



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**NEWS IN DYNAMIC MARKETS**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

MARICEL ESTELA VARGAS VERGARA

PROFESOR GUÍA:  
ALEJANDRO BERNALES SILVA

MIEMBROS DE LA COMISIÓN:  
MARCELA VALENZUELA BRAVO  
PATRICIO VALENZUELA AROS

SANTIAGO DE CHILE  
2017

RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL  
Y AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA  
POR: MARICEL ESTELA VARGAS VERGARA  
FECHA: 2017  
PROF. GUÍA: SR. ALEJANDRO BERNALES SILVA

## NEWS IN DYNAMIC MARKETS

Una de las interrogantes más interesantes asociadas a la economía tiene relación con el saber como se desencadenan las crisis, burbujas y movimientos bursátiles. De la misma forma, existe una gran pregunta sin responder asociada a la influencia de los medios de comunicación en las decisiones de los inversionistas. Es por lo anterior, que durante el desarrollo de esta investigación se generó una base de datos histórica de las noticias publicadas diariamente por el New York Times desde su puesta en circulación en el año 1851 hasta el 2014, junto con esto se caracterizó la relación existente entre el contenido de las noticias publicadas por dicho diario y la probabilidad de currency crisis, además de la relación entre la cantidad de noticias publicadas y los retornos del índice Dow Jones.

De esta forma para el primer caso fue posible ver que existe una relación positiva entre la cantidad de palabras positivas sobre el total de palabras publicadas por dicho diario y la probabilidad de currency crisis en Estados Unidos. Luego, si este mismo análisis se realiza para el resto del mundo, es posible observar que el efecto que tienen las noticias sobre la economía local difiere del efecto que podrían provocar sobre el resto del mundo, lo anterior, está estrechamente relacionado con el rol informativo que tienen las noticias de Estados Unidos en el mundo. Por otro lado para el caso de los retornos, estos pueden ser predichos por la cantidad de noticias publicadas, pero solo durante las recesiones y dicho efecto era más importante en la antigüedad.

Por otro lado se estudiaron los mecanismos que podrían relacionar la estabilidad con posibles episodios de crisis, en particular caracterizando la existencia de alguna relación entre la tasa de interés como un proxy del coeficiente de aversión al riesgo, el ratio inversión/capital y la current account como una medida de la salud de la economía, encontrando que un aumento en la variable asociada a la estabilidad llevaba a un aumento tanto en la current account como en el ratio inversión/capital, mientras que en el caso de la tasa de interés pasaba lo contrario.

Tomando en consideración los resultados es posible darse cuenta que tal cual lo plantea Minsky (1992) en su hipótesis de inestabilidad financiera, periodos de estabilidad podrían terminar desencadenando un episodio inestabilidad. Todo esto relacionado tanto con el exceso de información como con los sesgos de percepción asociados a los inversionistas, los cuales impiden que ellos puedan ver con claridad lo que está sucediendo, llevándolos a tomar decisiones de manera apresurada o considerando información que es irrelevante en la toma de decisiones y omitiendo aquella información que sí lo es.

*“El Mundo está en las manos de aquellos que tienen  
el coraje de soñar y correr el riesgo de vivir sus sueños”  
Paulo Coelho*

# Agradecimientos

Gracias, de todo corazón a mi profesor guía Alejandro Bernal y a la profesora Marcela Valenzuela, por su confianza, apoyo y dedicación durante todo este proceso. Además a Giorgio quien fue fundamental en el periodo más difícil de este trabajo. Ha sido un privilegio contar con su ayuda.

A mi familia, a mis padres por su apoyo, cariño y comprensión durante estos 8 años de universidad, aun cuando estábamos físicamente lejos, siempre los sentí conmigo. Este logro es en gran medida gracias a ustedes. A mis hermanos Ricardo y Marcelo, por la motivación y apoyo en todo momento. Un agradecimiento especial a Dominique, Marcelo y Franco, quienes durante los primeros años fueron mi principal apoyo en Santiago, sin ustedes todo hubiese sido mucho más difícil. A Teresa, mi hermana, amiga y compañera de viajes, gracias por la compañía, las risas y el apoyo. Gracias también a Gustavo, por su inmenso cariño, contención y paciencia durante este proceso.

Gracias a todos aquellos que conocí tanto en pregrado como en postgrado, por los momentos compartidos y por hacer de esta etapa universitaria una experiencia única.

Agradecer también el apoyo financiero para la realización de este trabajo proveniente del proyecto Fondecyt # 11140628 y al Instituto Milenio para la Investigación en Imperfecciones de Mercado y Políticas Públicas ICM IS130002.

Finalmente agradezco a todas esas personas que de una u otra manera fueron parte de este proceso, por la palabra de aliento, por las conversaciones de pasillo o simplemente por la buena energía. Con este trabajo termina una gran etapa de mi vida, que no hubiese sido posible sin el apoyo de todos ustedes.

# Tabla de Contenido

<b>Introducción</b>	<b>1</b>
<b>1. Datos</b>	<b>6</b>
1.1. Fuentes de Información . . . . .	6
1.2. Descarga de Datos . . . . .	7
1.3. Descarga Masiva de Información Diaria . . . . .	9
1.4. Base de Datos Diaria . . . . .	10
1.5. Limpieza de la Base de Datos Diaria . . . . .	10
1.6. Datos Adicionales . . . . .	11
<b>2. Relación Medios y Crisis</b>	<b>13</b>
2.1. Análisis Descriptivo . . . . .	13
2.2. Desarrollo de Hipótesis . . . . .	16
2.2.1. Relación Medios y Crisis . . . . .	16
2.2.2. Mecanismos de explicación . . . . .	17
2.3. Resultados . . . . .	19
2.3.1. Relación Medios y Crisis . . . . .	19
2.3.2. Mecanismos de explicación . . . . .	23
<b>3. Relación Medios, Retornos y Ciclos Económicos</b>	<b>27</b>
3.1. Análisis Descriptivo . . . . .	27
3.2. Desarrollo de Hipótesis . . . . .	31
3.3. Resultados . . . . .	32
3.3.1. Retornos Semanales . . . . .	32
3.3.2. Cantidad de Noticias, Retornos y Ciclos Económicos . . . . .	33
3.3.3. Contenido de los Medios, Retornos y Ciclos Económicos . . . . .	43
<b>4. Conclusiones</b>	<b>47</b>
4.1. Relación Medios y Crisis . . . . .	47
4.2. Relación Medios, Ciclos Económicos y Retornos . . . . .	48
<b>Bibliografía</b>	<b>49</b>
<b>A. Estadística de Noticias Diarias</b>	<b>54</b>
<b>B. Relación Medios y Currency Crisis</b>	<b>55</b>

<b>C. Modelo de Lucas</b>	<b>57</b>
<b>D. Behavioral finance</b>	<b>59</b>
D.1. Formación de creencias . . . . .	59
D.2. Procesamiento de la información . . . . .	60
D.3. Sesgos emocionales . . . . .	61
<b>E. Relación Medios, Retornos y Ciclos Económicos</b>	<b>64</b>
E.1. Cantidad de Noticias, Retornos y Ciclos Económicos . . . . .	64
E.2. Contenido de los Medios, Retornos y Ciclos Económicos . . . . .	66
<b>F. Datos de Panel</b>	<b>67</b>

# Índice de Tablas

2.1.	Estadística descriptiva anual currency crisis . . . . .	15
2.2.	Logit relación ratio $pos/(neg + pos)$ y currency crisis . . . . .	21
2.3.	Logit relación ratio $pos/(neg + pos)$ y currency crisis sin considerar guerras . . . . .	22
2.4.	Regresión ratio $pos/(pos + neg)$ - Current account . . . . .	24
2.5.	Regresión ratio $pos/(pos + neg)$ - ratio Inversión/Capital . . . . .	25
2.6.	Regresión ratio $pos/(pos + neg)$ - tasa de interés . . . . .	26
3.1.	Estadística sobre numero de noticias que contienen alguna de las palabras de interés publicadas semanalmente . . . . .	28
3.2.	Estadística descriptiva tono semanal . . . . .	30
3.3.	Estadística descriptiva Retornos . . . . .	33
3.4.	Regresión cantidad de noticias y retornos semanales . . . . .	34
3.5.	Regresión cantidad de noticias, ciclos económicos y retornos . . . . .	35
3.6.	Regresión medios y retornos semanales . . . . .	43
3.7.	Regresión medios y retornos semanales sin considerar guerras ni crisis 2008 . . . . .	44
3.8.	Regresión Medios, ciclos económicos y retornos . . . . .	45
1.	Estadística descriptiva anual expansiones y recesiones . . . . .	53
A.1.	Estadística sobre numero de noticias que contienen alguna de las palabras de interés publicadas diariamente . . . . .	54
B.1.	Relación simple Medios y Crisis considerando hasta 10 lag . . . . .	55
B.2.	Relación Medios y Crisis considerando hasta 10 lag . . . . .	56
E.1.	Regresión cantidad de noticias y retornos semanales sin considerar periodos de guerra ni la crisis del año 2008 . . . . .	64
E.2.	Regresión cantidad de noticias, ciclos económicos y retornos sin considerar guerras ni la crisis del año 2008 . . . . .	65
E.3.	Regresión Medios, ciclos económicos y retornos semanales sin considerar guerras ni crisis 2008 . . . . .	66
F.1.	Logit Relacion Ratio $pos/(pos + neg)$ y Currency Crisis . . . . .	68

# Índice de Ilustraciones

1.1. Chronicle . . . . .	7
2.1. Cantidad de Noticias publicadas anualmente . . . . .	14
2.2. Comportamiento de ratio $pos/(pos + neg)$ con agregación anual en el tiempo . . . . .	14
3.1. Cantidad de noticias que contienen alguna de las palabras de interés publicadas semanalmente . . . . .	28
3.2. Comportamiento del ratio $pos/(pos + neg)$ con agregación semanal en el tiempo . . . . .	29
3.3. Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en todo el Intervalo de Tiempo . . . . .	37
3.4. Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en todo el Intervalo de Tiempo . . . . .	38
3.5. Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en Periodo de Recesión . . . . .	39
3.6. Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en Periodo de Recesión . . . . .	40
3.7. Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en Periodo de Expansión . . . . .	41
3.8. Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en Periodo de Expansión . . . . .	42



# Introducción

Es conocido que la Economía y las Finanzas, como resultado de las nuevas tecnologías, cada día se vuelven más intangible, quedando como lo único inmutable el rol de la información, en este contexto una pregunta fundamental aun no resuelta, es la influencia generada por los medios de comunicación en las personas, y en particular, la influencia de las noticias en la toma de decisiones de los agentes y como esto se ve reflejado en el movimiento de los mercados y/o en la probabilidad de alguna crisis.

Los diarios, las radios, el cine, las revistas y en general la mayoría de los medios de comunicación han experimentado un gran crecimiento en las ultimas décadas, una de las razones que se cree que motiva este crecimiento es la influencia que ellos podrían llegar a tener sobre los lectores o espectadores. Si se ahonda un poco en este punto, aquellos países que poseen un régimen autoritario, tienen como primer objetivo controlar lo que se publica o más aun controlar directamente los medios de comunicación. De manera adicional, el crecimiento de los medios de comunicación ha hecho posible el acortamiento de las distancias, manteniendo vinculado a todo el mundo en cuestión de minutos.

Tal cual lo plantean Perse and Lambe (2016) en su libro, los medios de comunicación y sobre todo aquellos de carácter masivo contribuyen en gran medida a fijar las maneras en que piensa la sociedad; en algunos casos determinan ideas, hábitos y costumbres. Burton (2010) en su libro “Media and society: Critical perspectives” plantea que los medios de comunicación se caracterizan por introducir en las diversas capas sociales pautas de comportamiento y de consumo, además de que cada medio de comunicación ejerce a sus espectadores diferentes tipos de influencias que pueden ser desde la función persuasiva, la enajenante y la manipuladora hasta temas relacionados a la política y la publicidad. Más ligado con las finanzas, Kurtz (2000) postula que la información financiera ya dejó de ser asunto de especialistas para convertirse en un tipo de información seguida por millones de personas, él define al periodista financiero como un jugador más del mercado y comenta que las opiniones y noticias sobre compañías, ciertas o no, mejor o peor fundadas pueden generar grandes movimientos bursátiles.

Por otro lado Daniel et al. (1998) y Ritter (2003) estudiaron el efecto del behavioral finance en la toma de decisiones de los inversionistas y de como las decisiones tomadas bajo condiciones de euforia o pánico están directamente relacionadas con las emociones que se experimentan en esos momentos, descartando de esta forma supuestos tradicionales ligados a la economía tales como la maximización de la utilidad esperada con inversionistas racionales en mercados eficientes. De manera adicional la psicología del comportamiento financiero plantea que muchos inversores basan el precio que están dispuestos a pagar por un activo en la percepción que tienen de este,

centrándose en noticias y estimaciones a corto plazo. Dichas decisiones están basadas más que nada en percepciones de riesgo y estabilidad, que a su vez, están directamente relacionadas con tendencias mucho más amplias provocadas por sentimientos ligados al miedo o la ambición.

La hipótesis de mercado eficiente propone que en un momento dado en un mercado, los precios reflejan toda la información disponible. Muchos de los modelos tradicionales se basan en la creencia de que los participantes del mercado siempre actúan de una manera racional, o que maximizan su utilidad, limitando severamente la capacidad de estos modelos de hacer predicciones precisas o detalladas. El behavioral finance intenta hacerse cargo de este vacío mediante la combinación de conocimientos científicos en el razonamiento cognitivo con la teoría económica y financiera convencional. Más específicamente, estudia diferentes sesgos psicológicos a los cuales se ven expuestos los seres humanos, dichos sesgos, o atajos mentales, conducen a decisiones de inversión irracionales. Lo anterior, podría dar una explicación más clara de por qué se producen las burbujas, fluctuaciones abruptas del mercado y las crisis.

Si se habla de sesgos de percepción, en términos conductuales esto se refiere a aquellos fenómenos de la psicología y conducta que llevan a crear percepciones distorsionadas o equivocadas del entorno, es decir, que no corresponden de una manera relevante con la realidad que se enfrenta. En este ámbito Hens and Meier (2002), aportan información sobre las influencias emocionales y psicológicas que pueden tener impacto en las decisiones financieras, y cómo esto puede resultar en un comportamiento irracional.

Las observaciones claves documentadas en la literatura relacionada con behavioral finance como por ejemplo en Barberis and Thaler (2003) y Hur et al. (2010) incluyen la falta de simetría entre decisiones de adquirir o mantener recursos y la fuerte aversión a las pérdidas o arrepentimiento unido a cualquier decisión donde algunos recursos emocionalmente valiosos, como por ejemplo la casa, pueden ser totalmente perdidos. La aversión a las pérdidas parece manifestarse en la conducta de los inversores como una falta de inclinación a vender acciones.

En relación a lo mismo, es posible darse cuenta que algunos de los sesgos más importantes son: Exceso de confianza, ya que la mayoría de las personas sobrestima las habilidades propias y tiende a pensar que sus conocimientos se encuentra por encima de la media, contabilidad mental el cual consiste en hacer distinciones en la cabeza que no existen financieramente, los agentes tienden a vender acciones de su cartera demasiado pronto cuando obtienen una ganancia y demasiado tarde cuando sufren una pérdida, pasan por alto la interacción entre los activos.

Si se habla de volúmenes de información, Toffler (1990) muestra la primera evidencia que relaciona la cantidad de información proporcionada por la prensa y la toma de decisiones de los agentes bajo un concepto que el denomina: "information overload". En la misma línea de lo anterior, Cornella (1999) plantea que el exceso de información provoca ansiedad, angustia y confusión, ya que una vez superada la capacidad de procesamiento de información, toda la información extra a la cual se tenga acceso se convierte en ruido, lo que termina disminuyendo la calidad de la decisión.

La preocupación por la relación entre los mercados financieros y el riesgo económico tienen una larga historia en la literatura, Keynes (1937), Kemp (1960) y Minsky (1992), argumentan que los agentes económicos cambian su aversión al riesgo cuando cambia el riesgo del mercado fi-

nanciero. De forma paralela y relacionado con el comportamiento de los inversionistas, Minsky (1992) postula que la estabilidad podría llegar a ser desestabilizadora, es decir, que periodos de prosperidad o estabilidad prolongados podrían desencadenar la inestabilidad. Lo anterior, tal cual lo explica Bhattacharya et al. (2015) es sustentado en el hecho de que durante periodos de “bonanza” los inversores aumentan su apalancamiento y cambian sus carteras hacia proyectos que antes consideraban demasiado riesgosos, todo esto debido a que actualizan sus expectativas durante los buenos tiempos y se vuelven más optimistas sobre las perspectivas futuras de la economía. En esta misma materia Danielsson et al. (2016) descubren que las volatilidades inusualmente altas e inusualmente bajas permiten predecir una crisis, la baja volatilidad desencadena una acumulación de crédito, esto es un síntoma de que los agentes están dispuestos a tomar mayor riesgo en periodos de bajo riesgo financiero representado por la baja volatilidad. Lo anterior es consistente con lo planteado por Minsky (1992) y aumenta la probabilidad de crisis bancaria.

De la misma forma, la literatura psicológica argumenta que las emociones afectan la toma de decisiones y el como los inversionistas procesan la información. Por ejemplo, Tiedens and Linton (2001) muestran que la dependencia del procesamiento heurístico versus sistemático varía con las emociones. La literatura también muestra que la ansiedad, la esperanza y la tristeza están asociadas con un mayor sentido de la incertidumbre, Smith and Ellsworth (1985) y Ortony et al. (1990). De la misma forma se puede argumentar razonablemente que en períodos de expansión los inversionistas se sienten felices y optimistas, mientras que durante las recesiones se sienten temerosos y ansiosos, las pérdidas de empleo y la incertidumbre sobre el futuro que los inversores experimentan durante las recesiones ponen a la población en general en estados de ánimo negativos. Esta evidencia sugiere que los inversionistas usarán diferentes reglas de toma de decisiones en las recesiones que en las expansiones.

En cuanto a las investigaciones que ligan finanzas con noticias, Garcia (2013) analiza el efecto de los sentimientos en el precio de los activos entre los años 1905 y 1958, construyendo una variable que representa los sentimientos por medio del conteo de palabras positivas y negativas definidas por el diccionario de Loughran and McDonald (2011), en dos columnas financieras del New York Times, sus principales resultados tienen relación con que el tono de la noticia permitiría predecir de algún modo los rendimientos de algunas acciones, pero sólo durante periodos de recesiones.

De manera adicional el año 2014 se desarrollaron variadas publicaciones al respecto, una de ellas es la de McCarthy and Dolfsma (2014), quienes reúnen literatura relacionada con el tema a fin de poder analizar si el tono descriptivo con el cual se enfoca la cobertura de la noticia, tiene un impacto no neutral en la economía. Sus principales descubrimientos son que, eligiendo el tono a utilizar para describir un evento, los medios de comunicación pueden influir en las percepciones del público, de manera particular cuando se trata de riesgos, ya que la mayoría de los medios de comunicación hace hincapié en las probables consecuencias negativas del evento más que en las positivas.

De la misma forma, Garcia (2014) estudia el contenido de las noticias financieras en función de los retornos pasados. Para ello realiza un conteo de palabras negativas y positivas según el diccionario de Loughran and McDonald (2011) al igual que en Garcia (2013), estima un modelo de series de tiempo y concluye que rendimientos negativos del mercado tienden a dar noticias

negativas, mientras que los retornos positivos a penas lo hacen, además de que existe una fuerte no linealidad entre los rezagos de los rendimientos de mercado y el contenido de las noticias financiera.

Por otro lado, Shiller (2015) sostiene que los medios de comunicación desempeñan un papel importante en preparar el escenario para los movimientos del mercado e incluso provocarlos. Su conjetura es que los inversores siguen las palabras impresas a pesar de que gran parte de ellas sean solo hipótesis, lo que sugiere que el sentimiento del mercado está impulsado por el contenido de las noticias.

Tomando en consideración todo lo anteriormente mencionado y analizando la escasa literatura con que se cuenta hoy en día al respecto, así como también el poder definir el verdadero rol que están teniendo los medios de comunicación en la sociedad parece interesante estudiar la existencia de alguna relación entre el ratio  $pos/(pos + neg)$  planteado en este trabajo, donde  $pos$  es el número de palabras positivas pertenecientes al diccionario Loughran and McDonald (2011) que aparecen en el periodo de tiempo en el cual se trabajara, contando solo una aparición de cada palabra por noticia y  $neg$  se define de la misma forma pero considerando las palabras negativas proporcionadas por Loughran and McDonald (2011) y la probabilidad de currency crises o bien su efecto sobre los retornos.

Para el desarrollo de este estudio, se genero un proxy de la estabilidad del mercado utilizando todas aquellas noticias del New York Times que contenían al menos una de las 2709 palabras propuestas por Loughran and McDonald (2011), divididas en 354 positivas y 2355 negativas, entre los años 1851 y 2014. El principal interés del trabajo fue el hecho de poder explicar la currency crisis, para ello se utilizo la base de datos de Reinhart and Rogoff (2009), quienes entregan una variable dummy para dicha crisis que toma valor uno en caso de que el año presento la crisis y cero en caso contrario y un ratio anual  $pos/(pos + neg)$ , se utiliza agregación anual debido a que Minsky plantea que se necesitan periodos prolongados para ver efectos. Otro de los puntos de interés del trabajo es analizar la existencia de alguna relación entre la cantidad de noticias publicadas por el diario y los retornos asociados al Dow Jones así como también la existencia de alguna relación entre el ratio  $pos/(pos + neg)$  y los retornos, para lo anterior se utilizaron los retornos del Dow Jones proporcionados por su sitio web, la cantidad de noticias que publica el New York Times de manera semanal y el ratio mencionado anteriormente, en este caso se utiliza agregación semanal, porque se esta relacionando el posible efecto con sesgos de percepción, los cuales tienen un efecto en el mediano plazo.

Los resultados obtenidos durante el desarrollo de esta investigación aportan con nuevos conocimientos y al mismo tiempo son coherentes con la teoría que plantea la literatura hoy en día. De manera adicional es importante destacar que no se tiene conocimiento de algún documento que logre explicar la mayoría de los fenómenos que se descubrieron durante el desarrollo de esta investigación.

En primer lugar, se pudo demostrar que el contenido publicado por la prensa y de manera particular para este caso del New York Times, ayuda a predecir la currency crisis. De la misma forma, existe evidencia de que el ratio  $pos/(pos + neg)$ , asociado durante el desarrollo de este trabajo a estabilidad también es capaz de predecir la current account, la tasa de interés y el ratio inversión/capital. Por lo anterior, fue posible demostrar que un aumento en dicho ratio

trae consigo una disminución en la tasa de interés y si esta se utiliza esta como un proxy de la aversión al riesgo Lucas Jr (1978), es posible saber que los agentes al percibir una mayor estabilidad en el país, están dispuestos a tomar una mayor cantidad de riesgo.

En segundo lugar, se encontró evidencia asociada a la predictibilidad de los retornos, por medio del volumen de noticias publicadas de manera semanal por el New York Times. Lo anterior, confirma algunas cosas ya planteadas en la literatura relacionadas con el exceso de información al cual son expuestos los agentes durante el proceso de toma de decisiones y cuales son las consecuencias que esto trae consigo, debido a la complejidad que implica el procesamiento de altos volúmenes de información. Finalmente, se encontró evidencia que permite predecir el comportamiento de los retornos por medio de la variable asociada a la estabilidad, lo que es consistente con los efectos que podrían presentar los sesgos de percepción en los inversionistas, tal cual decía Benjamin Graham: “el peor enemigo del inversor es probablemente él mismo”.

En cuanto a la distribución del trabajo, esta organizado de la siguiente manera: En el capítulo uno se realizara una pequeña explicación de los datos, abordando las fuentes de información y la descarga de estos. En el capítulo dos se abordara la hipótesis 1: Relación Medios y Crisis y sus resultados, en este capítulo se hará un análisis descriptivo de los datos de interés, se desarrollara la hipótesis y posteriormente se mostraran los resultados asociados a esta. En el capítulo tres se analizan la hipótesis 2: Relación Medios, Ciclos Económicos y Retornos y sus resultados. El capítulo cinco abordan las conclusiones y algunos trabajos futuros que podrían ser interesantes de investigar.

# Capítulo 1

## Datos

### 1.1. Fuentes de Información

Tomando como referencia el diccionario de sentimientos de Loughran and McDonald (2011), ya que este fue creado con fines ligados a las finanzas y es el diccionario más utilizado por la literatura ligada a los temas que se pretenden abordar durante el desarrollo del trabajo, en primer lugar se extraerá el conteo anual de noticias que contienen alguna de las palabras de interés, con el fin de poder tener la base de datos anual de la cantidad de noticias que eran publicadas en el diario durante un año considerando solo una aparición de cada palabra por noticia. De manera adicional se realizara la extracción de los títulos y fechas de cada una de las noticias publicadas de manera diaria asociadas a las palabras tanto positivas como negativas de dicho diccionario, con el fin de obtener una base de datos que contenga todas las noticias de interés publicadas de manera diaria. Ambas bases de datos contendrán información en el periodo de tiempo que comprende entre los años 1851 y 2014.

Se utiliza como fuente de información el New York Time, dado que es el segundo periódico de mayor circulación en Estados Unidos, con casi dos millones de ejemplares emitidos diariamente, de manera adicional mensualmente tiene alrededor de 20 millones de visitantes únicos en su sitio web, y posee una infraestructura tecnológica acorde con el desafío de extracción de la información. El New York Times posee variadas plataformas que permiten acceder a su información, una de ellas es Chronicle <sup>1</sup>, la cual al proporcionarle una o más palabras entrega un gráfico con la cantidad de artículos por año que la contienen, desde su puesta en circulación el año 1851 hasta la actualidad. Otra de las plataformas que posee dicho periódico, es NYT Search <sup>2</sup>, el cual es un buscador que permite ingresar una palabra, un intervalo de tiempo y proporciona un listado con todos los artículos que contienen la palabra, ya sea en su título, encabezado, bajada o cuerpo. Finalmente New York Times proporciona APIs a todas aquellas personas que deseen tener acceso a su información, solo pidiendo un registro de API Key, el cual permite la realización de 1000 llamadas por día.

---

<sup>1</sup><http://chronicle.nytlabs.com>

<sup>2</sup><http://query.nytimes.com/search/sitesearch/>

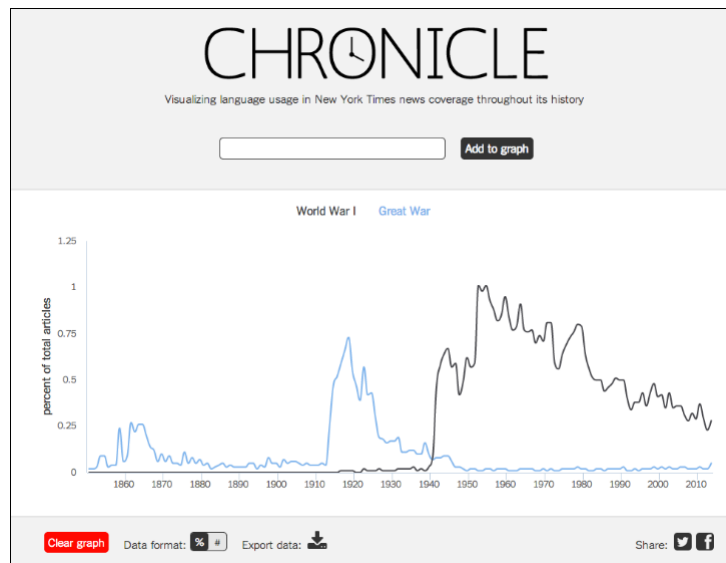


Figura 1.1: Chronicle

## 1.2. Descarga de Datos

### Datos Anuales

Para la descarga de los datos anuales se utilizó un programa en Python que permitía descargar la información desde el sitio web Chronicle, por medio de dicho programa era posible descargar el número de noticias que contenían la palabra de interés durante un año guardando en un archivo .csv el año y la cantidad de noticias que contenían al menos una vez alguna de las palabras de interés. Tomando en consideración lo anterior, se buscaron las 354 palabras con connotación positiva y las 2355 palabras con connotación negativa definidas según el diccionario de Loughran and McDonald (2011).

Luego de contar con la cantidad de noticias asociadas a cada una de las palabras, estas fueron agrupadas en positivas y negativas para ser posteriormente agregadas, con esto se obtuvo el número de palabras positivas o negativas según fuera el caso que aparecían en el diario durante un año considerando solo una aparición de cada palabra por noticia, lo cual era la información de interés.

### Datos Diarios

En primera instancia se intentó utilizar la misma fuente de información que para la descarga de datos anuales, para esto se desarrolló un programa en Python que permitía descargar la información desde el sitio web Chronicle, la principal ventaja asociada a esta plataforma tenía relación con lo estático de su código fuente, ya que esto permitía la extracción de manera sencilla de la información, mediante la localización de los datos de interés asociados a los “tag” del código

fuelle. Luego de realizar la descarga de un porcentaje considerable de la información se analizó la calidad de esta donde se pudo observar que era muy incompleta y la mayoría de los artículos carecían de su título o fecha.

Posterior a esto se comenzó a utilizar NYT Search, esta página tenía como principal inconveniente el dinamismo de su código fuente, lo que tornaba la extracción de datos un tanto más compleja que en el caso anterior, ya que no era posible utilizar los métodos que se habían desarrollado para la extracción de datos en Chronicle. Para esto se realizó un programa en Java que mediante el uso de Selenium, herramienta que permite el manejo de códigos dinámicos, el cual por medio del “className” permite extraer la información necesaria. Al momento de realizar la ejecución del programa, este tenía un rendimiento bastante deficiente ya que posee una alta dependencia de la calidad de la conexión a Internet, en varias ocasiones al demorarse más de lo habitual en cargar la página, la información que esta contenía era omitida en la descarga, dejando vacíos difíciles de detectar.

De forma paralela al desarrollo de dicho programa se desarrolló otro similar al anterior, pero que permitiera el manejo de la API proporcionada por el New York Times. La utilización de este sistema proporcionado por el New York Times elimina de cierta manera la fuerte dependencia con la conexión a Internet, ya que si bien es necesaria para realizar la descarga, la API genera un JSON, el cual es más liviano y se carga con mayor rapidez que el caso anterior, de la misma forma la velocidad de descarga aumenta de manera considerable. Por lo tanto para realizar la obtención de la información se utiliza una de las APIs anteriormente mencionadas, NYT Article Search <sup>3</sup>, la cual utiliza como URL base:

```
http : //api.nytimes.com/svc/search/v2/articlesearch.json?q = word&begin_date =  
20120101&end_date = 20120101&page = 2&sort = oldest&api - key = xxx
```

Donde *word* debe ser remplazada por la palabra que se desea buscar, *begin\_date* y *end\_date* son la fecha de inicio y término de la búsqueda respectivamente, *page* es la página en la cual se está posicionado, cada página contiene 10 artículos y en cada intervalo de tiempo la API permite visualizar como máximo 100 páginas, es decir, 1000 artículos. Si la palabra contiene más de 1000 artículos en el intervalo de tiempo a estudiar, es necesario fragmentar dicho intervalo para poder capturar todos los artículos, *sort* permite seleccionar la forma en que se desea ordenar la búsqueda, para el desarrollo de este trabajo se eligió la opción *oldest*, la cual ordena los artículos de más antiguo a más reciente, finalmente se tiene la *api-key* que es el identificador que otorga el New York Times para utilizar sus APIs, es necesario destacar en este punto que cada API Key permite realizar 1000 llamadas por día, se considera una llamada cada vez que se hace ingreso a la URL <sup>4</sup>.

El resultado asociado a la URL anterior es un JSON que contiene toda la información requerida, además del link con la noticia, el resumen, keywords, el tipo de documento, entre otros campos que ayudan a describir una noticia. Varios de los campos proporcionados en el JSON fueron

---

<sup>3</sup><https://developer.nytimes.com/>

<sup>4</sup>Más información sobre parámetros de la URL en [https://developer.nytimes.com/article\\_search\\_v2.json#/README](https://developer.nytimes.com/article_search_v2.json#/README)



omitidos durante la descarga debido a que la mayoría de las noticias presentaba estos como “null” al comienzo del estudio, lo que impedía la extracción fluida de información. A modo de contexto, mencionar que un JSON o JavaScript Object Notation es un formato utilizado para el intercambios de datos, básicamente el JSON describe los datos con una sintaxis dedicada que se usa para identificar y gestionar los datos, la forma en que se encuentra la información luego de parsearlo. Por lo que se ve más o menos así: “print\_page”:“1”, donde el primer string entre comillas representa la clave y el segundo el valor asociado a dicha clave.

Tomando lo anterior en consideración se realizó un programa en Java que lee la URL, toma el JSON que esta entrega como output, lo parsea para poder obtener la cantidad de artículos que existen en el intervalo de tiempo de la búsqueda además de la información ha utilizar, finalmente se escribe un archivo .csv con la fecha y el titulo del articulo. En primera instancia con el objetivo de minimizar el numero de llamadas a realizar por día se opto por descargar aquellas palabras que durante el periodo que contempla los años 1851 y 2014 tienen un numero menor o igual a 1000 artículos, ya que estas demandaban como máximo 16.300 llamadas, luego se decidió fragmentar el intervalo de tiempo en periodos más pequeños hasta llegar a realizar la descarga de manera mensual, es decir, las claves begin\_date y end\_date abarcaban un mes de cada año, en este caso para poder descargar una palabra eran necesarias hasta 195.600 llamadas.

### **1.3. Descarga Masiva de Información Diaria**

Una vez que se dimensiono la cantidad de llamadas necesarias para poder realizar la descarga de una palabra y tomando en consideración que se necesitaban 2709 palabras contando las positivas y las negativas además de que solo se disponía de 1000 llamadas diarias para descargar información, se decidió recurrir al uso de VPNs las cuales permiten crear una red local sin necesidad de que se este físicamente conectado entre si, sino que a través de Internet. De manera más sencilla, la dirección IP de la maquina que esta conectada a la VPN es la del servidor VPN, es decir, la maquina esta ubicada virtualmente en la dirección de la VPN. Lo anterior permite evadir los bloqueos de contenidos por ubicación geográfica, esto es muy útil para el desarrollo de este trabajo debido a que una vez que se cumplía la cuota diaria de 1000 llamadas de la API Key, la IP era bloqueada y se desbloqueaba al día siguiente. El uso de VPN permitió eliminar esta barrera, ya que cuando era bloqueada una IP se buscaba de manera manual otra VPN ubicada en alguna parte del mundo que no hubiese sido utilizada con anterioridad en el día para continuar con la descarga. Para poder implementar el sistema de VPNs fue necesaria la creación de tantas API Key como VPNs se utilizarían, el pago de un proveedor de VPNs y la utilización del programa que ellos proporcionaban para poder conectarse, esta mejora aumento en 8 veces la velocidad de descarga.

No obstante lo anterior, la estimación de los tiempos de descarga de las 2709 palabras eran muy elevados, por lo que se decidió automatizar el cambio de VPN y de API Key una vez que esta ultima cumpliera con el máximo de llamadas diarias con el fin de que el proceso corriera de manera continua día y noche sin interrupción. Para lo anterior fue necesaria la creación de maquinas virtuales con VirtualBox, en aquellos computadores que poseían sistema operativo Windows, de esta manera en un mismo computador era posible tener múltiples maquinas vir-

tuales con Linux que simulaban cada una un computador independiente que permitía correr el código. De manera adicional se debieron cambiar las rutas y se realizaron algunas mejoras en el código, con el fin de que este pudiese ser ejecutable en el sistema operativo Linux. Con este método, en el momento que se logro realizar la mayor cantidad de descargas, se contaba con 5 computadores corriendo el código. Esto se realizo antes de que el New York Times presentara problemas técnicos en su API, los cuales impidieron realizar la descarga fluida de información durante un periodo considerable de tiempo.

## **1.4. Base de Datos Diaria**

El programa creado para la descarga de la información generaba un archivo .csv para cada una de las palabras, es decir, los artículos que contenían la palabra happy entre los años 1851 y 2014 estaban guardados en el archivo happy.csv y así con todas las palabras de interés, por lo cual existían tantos archivos como palabras se iban descargando. Estos archivos no eran de mucha utilidad cada uno por si solo, ya que el principal objetivo del trabajo tenia relación con el efecto de los medios de comunicación en las persona y para esto era necesario analizar el conjunto de palabras positivas y/o negativas ya sea en la semana o en una noticia en particular.

Para lo anterior se creo una Base de Datos en MySQL llamada “NYTnew”, la cual esta formada de dos tablas, la primera es “words” la cual contiene las 2709 palabras del diccionario, un id asociado a cada una de estas palabras, una variable binaria que asigna el tono de la palabra; cero en caso de que tenga tono positivo y uno en caso de que su tono sea negativo, de manera adicional se creo una columna que muestra si la palabra ya fue cargada a la base de datos, esta variable toma el valor cero en caso de que aun no se ha cargado la palabra y el valor uno en caso de que ya ha sido cargada. La segunda tabla de la base de datos es “noticias”, esta tabla contiene un id de noticia, es decir, cada noticia ingresada a la base de datos tiene un id único, el titulo de la noticia, la fecha y hora de publicación de la misma y finalmente el id de la palabra a la cual corresponde dicha noticia.

Para realizar la inyección de la información en la base de datos, lo primero que se hizo fue crear un .csv con las palabras negativas y positivas para ingresarlas en la tabla “words”. Posterior a esto se realizo un programa que leía los archivos con las palabras positivas y negativas y las iba ingresando a la tabla, asignándoles una id y el efecto según fuese el caso. Adicionalmente a esto el programa permitía leer los archivos .csv de cada palabra e ir ingresando columna por columna su titulo y fecha, ademas según el nombre del archivo le asignaba la id de la palabra.

## **1.5. Limpieza de la Base de Datos Diaria**

En primera instancia se reviso que la información descargada fuese consistente, es decir, que las fechas tuviesen formato de fechas y que existieran la menor cantidad de registros sin información, en este proceso se descubrió que habían registros que en el campo fecha tenían un string en vez de números, es por lo anterior que se procedió a corregir de manera manual cada uno de

estos casos, recurriendo a la pagina del NYT Search <sup>5</sup>, buscando por el titulo del articulo con error y colocando de manera manual la fecha cuando era correspondiente. Al finalizar este proceso se logro que el 100 % de los registros tuviera su fecha. De manera adicional, se descubrió que algunos de los registros no contaban con un titulo, lo anterior no era posible de corregir, ya que para una determinada fecha existían tantos títulos como noticias se habían publicado durante ese día, por lo cual finalmente la base de datos cuenta con 130 noticias sin titulo.

Por otro lado se obtuvieron las noticias únicas por día, entiéndase por noticia única, aquellas que tienen igual fecha y titulo, recordando que la base de datos repite las noticias tantas veces como palabras del diccionario utilizado esta contenga. Una vez realizado este procedimiento se pudo observar que existía una noticia que al realizar el conteo de la cantidad de palabras positivas que esta mencionaba, poseía más de las que el diccionario de Loughran and McDonald (2011) definía como palabras positivas, de manera inmediata se busco el titulo asociado a esta noticia: “Why College Matters” el día 24 de Septiembre del año 2007 y se pudo observar que existan aproximadamente 125 noticias con ese titulo publicadas durante ese día. Es por lo anterior que se decidió eliminar esas noticias del conteo de noticias únicas.

Posterior a esto se realizo un agregado semanal de la data con la que se disponía. Realizando esto, se pudo observar que el año 2014 presenta semanalmente cantidades de noticias muy por sobre la media y de manera particular se encontró que días como el 11 de diciembre de 2014, 23 de diciembre de 2014 y 31 de diciembre de 2014 presentaban una gran cantidad de vídeos publicados, esta situación era bastante extraña por lo cual se decidió tomar esos días como outliers y no considerarlo para el análisis.

Finalmente, para el desarrollo de esta investigación se trabajo con el ratio  $\frac{pos}{\sum(neg+pos)}$ , donde *pos* es el numero de palabras positivas pertenecientes al diccionario Loughran and McDonald (2011) que aparecen en el periodo de tiempo en el cual se trabajara, contando solo una aparición de cada palabra por noticia y *neg* se define de la misma forma pero considerando las palabras negativas proporcionadas por Loughran and McDonald (2011). El objetivo de dicha medida, es poder tener un panorama del tono que tienen las noticias durante un periodo determinado de tiempo, con el fin de que este sea un proxis de la estabilidad que están observando los agentes de la economía.

## 1.6. Datos Adicionales

Como complemento a los datos anteriormente mencionados, se utilizaron otras variables económicas de interés, entre ellas se puede mencionar: datos históricos de las crisis de moneda obtenidas a partir de la base de datos de Reinhart and Rogoff (2009), donde por crisis de moneda se refiere a el abandono de tipo de cambio fijo o bien a la depreciación de la moneda de un 15 % o más respecto al Dolar. De igual forma se utilizo la información de ciclos económicos proporcionada por NBER, la cual se obtuvo de su sitio web <http://nber.org/cycles.html>.

De manera adicional se uso como variable de control el logaritmo natural del PIB per cápita,

---

<sup>5</sup><https://query.nytimes.com/search/sitesearch/>

el cual se obtuvo de la Maddison economic database de Bolt and Zanden (2014), el ratio deuda sobre PIB per cápita y la tasa de inflación, ambos índices se obtuvieron desde Reinhart and Rogoff (2010). Por otro lado, los datos asociados a la current account entre los años 1851 y 1970 se consiguieron del informe of the Census (1975) y fueron completados por los datos que proporciona el portal web de la Oficina de Análisis Económico <sup>6</sup> entre los años 1960 y 2014, estos datos fueron expresados en billones de dolares para poder obtener resultados que permitieran un mejor análisis. Para el calculo del ratio inversión/capital, se utilizaron datos de inversión provenientes de Federal Reserve Bank of St. Louis <sup>7</sup> entre los años 1929 y 2014, para los años entre 1901 y 2928 se utilizo la data que proporciona el portal web de la Oficina de Análisis Económico, finalmente se realizo la acumulación de capital de la misma forma que lo hace Cochrane (1991). En cuanto a la tasa de interés, esta es la utilizada por Schularick and Taylor (2012), los cuales provienen de una base de datos de Obstfeld et al. (2005) para los registros anteriores a la guerra mientras que los datos posguerra provienen de la IFS database. Finalmente los retornos promedios históricos de 1897 al 2013 del Dow Jones (DJIA)<sup>8</sup> utilizados fueron proporcionados por el sitio web <http://djaverages.com>.

---

<sup>6</sup><https://www.bea.gov/index.htm>

<sup>7</sup><https://fred.stlouisfed.org/>

<sup>8</sup> El Dow Jones Industrial Average (DJIA) refleja el comportamiento del precio de la acción de las 30 compañías industriales más importantes y representativas de Estados Unidos

# Capítulo 2

## Relación Medios y Crisis

### 2.1. Análisis Descriptivo

Para comenzar el análisis descriptivo de la data se realizó el conteo del número de noticias publicadas anualmente, el cual se puede ver en el gráfico 2.1, estas noticias fueron extraídas por medio de la página web Chronicle <sup>1</sup>. En el gráfico anteriormente mencionado es posible ver que la cantidad de noticias presenta una tendencia a ir aumentando con el transcurso del tiempo, lo anterior tiene bastante sentido, ya que con el paso de los años, las redes de comunicación entre países se han hecho más estrechas, permitiendo que se publiquen noticias de muchos más ámbitos y mucho más variadas que en la antigüedad. De manera adicional, es posible observar algunos años que presentan peak en la cantidad de noticias publicadas.

De la misma forma, se estudio el comportamiento gráfico del ratio  $pos/(pos + neg)$ . En el gráfico 2.2, se puede observar que existen ciertos saltos, en los cuales aumenta dicho ratio, de la misma forma, se ve que el ratio tiene una tendencia en el tiempo y no es tan ruidosa como se podría pensar. De manera adicional, es posible ver que posterior al año 1950 comienzan a disminuir considerablemente los valores, alcanzando los mínimos más bajos del intervalo de tiempo estudiado.

---

<sup>1</sup><http://chronicle.nytlabs.com>

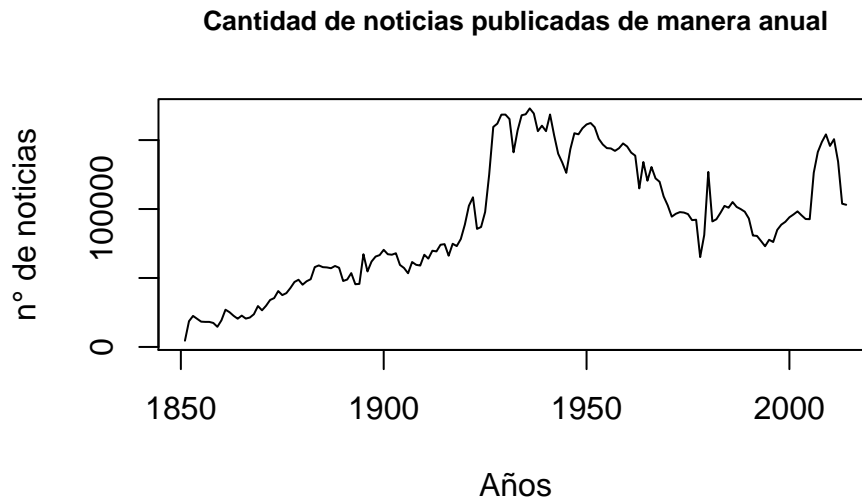


Figura 2.1: Cantidad de Noticias publicadas anualmente

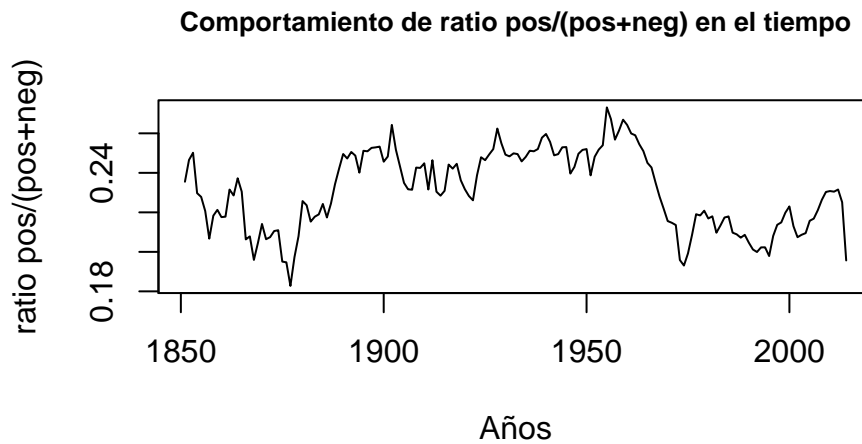


Figura 2.2: Comportamiento de ratio  $pos/(pos + neg)$  con agregación anual en el tiempo

De manera adicional se desarrollo la estadística descriptiva asociada a la cantidad de palabras negativas y positivas mencionadas durante un año, para eso se realizo el conteo de la cantidad de palabras, según el diccionario de Loughran and McDonald (2011), considerando todo el año independientemente si estas pertenecían o no a la misma noticia y contando solo una de las apariciones por palabra en cada noticia, la tabla 2.1 muestra los resultados de dicho análisis. De la

misma forma se dividió la data considerando los años con presencia de currency crisis y aquellos en que no, obteniéndose 29 años con presencia de currency crisis y 130 años sin presencia de la misma. En dicha tabla, es posible ver que en periodos de crisis se observan promedios más altos del ratio  $pos/(pos + neg)$  que en los periodos cuando no se esta en presencia de una crisis.

Tabla 2.1: Estadística descriptiva anual currency crisis

En esta tabla es posible observar la estadística descriptiva asociada al comportamiento del contenido de los medios durante currency crisis. La medida de cantidad  $pos$  es el numero de palabras positivas pertenecientes al diccionario Loughran and McDonald (2011) que aparecen en un año, contando solo una aparición de cada palabra por noticia,  $neg$  se define de la misma forma pero considerando las palabras negativas proporcionadas por Loughran and McDonald (2011). Por otro lado la medida de nivel se define de la siguiente forma  $\frac{pos}{(neg+pos)}$ . De manera adicional, se realizo una división de la data en periodos de currency crisis según la base de datos propuesta por Reinhart and Rogoff (2009).

Media measure	Mean	Median	Min	Max	25 %-quant	75 %-quant	Stand. dev
<b>All dates</b>							
<b>Medidas de cantidad</b>							
pos	235500	253700	13770	529600	151500	297200	107249.6
neg	537100	589200	29760	1176000	319500	1176000	248164
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.305	0.309	0.26	0.346	0.289	0.346	0.018
<b>Sin crisis</b>							
<b>Medidas de cantidad</b>							
pos	224200	250500	13770	502900	126300	286500	104770.2
neg	516800	569500	29760	1166000	295200	685900	246977.4
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.303	0.305	0.261	0.339	0.288	0.317	0.018
<b>Currency crisis</b>							
<b>Medidas de cantidad</b>							
pos	255200	255100	88080	383600	183900	334100	88146
neg	561500	639700	219400	824300	373300	734100	198472.9
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.313	0.329	0.26	0.346	0.306	0.371	0.020

De manera adicional en el anexo A, tabla 1, se muestran los resultados del análisis anterior, pero

considerando los ciclos económicos proporcionados por “NBER”, en ella es posible observar que durante periodos asociados a recesiones existe un menor promedio del ratio  $pos/(pos+neg)$  que en los periodos de expansión, es decir, durante recesiones es posible observar una menor cantidad de palabras positivas por sobre el total de palabras que en épocas de expansión.

## 2.2. Desarrollo de Hipótesis

### 2.2.1. Relación Medios y Crisis

En el desarrollo de esta hipótesis se busca poder caracterizar la existencia de alguna relación entre la estabilidad, definida como la cantidad de palabras positivas sobre el total de palabras negativas más positivas y la probabilidad de crisis. Todo lo anterior basándose en la hipótesis de inestabilidad financiera desarrollada por Minsky (1992), la cual plantea que conforme un sistema estable avanza en el tiempo, las condiciones financieras que operan bajo este, empiezan a incrementar el riesgo de inestabilidad, ya que por lo general, son las instituciones financiera y las regulaciones las que impiden estas inestabilidades, sin embargo, Minsky (1992) expone que un periodo de “bonanza” hace que las regulaciones se relajen, permitiendo la aparición de inestabilidad.

Para la validación de dicha hipótesis se decidió utilizar un modelo logístico, el cual se presenta en la ecuación 2.1. De manera adicional, la agregación utilizada en los datos es anual ya que cuando se habla de periodos de estabilidad es necesario que estos sean prolongados, ya que deben generar un “sentimiento” de estabilidad en la población y de manera particular en los inversionistas.

$$C_t^c = \beta_1 \overline{C}_{t-1 \rightarrow t-L}^c + \beta_2 \overline{M}_{t-1 \rightarrow t-L} + \beta_3 \overline{X}_{t-1 \rightarrow t-L} + \varepsilon_t \quad (2.1)$$

En el modelo logístico anterior se tiene que  $C_t^c$  es la dummy asociada a la variable crisis, en este caso en particular crisis de moneda, la cual toma el valor 1 en caso de que la fecha pertenezca a una crisis y 0 en caso contrario. La variable  $M$  es la que representa el ratio  $\frac{pos}{(neg+pos)}$ , el contenido de los medios. Finalmente,  $\overline{X}_{t-1 \rightarrow t-L}$  son las variables independientes que serán usadas como control, entre ellas se encuentra,  $\overline{volM}_{t-1 \rightarrow t-L}$  que es la volatilidad asociada al ratio  $\frac{pos}{(neg+pos)}$ ,  $\overline{inf}_{t-1 \rightarrow t-L}$  es la inflación,  $\overline{loggdpp}_{t-1 \rightarrow t-L}$  se define como el logaritmo del GDP per cápita y finalmente  $\overline{deb/gdpp}_{t-1 \rightarrow t-L}$  que es el ratio deuda sobre GDP per capita.  $L$  es la cantidad de lag a considerar en el modelo, para el desarrollo de esta hipótesis se utilizaron de 1 a 10 ya que estos son los números recurrentes que se utilizan en la literatura relacionada con temas similares al estudiado durante este trabajo como es el caso de Danielsson et al. (2016). De manera adicional las variables de promedio se construyeron como sigue:

$$\overline{z}_{t-1 \rightarrow t-L} = \frac{1}{L} \sum_{i=1}^L z_t, z = C, P, X \quad (2.2)$$



En cuanto a los resultados que se esperan con el desarrollo de esta hipótesis, estos tienen relación con el fenómeno explicado por Minsky (1992), ya que si se piensa en un diario de gran alcance como es el caso del New York Times, el cual anualmente es leído por millones de personas, el contenido que este publica en sus páginas tiene una penetración importante en la sociedad. Luego si el diario proyecta cierta estabilidad en sus noticias, las personas podrían verse tentadas a tomar decisiones en base a lo que están percibiendo, es decir, como el diario proyecta una situación de estabilidad y bonanzas, los inversionistas se sienten más relajados a la hora de tomar las decisiones, aun cuando esta estabilidad proyectada por el diario podría no estar relacionada con el ámbito en el cual ellos están tomando las decisiones. Por lo tanto como las decisiones que están tomando no son del todo serias y más bien están sustentadas en la irracionalidad, ansiedad y relajación que está generando este periodo de “bonanza”, se podría terminar desencadenando un episodio de inestabilidad financiera.

## 2.2.2. Mecanismos de explicación

### Relación Noticias y Current Account

La cuenta corriente (current account) es el punto de la balanza de pagos donde se recogen las operaciones reales (comercio de bienes y servicios) y rentas que se producen entre los residentes de un país y el resto del mundo, es por esto que se considera como un indicador importante de la salud de una economía. Esta se define como la suma de la balanza comercial (bienes y servicios exportados menos bienes y servicios importados) y los ingresos netos del exterior menos las transferencias netas. Una cuenta corriente positiva significa que el país es un prestamista para el resto del mundo, además de que sus exportaciones son mayores a sus importaciones.

Es por los puntos anteriormente planteados que pareció interesante estudiar la existencia de alguna relación entre la cantidad de palabras positivas sobre el total de palabras publicadas en un año por el New York Times y la cuenta corriente de los Estados Unidos, con el fin de entender de mejor manera la dinámica que podría desencadenar en una crisis de moneda. Para lo anterior se utilizó la regresión que se presenta en la ecuación 2.3.

$$CA_t = \beta_1 \bar{M}_{t-1 \rightarrow t-L} + \beta_2 \bar{X}_{t-1 \rightarrow t-L} + \varepsilon_t \quad (2.3)$$

Donde  $\bar{X}_{t-1 \rightarrow t-L}$  considera el término constante y otras variables independientes que podrían ser importantes de considerar, como es el caso de la tasa de interés, el tipo de cambio y el ratio deuda sobre GDP per capita, todas las anteriores con agregación anual. De manera adicional es importante tener en consideración que la  $L$  representa los lag

Parece esperable que la relación que se encuentre al regresionar la ecuación 2.3, muestre un efecto positivo entre el ratio  $pos/(pos + neg)$  y la current account, ya que el tener la percepción de “buenos tiempos” hace la economía local mucho más atractiva, de esta forma la gente tiene mayor disposición a invertir en ella y con esto aumentar la competitividad.

## Relación Noticias y Ratio Inversión/Capital

El ratio inversión/capital, permite tener una idea de cuanto del capital disponibles esta siendo utilizando en inversión, de esta forma, se buscara la existencia de alguna relación entre dicho ratio y el ratio  $pos/(pos+neg)$ . Los resultados que se esperan de esta relación, están ligados con que un sentimiento de estabilidad en los agentes podría llevarlos a tener una mayor confianza en la economía y con ello a aumentar la inversión en el país, con lo anterior aumentaría la competitividad explicando de esta forma un posible aumento en las exportaciones y con ello en la current account. La regresión a testear para el análisis de esta relación es la que se expresa en la ecuación 2.4.

$$I/C_t = \beta_1 \overline{M}_{t-1 \rightarrow t-L} + \beta_2 \overline{X}_{t-1 \rightarrow t-L} + \varepsilon_t \quad (2.4)$$

Donde  $\overline{X}_{t-1 \rightarrow t-L}$  considera el termino constante y otras variables de control que podrían ser importantes de considerar, como es el caso de la tasa de interés, el tipo de cambio y el ratio deuda sobre GDP per capita, todas las anteriores con agregación anual. De manera adicional es importante tener en consideración que la  $L$  representa los lag

## Relación Noticias y Tasa de Interés

De manera adicional y con el fin de entender de una manera más formal el mecanismo por el cual las variables asociadas a la current account y la inversión se podrían ver afectadas por la variable de estabilidad, se busco ver si existía una relación entre las “percepciones” que se generan por medio de las noticias y el grado a aversión al riesgo de los agentes. Debido a la complejidad que presenta el poder estimar la aversión al riesgo, se decidió usar la tasa de interés como proxi de esta. Lo anterior basado en el modelo de Lucas <sup>2</sup> Lucas Jr (1978), el cual plantea que:

$$(1 + r_t) = (1 + \rho)(1 + g_t)^\alpha \quad (2.5)$$

Tomando en consideración lo anterior, se utilizo la tasa de interés como un proxi de la aversión al riesgo con el fin de poder encontrar alguna relación entre la aversión al riesgo y la cantidad de noticias que contienen alguna de las palabras de interés publicadas de manera anual en el diario, como medida de estabilidad del país. Para lo anterior se realizo la siguiente regresión:

$$i_t = \beta_1 \overline{M}_{t-1 \rightarrow t-L} + \beta_2 \overline{X}_{t-1 \rightarrow t-L} + \varepsilon_t \quad (2.6)$$

Donde  $\overline{X}_{t-1 \rightarrow t-L}$  considera el termino constante y otras variables independientes que podrían ser de interés, como es el caso de la inflación, el logaritmo natural de GDP per cápita y el ratio

---

<sup>2</sup>mayor detalle disponible en el anexo D

deuda sobre GDP per cápita todas con agregación anual. De manera adicional al igual que en los casos anteriores  $L$  representa los lag.

Como se menciono anteriormente, es razonable pensar que debido a que se esta proyectando un episodio de bonanzas a través de los medios, los inversionistas presenten una mayor confianza sobre el mercado, lo anterior, los lleva a disminuir su aversión al riesgo y con ello a tomar decisiones que en otras condiciones no hubiesen pensado en tomar, de esta forma aumentarían su endeudamiento, ya que como se esta en un buen periodo, se privilegia el invertir, con ello en el mediano plazo debería observarse un aumento en la cantidad de empresas en el país, por lo que aumentaría la competitividad de este y junto a esto las exportaciones, ya que como el país se vuelve más competitivo sus productos se tornan atractivos para los otros países, pero como la competitividad que se formo no es del todo real, ya que se basa más que nada en esta estabilidad proyectada por lo medios, tal cual lo comenta Minsky (1992), se podría terminar desencadenando una crisis.

## 2.3. Resultados

Uno de los supuestos más grandes de la economía tiene relación con que se cree que las grandes decisiones económicas y financieras se toman desde el análisis y la observación de los datos micro y macroeconómicos. No obstante, lo cierto es que más a menudo de lo que se piensa, son factores psicológicos e irracionales, los que están detrás de los comportamientos de los inversores, economistas, financieros o políticos.

Por otro lado si se analiza lo postulado por Minsky (1992), es posible darse cuenta que periodos prolongados de estabilidad o bonanzas, para el caso de este estudio representada por la percepción que puede generar el New York Times con la porción de palabras positivas sobre el total de palabras que publica, podrían llegar a generar una idea relativamente falsa de lo que esta sucediendo en los mercados y con ello tomar decisiones equivocadas que terminen en un periodo de inestabilidad.

### 2.3.1. Relación Medios y Crisis

Luego de ejecutar el modelo logístico se obtuvieron los resultados que se muestran en la tabla 2.2, en ella se puede observar que en las regresiones simples, es decir, aquellas que solo consideran las variables de interés ( $\bar{C}_{t-1 \rightarrow t-L}$  y  $\bar{M}_{t-1 \rightarrow t-L}$ ) el ratio  $pos/(pos + neg)$  presenta significancia con 3, 5 y 7 lag, luego si se analizan 10 lag, como se puede ver en la tabla B.1 en anexos, entre el primero y el noveno lag se obtiene que el ratio  $pos/(pos + neg)$  es significativo. De manera adicional se puede observar que el efecto que esta variable presenta es positivo, lo anterior se interpreta como que un aumento en la porción de palabras positivas sobre el total de palabras (positivas más negativas), siempre pensando en las palabras de interés para el estudio, aumenta la probabilidad de currency crisis. Visto de otra forma se podría decir que a mayor estabilidad, representada por las palabras positivas publicadas en el diario, mayor probabilidad

de currency crisis.

De la misma forma se corrió la regresión completa, en este caso se consideraron otras variables económicas de interés que podían ayudar a explicar la probabilidad de currency crisis. Las variables independientes que se agregaron fueron: Volatilidad del ratio  $pos/(neg + pos)$ , inflación, logaritmo natural del GDP per cápita y el ratio deuda sobre GDP per capita. Los resultados de esta regresión no son muy distintos a los del caso simple, ya que es posible notar que el efecto mantiene tanto el signo como la significancia.

Si se observan los efectos marginales asociados al ratio de interés, los cuales se pueden ver en la tabla 2.2, es posible apreciar que estos se comportan de manera más bien cóncava, es decir, llegan a un peak cuando se consideran 4 lag si se observa la regresión simple y lo mismo sucede con la regresión completa, pero posterior a esto el efecto marginal comienza a disminuir. Por lo tanto entre más años sean considerados no significa que mayor será el efecto de aumentar en una unidad del ratio  $pos/(neg + pos)$ , sino que más bien hasta cierto punto los efectos marginales son crecientes y luego decrecen con la cantidad de lag considerados.

Finalmente con el fin de poder eliminar posibles sesgos que se presentaran por incluir la crisis del 2008 o las guerras, ya que los años anteriores a la Primera Guerra mundial presentan una menor frecuencia de crisis, mientras que los periodos de entre guerras están más propensos a crisis, se decidió correr el modelo sin considerar estos periodos, es decir, se eliminaron los años que comprenden entre 1914-1918, 1939-1945 y 2008. Al realizar esto, es posible observar que ni la significancia ni el signo asociado al efecto se ven alterados, solo se ve una disminución en los efectos marginales tal cual se muestra en la tabla 2.3.

Tabla 2.2: Logit relación ratio  $pos/(neg + pos)$  y currency crisis

En esta tabla es posible observar los resultados asociados a la ecuación 2.1, la cual relaciona el ratio  $pos/(pos + neg)$  y la crisis de moneda. De manera adicional, la tabla se divide en dos, los resultados que se muestran en la sección I son los que se obtuvieron a partir de las regresiones simples, mientras que los de la sección II son cuando se consideran variables independientes de interés como es el caso de  $\overline{volM}$  que representa la volatilidad de la variable asociada a los medios,  $\overline{inf}$  que es la tasa de inflación anual,  $\overline{loggdp}$  es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per capita. En esta tabla se utilizaron datos que contemplan entre 1852 y 2010.

Var. Dep: $C_t^{currency}$	I			II		
	3	5	7	3	5	7
$\overline{C}_{t-1 \rightarrow t-L}$	0.346 (0.673)	0.106 (0.919)	1.372 (1.153)	0.11 (0.741)	-0.372 (0.972)	0.73 (1.237)
$\frac{\overline{pos}}{\overline{(neg+pos)}}_{t-1 \rightarrow t-L}$	36.235*** (13.800)	35.391** (13.987)	24.53* (13.618)	37.634** (15.299)	38.36** (14.960)	24.164* (14.167)
$\overline{volM}_{t-1 \rightarrow t-L}$				17.372 (87.173)	71.53 (98.606)	46.809 (110.157)
$\overline{inf}_{t-1 \rightarrow t-L}$				-0.012 (0.047)	-0.05 (0.064)	-0.069 (0.070)
$\overline{loggdp}_{t-1 \rightarrow t-L}$				0.203 (0.522)	0.396 (0.516)	0.307 (0.485)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$				-0.009 (0.017)	-0.015 (0.017)	-0.016 (0.014)
Num. obs.	157	155	153	153	151	149
Pseudo $R^2$	0.066	0.055	0.056	0.074	0.076	0.073
Efectos Marginales						
$\frac{\overline{pos}}{\overline{(neg+pos)}}_{t-1 \rightarrow t-L}$	5.1	5.11	3.61	5.389	5.558	3.573

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Tabla 2.3: Logit relación ratio  $pos/(neg + pos)$  y currency crisis sin considerar guerras

En esta tabla es posible observar los resultados asociados a la ecuación 2.1, la cual relaciona el ratio  $pos/(pos + neg)$  y la crisis de moneda. De manera adicional, la tabla se divide en dos, los resultados que se muestran en la sección I son los que se obtuvieron a partir de las regresiones simples, mientras que los de la sección II son cuando se consideran variables independientes de interés como es el caso de  $\overline{volM}$  que representa la volatilidad de la variable asociada a los medios,  $\overline{inf}$  que es la tasa de inflación anual,  $\overline{loggdp}$  es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per capita. En esta tabla se utilizaron datos entre 1852 y 2010, pero se eliminaron los periodos de guerra (1914-1918 y 1939-1945) y la crisis del año 2008.

Var. Dep: $C_t^{currency}$	3	I		3	II	
		5	7		5	7
$\overline{C}_{t-1 \rightarrow t-L}$	0.807 (0.700)	1.163 (0.893)	1.648 (1.102)	0.514 (0.773)	0.38 (1.059)	0.589 (1.315)
$\frac{\overline{pos}}{\overline{(neg+pos)}}_{t-1 \rightarrow t-L}$	25.939** (12.138)	23.903** (11.757)	20.787* (12.174)	27.249** (13.497)	25.425** (12.744)	19.408 (12.539)
$\overline{volM}_{t-1 \rightarrow t-L}$				36.394 (89.490)	81.154 (101.894)	34.148 (126.893)
$\overline{inf}_{t-1 \rightarrow t-L}$				-0.019 (0.049)	-0.104 (0.084)	-0.107 (0.082)
$\overline{loggdp}_{t-1 \rightarrow t-L}$				0.376 (0.555)	0.665 (0.642)	0.509 (0.596)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$				-0.02 (0.018)	-0.028 (0.020)	-0.027 (0.017)
Num. obs.	144	142	140	140	138	136
Pseudo $R^2$	0.052	0.052	0.055	0.072	0.093	0.086
Efectos Marginales						
$\frac{\overline{pos}}{\overline{(neg+pos)}}_{t-1 \rightarrow t-L}$	3.57	3.33	2.93	3.75	3.43	6.69

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Los resultados que se encontraron en el desarrollo de esta hipótesis están muy relacionados con lo que aborda Minsky (1992) en su hipótesis de inestabilidad financiera esto debido a que si se piensa en el ratio  $pos/(pos + neg)$  como una medida de estabilidad en el tiempo, se tendrá que periodos prolongados de estabilidad, representados en el modelo por los lag considerados, podrían desencadenar en una crisis.

Por otro lado si lo anterior lo relacionamos con los sesgos de percepción, como lo son el de disponibilidad y el de representatividad <sup>3</sup>. Estos podrían llegar a justificar las decisiones que están tomando los inversionistas. Lo anterior debido a que los agentes perciben que el país esta en una buena situación ya que se presentan noticias con contenido más positivo, entonces tienen una mayor “confianza” y están más dispuestos a tomar riesgos. Aun cuando esta situación de “estabilidad” no sea tan real como lo parece o más aun no este en absoluto relacionada con la decisión que ellos están tomando.

### **2.3.2. Mecanismos de explicación**

#### **Relación Noticias y Current Account**

Una de las explicaciones interesantes por las cuales el aumento en la porción de palabras positivas por sobre el total de palabras publicadas por el New York Times aumenta la probabilidad de crisis, podría estar relacionada con la cuenta corriente o “current account”, ya que esta es un indicador importante de la “salud” de la economía. En la misma línea, su principal característica tiene que ver con que es un signo de competitividad relativa.

Tal cual se puede ver en la tabla 2.4 un aumento en el ratio  $pos/(pos + neg)$  aumenta la current account, la cual se define como la diferencia entre las exportaciones e importaciones de un país. Un aumento en la cuenta corriente es indicativo de una mayor demanda externa de los productos del país, ya que se están aumentando las exportaciones, es decir, se esta siendo más competitivo, lo anterior aumenta el dinero circulante en la economía pudiendo desencadenar la depreciación de este.

---

<sup>3</sup>mayor información de los sesgos de percepción en anexo E

Tabla 2.4: Regresión ratio  $pos/(pos + neg)$  - Current account

En esta tabla es posible observar los resultados asociados a la ecuación 2.3, la cual relaciona el ratio  $pos/(pos + neg)$  y current account. La tabla muestra los resultados en regresiones simples y considerando variables independientes que podrían ser de interés para 1 y 5 lag. Las variables  $\overline{TC}$  representa el tipo de cambio anual,  $\bar{i}$  es la tasa de interés anual y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda publica sobre el GDP per capita. En esta tabla se utilizaron datos entre 1852 y 2010.

Var. Dep: $CA$	L=1		L=5	
$\frac{pos}{(neg+pos)}_{t-1 \rightarrow t-L}$	2.228***	1.635*	2.852***	2.354***
	(3.595)	(1.884)	(4.365)	(2.484)
$\overline{TC}_{t-1 \rightarrow t-L}$		-0.737***		-1.142***
		(-5.672)		(-6.730)
$\bar{i}_{t-1 \rightarrow t-L}$		3.018***		6.098***
		(5.083)		(7.213)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$		0.0007		0.002***
		(1.183)		(3.296)
Num. obs.	163	123	159	123
$R^2$	0.074	0.428	0.108	0.538
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

### Relación Noticias y Ratio Inversión/Capital

De la misma forma, uno de los mecanismos que podrían ayudar a explicar el aumento de la current account, tiene relación con la inversión, lo anterior debido a que como se puede ver en la tabla 2.5, un aumento en el ratio  $pos/(pos + neg)$  provoca un aumento en el ratio inversión/capital, esto se puede ver reflejado en una mayor competitividad por parte del país, ya que al haber mayor inversión se podrían estar generando más industrias, con ello los productos del país se vuelven más atractivos, lo que explica el aumento que se observo con anterioridad en la current account.



Tabla 2.5: Regresión ratio  $pos/(pos + neg)$  - ratio Inversión/Capital

En esta tabla es posible observar los resultados asociados a la ecuación 2.4, la cual relaciona el ratio  $pos/(pos + neg)$  y el ratio inversión/capital. La tabla muestra los resultados considerando regresiones simples y luego con la incorporación de algunas variables independientes que podrían ser de interés para 1 y 5 lag. Las variables  $\overline{inf}$  que representa la inflación anual anual,  $\overline{loggd\dot{p}}$  que es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per capita. En esta tabla se utilizaron datos entre 1901 y 2014.

Var. Dep: $I/k$	L=1		L=5	
$\frac{\overline{pos}}{(\overline{neg+pos})_{t-1 \rightarrow t-L}}$	7.397***	5.391**	8.311***	6.119**
	(4.154)	(2.477)	(4.540)	(2.398)
$\overline{inf}_{t-1 \rightarrow t-L}$		-0.004		-0.004
		(-0.443)		(-0.487)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$		-0.009***		-0.008***
		(-5.022)		(-4.693)
$\overline{Tint}_{t-1 \rightarrow t-L}$		-3.523**		-2.887*
		(-2.404)		(-1.766)
Num. obs.	114	114	114	114
$R^2$	0.17	0.419	0.197	0.416
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

### Relación Noticias y Tasa de Interés

En cuanto a los resultados encontrados para la relación entre el ratio  $pos/(pos + neg)$  y la tasa de interés, como proxy de la aversión al riesgo, es posible ver en la tabla 2.6, que existe una relación inversa entre la cantidad de palabras positivas sobre el total de palabras, positivas más negativas y la tasa de interés, es decir, una mayor percepción de estabilidad en el país disminuye la tasa de interés. Si se analiza lo anterior haciendo un equivalente al coeficiente de aversión al riesgo, se podría interpretar como que una mayor percepción de estabilidad hace que los agentes disminuyan su aversión al riesgo, de esta forma están dispuestos a tomar posiciones que en otras condiciones de mercado no estarían dispuestos a tomar, lo que se podría traducir como un aumento de confianza.

Tabla 2.6: Regresión ratio  $pos/(pos + neg)$  - tasa de interés

En esta tabla es posible observar los resultados asociados a la ecuación 2.6, la cual relaciona el ratio  $pos/(pos + neg)$  y tasa de interés e implícitamente con el coeficiente de aversión al riesgo. La tabla muestra los resultados en regresiones simples y considerando variables independientes que podrían ser de interés para 1 y 5 lag. Las variables  $\overline{inf}$  que representa la inflación anual,  $\overline{loggd\dot{p}}$  que es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per capita. En esta tabla se utilizaron datos entre 1880 y 2010.

Var. Dep: i	L=1		L=5	
$\frac{\overline{pos}}{(\overline{neg+pos})_{t-1 \rightarrow t-L}}$	-0.773***	-0.367***	-0.865***	-0.441***
	(-7.135)	(-3.259)	(-7.913)	(-3.866)
$\overline{inf}_{t-1 \rightarrow t-L}$		0.001***		0.001***
		(3.856)		(3.732)
$\overline{loggd\dot{p}}_{t-1 \rightarrow t-L}$		0.022***		0.021***
		(5.849)		(5.714)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$		-0.0007***		-0.0007***
		(-7.944)		(-7.681)
Num. obs.	123	123	123	123
$R^2$	0.296	0.581	0.341	0.595

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Como fue posible evidenciar durante el transcurso de este capítulo, la medida de estabilidad generada por medio de las publicaciones realizadas por el New York Times permite realizar la predicción de currency crisis. De manera adicional fue posible mostrar que un aumento en el ratio  $pos/(pos + neg)$  lleva consigo una disminución en el coeficiente de aversión al riesgo lo que permite explicar el hecho de que una mayor estabilidad trae consigo un aumento tanto en la inversión como en la current account.

La evidencia anterior muestra que periodos de estabilidad prolongados, representados por el ratio  $pos/(pos + neg)$ , llevan a los inversionistas a disminuir su coeficiente de aversión al riesgo, esto tiene un efecto en las decisiones que ellos toman y como es de esperar, si este disminuye, tendrán una mayor disposición a tomar riesgos, lo que es consistente con la hipótesis planteada en Minsky (1992). Tomando en consideración lo mencionado, una disminución del coeficiente de aversión al riesgo lleva a un aumento de la inversión en el país lo que lo vuelve más competitivo ya que presentaría un incremento en la cantidad de industrias, esto se traduciría en que sus productos se tornan más cotizados, aumentan las exportaciones y con ello la current account del país. Lo anterior aumenta la probabilidad de una currency crisis ya que un aumento en la current account significa que aumenta el dinero circulante y con ello se experimenta la depreciación de este.

## Capítulo 3

# Relación Medios, Retornos y Ciclos Económicos

### 3.1. Análisis Descriptivo

En primer lugar se realizó el conteo del número de noticias que contenían alguna de las palabras de interés publicadas semanalmente por el New York Times, la figura 3.1 muestra su comportamiento, en el es posible ver que existen algunas anomalías, como es el caso de la última semana de Diciembre de 2014, donde se produce un salto, alcanzándose las 12000 noticias que contienen alguna de las palabras buscadas o la segunda semana de mayo de 1964, donde se produce una disminución abrupta en la cantidad de noticias publicadas que contienen alguna de las palabras buscadas, teniendo menos de 2000 noticias.

De manera adicional si se realiza una comparación por periodo de la cantidad de noticias que contienen alguna de las palabras de interés publicadas de manera semanal, el cual se muestra en tabla 3.1, en periodos de expansión se observa que en promedio se publican más noticias que en periodos de contracción. Lo mismo pasa con los mínimos y los máximos, en periodos de expansión se alcanzan mínimos más bajos y máximos más altos que en periodos de contracción, lo anterior podría ser explicado con el hecho de que en periodos de contracción se busca tener en el contenido del diario todo lo que esta en la palestra en ese momento, ahondando en la mayor cantidad de detalles posibles, sin importar la cantidad de noticias, mientras que en periodos de expansión puede pasar que se intente complementar el diario con noticias de menor importancia.

### Cantidad de noticias publicadas semanalmente

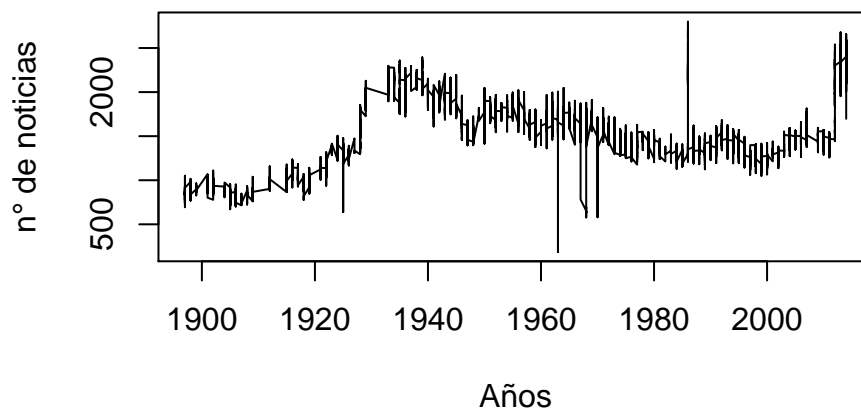


Figura 3.1: Cantidad de noticias que contienen alguna de las palabras de interés publicadas semanalmente

Tabla 3.1: Estadística sobre numero de noticias que contienen alguna de las palabras de interés publicadas semanalmente

Media measure	Mean	Median	Min	Max	25 %-quant	75 %-quant	Stand. dev
<b>All dates</b>	1471	1457	185	2799	1230	1729	389.9
<b>Expansiones</b>	1497	1475	185	2799	1285	1731	375.44
<b>Recesiones</b>	1387	1363	691	2374	984	1723	423.79

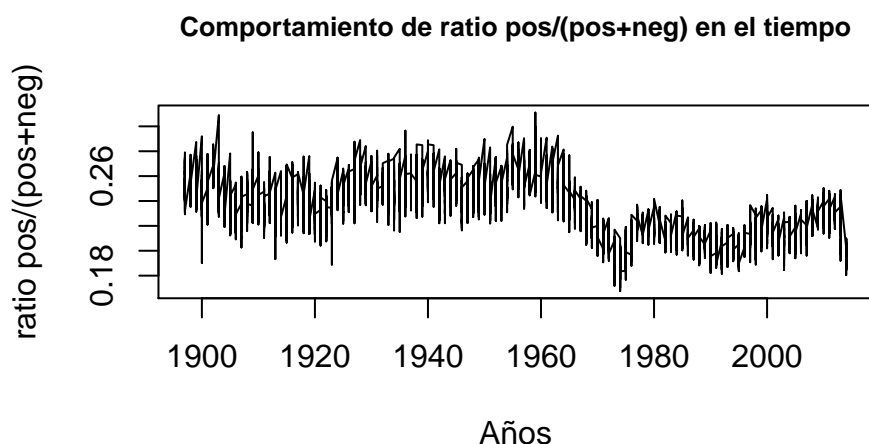


Figura 3.2: Comportamiento del ratio  $pos/(pos + neg)$  con agregación semanal en el tiempo

De la misma forma se estudio el comportamiento gráfico del ratio  $pos/(pos + neg)$  en el tiempo. En el gráfico 3.2, se puede observar que existen ciertos saltos, en los cuales aumenta dicho ratio, de la misma forma, se ve que el ratio es algo ruidoso en el tiempo, se mueve hacia ambos sentidos de forma abrupta, así como también es posible ver que posterior al año 1950 disminuye sus valores considerablemente.

Finalmente la tabla 3.2 muestra la estadística descriptiva asociada a la cantidad de palabras negativas y positivas mencionadas durante una semana, es decir, se realizo un conteo de la cantidad de palabras, según el diccionario de Loughran and McDonald (2011), considerando toda la semana independientemente si estas pertenecían o no a la misma noticia y contando solo una de las apariciones de cada palabra por noticia. De manera adicional la data se separo según la información de los ciclos económicos proporcionados por “NBER” obteniéndose 4478 semanas en periodos de expansión y 1350 semanas en periodos de contracción. En dicha tabla es posible observar que durante periodos asociados a recesiones existe un leve aumento en el promedio del ratio  $\frac{pos}{(neg+pos)}$  en relación a los periodos de expansión, es decir, durante recesiones es posible observar una mayor cantidad de palabras positivas por sobre el total de palabras que en periodos de expansión.

Tabla 3.2: Estadística descriptiva tono semanal

En esta tabla es posible observar la estadística descriptiva asociada al tono semanal, es decir, en el caso de las medidas de cantidad *pos* es el número de palabras positivas pertenecientes al diccionario Loughran and McDonald (2011) que aparecen en una semana, contando solo una aparición de cada palabra por noticia, *neg* se define de la misma forma pero considerando las palabras negativas proporcionadas por Loughran and McDonald (2011). Por otro lado la medidas de nivel se define de la siguiente forma  $\frac{pos}{(neg+pos)}$ , representando de esta forma el sentimiento de “positivismo”.

Media measure	Mean	Median	Min	Max	25 %-quant	75 %-quant	Stand. dev
<b>All dates</b>							
<b>Medidas de cantidad</b>							
pos	2125	2053	178	7259	1662	2390	785
neg	7014	6681	549	28072	5152	28072	3062.23
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.235	0.237	0.18	0.31	0.218	0.252	0.021
<b>Expansiones</b>							
<b>Medidas de cantidad</b>							
pos	2170	2071	178	7259	1782	7259	816.04
neg	7237	6889	5429	28072	5429	8094	3260.16
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.234	0.235	0.18	0.311	0.216	0.252	0.022
<b>Recesiones</b>							
<b>Medidas de cantidad</b>							
pos	1994	1966	660	4474	1417	2396	671.03
neg	6372	6046	2605	16001	4491	7813	2285.22
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.239	0.241	0.182	0.309	0.227	0.252	0.018

## 3.2. Desarrollo de Hipótesis

La segunda hipótesis a testear tiene relación con dos cosas, la primera de ellas es si la cantidad de noticias publicadas por el New York Times de manera semanal tiene algún efecto en la toma de decisiones de los agentes. La segunda tiene que ver con el hecho de si el contenido de los medios de comunicación y de manera particular del New York Times, influye en como los inversionistas toman las decisiones, generando un efecto a corto plazo medible en los retornos. Para lo anterior se tomaran como antecedente algunas de las investigaciones en torno a la cantidad de información a la cual son expuestos los inversionistas, los sesgos de percepción y como estas cosas pueden llegar a influir en las decisiones que toman los agentes en la economía.

Para el primer caso se buscara encontrar algún efecto que muestre si el exceso de información tiene efecto en como los inversionistas están tomando las decisiones y si esto a su vez trae consigo un efecto en los retornos de mercado. De la misma forma, el punto dos de esta hipótesis se buscara relacionarlo con lo planteado por Minsky (1992), ya que se tendría que cuando los inversionistas tienen una mejor percepción de su entorno, esto los podría llevar a tomar decisiones mucho más irracionales, es decir, decisiones sin un mayor sustento económico, sino más bien abaladas por su estado de animo.

Para testear esta hipótesis y los dos puntos a estudiar, se realizara un análisis descriptivo de los retornos asociados de Dow Jones y como estos se comportan en los diferentes ciclos económicos. Posterior a ello por medio de la ecuación 3.1 se buscara ver si existe alguna relación entre las variables asociadas a los medios, es decir, el ratio  $\frac{pos}{(neg+pos)}$ , *Cantidad* que representara la cantidad de noticias publicadas por el New York Times de manera semanal y *Normalizacion* que es una normalización de la cantidad de noticias por el promedio de noticias publicadas semanalmente en un año y los retornos.

$$R_t = \beta_1 \mathcal{L}_s(M_t) + \gamma_1 \mathcal{L}_s(R_t) + \psi_1 \mathcal{L}_s(R_t^2) + \varepsilon_t \quad (3.1)$$

En la ecuación 3.2 se muestra la regresión que se utilizara para relacionar las variables de interés en la validación final de la hipótesis, donde se consideraron los ciclos de expansión y recesión proporcionados por NBER, además en ella es posible notar que  $R_t$  son los retornos logarítmicos del Dow Jones,  $D_t$  es la variable dummy asociada a los periodos de recesión y expansión proporcionados por la pagina de NBER <sup>1</sup>,  $M_t$  es la variable que representa a los medios de comunicación, esta puede tomar el valor de  $\frac{pos}{(neg+pos)}$ , *Cantidad* o *Normalizacion* según se prefiera, además se consideran los retornos cuadrados, los cuales en la ecuación 3.2 están representados por  $R_t^2$ ,  $\mathcal{L}_s$  es el operador lag, donde  $s$  es el numero de lag considerados, para el desarrollo de este estudio se utilizaron 5 lag, los cuales se eligieron en base a la literatura relacionada con el tema. La agregación de los datos usada durante el desarrollo de esta hipótesis es semanal, lo anterior para reducir el ruido que pudiera existir al utilizar datos diarios.

---

<sup>1</sup><http://nber.org/cycles.html>

$$R_t = (1 - D_t)(\beta_1 \mathcal{L}_s(M_t) + \gamma_1 \mathcal{L}_s(R_t) + \psi_1 \mathcal{L}_s(R_t^2)) + D_t(\beta_2 \mathcal{L}_s(M_t) + \gamma_2 \mathcal{L}_s(R_t) + \psi_2 \mathcal{L}_s(R_t^2)) + \varepsilon_t \quad (3.2)$$

De manera adicional, para el desarrollo de la hipótesis se corrió el modelo considerando todo el periodo de estudio, es decir, entre 1897 y 2014. Además, al igual que en el caso de la hipótesis 1 se decidió correr una segunda versión del modelo, pero esta vez eliminando las guerras y la crisis del año 2008, es decir, se eliminaron los años que comprenden entre 1914-1918, 1939-1945 y 2008 con el fin de observar si los efectos encontrados dependían de estos eventos en particular.

### 3.3. Resultados

Durante los últimos años, se han hecho importantes descubrimientos en el campo de la psicología que explican por qué el mercado es tan proclive a periodos de volatilidad, formación de burbujas o grandes desplomes. Lo anterior, está estrechamente relacionado con los resultados obtenidos a partir de los modelos postulados durante este capítulo, por lo que en esta sección se busca poder entender cómo las barreras psicológicas pueden afectar las decisiones de inversión convirtiéndolas muchas veces en un desastre.

#### 3.3.1. Retornos Semanales

En primer lugar se analizará el comportamiento de los retornos asociados al DJIA semanales en el tiempo, en la tabla 3.3 se muestra la estadística descriptiva asociada a lo anterior y es posible ver que el rendimiento medio del DJIA fue de nueve puntos básicos por semana si se considera todo el periodo de estudio, con una volatilidad de 216 puntos básicos. Si se observa solo las recesiones, la rentabilidad promedio fue de -20 puntos básicos, mientras que en expansiones es de 15. De la misma forma, es importante mencionar las diferencias en las volatilidades durante los ciclos económicos, durante las recesiones, la volatilidad semanal es de 299 puntos básicos, mientras que en las expansiones esta fue de 184 puntos básicos. Lo anterior, es esperable, ya que en épocas de recesión los rendimientos asociados a las acciones o como en este caso al índice disminuyen. Si se mira la skewness, es posible darse cuenta que en todos los periodos la cola más larga se encuentra hacia la izquierda, mientras que si se mira el primer cuarto, es posible ver que en recesiones este es mucho más negativo que en expansión. Finalmente si se observa la curtosis, en periodos de expansión se observa una curtosis mucho más grande, lo que quiere decir que los datos se encuentran más cercanos a la media (la curva es más puntiaguda) que en episodios de expansión, que la curva está más achatada.



Tabla 3.3: Estadística descriptiva Retornos

Media measure	Mean	Median	25 %-quant	75 %-quant	Stand. dev	Skewness	Kurtosis
<b>Semanales</b>							
All dates	0.07	0.18	-0.91	1.22	2.16	-0.57	9.34
Expansions	0.15	0.22	-0.74	1.19	1.84	-0.74	9.72
Recessions	-0.2	-0.011	-1.57	1.33	2.99	-0.19	6.45

### 3.3.2. Cantidad de Noticias, Retornos y Ciclos Económicos

#### Cantidad de Noticias y Retornos Semanales

Al realizar un análisis de la relación entre los retornos y la cantidad de información publicada por el New York Times en todo el periodo de estudio, sin discriminar por recesión o expansión, tabla 3.4, es posible ver que un aumento en las variables *Cantidad* y *Normalizacion* predicen una disminución de los retornos. lo anterior puede ser explicado por medio de information overload, que es algo así como una sobrecarga de información. Toffler (1990) y Cornella (1999) plantean que las personas tienen una capacidad máxima de procesamiento de información, luego una vez que esta capacidad es sobrepasada cualquier información adicional termina convirtiéndose solo en ruido y disminuye la calidad de la decisión, ya que los inversionistas necesitan más tiempo para poder hacer predicciones, pero aun así estas por lo general son de una menor calidad.

De manera adicional, es posible ver en la tabla E.1, la cual esta disponible en el anexo E, que cuando se eliminan las guerras y la crisis del año 2008, el efecto no cambia ni en signo ni en significancia, solo disminuye un poco su intensidad. Los coeficientes de la variable de medios retardada entre  $t - 2$  y  $t - 5$  dan un indicio de si el shock sobre los retornos causados por el exceso de información en la fecha  $t - 1$  es permanente o temporal. De la tabla es posible darse cuenta que desde el tercer lag no existe significancia en los resultados.

Tabla 3.4: Regresión cantidad de noticias y retornos semanales

En esta tabla es posible observar los resultados asociados a la ecuación 3.1, la cual en este caso relaciona las variables de cantidad de noticias *Cantidad* y *Normalizacion*, donde *Cantidad* es la cantidad de noticias publicadas de manera semanal por el New York Times y *Normalizacion* es una normalización *Cantidad* por el promedio de noticias publicadas de manera semanal durante un año, con los retornos. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014.

Variable Medios	<i>Cantidad</i>		<i>Normalizacion</i>	
	$\beta$	t-stat	$\beta$	t-stat
$M_{t-1}$	-0.764**	(-2.491)	-1.031**	(-2.287)
$M_{t-2}$	1.122***	(2.847)	1.449***	(2.627)
$M_{t-3}$	-0.628	(-1.481)	-0.770	(-1.334)
$M_{t-4}$	0.0973	(0.229)	0.134	(0.236)
$M_{t-5}$	0.0472	(0.151)	-0.0665	(-0.147)

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

### Cantidad de Noticias, Ciclos Económicos y Retornos Semanal

En la misma línea de lo anterior, se separo la data en ciclos de expansión y ciclos de recesión según lo publicado por NBER en su pagina web. El panel A incluye las estimaciones de los coeficientes  $\beta_1$ , los cuales miden el efecto de la cantidad de noticias sobre los retornos de mercado durante periodos de expansión. Como es posible ver en la tabla 3.5, en estos periodos no se observo significancia.

El panel B presenta las estimaciones de  $\beta_2$ , el cual mide el efecto de la cantidad de noticias sobre el rendimiento de las acciones durante las recesiones. La estimación puntual de la cantidad de noticias es -3.083, con un estadístico t de -4.085. Un patrón similar se observa para la variable *Normalizacion* donde el valor de  $\beta$  para el primer lag es de -3.178 y su estadístico t es de -2.995. Es importante mencionar que el efecto de la cantidad de noticias es significativo solo en periodos de recesión y el signo de este es negativo para ambos casos, es decir, un aumento en la cantidad de noticias publicadas de manera semanal por el diario genera una disminución en los retornos cuando se esta en presencia de periodos de recesión.

En el anexo E, tabla E.2, es posible ver que tanto la significancia como el signo del efecto permanece al momento de sacar las guerras y la crisis del año 2008, mientras que se observa una pequeña disminución en la magnitud asociada al efecto.

Tabla 3.5: Regresión cantidad de noticias, ciclos económicos y retornos

En esta tabla es posible observar los resultados asociados a la ecuación 3.2, la cual en este caso relaciona las variables de cantidad de noticias *Cantidad* y *Normalizacion*, donde *Cantidad* es la cantidad de noticias publicadas de manera semanal por el New York Times y *Normalizacion* es una normalizacion *Cantidad* por el promedio de noticias publicadas de manera semanal durante un año, con los retornos y los ciclos económicos proporcionados pro NBER. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014.

	<i>Cantidad</i>		<i>Normalizacion</i>	
	$\beta$	t-stat	$\beta$	t-stat
<b>A. Expansión (<math>\beta_1</math>)</b>				
$(1 - D_t)xM_{t-1}$	-0.159	(-0.439)	-0.331	(-0.584)
$(1 - D_t)xM_{t-2}$	0.421	(0.944)	0.649	(0.952)
$(1 - D_t)xM_{t-3}$	-0.404	(-0.907)	-0.667	(-0.979)
$(1 - D_t)xM_{t-4}$	0.395	(0.884)	0.625	(0.917)
$(1 - D_t)xM_{t-5}$	-0.274	(-0.749)	-0.607	(-1.068)
<b>B. Recesión (<math>\beta_2</math>)</b>				
$D_t x M_{t-1}$	-3.083***	(-4.085)	-3.178***	(-2.995)
$D_t x M_{t-2}$	3.322***	(3.919)	3.518***	(3.005)
$D_t x M_{t-3}$	-1.341	(-1.593)	-1.001	(-0.857)
$D_t x M_{t-4}$	-0.556	(-0.652)	-0.834	(-0.707)
$D_t x M_{t-5}$	1.381*	(1.826)	1.682	(1.575)
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Tal cual se ha comentado anteriormente, un aumento en la cantidad de noticias publicadas genera confusión en los agentes, ya que estos no son capaces de procesar toda la información disponible. Lo anterior se ve reflejado en que los inversionistas tienden a ponerse ansiosos, se sienten impotentes y se confunden, por lo cual reducen sus esfuerzos en la toma de decisiones, llegando así a invertir de forma incorrecta, ya que se resaltan en mayor medida sus sesgos de percepción y tienden a invertir de manera muy arriesgada, siendo excesivamente conservadores o simplemente no diversifican de la manera correcta. De manera adicional, la ansiedad los hace resaltar su sesgo de aversión miope a las perdidas, por lo cual venden aquellas inversiones que han sido rentables en el tiempo y mantienen aquellas que están generando perdidas. De esta forma en el mediano plazo sus retornos se verán disminuidos

## Comportamiento de los Lag del Contenido de los Medios en el Tiempo

Tomando en consideración los resultados que se mostraron en los puntos anteriores, pareció relevante estudiar la significancia del coeficiente asociado al lag del contenido de los medios en el tiempo, es decir,  $M_{t-1}$ ,  $M_{t-2}$ ,  $M_{t-3}$  tanto en periodos de recesión y expansión como sin hacer esa distinción. Para lo anterior se corrieron las regresiones 3.1 y 3.2 considerando un intervalo móvil de 5 años entre cada regresión.

Los gráficos 3.3 y 3.4 muestran el comportamiento del coeficiente asociado a la variable que representa el contenido de los medios, *Cantidad* y *Ratio* en todo el intervalo de tiempo. En ellos es posible observar que el efecto de los medios de comunicación es significativo en pocas de las regresiones cuando se analiza todo el intervalo de tiempo, esto se puede ver ya que el intervalo de confianza en la mayoría de los casos incluye al cero.

Por otro lado si se separa la data en periodos de expansión y recesión, gráficos 3.5, 3.6 3.7 y 3.8, si se se analiza el periodo de recesión, es posible ver que existe mayor significancia en los betas y que esta se encuentra en mayor medida entre los años 1925 y 1950, sobretodo cuando se observan los resultados asociados a la variable *Normalizacion* la cual representa la normalización de la cantidad de noticias publicadas de manera semanal por el promedio de noticias publicadas semanalmente en un año, lo anterior se puede relacionar con que en la antigüedad el rol que podía llegar a cumplir el diario sobre la sociedad tenía mayor importancia que en la actualidad, ya que existían una menor cantidad de medios de comunicación que los que se pueden observar hoy en día.

En cuanto a los resultados asociados con el segundo y tercer lag de la variable de los medios, en ellos se puede ver un comportamiento similar al del primer lag, ya que se observa una mayor cantidad de coeficientes significativos en periodos de recesión y asociados la mayoría de estos al periodo que comprende entre 1925 y 1950. Mientras que en periodos de expansión, casi no existen coeficientes significativos asociados al contenido de los medios.

Todo lo anteriormente mencionado permite confirmar el hecho de que el efecto de los medios de comunicación se torna especialmente importante en periodos de recesión, ya que en estos momentos los inversionistas están mucho más preocupados y existe mayor cantidad de información circulante que les impide enfocarse en lo importante, de la misma forma, los gráficos permiten darse cuenta que el efecto de los medios era especialmente importante en la antigüedad. Lo anterior esta estrechamente relacionado con la información disponible durante cada periodo de tiempo, ya que anteriormente existía menor cantidad de medios por los cuales se pudiera acceder a información, esto hacia que medios tales como el New York Times tuvieran mayor importancia de la que pueden tener hoy en día y más aun, como existe una mayor cantidad de medios por los cuales es posible acceder a información, los mercados de cierta forma se completan con mayor facilidad evitando anomalías.

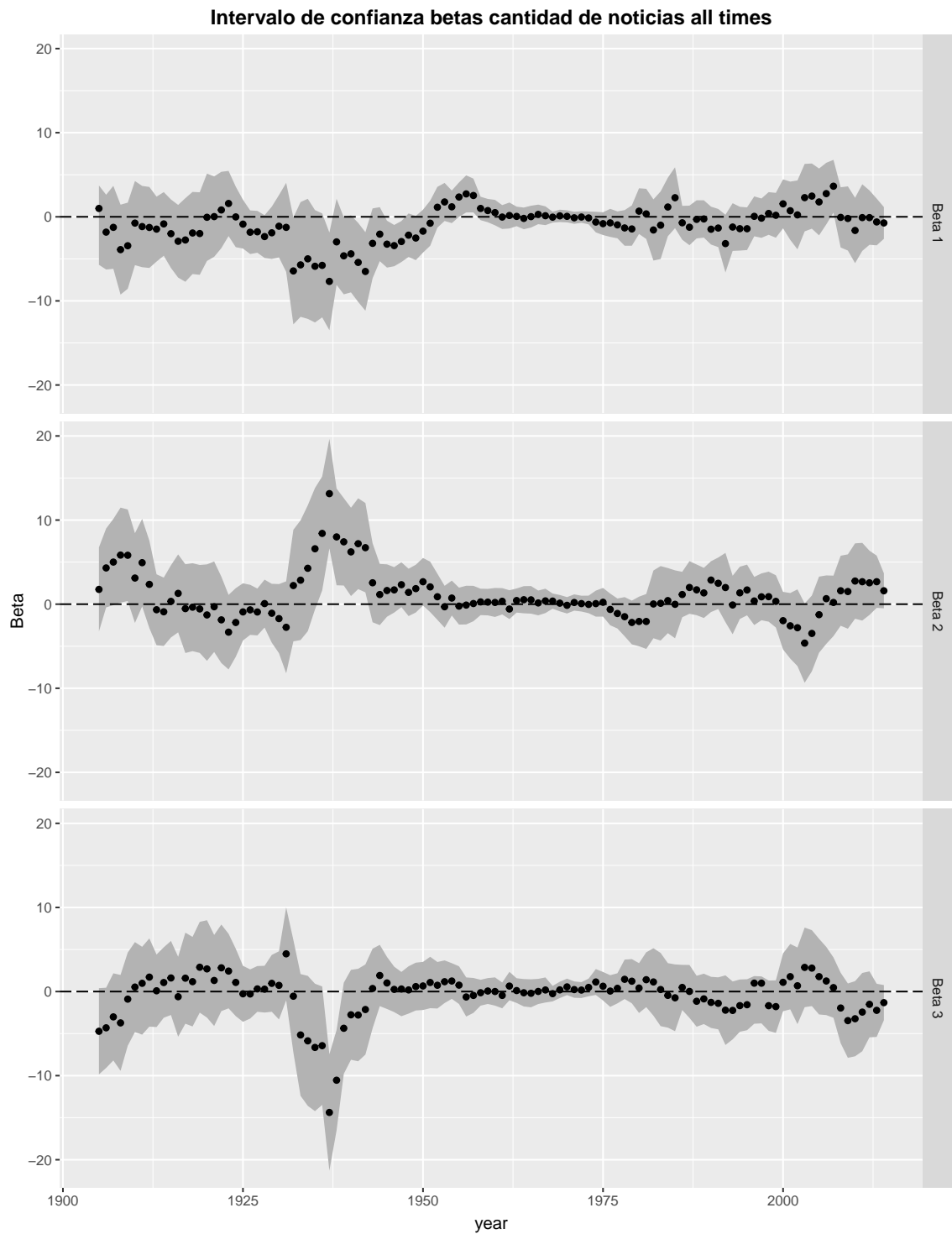


Figura 3.3: Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en todo el Intervalo de Tiempo

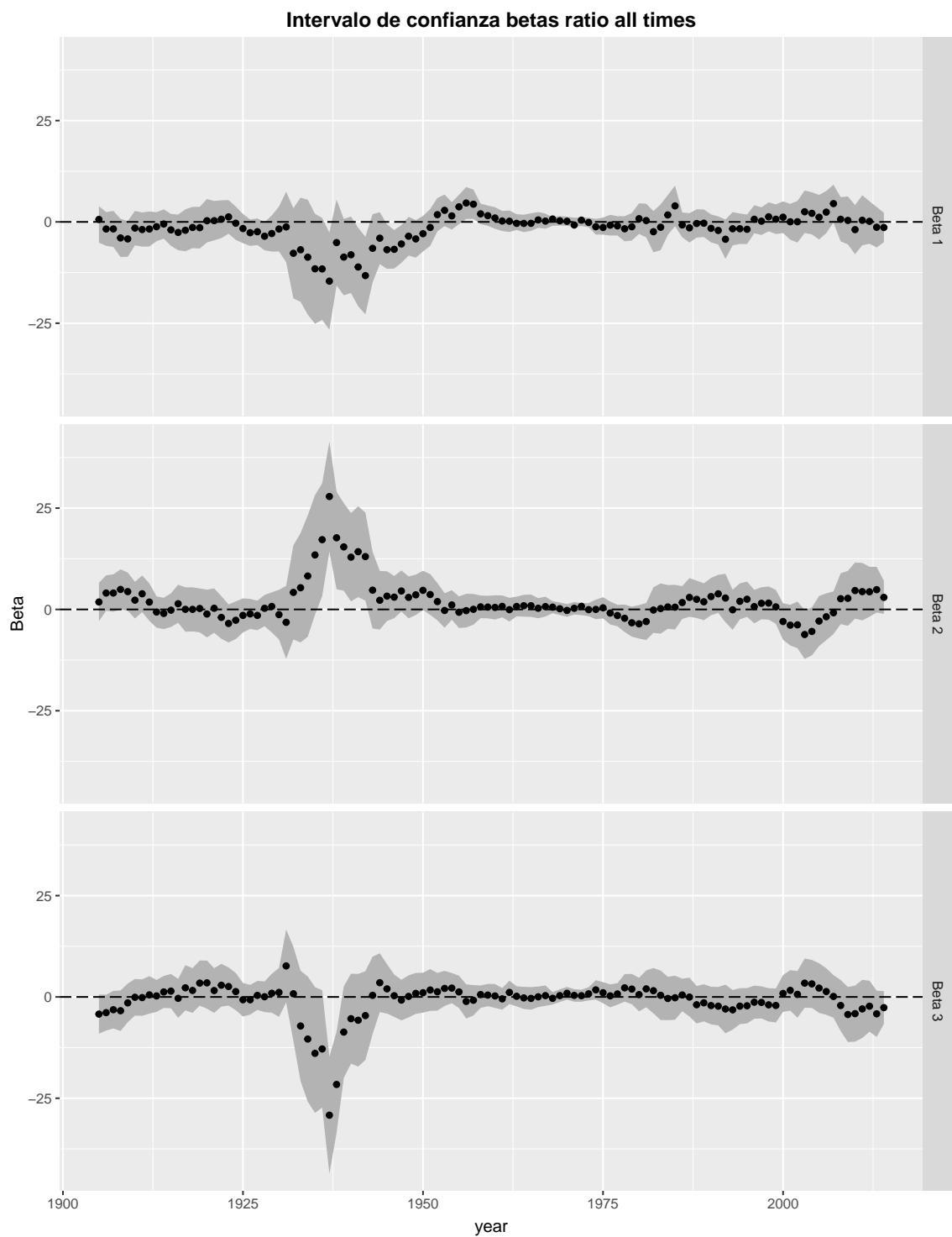


Figura 3.4: Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en todo el Intervalo de Tiempo

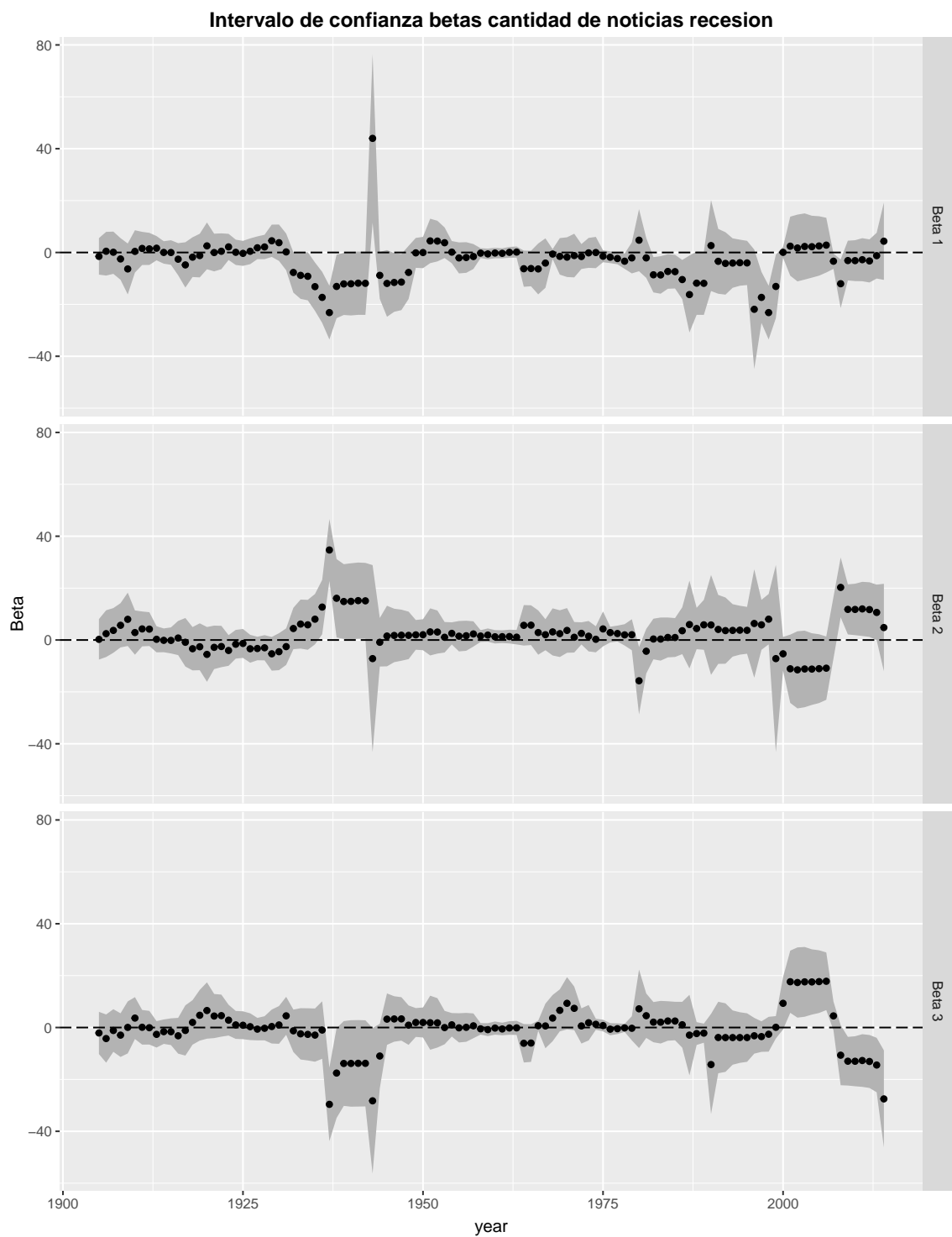


Figura 3.5: Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en Periodo de Recesión

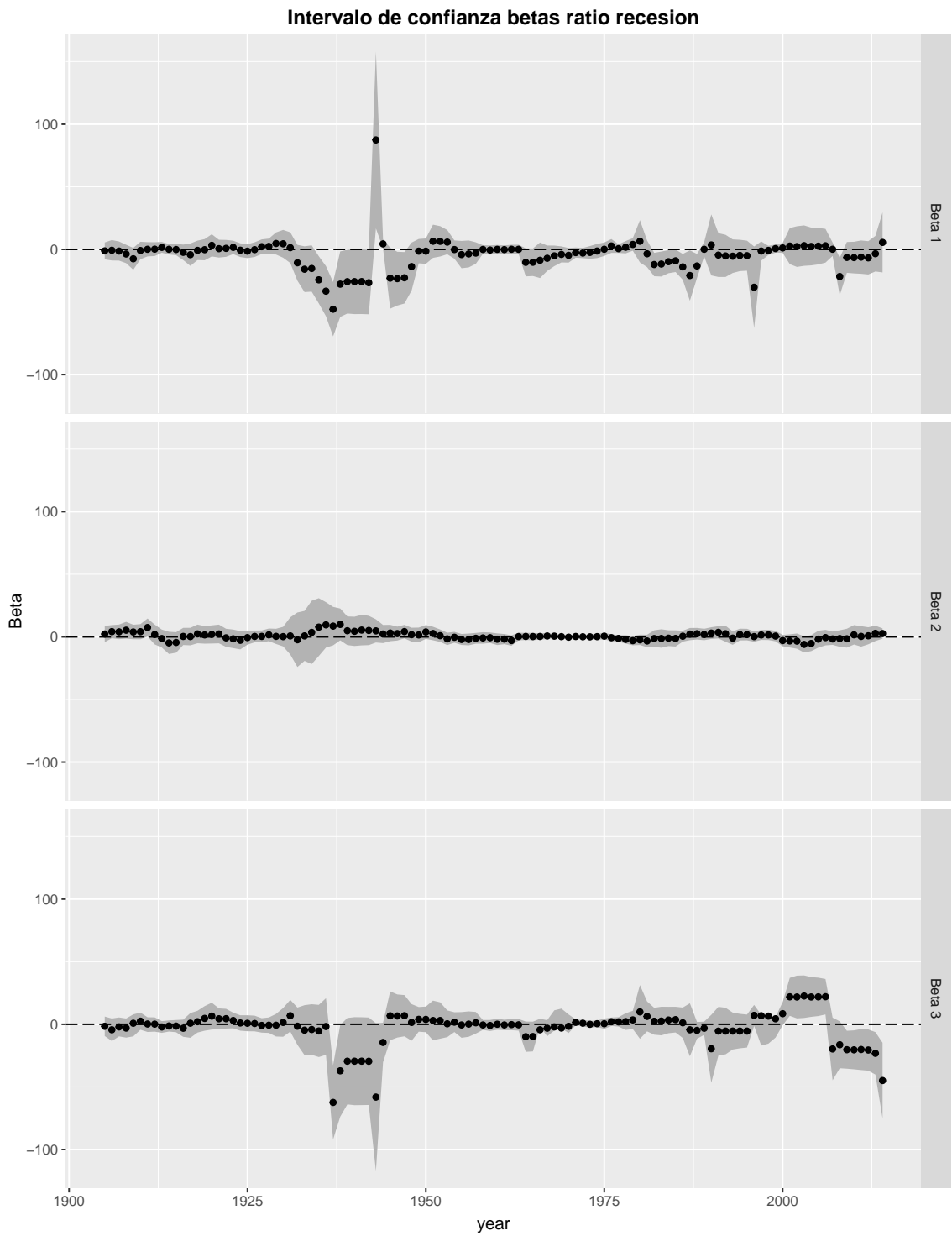


Figura 3.6: Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en Periodo de Recesión



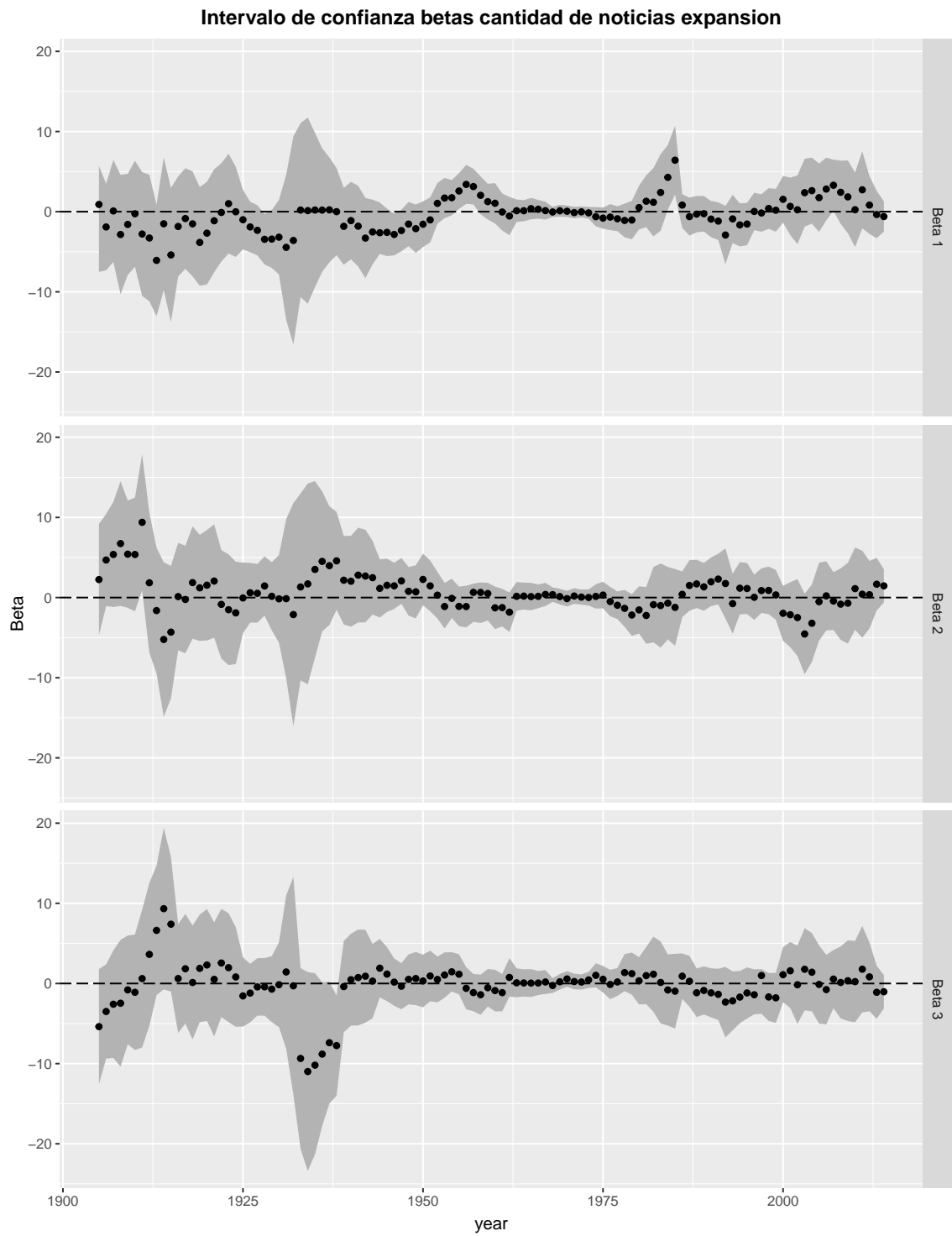


Figura 3.7: Comportamiento Coeficiente Asociados al Contenido de los Medios (cantidad) en Periodo de Expansión

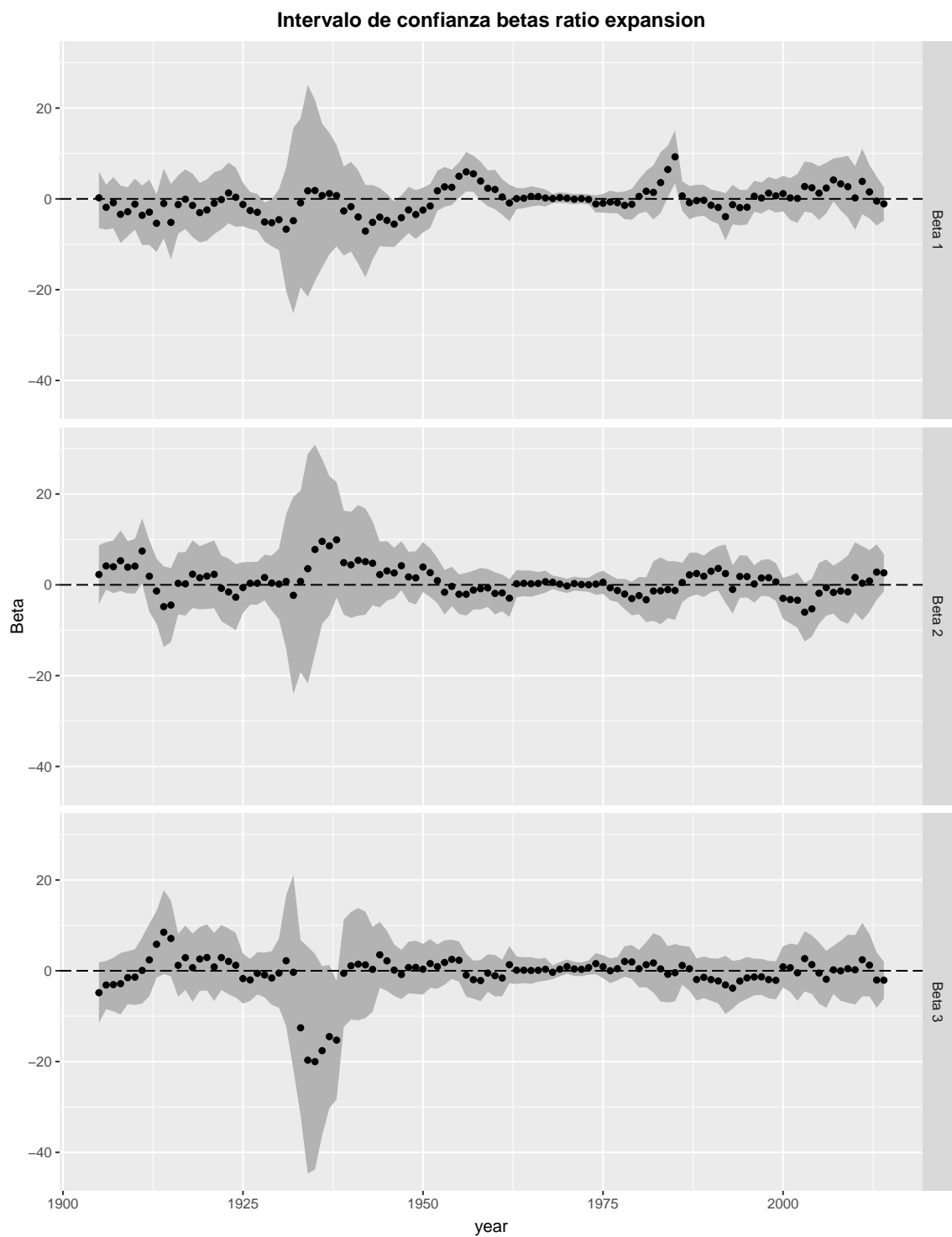


Figura 3.8: Comportamiento Coeficiente Asociados al Contenido de los Medios (ratio) en Período de Expansión

### 3.3.3. Contenido de los Medios, Retornos y Ciclos Económicos

#### Medios y Retornos Semanales

Al realizar un análisis de la segunda parte de la hipótesis, donde se buscaba encontrar alguna relación entre los retornos y la información publicada por el New York Times, es posible ver en la tabla 3.6, que un aumento en el ratio  $pos/(neg + pos)$  predice una disminución de los retornos. Este resultado puede ser explicado mediante los sesgos de percepción a los cuales están expuestos los agentes, ya que ellos se hacen una idea de como esta el mercado o la economía viendo lo publicado por los medios de comunicación aun cuando lo que estén viendo no tenga relación alguna con las decisiones que están tomando. De la misma forma un aumento del ratio  $\frac{pos}{(neg+pos)}$ , que puede ser visto como una medida de estabilidad, les da una idea de mayor seguridad a los agentes. En este sentido la literatura plantea que cuando los inversionistas se sienten más seguros, están dispuestos a pagar un precio mayor por los activos, de la misma forma, tienden a vender a precios muy bajos cuando sienten temor, por lo que en el mediano plazo observan bajos retornos.

Tabla 3.6: Regresión medios y retornos semanales

En esta tabla es posible observar los resultados asociados a la ecuación 3.1, la cual relaciona las variables asociadas a los medios  $\frac{pos}{(neg+pos)}$  con los retornos. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014.

Variable Medios	$pos/(neg + pos)$	
	$\beta$	t-stat
$M_{t-1}$	-7.067**	(-2.204)
$M_{t-2}$	4.850	(1.376)
$M_{t-3}$	-2.359	(-0.664)
$M_{t-4}$	-0.298	(-0.084)
$M_{t-5}$	3.686	(1.151)

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Los coeficientes de la variable de medios retardada entre  $t - 2$  y  $t - 5$  dan un indicio de si el shock sobre los retornos causados por el contenido de los medios en la fecha  $t - 1$  es permanente o temporal. De la tabla es posible darse cuenta que desde el segundo lag no existe significancia en los resultados.

De manera adicional cuando se eliminan las guerras y la crisis del 2008, como se puede ver en la tabla 3.7 los efectos asociados a cada una de las variables no cambian ni en signo ni en

significancia.

Tabla 3.7: Regresión medios y retornos semanales sin considerar guerras ni crisis 2008

En esta tabla es posible observar los resultados asociados a la ecuación 3.1, la cual relaciona las variables asociadas a los medios  $\frac{pos}{(neg+pos)}$  con los retornos. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014, pero se eliminaron los periodos relacionados con las guerras y la crisis del 2008, es decir, entre 1914 y 1918, 1939 y 1945 así como también el año 2008.

Variable Medios	$pos/(neg + pos)$	
	$\beta$	t-stat
$M_{t-1}$	-8.112**	(-2.384)
$M_{t-2}$	7.009*	(1.870)
$M_{t-3}$	-3.054	(-0.808)
$M_{t-4}$	0.574	(0.153)
$M_{t-5}$	2.291	(0.674)

t-statistics in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

### Medios, Ciclos Económicos y Retornos Semanal

En la misma línea de lo anterior, se separó la data en ciclos de expansión y ciclos de recesión según lo publicado por NBER en su página web. El panel A de la tabla 3.8 incluye las estimaciones de los coeficientes  $\beta_1$ , los cuales miden el efecto del contenido de los medios sobre los retornos de las acciones durante periodos de expansión. Como es posible ver en la tabla, en estos periodos no es posible observar significancia.

El panel B presenta las estimaciones de  $\beta_2$ , el cual mide el efecto de las medidas asociadas al contenido de las noticias sobre el rendimiento de las acciones durante las recesiones. El coeficiente asociado al ratio  $\frac{pos}{(neg+pos)}$  es -14.33, con un estadístico t de -2.4. Como se observa en la tabla, si existe significancia durante los periodos de recesión. En cuando a los signos asociados a los efectos encontrados, es posible ver que una mayor porción de palabras positivas sobre el total de palabras o una mayor estabilidad implica una disminución en los retornos. Tal cual se ha comentado anteriormente, un aumento del ratio  $\frac{pos}{(neg+pos)}$  puede ser interpretado por los agentes como un periodo de mayor estabilidad y calma, por lo anterior se espera que los inversionistas reaccionen con mayor relajación y confianza.

Al sacar las guerras y la crisis del año 2008 el efecto permanece, por lo cual se puede decir que no es un efecto que este altamente influenciado por estos episodios, los resultados de dicha regresión se pueden observar en la tabla E.3, disponible en el anexo E.

Tabla 3.8: Regresión Medios, ciclos económicos y retornos

En esta tabla es posible observar los resultados asociados a la ecuación 3.2, la cual relaciona las variables asociadas a los medios  $\frac{pos}{(neg+pos)}$  con los retornos y los ciclos económicos proporcionados por NBER. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014.

	$pos/(neg + pos)$	
	$\beta$	t-stat
<b>A. Expansión (<math>\beta_1</math>)</b>		
$(1 - D_t)xM_{t-1}$	-3.543	(-0.930)
$(1 - D_t)xM_{t-2}$	0.626	(0.149)
$(1 - D_t)xM_{t-3}$	0.06	(0.014)
$(1 - D_t)xM_{t-4}$	-1.347	(-0.321)
$(1 - D_t)xM_{t-5}$	3.106	(0.820)
<b>B. Recesión (<math>\beta_2</math>)</b>		
$D_t x M_{t-1}$	-14.333**	(-2.408)
$D_t x M_{t-2}$	14.467**	(2.250)
$D_t x M_{t-3}$	-7.602	(-1.156)
$D_t x M_{t-4}$	1.001	(0.155)
$D_t x M_{t-5}$	4.526	(0.756)
t-statistics in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Tal cual se ha comentado con anterioridad, este estudio comprende todas las noticias publicadas por el New York Times que contienen alguna de las palabras de interés, de este modo el inversionista esta expuesto a un sin numero de noticias relacionadas con diversos temas que le impiden darse cuenta cual de ellas es la realmente relevante para la decisión que el debe tomar. A modo de ejemplo, cuando se esta frente a un desastre natural, algún ataque terrorista o a cambios de gobierno, hay muchas noticias relacionadas con estos temas, con diversos datos, de esta forma existe mucho ruido en el periódico que impide hacerse una perspectiva clara de la situación. Por lo general los inversionistas responden con una mayor cantidad de emoción y de forma mucho más exagerada durante estos episodios, lo que los lleva a comprar o vender activos sin considerar el verdadero valor de ellos. De forma particular, durante estas situaciones se exagera en gran medida el sesgo de aversión miope a la perdida, donde los inversionistas tienden a pagar más por un activo si se esta optimista y a vender muy barato si se esta frente a una situación pesimista.

Considerando que en periodos de recesión los inversionistas están mucho más expuestos a si-

tuaciones de ansiedad y miedo, es que el efecto que las noticias pueden causar sobre ellos es mucho más importante en estos periodos que cuando se está en un ambiente más controlado. De esta forma, la disminución en los retornos debido a un aumento en la cantidad de palabras positivas sobre el total de palabras, puede ser explicado por la compra de activos a un precio más alto del que se debía, ya que las noticias publicadas por el diario, representadas en este caso por el ratio  $\frac{pos}{(neg+pos)}$ , daban la sensación de estabilidad y la posterior venta de los activos a un precio mucho menor que el original de venta, esto debido a la sensación de pánico a la que están expuestos los agentes frente a las crisis, la cual resalta en mayor medida su aversión miope a las pérdidas.

Los resultados que se han observado durante este capítulo están estrechamente relacionados con information overload o exceso de información y con los sesgos de percepción. En el primer caso debido a que el exceso de noticias publicadas tiene un efecto en la toma de decisiones de los agentes, ya que una vez que se alcanza el máximo de información que son capaces de procesar, la información adicional se transforma en ruido. En el segundo caso se cuenta con todas las noticias publicadas, independientemente si estas tienen o no un contenido ligado a la economía, de esta forma, el común de los inversionistas no es capaz de filtrar el ruido presente en ellas. Ambos casos provocan en el inversionista ansiedad y confusión, llevándolos a invertir de forma incorrecta, ya que se resaltan en mayor medida los sesgos de percepción.

# Capítulo 4

## Conclusiones

Durante el desarrollo de esta investigación, se trabajó con las noticias asociadas al New York Times desde su puesta en circulación el año 1851 hasta el año 2014, que contenían alguna palabra de interés para el trabajo. De la misma forma se utilizaron los datos de los retornos asociados al Dow Jones entre los años 1897 y 2014. Todo esto con el fin de poder explicar de mejor manera el posible efecto que podrían tener los medios de comunicación sobre las decisiones de los agentes en la economía.

### 4.1. Relación Medios y Crisis

Respecto a los efectos encontrados, fue posible hallar evidencia de la existencia de una relación negativa entre la probabilidad de crisis y la cantidad de palabras positivas sobre el total de palabras publicadas por el New York Times. De la misma forma, se observó que existe una relación negativa entre la tasa de interés, usándola como un proxy del coeficiente de aversión al riesgo y el ratio  $pos/(pos + neg)$ . Por otro lado, se pudo ver que existe una relación positiva entre la variable asociada a la estabilidad y el ratio inversión/capital. Finalmente cuando se habla de current account se obtuvo una relación positiva entre el ratio  $pos/(pos + neg)$  y esta.

En base a los efectos encontrados, es posible hablar de Minsky y en particular de un Momento Minsky, el cual es en cierto modo previsible, ya que se sabe que ocurrirá porque tiene su génesis en una inestabilidad endógena, derivada del propio sistema y no de circunstancias externas, es decir, se trata de una fragilidad sistémica que no es posible saber cuando ocurrirá, pero que con certeza se puede decir que ocurrirá.

En cuanto a posibles mejoras que se podrían realizar al modelo estudiado, parece interesante realizar la agrupación de las noticias por temática, es decir, como la base de datos entrega el título de ellas, sería de gran utilidad realizar un análisis semántico de los títulos, el cual permitiría dividirlos según sus temáticas, ya sea política, deporte, economía, etc. Para de esta forma analizar que secciones del diario influyen de mayor medida en la toma de decisiones de los inversionistas y cuáles no. Por otro lado, es importante realizar una estimación directa del

coeficiente de aversión al riesgo de los agentes y como este cambia según el contenido que presentan las noticias publicadas por el diario.

## 4.2. Relación Medios, Ciclos Económicos y Retornos

Si en primer lugar se mira el efecto que tiene la cantidad de noticias publicadas por el diario sobre los retornos, es posible ver que una mayor cantidad de noticias trae consigo una disminución en los retornos. Dicho efecto, es significativo solo durante periodos de recesión. Esta disminución puede ser explicada por medio del efecto infomation overload y de los sesgos de percepción a los cuales están expuestos los inversionistas.

En la relación entre medios y retornos, fue posible encontrar una relación negativa entre el ratio  $pos/(pos + neg)$  y los retornos. De manera adicional, al realizar la división en periodos de recesión y expansión, es posible observar que el efecto de los medios es significativo solo en periodos de recesión, en periodos de expansión, los medio no tienen una mayor influencia.

Es importante notar como los sesgos de percepción se ven reflejados en los resultados y afectan a los inversionista, teniendo consecuencias mucho más graves de las esperadas, ya que como se ha mencionado anteriormente, los inversionistas están muy lejos de tomar las decisiones de manera racional.

De manera adicional si se realiza un análogo con los resultados que muestra Garcia (2013) en sus paper, es posible ver que si bien los resultados son distintos, los datos utilizados también lo son, ya que él utiliza una medida del tono de la noticia, generada por medio de dos columnas financieras y de como estas influyen en los sentimientos del inversionista. Por otro lado para el desarrollo de este trabajo se utilizo la información proporcionada por todo el New York Times, sin importar a la sección del diario a la cual esta perteneciera, como un proxy de la estabilidad que se podría estar observando en el país. Tomando en consideración lo anterior, no parece extraña la diferencia en los resultados y no es posible descartar la veracidad de ninguno de los estudios.

Finalmente, es posible concluir que el contenido de las noticias ayuda a predecir las crisis, mientras que si se observa los retornos, estos también pueden ser predichos tanto por la cantidad de noticias publicadas por el New York Times como por el contenido de estas, pero solo durante las recesiones. La interpretación más natural ligada a este resultado tiene relación con el *Disposition effect* o *efecto de disposición*, el que dice que la mayoría de las personas temen en mayor medida a los efectos de las posibles pérdidas que a la anticipación de las posibles ganancias. Es por esto que ante una situación de crisis se experimenta mucho más temor, lo que resalta los sesgos de percepción, tomando muchas veces decisiones irracionales.



# Bibliografía

- Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128.
- Ben-David, I. and Doukas, J. (2006). Overconfidence, trading volume, and the disposition effect: Evidence from the trades of institutional investors. Technical report, Working Paper, University of Chicago and Old Dominion University.
- Bhattacharya, S., Goodhart, C. A., Tsomocos, D. P., and Vardoulakis, A. P. (2015). A reconsideration of minsky's financial instability hypothesis. *Journal of Money, Credit and Banking*, 47(5):931–973.
- Bloomfield, R. (2010). Behavioural finance. In *Behavioural and Experimental Economics*, pages 32–41. Springer.
- Bolt, J. and Zanden, J. L. (2014). The maddison project: collaborative research on historical national accounts. *The Economic History Review*, 67(3):627–651.
- Bosman, R., Kräussl, R., and Mirgorodskaya, E. (2014). The 'tone effect' of news on investor beliefs: An experimental approach. Available at SSRN 2528520.
- Burton, G. (2010). *Media and society: Critical perspectives*. McGraw-Hill Education (UK).
- Canto, J. V. and Mejía, C. B. (2014). Los indicadores financieros y el valor económico agregado (eva) en la creación de valor. *Industrial Data*, 10(1):042–047.
- Chan, W. S. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, 70(2):223–260.
- Chen, G., Kim, K. A., Nofsinger, J. R., and Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*, 20(4):425–451.
- Chevapatrakul, T. and Tee, K.-H. (2014). The effects of news events on market contagion: evidence from the 2007–2009 financial crisis. *Research in International Business and Finance*, 32:83–105.
- Cochrane, J. H. (1991). Production-based asset pricing and the link between stock returns and

- economic fluctuations. *The Journal of Finance*, 46(1):209–237.
- Cornella, A. (1999). Cómo sobrevivir a la infoxicación. In *Trascripción de la conferencia del acto de entrega de títulos de los programas de Formación de Posgrado del año académico*, volume 2000.
- Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance*, 53(6):1839–1885.
- Danielsson, J., Valenzuela, M., and Zer, I. (2016). Learning from history: volatility and financial crises.
- Ederington, L. H. and Lee, J. H. (1993). How markets process information: News releases and volatility. *The Journal of Finance*, 48(4):1161–1191.
- Edmans, A., Garcia, D., and Norli, Ø. (2007). Sports sentiment and stock returns. *The Journal of Finance*, 62(4):1967–1998.
- Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, 49(3):283–306.
- Fang, L. and Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5):2023–2052.
- Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, 61(4):2017–2046.
- Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.
- Garcia, D. (2014). The kinks of financial journalism. *Available at SSRN 2517591*.
- Hagenau, M., Liebmann, M., and Neumann, D. (2013). Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision Support Systems*, 55(3):685–697.
- Hens, T. and Meier, A. (2002). Behavioral finance: The psychology of investing.
- Heston, S. L. and Sinha, N. R. (2014). News versus sentiment: Comparing textual processing approaches for predicting stock returns. *Robert H. Smith School Research Paper*.
- Hur, J., Pritamani, M., and Sharma, V. (2010). Momentum and the disposition effect: the role of individual investors. *Financial Management*, 39(3):1155–1176.
- Kemp, A. (1960). The constitution of liberty. *Journal of the American Medical Association*, 172(8):825–826.
- Keynes, J. M. (1937). The general theory of employment. *The quarterly journal of economics*, 51(2):209–223.

- Klibanoff, P., Lamont, O., and Wizman, T. A. (1998). Investor reaction to salient news in closed-end country funds. *The Journal of Finance*, 53(2):673–699.
- Kurtz, H. (2000). *The fortune tellers: Inside Wall Street's game of money, media, and manipulation*. Simon and Schuster.
- Li, X., Xie, H., Chen, L., Wang, J., and Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, 69:14–23.
- Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.
- Lucas Jr, R. E. (1978). Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, pages 1429–1445.
- McCarthy, K. J. and Dolfsma, W. (2014). Neutral media? evidence of media bias and its economic impact. *Review of Social Economy*, 72(1):42–54.
- Minsky, H. P. (1992). The financial instability hypothesis.
- Mitchell, M. L. and Mulherin, J. H. (1994). The impact of public information on the stock market. *The Journal of Finance*, 49(3):923–950.
- Mun, M. and Brooks, R. (2012). The roles of news and volatility in stock market correlations during the global financial crisis. *Emerging markets review*, 13(1):1–7.
- Obstfeld, M., Shambaugh, J. C., and Taylor, A. M. (2005). The trilemma in history: tradeoffs among exchange rates, monetary policies, and capital mobility. *Review of Economics and Statistics*, 87(3):423–438.
- of the Census, U. S. B. (1975). *Historical statistics of the United States, colonial times to 1970*. Number 93. US Department of Commerce, Bureau of the Census.
- Ortony, A., Clore, G. L., and Collins, A. (1990). *The cognitive structure of emotions*. Cambridge university press.
- Perse, E. M. and Lambe, J. (2016). *Media effects and society*. Routledge.
- Pliska, S. (1997). *Introduction to mathematical finance*. Blackwell publishers Oxford.
- Reinhart, C. M. and Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly*. princeton university press.
- Reinhart, C. M. and Rogoff, K. S. (2010). Growth in a time of debt (digest summary). *American Economic Review*, 100(2):573–578.
- Ritter, J. R. (2003). Behavioral finance. *Pacific-Basin finance journal*, 11(4):429–437.
- Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: monetary policy, leverage

- cycles, and financial crises, 1870–2008. *The American Economic Review*, 102(2):1029–1061.
- Shiller, R. J. (2015). *Irrational exuberance*. Princeton university press.
- Smith, C. A. and Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology*, 48(4):813.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.
- Thaler, R. H. (2005). *Advances in behavioral finance*, volume 2. Princeton University Press.
- Tiedens, L. Z. and Linton, S. (2001). Judgment under emotional certainty and uncertainty: the effects of specific emotions on information processing. *Journal of personality and social psychology*, 81(6):973.
- Toffler, A. (1990). *Future shock*. Bantam.
- Tversky, A. and Kahneman, D. (1985). The framing of decisions and the psychology of choice. In *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, pages 107–129. Springer.

# Estadística de Noticias Anual

Tabla 1: Estadística descriptiva anual expansiones y recesiones

En esta tabla es posible observar la estadística descriptiva asociada al comportamiento del contenido de los medios durante los distintos ciclos económicos. En el caso de las medidas de cantidad *pos* es el número de palabras positivas pertenecientes al diccionario Loughran and McDonald (2011) que aparecen en un año, contando solo una aparición de cada palabra por noticia, *neg* se define de la misma forma pero considerando las palabras negativas proporcionadas por ellos mismos. Por otro lado la medida de nivel se define como:  $\frac{pos}{(neg+pos)}$ , representando de esta manera un proxy de la estabilidad. De manera adicional, se dividió la data en periodos de expansión y de recesión propuestos por NBER.

Medía measure	Mean	Median	Min	Max	25 %-quant	75 %-quant	Stand. dev
<b>All dates</b>							
<b>Medidas de cantidad</b>							
pos	235500	253700	13770	529600	151500	297200	107249.6
neg	537100	589200	29760	1176000	319500	1176000	248164
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.305	0.309	0.26	0.346	0.289	0.346	0.018
<b>Expansiones</b>							
<b>Medidas de cantidad</b>							
pos	234700	255500	13770	482900	167000	298600	9756.6
neg	537500	618600	29760	1075000	354300	714400	229932.16
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.305	0.306	0.263	0.346	0.289	0.321	0.018
<b>Recesiones</b>							
<b>Medidas de cantidad</b>							
pos	220100	206000	46860	502900	122600	283800	110940.9
neg	499900	435500	121800	1166000	292600	660300	254385.2
<b>Medidas de nivel</b>							
$\frac{pos}{(neg+pos)}$	0.304	0.309	0.26	0.337	0.288	0.319	0.02

# Apéndice A

## Estadística de Noticias Diarias

Tabla A.1: Estadística sobre numero de noticias que contienen alguna de las palabras de interés publicadas diariamente

Media measure	Mean	Median	Min	Max	25 %-quant	75 %-quant	Stand. dev	obs
<b>All dates</b>	180.7	168	1	976	102	226	112.79	10665903
<b>Expansiones</b>	197	184	7	976	126	235	113.95	7644997
<b>Recesiones</b>	158.5	132	14	647	88	212	101.47	2947514

# Apéndice B

## Relación Medios y Currency Crisis

Tabla B.1: Relación simple Medios y Crisis considerando hasta 10 lag

En esta tabla es posible observar los resultados asociados a la ecuación 2.1, la cual relaciona el ratio  $\frac{pos}{(neg+pos)}$  y la crisis de moneda. En esta tabla se utilizaron datos que contemplan entre 1852 y 2010.

Var. Dep: $C_t^{currency}$	1	2	3	4	5	6	7	8	9	10
$\bar{C}_{t-1 \rightarrow t-L}$	0.886** (0.444)	0.529 (0.563)	0.346 (0.673)	-0.023 (0.859)	0.106 (0.919)	1.219 (1.095)	1.372 (1.153)	1.301 (1.136)	1.437 (1.215)	1.556 (1.322)
$\frac{pos}{(neg+pos)}_{t-1 \rightarrow t-L}$	22.537* (12.853)	30.415** (13.656)	36.235*** (13.800)	38.72*** (14.573)	35.391** (13.987)	28.235** (13.753)	24.53* (13.618)	23.649* (12.910)	22.111* (12.788)	20.738 (13.493)
Num. obs.	159	158	157	156	155	154	153	152	151	150
Pseudo $R^2$	0.062	0.057	0.066	0.064	0.055	0.063	0.056	0.05	0.049	0.048
Efectos Marginales										
$\frac{pos}{(neg+pos)}_{t-1 \rightarrow t-L}$	3.2	4.33	5.1	5.49	5.11	4.09	3.61	3.52	3.31	3.13

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Tabla B.2: Relación Medios y Crisis considerando hasta 10 lag

En esta tabla es posible observar los resultados asociados a la ecuación 2.1, la cual relaciona el ratio  $\frac{pos}{(neg+pos)}$  y la crisis de moneda. De manera adicional, se consideran variables independientes de interés como es el caso de  $\overline{volP}$  que representa la volatilidad de la variable asociada a la estabilidad,  $\overline{inf}$  que es la tasa de inflación anual,  $\overline{loggdp}$  es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per cápita. En esta tabla se utilizaron datos que contemplan entre 1852 y 2010.

Var. Dep: $C_t^{currency}$	1	2	3	4	5	6	7	8	9	10
$\overline{C}_{t-1 \rightarrow t-L}$	0.864*	0.413	0.11	-0.386	-0.372	0.711	0.73	0.811	0.882	1.052
	(0.464)	(0.615)	(0.741)	(0.933)	(0.972)	(1.155)	(1.237)	(1.267)	(1.359)	(1.431)
$\frac{pos}{(neg+pos)}_{t-1 \rightarrow t-L}$	21.392	30.851**	37.634**	40.882***	38.36**	30.804**	24.164*	23.759*	20.617	19.246
	(13.357)	(14.394)	(15.299)	(15.694)	(14.960)	(14.713)	(14.167)	(13.723)	(13.466)	(14.227)
$\overline{volM}_{t-1 \rightarrow t-L}$	10.917	14.529	17.372	39.223	71.53	79.25	46.809	38.805	32.445	44.79
	(68.531)	(79.289)	(87.173)	(91.583)	(98.606)	(109.424)	(110.157)	(115.834)	(125.209)	(136.716)
$\overline{inf}_{t-1 \rightarrow t-L}$	0.022	0.001	-0.012	-0.029	-0.05	-0.055	-0.069	-0.104	-0.121	-0.113
	(0.038)	(0.042)	(0.047)	(0.054)	(0.064)	(0.067)	(0.070)	(0.085)	(0.092)	(0.092)
$\overline{loggdp}_{t-1 \rightarrow t-L}$	-0.034	0.083	0.203	0.313	0.396	0.355	0.307	0.437	0.458	0.446
	(0.458)	(0.497)	(0.522)	(0.524)	(0.516)	(0.495)	(0.485)	(0.497)	(0.503)	(0.507)
$\overline{deb/gdp}_{t-1 \rightarrow t-L}$	-0.005	-0.006	-0.009	-0.013	-0.015	-0.014	-0.016	-0.016	-0.0181	-0.018
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)
Num. obs.	155	154	153	152	151	150	149	148	147	146
Pseudo $R^2$	0.069	0.063	0.074	0.077	0.076	0.079	0.073	0.078	0.08	0.077
Efectos Marginales										
$\frac{pos}{(neg+pos)}_{t-1 \rightarrow t-L}$	3.085	4.469	5.389	5.861	5.558	4.490	3.573	3.418	2.973	2.801

t-statistics in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1



# Apéndice C

## Modelo de Lucas

Si se toman dividendos que se mueven de acuerdo a un paseo aleatorio con volatilidad constante  $\sigma$  y tendencia  $\mu_{t+1}$  que a su vez se mueve con volatilidad constante  $\alpha$  y tendencia  $\omega_{t+1}$ , es decir,

$$\ln\left(\frac{D_{t+1}}{D_t}\right) = \mu_{t+1} + \sigma\varepsilon_{t+1} \quad (\text{C.1})$$

Donde el termino  $\varepsilon_{t+1}$  es homocedastico y no esta relacionado con la serie de tiempo, sin embargo,  $\mu_{t+1}$  presenta cambios en el tiempo, ya que esta relacionado con  $g_{t+1}$  por  $g_{t+1} = \exp(\mu_{t+1} + \sigma^2/2)$ . De manera adicional se asume un mercado perfecto, sin fricciones y completo. De la misma forma, el agente representativo tiene una función de utilidad isoelastica, es decir, una función de utilidad con aversión relativa al riesgo constante (CRRA);

$$u(c_t) = \begin{cases} \frac{c_t^{1-\alpha}-1}{1-\alpha} & \alpha \geq 0 \\ \ln(c_t) & \alpha = 1 \end{cases} \quad (\text{C.2})$$

Donde  $C_t$  es el consumo y  $\alpha$  el coeficiente de aversión al riesgo relativo. Se asumirá que los dividendos son la única fuente de ingresos para el agente representativo y como es típico en el modelo de Lucas los dividendos son consumidos cuando se reciben, es decir,  $C_{t+k} = D_{t+k}$ . Con lo anterior, el agente maximiza su utilidad teniendo en consideración su restricción presupuestaria, lo que se expresa como:

$$\begin{aligned} & \underset{D_{t+k}, W_{t+k}^S, W_{t+k}^B}{\text{máx}} && \sum_{k=0}^{\infty} \beta^k u(D_{t+k}) && (\text{C.3}) \\ \text{s.a} &&& c_{t+k} + w_{t+k}^S S_{t+k} + w_{t+k}^B B_{t+k} = w_{t+k-1}^S (S_{t+k} + D_{t+k-1}) + w_{t+k-1}^B \end{aligned}$$

Donde  $\beta = 1/(1 + \rho)$  con  $\rho$  la tasa de impaciencia,  $w_{t+k}^S$  y  $w_{t+k}^B$  las porciones de acciones y bonos en el portafolio respectivamente. Utilizando métodos de programación dinámica, es

posible obtener las condiciones de primer orden y con ellas las ecuaciones de Euler:

$$S_t = E_t[\beta(\frac{D_{t+1}}{D_t})^{-\alpha}(S_{t+1} + D_{t+1})] \quad (\text{C.4})$$

$$B_t = E_t[\beta(\frac{D_{t+1}}{D_t})^{-\alpha}] \quad (\text{C.5})$$

Luego si se asume total conocimiento de los parámetros que aparecen en el proceso de dividendos, es decir,  $\mu_{t+1}$  y  $\sigma$  conocidos y la tasa de crecimiento  $g_{t+1}$  conocida por el agente representativo se tendrá que la ecuación C.5 quedan como sigue:

$$B_t = \frac{(1 + g_t)^{-\alpha}}{(1 + \rho)} \quad (\text{C.6})$$

De manera adicional se sabe que el precio de un bono cero cupón esta dado por:  $B_t = 1/(1+r_t)$  con  $r_t$  la tasa de interés. Con esto se tiene que:

$$B_t = \frac{(1 + g_t)^{-\alpha}}{(1 + \rho)} = \frac{1}{(1 + r_t)} \quad (\text{C.7})$$

Encontrando de esta forma una relación para la tasa de interés  $r_t$  y  $\alpha$  el coeficiente de aversión al riesgo, la cual se muestra en la ecuación C.8 y en la cual se puede ver que se relacionan de manera directa, es decir, a mayor coeficiente a aversión al riesgo mayor tasa de interés.

$$(1 + r_t) = (1 + \rho)(1 + g_t)^\alpha \quad (\text{C.8})$$

# Apéndice D

## Behavioral finance

El behavioral finance aplica la investigación científica en las tendencias cognitivas y emocionales humanas y sociales, para una mejor comprensión de la toma de decisiones económicas. Los campos de estudio están principalmente relacionados con la racionalidad o con la ausencia de esta, por parte de los agentes económicos. Este campo propone teorías basadas en la psicología para explicar las anomalías del mercado de valores, tales como subidas o caídas graves en el precio de las acciones. Dentro de Behavioral finance, se supone que la estructura de la información y las características de los participantes del mercado influyen en las decisiones de inversión sistemática de los individuos, así como los resultados del mercado.

Los sesgos más importantes relacionados con el Behavioral finance, según Barberis and Thaler (2003), se detallan continuación:

### D.1. Formación de creencias

- **Sesgo de representatividad:** Después de un breve período de retornos positivos en los mercados financieros, se tiende a pensar que el mundo ha cambiado para volverse mejor. Las personas tienden a pensar en los esquemas y estereotipos que aprendieron en el pasado. Llegan a un resultado demasiado rápido y en base a información imprecisa.
  - **Efecto sobre los retornos:** Negativo, debido a que se da mucha importancia a la información más nueva, existe una rotación de cartera excesiva y se toman las decisiones basadas en clasificaciones superficiales.
- **Sesgo de confirmación:** Se refiere al fenómeno de la búsqueda de información selectiva para respaldar las opiniones propias o interpretar los hechos de manera que se adapten a la visión personal del mundo. Los inversores buscan una confirmación de sus suposiciones. Evitan opiniones e informes críticos, leyendo únicamente aquellos artículos que ponen hincapié en su punto de vista.

- **Efecto sobre los retornos:** Negativo, ya que puede llevar a los inversores a tener un exceso de confianza, ignorando pruebas de que sus estrategias les harán perder dinero.
- **Sesgo de anclaje:** Al tomar decisiones, los inversores tienden a basar su decisión en el precio al cual se compró la posición original o la última posición de una acción. Este precio de compra se transforma en el ancla y provoca decisiones irracionales. A diferencia del costo de adquisición, el nuevo precio parece barato para el inversor. El anclaje influye en las decisiones individuales en base al hecho de que los inversores no se dan cuenta de cómo se presenta la información. Cuando se trata de tomar decisiones, la gente parece estar influenciada por datos aleatorios, incluso si saben que los datos no tienen ningún valor informativo o que son escandalosamente altos o bajos. Existe una tendencia a confiar de manera excesiva en un dato específico a la hora de tomar la decisión.
  - **Efecto sobre los retornos:** Negativo, ya que tienden a mantener las inversiones que han perdido valor debido a que han anclado su estimación del valor del mercado para el precio original en lugar de a los fundamentos de la inversión.
- **Sesgo de Disponibilidad:** Este sesgo afirma que las cosas como productos, empresas y emisores que se presentan con mayor frecuencia en los medios de comunicación serán recordadas más rápidamente por los inversores cuando buscan un instrumento de inversión adecuado. La información mala o apenas accesible no es considerada (inconscientemente).
  - **Efecto sobre los retornos:** Negativo, ya que estarían tomando la decisión basados en la frecuencia de publicación, sin importar el motivo por el que aparecen.
- **Sesgo retrospectivo:** Es un claro reflejo de la afirmación : “yo sabía que esto iba a pasar”, es decir, existe una explicación para todo después de que el hecho ha ocurrido. Lo anterior impide a las personas aprender de los errores.
  - **Efecto sobre los retornos:** Negativa, ya que se sobrestima el ratio de acierto de las predicciones y se es excesivamente críticos con las decisiones tomadas por otros.

## D.2. Procesamiento de la información

- **Contabilidad mental:** Consiste en hacer distinciones en la cabeza que no existen financieramente. A menudo, las pérdidas incurridas son consideradas de forma separada de las pérdidas en papel. Esto significa que la gente vende acciones de su cartera demasiado pronto cuando obtienen una ganancia y demasiado tarde cuando sufren una pérdida. Se tiende a pasar por alto la interacción entre los activos.
  - **Efecto sobre los retornos:** Negativo, el problema de la separación es que la riqueza

neta no cambiaría si el inversor mantuviera una cartera, pero al separarla en dos, el inversor podría sentirse tentado a reducir el peso de los ganadores en la cuenta especulativa para reforzar la cuenta segura.

- **Narrow framing:** Consiste en la propensión de algunos inversionistas a seleccionar los activos de forma individual, pasando por alto el impacto sobre el portafolio en conjunto.
  - **Efecto sobre los retornos:** Negativa, al no pensar en la totalidad de la cartera se podría estar omitiendo la correlación entre activos y tomar decisiones apresuradas.
- **Sesgo de encuadre:** Las decisiones se basan en gran medida en cómo los hechos son planteados en términos estadísticos. Por ejemplo, no se considera que “cuatro de cada diez ganaron” y “seis de cada diez perdieron” tienen el mismo significado. Las afirmaciones anteriormente hechas dicen lo mismo, pero la mayoría de las personas no se dan cuenta de ello.
  - **Efecto sobre los retornos:** Negativo, ya que se tiende a construir carteras con demasiada aversión al riesgo o, por el contrario, carteras que proporcionan más riesgo del deseado.
- **Information overload (infoxicación):** Sobresaturación de información, una vez que se supera la capacidad de procesamiento de información, la información adicional se convierte en ruido y disminuye la calidad de la decisión, ya que los inversionistas necesitan más tiempo para realizar sus predicciones, pero estas por lo general son de menor calidad.
  - **Efecto sobre los retornos:** Negativo, ya que lleva a los agentes a invertir de forma incorrecta, siendo excesivamente conservadores, más arriesgados o no diversifican de la manera correcta. Este sesgo genera en los agentes ansiedad, confusión e impotencia, lo que hace que muchos otros sesgos aparezcan junto a él.

### D.3. Sesgos emocionales

- **Aversión miope a las pérdidas:** La mayoría de los inversores es más temeroso de las pérdidas de lo feliz que ponen por las utilidades. Tener una visión a largo plazo sería mejor. Deben verificar la evolución de las acciones con una menor frecuencia.
  - **Efecto sobre los retornos:** Negativo, ya que se tienden a vender las acciones que han incrementado de valor y mantener aquellas que lo han perdido, con la esperanza de que podrían remontar. A la larga las que han perdido valor pueden seguir haciéndolo.
- **Exceso de confianza:** La mayoría de las personas sobrestima las habilidades propias y tiende a pensar que se encuentra por encima de la media. En particular, la mayoría de los

expertos también se sobrestima a ellos mismos, con frecuencia en un mayor grado que el común de las personas. El exceso de confianza se observa a menudo cuando los mercados están al alza.

- **Efecto sobre los retornos:** Negativa, las ganancias del mercado aumentan el exceso de confianza de los inversionistas y por lo tanto se comportan de forma más agresiva en los siguientes periodos y este exceso de confianza contribuye a la excesiva volatilidad observada. Asimismo, al existir exceso de confianza, los inversionistas subestiman el riesgo e invierten en activos más riesgosos.
- **Sesgo nacional:** Las estadísticas muestran que la mayoría de los inversores tienden a comprar acciones de las compañías de su país de origen. Estas acciones parecen más dignas de confianza y los inversores crecieron con los nombres de éstas empresas. Además es importante saber que estas se mencionan con más frecuencia en los medios locales de comunicación.
  - **Efecto sobre los retornos:** Negativo, que el inversionista no es capaz de diversificar de manera óptima su cartera.
- **Efecto de disposición:** Las ganancias se realizan muy temprano y, como resultado, las pérdidas muy tarde. Es lo que convierte las utilidades de papel en ganancias reales que generan una mayor felicidad, mientras que se evita convertir una pérdida de papel en una pérdida real. Una posible explicación es la contabilidad mental explicada anteriormente.
  - **Efecto sobre los retornos:** Negativo, la rentabilidad de los valores que venden es superior a la de aquellos que mantienen en su cartera, actuando de esta forma, obtienen una rentabilidad inferior a la que obtendría si, efectivamente, vendieran antes valores que están en pérdidas y más tarde los que están en ganancias.
- **Necesidad de alcanzar el equilibrio:** Una vez que se ha perdido dinero se toma un mayor riesgo para recuperarlo. La necesidad de alcanzar el equilibrio puede hacer que se ponga todo en una sola canasta y probablemente se llegue a perder aún más dinero.
  - **Efecto sobre los retornos:** Negativo, el no asumir las pérdidas y tomar más riesgo para recuperar, puede llevar a pérdidas mucho mayores.
- **Falacia del jugador:** En este punto, las probabilidades efectivas son ampliamente subestimadas o sobre estimadas. Esto puede significar que, en base al supuesto (falso) de que los precios están a punto de caer, vendemos demasiado pronto y viceversa (asumiendo que los precios se recuperarán pronto, aunque (todavía) no lo van a hacer).
  - **Efecto sobre los retornos:** Negativo, se piensa erróneamente que los sucesos pasados afectan a los futuros en lo relativo a actividades aleatorias.

- **Evitar el arrepentimiento:** Si se invirtió en una acción “confiable” y esta no se comporto como se esperaba, se trata de mala suerte. Sin embargo, si se invierte en un producto más bien de nicho y este no se comporta correctamente, se tiende a lamentar esto más que el fracaso de las acciones “confiables”. Esto es porque muchas otras personas han cometido el mismo error y, por lo tanto, la decisión de comprar no parece tan errada.
- **Efecto sobre los retornos:** Negativo, ya que se tiende a ser excesivamente conservador con las inversiones y, por ello, no alcanzar el retorno deseado.

# Apéndice E

## Relación Medios, Retornos y Ciclos Económicos

### E.1. Cantidad de Noticias, Retornos y Ciclos Económicos

Tabla E.1: Regresión cantidad de noticias y retornos semanales sin considerar periodos de guerra ni la crisis del año 2008

En esta tabla es posible observar los resultados asociados a la ecuación 3.1, la cual en este caso relaciona las variables de cantidad de noticias *Cantidad* y *Normalizacion*, donde *Cantidad* es la cantidad de noticias publicadas de manera semanal por el New York Times y *Normalizacion* es una normalizacion *Cantidad* por el promedio de noticias publicadas de manera semanal durante un año, con los retornos. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014 eliminando los episodios de guerra, es decir, entre 1914 y 1918 y entre 1939 y 1945, de manera adicional se elimino la crisis del año 2008.

Variable Medios	<i>Cantidad</i>		<i>Normalizacion</i>	
	$\beta$	t-stat	$\beta$	t-stat
$M_{t-1}$	-0.677**	(-2.134)	-0.948**	(-2.056)
$M_{t-2}$	1.137***	(2.769)	1.389***	(2.444)
$M_{t-3}$	-0.703	(-1.639)	-0.917	(-1.58)
$M_{t-4}$	0.03	(0.068)	0.032	(0.055)
$M_{t-5}$	0.204	(0.623)	0.015	(0.032)

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1



Tabla E.2: Regresión cantidad de noticias, ciclos económicos y retornos sin considerar guerras ni la crisis del año 2008

En esta tabla es posible observar los resultados asociados a la ecuación 3.2, la cual en este caso relaciona las variables de cantidad de noticias *Cantidad* y *Normalizacion*, donde *Cantidad* es la cantidad de noticias publicadas de manera semanal por el New York Times y *Normalizacion* es una normalizacion *Cantidad* por el promedio de noticias publicadas de manera semanal durante un año, con los retornos y los ciclos económicos proporcionados pro NBER. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014, eliminando los episodios de guerra, es decir, entre 1914 y 1918 y entre 1939 y 1945, de manera adicional se elimino la crisis del año 2008.

	<i>Cantidad</i>		<i>Normalizacion</i>	
	$\beta$	t-stat	$\beta$	t-stat
<b>A. Expansión (<math>\beta_1</math>)</b>				
$(1 - D_t)xM_{t-1}$	-0.076	(-0.206)	-0.182	(-0.315)
$(1 - D_t)xM_{t-2}$	0.396	(0.865)	0.515	(0.737)
$(1 - D_t)xM_{t-3}$	-0.48	(-1.050)	-0.759	(-1.086)
$(1 - D_t)xM_{t-4}$	0.311	(0.678)	0.504	(0.722)
$(1 - D_t)xM_{t-5}$	-0.096	(-0.258)	-0.457	(-0.79)
<b>B. Recesión (<math>\beta_2</math>)</b>				
$D_t x M_{t-1}$	-3.081***	(-3.958)	-3.197***	(-2.899)
$D_t x M_{t-2}$	3.518***	(4.016)	3.783***	(3.097)
$D_t x M_{t-3}$	-1.473*	(-1.68)	-1.172	(-0.963)
$D_t x M_{t-4}$	-0.559	(-0.632)	-0.706	(-0.572)
$D_t x M_{t-5}$	1.322*	(1.682)	1.777	(1.589)
t-statistics in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

## E.2. Contenido de los Medios, Retornos y Ciclos Económicos

Tabla E.3: Regresión Medios, ciclos económicos y retornos semanales sin considerar guerras ni crisis 2008

En esta tabla es posible observar los resultados asociados a la ecuación 3.2, la cual relaciona las variables asociadas a los medios  $\frac{pos}{(neg+pos)}$  con los retornos y los ciclos económicos proporcionados pro NBER. De manera adicional, la tabla muestra los test t de cada uno de los lag asociados a la variable de los medios. En esta tabla se utilizaron datos que contemplan entre 1897 y 2014, pero se eliminaron los periodos relacionados con las guerras y la crisis del 2008, es decir, entre 1914 y 1918, 1939 y 1945 así como también el año 2008.

	$pos/(neg + pos)$	
	$\beta$	t-stat
<b>A. Expansión (<math>\beta_1</math>)</b>		
$(1 - D_t)xM_{t-1}$	-3.380	(-0.834)
$(1 - D_t)xM_{t-2}$	4.113	(0.917)
$(1 - D_t)xM_{t-3}$	-0.003	(-0.001)
$(1 - D_t)xM_{t-4}$	-0.892	(-0.200)
$(1 - D_t)xM_{t-5}$	1.205	(0.299)
<b>B. Recesión (<math>\beta_2</math>)</b>		
$D_t x M_{t-1}$	-16.633***	(-2.675)
$D_t x M_{t-2}$	14.044**	(2.086)
$D_t x M_{t-3}$	-9.233	(-1.340)
$D_t x M_{t-4}$	3.544	(0.521)
$D_t x M_{t-5}$	4.302	(0.688)
t-statistics in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

# Apéndice F

## Datos de Panel

En la tabla F.1 es posible observar el efecto de las noticias publicadas por el New York Times en la probabilidad de currency crisis en el resto del mundo, en este caso se utilizaron los datos asociados a 68 países. Dicha tabla muestra que un aumento en la cantidad de palabras positivas, sobre el total de palabras publicadas por el New York Times trae consigo una disminución de la probabilidad de crisis, esto sucede tanto cuando se corre el modelo simple como cuando se agregan variables de control.

Es esperable que las noticias ligadas fundamentalmente con Estados Unidos, como lo son aquellas publicadas por el New York Times, tengan un efecto distinto en la economía local que en el resto del mundo. Lo anterior tiene relación con el hecho de que las noticias de Estados Unidos cumplen un rol más que nada informativo en el resto del mundo.

Tabla F.1: Logit Relacion Ratio  $pos/(pos + neg)$  y Currency Crisis

En esta tabla es posible observar los resultados asociados a la ecuación 2.1, la cual relaciona el ratio  $pos/(pos + neg)$  y la crisis de moneda utilizando datos de panel. De manera adicional, la tabla de divide en dos, los resultados que se muestran en la sección I son los que se obtuvieron a partir de las regresiones simples, mientras que los de la sección II son cuando se consideran variables independientes de interés como es el caso de  $\overline{volM}$  que representa la volatilidad de la variable asociada los medios,  $\overline{inf}$  que es la tasa de inflación anual,  $\overline{loggdp}$  es el logaritmo natural del GDP per cápita y  $\overline{deb/gdp}$  que es el cambio en el ratio deuda pública sobre el GDP per cápita. El panel considera datos de 68 países entre los años 1852 y 2010.

Var. Dep: $C_{i,t}^{currency}$	I			II		
	3	5	7	3	5	7
$\overline{C}_{i,t-1 \rightarrow t-L}$	3.416*** (0.165)	3.851*** (0.196)	3.978*** (0.216)	2.295*** (0.244)	2.651*** (0.318)	2.879*** (0.372)
$\frac{\overline{pos}}{(\overline{neg+pos})}_{i,t-1 \rightarrow t-L}$	-12.174*** (4.118)	-11.396*** (4.200)	-12.034*** (4.288)	-11.335* (5.826)	-12.424** (6.168)	-15.51** (6.926)
$\overline{volM}_{i,t-1 \rightarrow t-L}$				-47.828 (36.895)	-56.702 (34.565)	-59.633 (37.854)
$\overline{inf}_{i,t-1 \rightarrow t-L}$				0.0004 (0.001)	0.0001 (0.001)	-0.0002 (0.001)
$\overline{loggdp}_{i,t-1 \rightarrow t-L}$				0.223 (0.166)	0.143 (0.18)	0.076 (0.192)
$\overline{deb/gdp}_{i,t-1 \rightarrow t-L}$				0.0001 (0.001)	-0.0004 (0.001)	-0.0008 (0.001)
Num. obs.	10511	10375	10239	5030	4841	4654
Pseudo $R^2$	0.232	0.230	0.222	0.219	0.217	0.217

t-statistics in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1