, esto se puede notar en la Gráfica 2, dónde se aprecian valores de cola en los residuales, lo que puede explicar el signo no esperado en el parámetro.

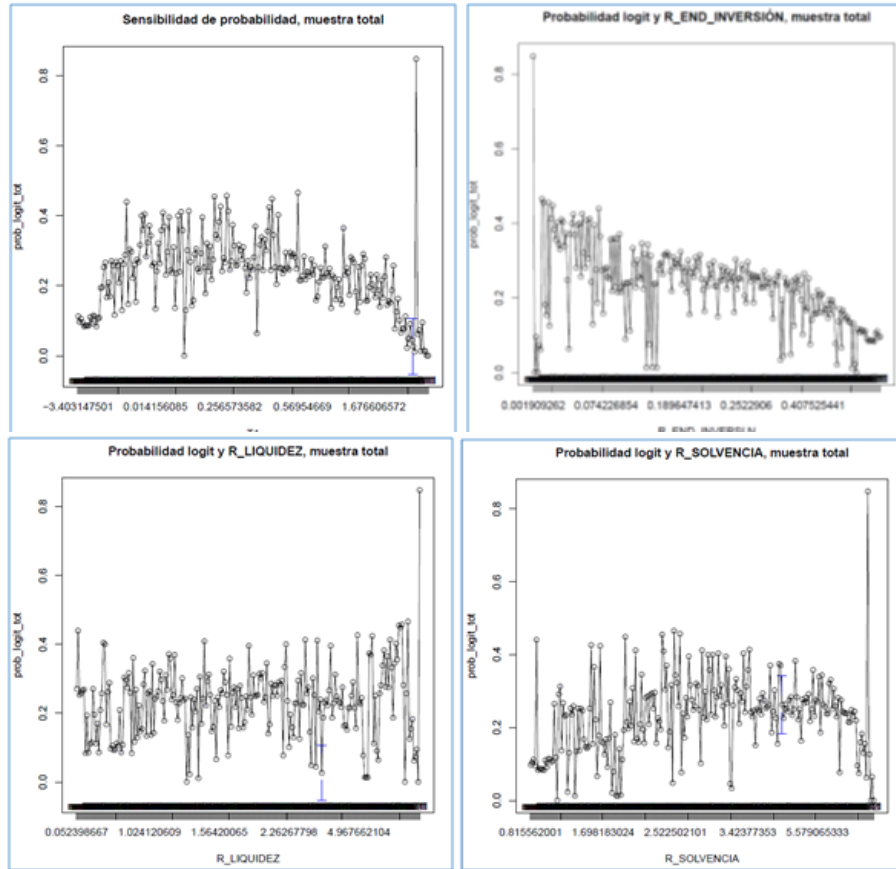
Gráfica 2. Gráfico de ajuste y residuales del modelo Logit aplicado a la muestra completa.



Fuente: Elaboración propia con datos del BOLETÍN: Evolución del Producto Interno Bruto (CEFP, 2015).

En la siguiente gráfica (Gráfica 3) se muestra la evolución de las probabilidades de incumplimiento estimadas a lo largo de los valores de las variables independientes. De estos gráficos se desprende que el modelo Logit predice probabilidades de incumplimiento similares a diferentes niveles de todas las variables independientes, excepto el cociente de endeudamiento sobre la inversión (R_END_INVERSI.N), el cual es decreciente.

Gráfica 3. Evolución de las probabilidades estimadas por el modelo Poisson para toda la muestra a lo largo de las variables significativas del modelo.



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

Con la finalidad de evaluar si el modelo es robusto y buscando examinar si los valores extremos que afectan al modelo están asociados con la estacionalidad de los datos, se ha dividido la muestra en trimestres. A cada una de estas submuestras se le ha aplicado el mismo modelo logístico partiendo de la totalidad de las variables de la muestra y depurándolo al eliminar la variable independiente con mayor probabilidad de ser cero. Como resultado de este ejercicio, se obtuvieron distintas variables explicativas para cada trimestre,⁶ cuyos resúmenes del análisis de regresión logística se muestran en la Tabla 3.

⁶ Para el trimestre 1 el Rendimiento sobre Activos (ROA) y T3 (Utilidad de operación / Activo Total), para el trimestre 2 el R_END_INVERSIÓN (Cociente de Endeudamiento Sobre la Inversión a CP) y T4 (dada por el cociente de Capital Contable entre Pasivo Total), para el trimestre 3 R_ACTIVOS_FIJO (Rotación de Activos Fijos) y para el trimestre 4 V_AT (Rotación Total de Activos) y T2 (el cociente de la diferencia de Capital Contable del año t menos Capital Contable del año t-1 entre Activo Total).

Tabla 3. Resumen de los resultados de la regresión logística para las 4 sub muestras trimestrales.

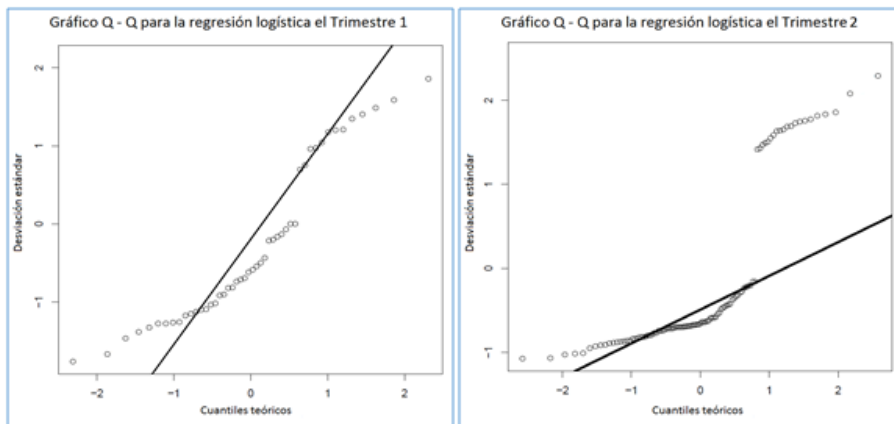
Trimestre	Variable	Estimador	Pr(> z)	AIC	Pendiente del Modelo	Estimador ajustado por la pendiente
Trimestre 1	ROA	49.4800	0.0104	54.374	0.7831	38.7502
Trimestre 1	T3	-42.0200	0.0107	54.374	0.7831	-32.9080
Trimestre 2	R_END_INVERSI.N	-3.8660	0.0002	102.54	0.8135	-3.1450
Trimestre 2	T4	-0.1208	0.0659	102.54	0.8135	-0.0983
Trimestre 3	R_ACTIVOS_FIJO	-0.2645	0.0097	57.915	0.9047	-0.2393
Trimestre 4	V_AT	-0.6867	0.0036	59.193	0.7726	-0.5305
Trimestre 4	T2	5.9167	0.0452	59.193	0.7726	4.5709

Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

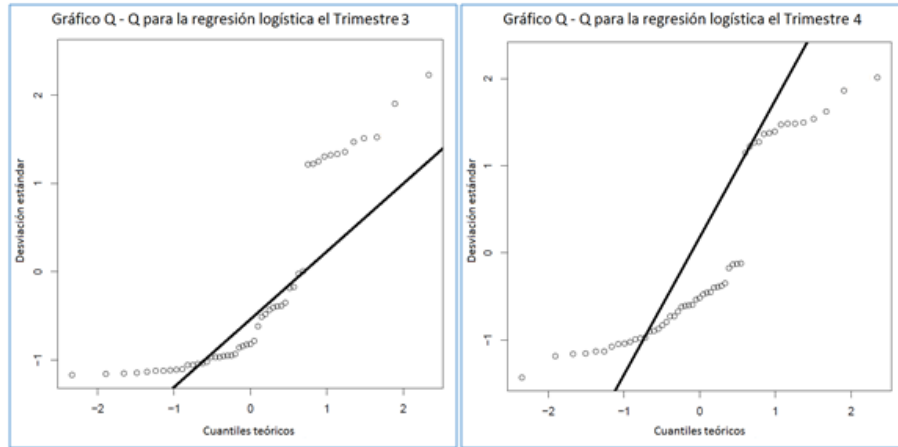
En la Tabla 3 se muestra que la regresión logística no es estable a lo largo de los trimestres. Esta inestabilidad resulta problemática para una empresa otorgante de crédito pues implicaría que las variables determinantes de la capacidad de pago de las empresas mineras cambian con cada trimestre, lo cual resulta ilógico.

Adicionalmente se muestran en la Gráfica 4 las comparaciones del cuantil normal *versus* el cuantil empírico de los cuatro trimestres. En esta gráfica se muestra el mismo problema de ajuste para los valores extremos de cada sub muestra trimestral.

Gráfica 4. Gráficas de cuantiles normales versus cuantiles empíricos para los residuales de las regresiones logísticas trimestrales.



Gráfica 4. Gráficas de cuantiles normales versus cuantiles empíricos para los residuales de las regresiones logísticas trimestrales.
continuación



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

Algunos de los signos parecieran ser contrarios a lo que indica la teoría financiera, como son el ROA (relación directa) y el cociente de endeudamiento e inversión (relación inversa). Con la finalidad de hacer el modelo comparable con otros en la literatura, se construyeron las variables de la Z^7 de (Altman E. I., 1968) para los datos de empresas mineras, incluyendo la Z propuesta para empresas pequeñas en (Altman & Sabato, 2007), aunque estas no resultaron estadísticamente significativas en todos los casos. El caso más llamativo es el de T2 (Ingresos Retenidos / Activo Total) con un signo negativo que podría estar asociado con el que las firmas usan los recursos para crecer y no para liquidar pasivos. Lamentablemente, la inestabilidad de la regresión impide hacer mayores análisis como el practicado en (Eidleman, 1995) o (Diakomihalis, 2012).

b. Aplicación del Modelo Poisson

Ante la poca robustez del modelo Logístico tradicionalmente usado en la medición de las probabilidades de incumplimiento de las empresas, se decidió explorar la posibilidad de que la distribución usada no fuese la adecuada, dado el tamaño de las colas visto en los residuales de las regresiones logísticas. Para enfrentar este problema se recurrió al supuesto de una distribución Poisson (dada su plasticidad) para dichas probabilidades. Ejemplos de su uso en análisis de riesgo crédito se encuentran en (Nelson, Lust, Story, & Ehlinger, 2008) o (Löeffler & Posch, 2011). En particular la distribución Poisson está dada por:

$$P(x, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \text{ con } x = 0, 1, 2, \dots, e \approx 2.718228 \dots \quad (2)$$

⁷ T1: (Capital de Trabajo / Activo Total), T2: (Ingresos Retenidos / Activo Total), T3: (EBITDA / Activo Total), T4: (Valor de mercado del capital / Pasivo Total), T5: (Ventas / Activos Totales).

dónde el parámetro de la distribución λ es igual a la media y a la varianza de la variable.

Al igual que en el caso de las regresiones logísticas, se inició la depuración del modelo con el conjunto completo de variables, para después eliminar la variable independiente con mayor probabilidad de ser estadísticamente cero hasta que todas las variables del modelo fueron significativas y se obtuviera el menor criterio de Akaike posible.⁸

En la Tabla 4, se muestra la regresión Poisson aplicada a toda la muestra. En ella se sigue encontrando un problema de signo en la variable del cociente de liquidez (R_Liquidez), aunque las probabilidades de ser estadísticamente cero mejoran, así como el ajuste en los residuales mostrados en la Gráfica 4.

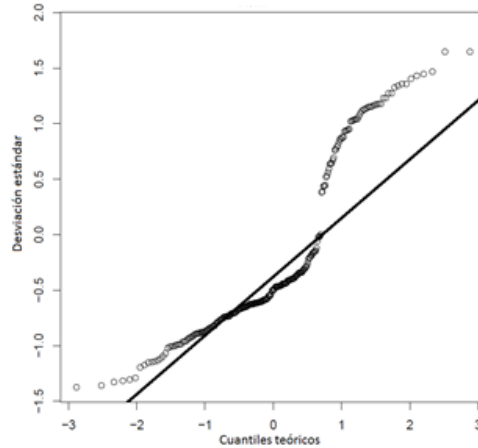
Tabla 4. Parámetros de la regresión Poisson para toda la muestra.

variable	Variable	Pr(> z)	Estimador ajustado por la pendiente
R_LIQUIDEZ	0.011474	0.000757	0.000455
V_AT	0.48455	0.003458	0.019198
Z	-0.36683	1.47E-06	-0.014534
R_END_INVERSI.N	-3.357595	6.47E-07	-0.133031

Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

En la Gráfica 5 se aprecia la forma en que la distribución Poisson captura de mejor manera los efectos de cola de la muestra, aunque su ajuste no es el mejor en el intervalo de 0.5 a 1.5 desviaciones estándar de los errores, lo que provoca los problemas antes señalados.

Gráfica 5. Gráfica de cuantiles normales versus cuantiles empíricos para los residuales de la regresión Poisson para la muestra total.

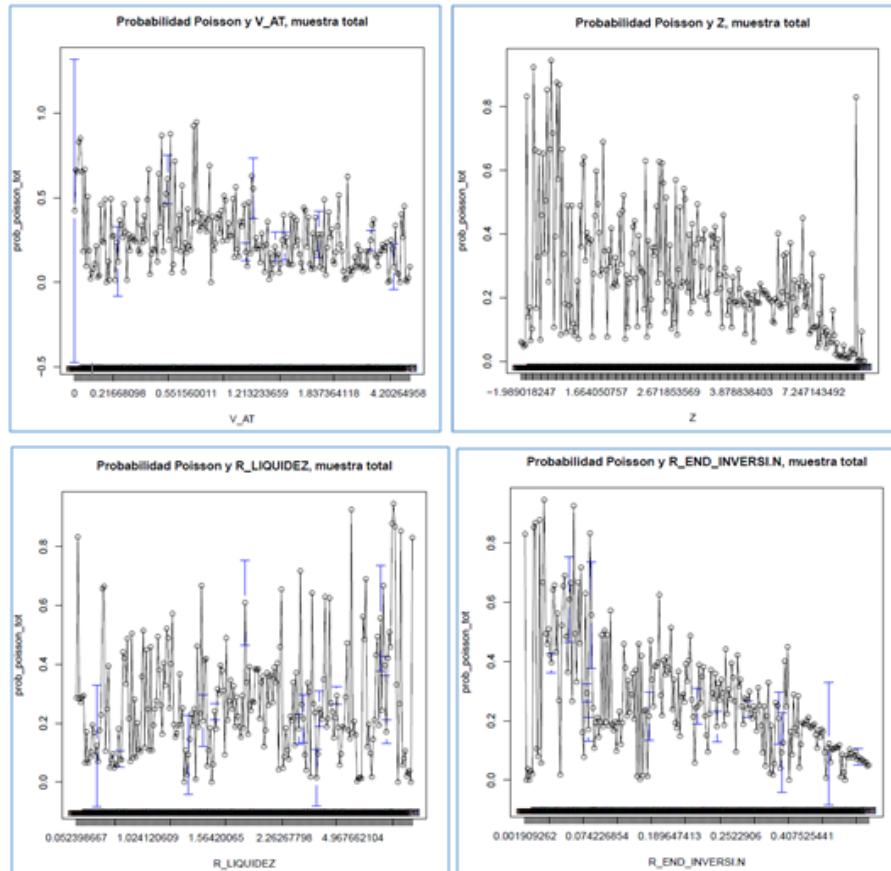


Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

⁸ Este procedimiento pretende reducir la posibilidad de sesgar el modelo por problemas de variable omitida

Con la finalidad de analizar el comportamiento de la probabilidad de incumplimiento ante cambios en las variables independientes, se muestra en la Gráfica 6 la evolución de la probabilidad estimada por el modelo respecto a los movimientos de las variables estadísticamente significativas del modelo Poisson. Es necesario hacer notar la tendencia decreciente en las gráficas de las variables Z y cociente del endeudamiento sobre inversión y las pendientes positivas aunque pequeñas en el resto de las variables; también se hace notar la existencia de valores aberrantes en los extremos de las gráficas, los cuales ocasionan los problemas de bondad de ajuste.

Gráfica 6. Evolución de las probabilidades estimadas por el modelo Poisson para toda la muestra a lo largo de las variables significativas del modelo.



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

Es posible que los problemas de signo presentados en la regresión Poisson aplicada a toda la muestra (Tabla 4) se deban al ajuste forzado en el intervalo de la muestra donde ésta se aleja del cuantil teórico normal (0.5 a 1.5 desviaciones estándar), por lo que es preciso subdividir la muestra para probar esta hipótesis.

Por tanto, también se ha aplicado una regresión Poisson a cada uno de los trimestres de la misma. Los resultados de este ejercicio se muestran en la Tabla 5.

Tabla 5. Resumen de los resultados de la regresión Poisson para las 4 sub muestras trimestrales.

Trimestre	Variable	Estimador	Pr(> z)	AIC	Pendiente del Modelo	Estimador ajustado por la pendiente
Trimestre 1	R_ENDEUDAMIENTO	-3.1957	0.0000361	62.351	0.04094	-0.13083
Trimestre 2	R_ENDEUDAMIENTO	-3.6531	3.85E-10	121.52	0.02591	-0.09465
Trimestre 3	R_ENDEUDAMIENTO	-3.9545	0.0000081	62.918	0.01917	-0.07580
Trimestre 4	R_ENDEUDAMIENTO	-2.153	0.0047	74.022	0.00803	-0.01729
Trimestre 4	T3	-2.6716	0.0962	74.022	0.00803	-0.02145

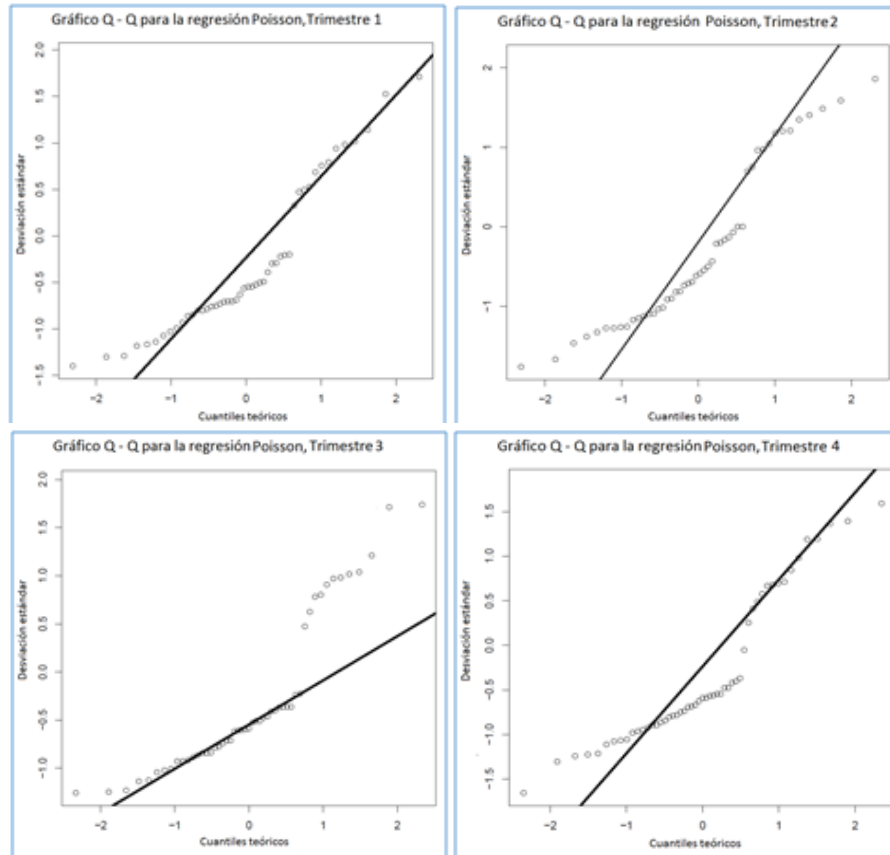
Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

En la Tabla 5 se observa que el uso de regresiones Poisson estabiliza significativamente tanto los regresores como sus valores, aunque aún se tiene el problema de signos no esperados, pues la regresión Poisson implica que a mayor cociente de endeudamiento,⁹ menor es la probabilidad de incumplimiento. Este problema se mantiene a lo largo de los tres primeros trimestres con estimadores estadísticamente significativos que también son estadísticamente iguales.

En la Gráfica 7, se muestran los análisis de cuantil normal versus cuantil empírico de los residuales de las regresiones Poisson de los cuatro trimestres. En estas gráficas es posible apreciar que, al igual que en los intentos anteriores, los signos contrarios a los esperados están relacionados con separaciones importantes de los residuales respecto a la distribución normal, lo que indica que no se ha logrado aislar la característica de la muestra que causa estos valores extremos:

⁹ Calculado como el cociente del Pasivo total entre los Activos Totales.

Gráfica 7. Gráficas de cuantiles normales versus cuantiles empíricos para los residuales de las regresiones Poisson trimestrales.



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

Ante esta evidencia, se dividió la muestra en cuantiles siguiendo la distribución empírica del cociente de endeudamiento (la variable es significativa en las regresiones Poisson de tres de los cuatro trimestres y en la regresión Poisson para toda la muestra). Con esta nueva forma de partición de la muestra, se pretenden separar los valores de cola inferior y superior que afectan los ajustes alrededor de la media. En la Tabla 6 se muestran los resultados de las regresiones Poisson practicadas para cada uno de los cuatro cuantiles de la muestra.

Tabla 6. Resumen de los resultados de la regresión Poisson para los 4 cuartiles de la muestra.

Trimestre	Variable	Estimador	Pr(> z)	AIC
Cuartil 1	WC_D	1.5725	4.15E-08	74.238
Cuartil 1	R_LIQUIDEZ	-1.5708	3.86E-08	74.238
Cuartil 2	WC_D	0.9397	0.004215	100.21
Cuartil 2	R_LIQUIDEZ	-0.933	0.000171	100.21
Cuartil 3	WC_D	1.8789	0.0000138	56.968
Cuartil 3	R_LIQUIDEZ	-1.9001	0.000000594	56.968
Cuartil 4	WC_D	2.1172	0.000000199	66.623
Cuartil 4	R_LIQUIDEZ	-1.8874	3.66E-08	66.623

Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

Para este grupo de regresiones, se tienen los signos teóricos esperados en todos los cuartiles (directamente proporcional para el capital de trabajo neto sobre deudas¹⁰ e inversamente proporcional para el cociente de liquidez¹¹) con parámetros estimados estadísticamente significativos a menos del 1%, los cuales son estadísticamente diferentes entre sí (cada cuartil tiene su propio estimador, distinto del resto, aun cuando las variables explicativas sean las mismas).

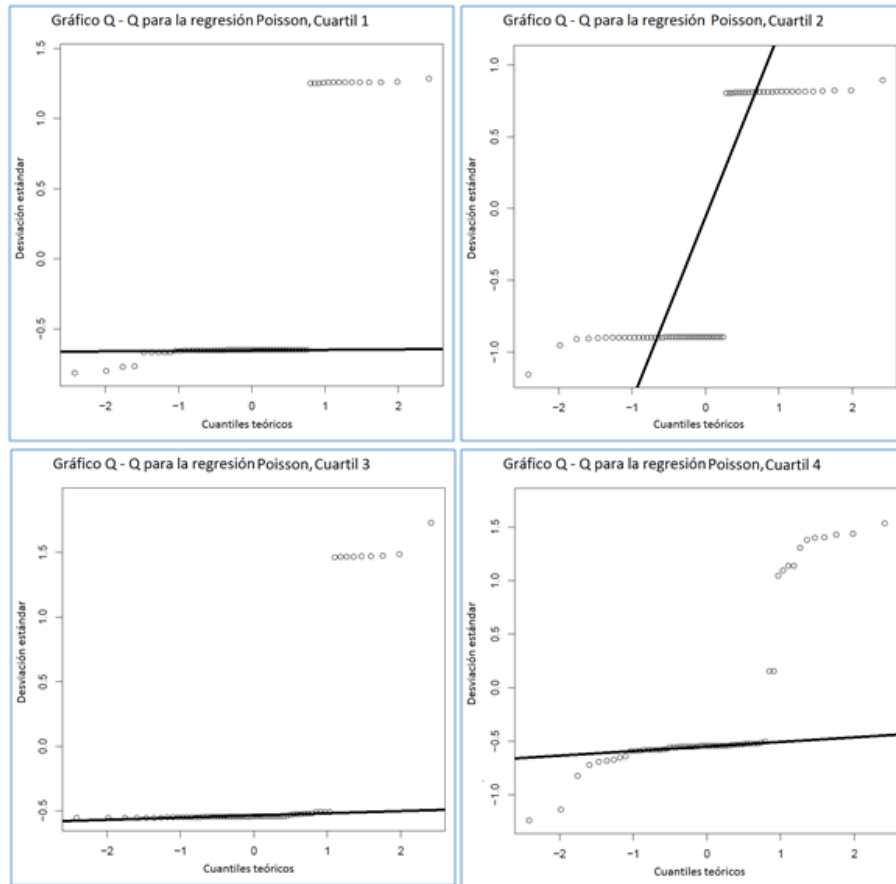
Resulta llamativo el hecho de que en este grupo de regresiones sean indicadores financieros de corto plazo (el capital de trabajo dedicado a los pasivos de corto plazo y el cociente de liquidez) los que conduzcan la probabilidad de incumplimiento de las empresas del ramo minero. Esto parece indicar que en general, estas empresas no tienen una planeación financiera de largo plazo y pagan sus créditos con los flujos de efectivo de corto plazo.

También es necesario hacer notar el que los criterios de Akaike (bondad de ajuste) son mejores para los dos primeros cuartiles que para el tercero y cuarto (donde todas las regresiones anteriores perdían su ajuste y generaban cambios en los signos). Con la finalidad de revisar el ajuste del modelo, se muestra en la Gráfica 8 el ajuste de los cuantiles teóricos de la normal contra los empíricos de los residuales de las regresiones Poisson aplicada a cada uno de los cuartiles de la muestra.

¹⁰ (Activos Circulantes - Pasivo Circulante) / Pasivo Circulante.

¹¹ Activos Circulantes - Pasivo Circulante.

Gráfica 8. Gráficas de cuantiles normales versus cuantiles empíricos para los residuales de las regresiones Poisson por cuantiles.



Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

En la Gráfica 8 se muestra que aunque siguen existiendo algunos valores aberrantes al interior de los cuantiles, estos son menores en cantidad en la probabilidad, lo que termina por validar los estimadores obtenidos en los cuantiles.

Con la finalidad de medir el desempeño de los modelos propuestos, se realizaron pruebas de efectividad predictiva. Estas pruebas toman como umbrales de conteo el 20%, 50% y 80% de probabilidad, cualquier probabilidad superior es considerada un incumplimiento, mientras que cualquier probabilidad menor como cumplimiento, los resultados de estas pruebas de efectividad se muestran en la Tabla 7.

Tabla 7. Efectividad de cada estrategia de modelado a diferentes umbrales de sensibilidad..

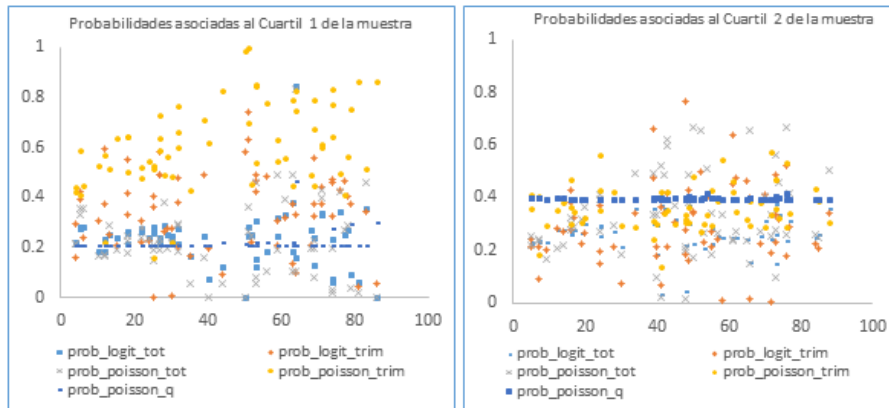
Umbral	Efectividad Logit Total	Efectividad Logit Trimestral	Efectividad Poisson Total	Efectividad Poisson Trimestral	Efectividad Poisson por cuantiles
0.2	0.5217391	0.498023715	0.604743083	0.482213439	0.553359684
0.5	0.7628458	0.750988142	0.727272727	0.644268775	0.758893281
0.8	0.7628458	0.76284585	0.739130435	0.731225296	0.766798419

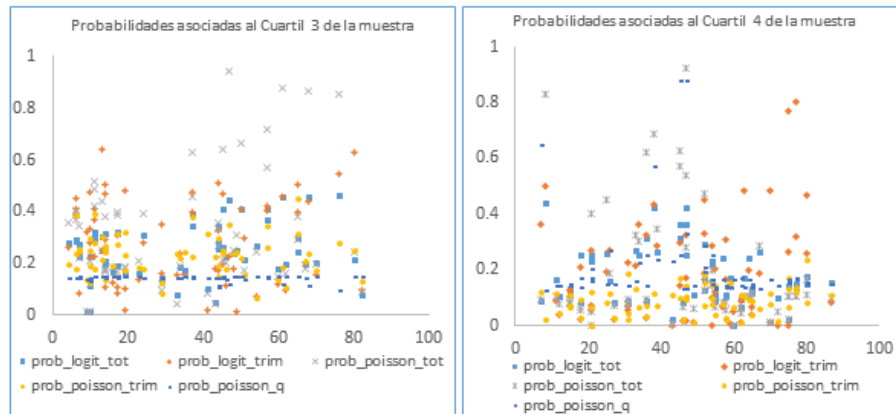
Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

En la Tabla 7 se muestran los mejores modelos para cada umbral de probabilidad. Es posible observar que, aunque la diferencia en efectividad entre los modelos Logit para la muestra entera y los Poisson aplicados a los cuantiles no son muy diferentes (menos de 2% en algunos casos), el ajuste de los modelos, su interpretación y por tanto recomendaciones de política de asignación de créditos pueden llegar a ser muy diferente.

En efecto, aunque el modelo Logit para la muestra entera es casi tan efectivo como el Poisson por cuantiles, sus problemas con los signos de las variables independientes y la no normalidad de sus residuales hacen difícil tanto su interpretación como su uso para determinar la asignación de créditos. Todo esto se puede observar mejor en la Gráfica 9, donde se muestran las probabilidades generadas por cada uno de los modelos propuestos a lo largo de los cuatro cuantiles de la muestra.

Gráfica 9. Probabilidad de los distintos modelos propuestos, por cuantiles.



Gráfica 9. Probabilidad de los distintos modelos propuestos, por cuartiles.
(continuación)

Fuente: Elaboración propia con R Statistical Package, con datos del FIFOMI, trimestres II-2012 al III-2013.

En la gráfica 9 se puede ver que las probabilidades obtenidas mediante los modelos Poisson por cuantiles son las más estables de todas las obtenidas a lo largo de la investigación. También se puede observar que las probabilidades basadas en modelos logísticos tienen una gran dispersión, aunque casi siempre se sitúan por encima de las probabilidades Poisson.

Es de notarse el hecho de que las probabilidades de los distintos modelos parecen agruparse en estratos. Estos estratos parecen ser más compactos para los modelos Poisson por cuantiles y más dispersos para los Logísticos. Queda como posible línea de investigación el análisis de la distribución Weibull u otras de la familia de Pareto.

4. Conclusiones

En este trabajo se analizaron modelos que pretenden evaluar la posibilidad de un cambio en la calidad crediticia de las empresas mineras mexicanas y con ello determinar si son candidatas o no para un crédito; mediante un análisis de estabilidad y confiabilidad de las probabilidades de incumplimiento de pago, utilizando modelos Logit y Poisson; así como variables financieras internas de las empresas. Los resultados indican que los modelos Logísticos no están diseñados para analizar óptimamente las variables independientes con valores extremos; en el caso de los datos de la minería mexicana los valores extremos no están asociados al trimestre de operación, sino que son propios de cada empresa. Los modelos Poisson resultaron ser capaces de captar los valores extremos de la distribución; por lo que son más adecuados para determinar la calidad crediticia de las empresas mineras y con ello determinar si se les otorga un préstamo.

Se presentaron problemas de signos contrarios, respecto a la teoría, esto se debe a una forma funcional del modelo erróneo y/o a la incorrecta segmentación de la muestra; se eliminó gran parte de la probabilidad de tener problemas de variables omitidas al introducir todas las variables de la base de datos e ir eliminando las menos significativas. Se queda como línea de investigación el incluir la afectación de las variables macro a cada empresa.

Se requiere de un modelo estable, con signos correctos y errores lo más cercanos posible a la normal para asignar probabilidades de incumplimiento parsimoniosas a los acreditados, de tal manera que la política interna de asignación de crédito no esté sesgada por una elección incorrecta del modelo y con esto garantizar acreditados de mejor calidad.

En el presente estudio realizado para las empresas mineras mexicanas, el único modelo que resultó parsimonioso y estable fue el Poisson aplicado a la muestra segmentada por cuantiles, ya que las mismas variables fueron significativas en todos los cuantiles y los parámetros son estadísticamente distintos entre sí. Los resultados de la regresión Poisson por cuantiles, indican que los indicadores financieros de corto plazo (el capital de trabajo dedicado a los pasivos de corto plazo y el cociente de liquidez) son los que conducen la probabilidad de incumplimiento de pago; esto indica que dichas empresas no tienen una planeación financiera de largo plazo y pagan sus créditos con los flujos de efectivo de corto plazo.

La relevancia de esta investigación radica, en que no existen otros estudios en donde se realicen comparaciones similares entre modelos de *scoring* para la industria minera (en general para las micro, pequeña y mediana empresa mexicana); también ofrece una visión estadística y econométrica de la forma en que estos modelos pueden ser contrastados y las muestras segregadas con la finalidad de ser más estables y fidedignos, con lo que se puede obtener mejor calidad en los acreditados aceptados.

Bibliografía

- Acharya, V., Davydenko, S., & Strebulaev, I. (2012). Cash Holdings and Credit Risk. *Review of Financial Studies*, 25 (12), 3572-3609.
- Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23 (4), 589-609.
- Altman, E., & Sabato, G. (2007). Modelling Credit Risk for SMEs: Evidence from the US Market. *Abacus*, 43 (3), 332-357.
- Baesens, B., Van Gestel, T., & Vanthienen, J. (2003). Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring. *Journal of the Operational Research Society*, 54(6), 627-635.
- Bellotti, T., & Crook, J. (2009). Support Vector Machines for Credit Scoring and Discovery of Significant Features. *Expert Systems with Applications*, 36 (2), 3302-3308.
- Bharath, S., & Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. *Review of Financial Studies*, 21(3), 1339-1369.
- Bottazzi, G., Grazi, M., Secchi, A., & Tamagni, F. (2011). Financial and Economic Determinants of Firm Default. *Journal of Evolutionary Economics*, 21(3), 373-406.
- Cardona, H. (2004). Aplicación de árboles de decisión en modelos de riesgo crediticio. *Revista Colombiana de Estadística*, 27 (2), 139 - 151.
- Carey, M., & Hrycay, M. (2001). Parameterizing Credit Risk Models with Rating Data. *Journal of Banking & Finance*, 25 (1), 197 - 270.
- CEFP, C. d. (2015). BOLETÍN: *Evolución del Producto Interno Bruto. México D.F.: Centro de Estudios de las Finanzas Públicas (CEFP)*, LXII Legislatura, Camara de Diputados.
- Chocce, G., González, P., Aguirre, N., Vega, P., & Carrazana, J. (2002). Modelos de clasificación y predicción de quiebra de empresas: una aplicación a empresas chilenas. *Forum Empresarial*, 7 (1), 33 - 50.
- Cuadras, C. (2014). *Nuevos Métodos de Análisis Multivariante*. Barcelona, España: CMC Editions.

- Daniélsson, J. (2008). Blame the Models. *Journal of Financial Stability*, 4 (4), 321 - 328.
- Delen, D., Kuzey, C., & Uyar, A. (2013). Measuring Firm Performance using Financial Ratios: A Decision Tree Approach. *Expert Systems with Applications*, 40 (10), 3970-3983.
- Diakomihalis, M. (2012). The Accuracy of Altman's Models in Predicting Hotel Bankruptcy. *International Journal of Accounting and Financial Reporting*, 2 (2), 96 - 113.
- Douglas, A., Huang, A., & Vetzal, K. (2016). Cash flow Volatility and Corporate Bond Yield Spreads. *Review of Quantitative Finance and Accounting*, 46 (2), 417-458.
- Duffie, D., & Singleton, K. (2003). *Credit Risk, Pricing, Measurement, and Management*. Princeton, New Jersey: Princeton Series in Finance.
- Eidleman, G. (1995). Z Scores-A Guide to Failure Prediction. *The CPA Journal*, 65 (2), 52 - 55.
- Gordy, M. (2000). A Comparative Anatomy of Credit Risk Models. *Journal of Banking & Finance*, 24(1), 119-149.
- Haque, M., Topal, E., & Lilford, E. (2014). A numerical study for a mining project using real options valuation under commodity price uncertainty. *Resources Policy*, 39, 115-123.
- Hosmer, J., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Hoboken, NJ: John Wiley & Sons.
- Huang, L., Chen, M., & Wang, C. (2007). Credit Scoring with a Data Mining Approach based on Support Vector Machines. *Expert systems with applications*, 33 (4), 847-856.
- Morgan, J. (1997). CreditMetrics-Technical Document. New York: J.P. Morgan & Co. Incorporated.
- Jarrow, R., Lando, D., & Turnbull, S. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *The Review of Financial Studies*, 10 (3), 481 - 523.
- Khan, F., Anuar, M., Choo, L., & Tahir, M. (2014). Lagged Macroeconomic Variables Sensitivity of Firm Stock Returns on an Emerging Stock Market. *International Journal of Information Processing and Management*, 5(2), 40.
- King, G., & Zeng, L. (2001). Logistic Regression in Rare Events Data. *Political analysis*, 9 (2), 137-163.
- Lee, S., & Choi, W. (2013). A Multi-Industry Bankruptcy Prediction Model using Back-Propagation Neural Network and Multivariate Discriminant Analysis. *Expert Systems with Applications*, 40 (8), 2941-2946.
- Löffler, G., & Posch, P. (2011). *Credit Risk Modeling using Excel and VBA*. West Sussex: John Wiley and Sons.
- Maslyuk, S., & Dharmaratna, D. (2013). Impact of Shocks on Australian Coal Mining. *In Global Energy Policy and Security*, 231 - 255.
- Merton, R. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29 (2), 449-470.
- Mining, S. (2015). Mining Financing by Region, *January 2013-March 2015*. PDAC International Convention.
- Nelson, M., Lust, K., Story, M., & Ehlinger, E. (2008). Credit Card Debt, Stress and Key Health Risk Behaviors among College Students. *American Journal of Health Promotion*, 22 (6), 400-407
- Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18 (1), 109-131.
- Peng, C., Lee, K., & Ingersoll, G. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, 96 (1), 3 - 14.
- Perez, J., Marmol, P., & Picapiedra, P. (1968). Minería Prehistórica. *ABC*, 15 (2), 50 - 60.
- Ping, Y., & Yongheng, L. (2011). Neighborhood Rough Set and SVM based Hybrid Credit Scoring Classifier. *Expert Systems with Applications*, 38 (9), 11300-11304.
- R Core Team. (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Rodríguez, G. (2015). Modelo de equilibrio general aplicado para México y análisis de impuestos a la extracción de hidrocarburos. *Ensayos Revista de Economía*, 34 (1), 35-74.

- Tetreault, D. (2013). La megaminería en México. Reformas estructurales y resistencia. Letras Verdes, *Revista Latinoamericana de Estudios Socioambientales*, 14, 214-234.
- Wilde, T. (1997). CreditRisk+: A Credit Risk Management Framework. London: Credit Suisse First Boston International.
- Wise, R., y Pozo, R. (2002). Minería, Estado y gran capital en México. Ciudad de México: UNAM.
- Zhang, K., Nieto, A., & Kleit, A. N. (2015). The Real Option Value of Mining Operations using Mean-Reverting Commodity Prices. *Mineral Economics*, 28 (1), 11 - 22
- Zhou, L. (2013). Performance of Corporate Bankruptcy Prediction Models on Imbalanced Dataset: The effect of Sampling Methods. *Knowledge-Based Systems*, 41, 16 - 25.