

Reporte de Riesgo

Cívico - Equipo DATA

2022-07-06

Introducción

Este informe tiene dos objetivos puntuales:

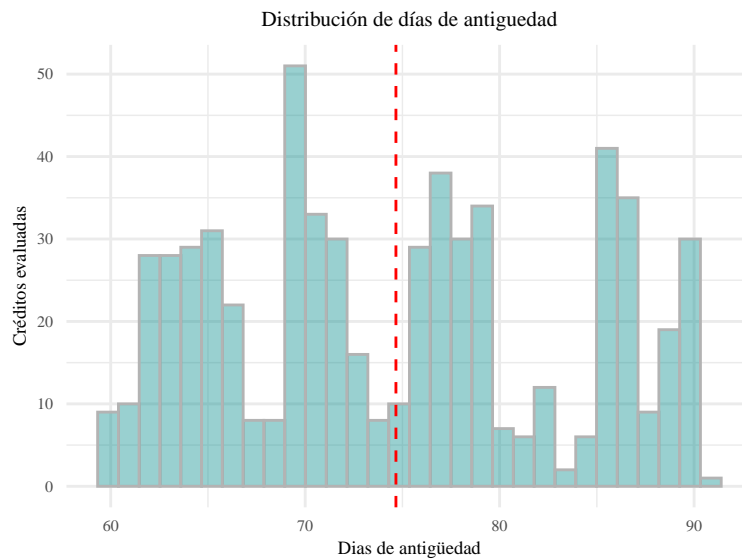
- Revisar cuál es el aporte de los diferentes conjuntos de variables al mejoramiento de la predicción del riesgo.
- Revisar cuales son las variables que pueden ser utilizadas como **marcadores** directos de riesgo

Muestra empleada

El análisis parte de los créditos que cumplan las siguientes condiciones:

- Más de 60 días en la cartera
- Que hallan llenado el formulario nuevo de solicitud del crédito (45 preguntas)
- Que tengan todos los campos de información completa y con respuestas coherentes

Con estas **condiciones muestrales**, el tamaño de la población evaluada fue de 620 créditos.



Dado el tamaño de la muestra más su maduración (antigüedad de los créditos en la cartera), es posible que aquellos efectos pequeños (pero existentes) no sean detectados. La meta para la transición hacia la nueva solicitud como base de datos es de 1,001 observaciones validas.

Metodología

Se clasificaron los variables en 10 conjuntos como se ve en la siguiente tabla:

Table 1: Conjuntos temáticos de variables

Vriable.	Grupo.	Variable	Grupo
Antigüedad	1. Control	Tipo de inventario	6. Gestion
Semana	1. Control	Tipo de contabilidad	6. Gestion
Género	2. Persona	Fía o no	6. Gestion
Edad	2. Persona	Tipo de persona	7. Formalidad
Estado civil	2. Persona	Registro en C.comercio (1 = Sí)	7. Formalidad
Nivel educativo	2. Persona	Cotiza SS.Salud (1 = Sí)	7. Formalidad
Tamaño del hogar	2. Persona	Cotiza SS.Pensión (1 = Sí)	7. Formalidad
Contribuidores en el hogar	2. Persona	Cotiza SS.Riesgo (1 = Sí)	7. Formalidad
Dependencia	2. Persona	Coiza SS.Caja Familiar (1 = Sí)	7. Formalidad
Ingresos Negocio	3. Ingresos	Trato de proveedores	8. Comunidad
Ingresos Hogar	3. Ingresos	# de clientes que conoce	8. Comunidad
% de ventas en Nequi	3. Ingresos	Comparado con año anterior	9. Autopercepcion
Tipo de negovcio	4. Negocio	Como le irá el otro año	9. Autopercepcion
Experiencia en el negocio	4. Negocio	Comparado con otros	9. Autopercepcion
Tipo de operación	4. Negocio	Dificultad de llegar al mismo estado	9. Autopercepcion
Cambio de actividad (1 = Sí)	4. Negocio	Calificación como comerciante	9. Autopercepcion
Número de empleados	4. Negocio	Prjde pagar tarde	9. Autopercepcion
Número de empleados familiares	4. Negocio	Destino del crédito	10. Carácter
Créditos con amigos (1 = Sí)	5. Bancario	Situación: Acueducto	10. Carácter
Créditos gota a gota (1 = Sí)	5. Bancario	Situación: Transmilenio	10. Carácter
Créditos con bancos (1 = Sí)	5. Bancario	Acciones para crecer	10. Carácter
Créditos "otros" (1 = Sí)	5. Bancario	Metas	10. Carácter
Créditos vigentes (1 = Sí)	5. Bancario	.	.
Reportado en centrales (1 = Sí)	5. Bancario	.	.

Se utilizaron modelos logit para evaluar el efecto de cada una de las variables, la modelación de realizo de manera aditiva como se muestra en la siguiente fórmula:

$$\log \left(\frac{Pr|PAR_n}{1 - Pr|PAR_n} \right) = \alpha + \beta(C_j)_i' + \theta(C_{k:j-1})_i' + \epsilon_i$$

De esta manera, la probabilidad de entrar en un PAR cualquiera (15, 30 o 60), está determinada por el conjunto de variables a evaluar, más los conjuntos de variables anteriores.

Modelo	Variables explicataorias
01. Básico	Control
02. Persona	Control, Persona
03. Ingresos	Control, Persona, Ingresos
04. Negocio	Control, Persona, Ingresos, Neogcio
(...)	(...)

Limitaciones

Las limitaciones de este enfoque son:

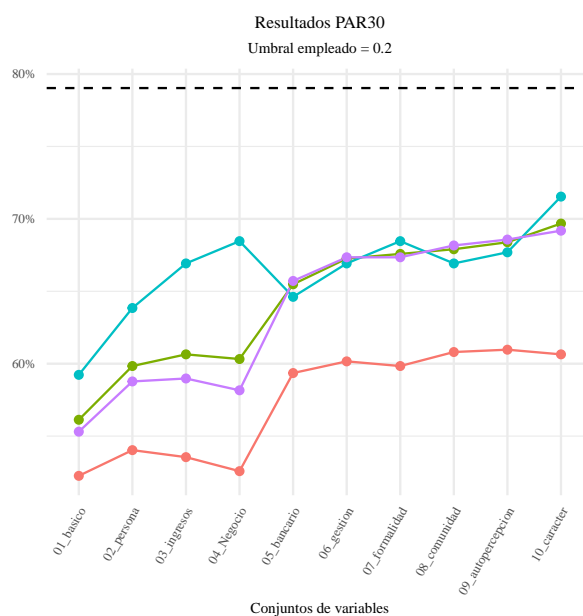
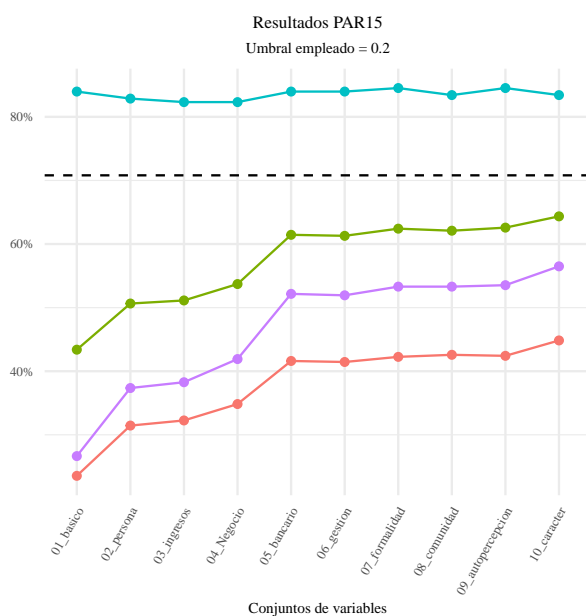
- El orden de los conjuntos de variables no es conmutativo respecto del desempeño de los modelos.
- Se asume un umbral constante para la evaluación de todo los PAR's
- Se emplea un modelo logit con las variables nuevas, pero el que está en uso es un RM con las variables viejas. Es un simulación no una evaluación.
- Por la maduración de la cartera, los resultados son más confiables para PAR's juveniles que para PAR's más maduros.

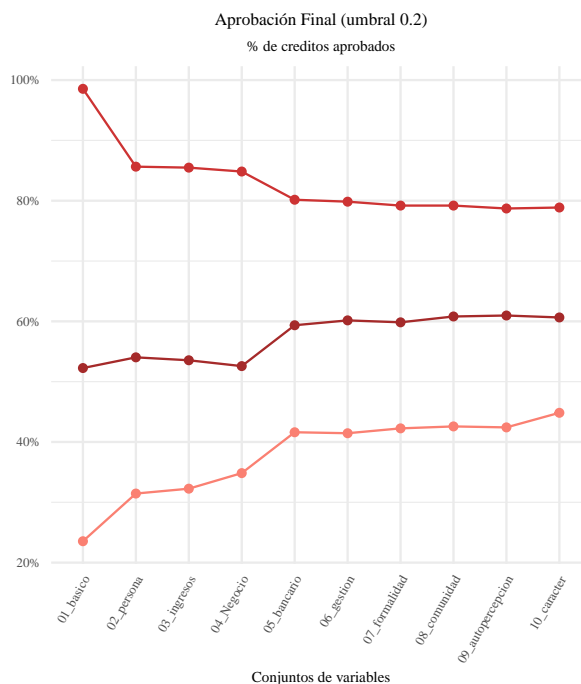
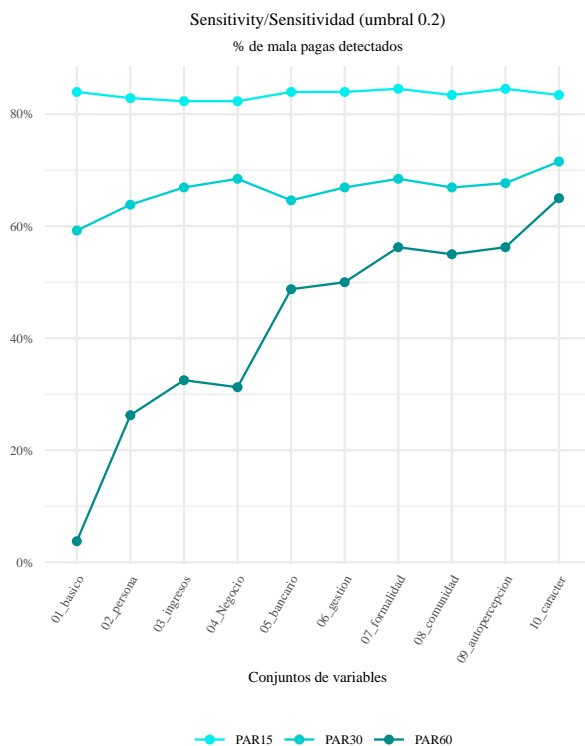
Resultados I: Aporte estadístico

Antes de analizar los resultados del aporte estadístico de las variables es necesario tener en cuenta unas definiciones estadísticas aplicadas a nuestro contexto.

- **Sensibilidad (Sensitivity)**: es la proporción de mala pagas detectados (si es igual a 100% quiere decir que detecte al 100% de los mala pagas).
- **Especificidad (Specificity)**: es la proporción de buena pagas detectados (si es igual a 100% quiere decir que detecte al 100% de los buena pagas).
- **Exactitud (Acuarcy)**: es la proporción de gente que se clasifico como buena paga o mala paga de manera correcta.
- **Aprobación final**: Dado un umbral y un modelo, que porcentaje de aprobación puede esperar.
- **Aprobación real (línea punteada negra)** : Si mi modelo fuera perfecto (Exactitud = 100%), cuanto debería aprobar.

`SUPER_UMBRAL = 0.2` # Defina un umbral de riesgo para la evaluación





Resultados II: Variables relevantes (Marcadores de Riesgo)

Las tablas de los resultados estadísticos pueden consultarse en la sección de anexos.

PAR 15 (marcadores de mora de 15 días)

A continuación, se detallan las conclusiones del análisis sobre la mora de 15 días:

- **Semana en la que se realiza la solicitud:** Los créditos solicitados en la 1ra semana del mes tiene una probabilidad más alta de entrar en mora de 15 días que los solicitados en la 3ra y 4ta semana. Dicho de manera más específica, los créditos solicitados en de la 3ra semana son tiene 31% menos probabilidad de entrar en PAR15 que los solicitados la 1ra semana, y los de la 4ta semana 38% menos de probabilidad comparados igualmente contra los de la 1ra. Este efecto es estadísticamente significativo y robusto.

Ahora bien, la semana de solicitud no es determinante en moras superiores a 15 días. Es decir, la semana no me permite diferenciar quien se va a ser un “super moroso”, pero si me permite discriminar entre quienes tendrán problemas con el pago de la primera cuota y quienes no. quizá esto pueda indicar un grupo de clientes que necesitan mayor acompañamiento en métodos de pago y planeación financiera.

- **Género:** Aunque no es un resultado totalmente robusto (el género es estadísticamente relevante en 9 de 10 modelos), pareciera que las mujeres tienen entre 65% y 70% más de probabilidades de entrar en una mora de 15 días. Es importante tener en cuenta la comisión de variable como la pobreza y el estrato en el ejercicio, y también otros fenómenos sociales como las brechas salariales y laborales de género. Es decir, en general el en la literatura y en los datos económicos nacionales, la mujer tiene mayores retos económicos por el hecho de ser mujer. Hay que tener precaución ética en este punto. En la muestra 47% son mujeres y 53% hombres.

- **Tamaño del hogar:** El tamaño del hogar es una variable no robusta (estadísticamente relevante en 6 de los 10 modelos), por cada miembro que tiene el hogar la probabilidad de entrar en mora de 15 días aumenta entre 129% y 146%. El tamaño promedio del hogar en la muestra es de 3.02 personas.
- **Miembros del hogar que aportan económicamente y la dependencia:** Aunque no son resultados robustos, parece existir una relación (no muy fuerte) entre dependencia (% de miembros del hogar que aportan) y el riesgo de mora de 15 días en el sentido esperado: más miembros que aportan en el hogar, menor riesgo moratorio. En promedio en la muestra, los hogares están compuestos por un 50,6% de miembros aportantes: es decir, en promedio en los hogares asociados a la muestra hay 1 dependiente por cada aportante.
- **Créditos vigentes:** Una de las variables más relevantes (tanto por su robustez como por su significancia estadística) es si la persona reporta tener créditos vigentes. Las personas que reportan tener créditos vigentes tienen una probabilidad entre 248% a 266% más alta de entrar en mora de 15 días. En la muestra el 27.4% tienen reportaron tener créditos vigentes, contra el 72% que reportaron no tener. Esta situación no discrimina con quien los tenga. Es decir, **NO es relevante si el crédito es con un banco, prestamista, amigo o familiar**, la existencia de obligaciones crediticias pendientes representa un factor de riesgo.
- **Si fía a sus vecinos:** Las personas que reportaron fiar a sus vecinos y llevar registro en un cuaderno tiene entre 34% a 43% menos probabilidad de entrar en mora de 15 días, con respecto a los que fian y reportan llevar cuentas en un “Excel”. El 18% de las personas reportaron fiar y llevar registro en un cuaderno, frente al 7.6% que reportaron llevar registro en un Excel o programa. Esta diferencia es robusta (estadísticamente significativa en los 10 modelos).
- **Destino del crédito (publicidad):** Las personas que reportaron la intención de destinar el crédito a publicidad tienen 231% más probabilidades de entrar en mora de 15 días respecto a las personas que reportaron tener interés de destinar el crédito a inventario. En la muestra 7.2% destinarían los fondos a publicidad, 64.3% a inventario, 17,8% a maquinaria, 8.5% a remodelaciones y el restante 2.1% a personal.

PAR 30 (marcadores de mora de 30 días)

Anexos

Table 3: Variables relevantes en PAR15 - Primera cuota

	Dependent variable:									
	PAR15									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Semana 2	0.28 (0.54)	0.21 (0.55)	0.20 (0.56)	0.17 (0.56)	0.28 (0.59)	0.20 (0.60)	0.15 (0.59)	0.17 (0.60)	0.13 (0.62)	0.32 (0.65)
Semana 3	-0.83* (0.43)	-0.90** (0.44)	-0.90** (0.44)	-0.98** (0.45)	-1.14** (0.47)	-1.22** (0.48)	-1.28*** (0.48)	-1.27*** (0.49)	-1.29** (0.50)	-1.17** (0.52)
Semana 4	-0.77** (0.31)	-0.82** (0.32)	-0.81** (0.32)	-0.85*** (0.32)	-0.92*** (0.34)	-0.97*** (0.34)	-0.98*** (0.34)	-0.96*** (0.34)	-0.97*** (0.35)	-0.95*** (0.36)
Genero (Hombre)		-0.37* (0.19)	-0.36* (0.19)	-0.36* (0.20)	-0.35* (0.21)	-0.39* (0.21)	-0.37* (0.21)	-0.35* (0.21)	-0.35 (0.22)	-0.42* (0.23)
Tamaño del hogar		0.38*** (0.14)	0.38*** (0.15)	0.35** (0.15)	0.24 (0.15)	0.25 (0.16)	0.26* (0.16)	0.26 (0.16)	0.30* (0.16)	0.31* (0.17)
Aportantes en el hogar		-0.46* (0.24)	-0.46* (0.24)	-0.44* (0.24)	-0.26 (0.26)	-0.26 (0.26)	-0.30 (0.26)	-0.26 (0.27)	-0.31 (0.27)	-0.35 (0.29)
Dependencia		1.38** (0.63)	1.41** (0.63)	1.33** (0.64)	1.00 (0.66)	1.00 (0.67)	1.08 (0.68)	1.02 (0.68)	1.09 (0.70)	1.08 (0.73)
Créditos vigentes (Sí)					0.98*** (0.25)	0.98*** (0.26)	0.95*** (0.26)	0.96*** (0.26)	0.91*** (0.27)	0.94*** (0.29)
Fía. memoria						-0.88 (1.05)	-0.70 (1.05)	-0.72 (1.05)	-0.80 (1.06)	-1.20 (1.09)
Fía.cuaderno						-0.84* (0.45)	-0.85* (0.46)	-0.84* (0.46)	-0.88* (0.47)	-1.05** (0.50)
Fía. (NO)						-0.53 (0.39)	-0.50 (0.40)	-0.53 (0.40)	-0.60 (0.42)	-0.74* (0.43)
Dificultad re-empezar									-0.004 (0.003)	-0.01* (0.003)
Destino: Maquinaria										-0.20 (0.32)
Destino: Personal										-0.92 (0.78)
Destino: Publicidad										0.84** (0.42)
Destino: Remodelaciones										-0.15 (0.40)
Observations	620	620	620	620	620	620	620	620	620	620
Log Likelihood	-352.40	-343.50	-342.74	-339.46	-324.18	-320.63	-317.64	-316.20	-309.47	-299.21
Akaike Inf. Crit.	714.80	723.00	727.47	740.93	722.37	733.26	739.27	746.39	750.95	768.43

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4: Variables relevantes en PAR30 - Primera y segunda cuota

	<i>Dependent variable:</i>									
	PAR30									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Semana 2	0.39 (0.58)	0.36 (0.59)	0.32 (0.59)	0.28 (0.59)	0.32 (0.63)	0.34 (0.64)	0.28 (0.63)	0.33 (0.64)	0.30 (0.66)	0.41 (0.70)
Semana 3	-0.39 (0.47)	-0.39 (0.47)	-0.42 (0.48)	-0.47 (0.48)	-0.67 (0.51)	-0.59 (0.52)	-0.65 (0.52)	-0.62 (0.52)	-0.63 (0.54)	-0.59 (0.57)
Semana 4	-0.30 (0.34)	-0.30 (0.34)	-0.30 (0.34)	-0.32 (0.34)	-0.41 (0.36)	-0.37 (0.37)	-0.41 (0.37)	-0.38 (0.37)	-0.39 (0.38)	-0.37 (0.40)
Genero (Hombre)		-0.32 (0.21)	-0.31 (0.21)	-0.33 (0.22)	-0.33 (0.23)	-0.36 (0.23)	-0.32 (0.23)	-0.31 (0.24)	-0.31 (0.24)	-0.42 (0.26)
EC: Soltero/a		0.41 (0.34)	0.43 (0.34)	0.46 (0.34)	0.39 (0.35)	0.38 (0.36)	0.36 (0.36)	0.31 (0.37)	0.27 (0.37)	0.26 (0.40)
EC: Unión libre		0.60* (0.32)	0.61* (0.32)	0.63* (0.33)	0.48 (0.34)	0.49 (0.35)	0.44 (0.35)	0.40 (0.35)	0.34 (0.36)	0.27 (0.38)
EC: Viudo/a		0.82 (0.77)	0.81 (0.77)	0.89 (0.78)	0.76 (0.80)	0.76 (0.81)	0.73 (0.82)	0.72 (0.83)	0.68 (0.84)	0.49 (0.93)
Tamaño del hogar		0.35** (0.16)	0.35** (0.16)	0.30* (0.16)	0.19 (0.17)	0.19 (0.17)	0.20 (0.17)	0.21 (0.17)	0.24 (0.18)	0.26 (0.19)
Créditos vigentes (Sí)					1.26*** (0.26)	1.22*** (0.27)	1.17*** (0.27)	1.18*** (0.27)	1.13*** (0.28)	1.16*** (0.30)
Cotiza SS.Riesgo (Sí)							-0.73* (0.41)	-0.73* (0.42)	-0.71* (0.42)	-0.71 (0.45)
Destino: Maquinaria										0.03 (0.35)
Destino: Personal										-1.32 (0.93)
Destino: Publicidad										0.87** (0.43)
Destino: Remodelaciones										-0.20 (0.45)
Observations	620	620	620	620	620	620	620	620	620	620
Log Likelihood	-312.36	-303.51	-302.83	-300.85	-284.15	-281.63	-277.58	-276.66	-272.82	-256.61
Akaike Inf. Crit.	634.72	643.02	647.66	663.70	642.30	655.27	659.16	667.31	677.65	683.21

Note:

* p<0.1; ** p<0.05; *** p<0.01

Table 5: Variables relevantes en PAR60 - Segunda y tercera cuota*(!!)

	<i>Dependent variable:</i>									
	PAR60									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Genero (Hombre)		-0.48* (0.25)	-0.56** (0.26)	-0.59** (0.27)	-0.66** (0.28)	-0.73** (0.29)	-0.73** (0.30)	-0.74** (0.30)	-0.72** (0.31)	-0.80** (0.33)
TL: Rentado				-0.02 (0.45)	0.02 (0.46)	0.03 (0.48)	0.13 (0.49)	0.16 (0.50)	0.13 (0.50)	0.14 (0.53)
TL: Virtual/Digital				0.78 (0.57)	0.74 (0.59)	0.99 (0.61)	1.22* (0.62)	1.35** (0.64)	1.23* (0.65)	1.46** (0.70)
TP: Vivenda propia				0.47 (0.55)	0.35 (0.56)	0.44 (0.59)	0.64 (0.61)	0.77 (0.63)	0.69 (0.63)	0.68 (0.67)
PT: Vivienda rentada				-0.23 (0.57)	-0.44 (0.59)	-0.49 (0.62)	-0.15 (0.64)	-0.07 (0.65)	-0.11 (0.65)	-0.20 (0.70)
Cambio de actividad (Sí)				0.64 (0.51)	0.78 (0.52)	1.01* (0.54)	1.07* (0.55)	1.10** (0.56)	1.13** (0.56)	1.11* (0.60)
Créditos vigentes (Sí)					1.26*** (0.32)	1.21*** (0.33)	1.12*** (0.34)	1.13*** (0.34)	1.06*** (0.35)	1.03*** (0.38)
Iventario: Aplicación						-1.98** (0.86)	-2.02** (0.89)	-2.09** (0.89)	-2.09** (0.91)	-1.94** (0.94)
Inventario: Cuaderno						0.03 (0.71)	0.13 (0.72)	0.08 (0.73)	0.02 (0.75)	0.03 (0.80)
Inventario: (No)						-1.71 (1.21)	-1.65 (1.22)	-1.79 (1.23)	-2.03 (1.26)	-2.09 (1.33)
Contabilidad: Aplicación						1.80** (0.77)	1.80** (0.81)	1.88** (0.82)	1.87** (0.83)	1.86** (0.87)
Contabilidad: Cuaderno						0.91 (0.73)	0.76 (0.74)	0.85 (0.75)	0.90 (0.77)	1.05 (0.83)
Contabilidad: (No)						2.47** (1.04)	2.33** (1.05)	2.42** (1.07)	2.43** (1.08)	2.73** (1.17)
Cotiza SS.Riesgo (Sí)							-0.81 (0.50)	-0.85* (0.51)	-0.81 (0.51)	-0.74 (0.55)
Destino: Maquinaria										0.18 (0.43)
Destino: Personal										-1.55 (1.31)
Destino: Publicidad										0.89* (0.51)
Destino: Remodelaciones										0.23 (0.55)
Observations	620	620	620	620	620	620	620	620	620	620
Log Likelihood	-234.22	-227.08	-225.82	-221.10	-209.16	-200.77	-197.17	-195.31	-193.51	-183.65
Akaike Inf. Crit.	478.44	490.17	493.63	504.20	492.31	493.53	498.35	504.61	519.01	537.31

Note:

*p<0.1; **p<0.05; ***p<0.01