



Facultad de Ciencias Económicas y Empresariales

DISCREPANCIAS ENTRE EV/EBITDA Y P/E EN LA VALORACIÓN RELATIVA DE EMPRESAS

Autor: Santiago Álvarez-Sala Díaz-Caneja
Director: Carlos Bellón Núñez-Mera

MADRID | Junio 2022

RESUMEN

El propósito de este trabajo es la realización de un estudio estadístico de la anómala circunstancia que se da cuando dos múltiplos de valoración, el P/E y el EV/EBITDA, otorgan señales contradictorias, al indicar uno de ellos que la empresa en cuestión ofrece una oportunidad de compra, mientras que el otro sugiere una decisión de venta. Esta situación representa una contradicción respecto de la lógica fundamental de los múltiplos de valoración, pues genera una coyuntura en la que, en función del múltiplo utilizado como patrón de referencia, se estaría obteniendo una valoración relativa en un sentido o en otro.

Así, a lo largo del siguiente trabajo, se hace uso de un registro histórico de los constituyentes del índice S&P 500 para tratar de vislumbrar los motivos por los que se puede dar esta situación y descifrar posibles soluciones u oportunidades. Para ello, se describe la recurrencia con que los múltiplos de valoración sugieren decisiones contrarias a lo largo del tiempo, y se trata de explicar si ello guarda relación o bien con los determinantes de los múltiplos, entendiendo por tal el crecimiento, la rentabilidad y el riesgo; o bien con la industria a la que pertenece la empresa de acuerdo con la actividad que desarrolla.

ABSTRACT

The purpose of this paper is to carry out a statistical study of the anomalous circumstance that occurs when two valuation multiples, P/E and EV/EBITDA, give contradictory signals, one of them indicating that the company in question offers a buying opportunity, while the other suggests a decision to sell. This situation represents a contradiction with respect to the fundamental logic of valuation multiples, since it generates a situation in which, depending on the multiple used as a benchmark, a relative valuation would be obtained in one direction or the other.

Thus, throughout the following paper, we use a historical record of the constituents of the S&P 500 index to try to glimpse the reasons for this situation and discover possible solutions or opportunities. To this end, we describe the recurrence with which valuation multiples suggest contrary decisions over time and try to explain whether this is due to the drivers of valuation multiples, that include growth, profitability and risk; or to the industry to which the company belongs according to its main activity.

PALABRAS CLAVE

Valoración relativa, múltiplos de valoración, análisis fundamental, determinantes de los múltiplos, EV/EBITDA, P/E, regresión logística.

KEY WORDS

Relative valuation, valuation multiples, fundamental analysis, drivers of valuation multiples, EV/EBITDA, P/E, logistic regression.

ÍNDICE

<i>INTRODUCCIÓN</i>	5
<i>I. INTRODUCCIÓN TEÓRICA A LA VALORACIÓN POR MÚLTIPLOS</i>	7
I.1. Valoración de empresas.....	7
I.2. Los distintos tipos de múltiplos de valoración	9
<i>II. DESCRIPCIÓN Y TRATAMIENTO DE DATOS</i>	13
II.1. Base de datos de COMPUSTAT para información sobre fundamentales	13
II.2. Extracción de información de mercado	20
II.3. Cálculo de múltiplos y valoración relativa	21
<i>III. HIPÓTESIS QUE JUSTIFICAN LA DIFERENCIA VALORATIVA ENTRE MÚLTIPLOS</i>	23
III.1. Hipótesis 1: Por determinantes de los múltiplos.....	23
III.2. Hipótesis 2: Por industria	41
<i>CONCLUSIONES</i>	49
<i>ANEXOS</i>	53
ANEXO 1: WEB SCRAPING DE YAHOO FINANCE	53
ANEXO 2: REGRESIÓN LOGÍSTICA PARA EV/EBITDA	55

INTRODUCCIÓN

La valoración por múltiplos está enormemente asentada a lo largo de todas las disciplinas financieras, y se concibe como una forma de averiguar el valor que tiene una empresa mediante su comparación con otras que, compartiendo una serie de características con la primera, ofrecen un marco de contraste relativo idóneo para averiguar si merece la pena realizar o no una inversión. A diferencia de otros métodos que consideran aisladamente al empresa como activo generador de flujos de caja, la valoración por múltiplos permite valorar una compañía en base la percepción que tiene el mercado de otras similares.

Son varios los múltiplos a los que se puede recurrir para realizar este ejercicio, y la elección dependerá de la finalidad perseguida, pero en este trabajo se ha optado por acudir al EV/EBITDA y al P/E por ser los comúnmente empleados y por ser generalizables para la mayoría de las industrias. Así, en función de la posición que ocupen estos múltiplos referidos a una empresa concreta en relación con aquéllos de sus competidores, se puede inferir si ésta está sobrevalorada o infravalorada y tomar una decisión acorde.

No obstante, la problemática se origina cuando cada múltiplo proporciona una señal distinta; esto es, uno de los múltiplos sugiere que la empresa está sobrevalorada con respecto a sus comparables, implicando por tanto que la decisión ideal sería la de vender el activo; mientras que el otro hace una proclama en sentido contrario, entendiendo que la empresa está infravalorada y que, por tanto, representa una oportunidad de inversión por esperarse que se dé un aumento de su precio de mercado. Todo ello se acentúa por el hecho de que ambos múltiplos guardan una relación matemática que permite la obtención de uno a través del otro, ahondándose por tanto aún más en lo incongruente de la situación. Es por todo ello por lo que este tipo de situaciones merecen un interés detenido y, precisamente por eso, el trabajo que sigue se ocupará de esta cuestión.

Así, en un primer lugar, se acudirá a la base de datos de Compustat para obtener información contable histórica de las empresas que constituyen el S&P 500. Seguidamente, se utilizará a Bloomberg para acceder a la información de mercado de estas empresas; y, una vez se tenga todo el contenido necesario para ello, se calcularán los múltiplos, tanto de cada observación como la mediana de la industria a la que pertenecen, para a continuación comparar ambos valores.

Hecho todo lo anterior, y explicada la coherencia teórica que debería existir entre P/E y EV/EBITDA, se tratará de responder a las siguientes cuestiones:

En primer lugar, la recurrencia a lo largo del tiempo con que se da simultáneamente una sobrevaloración de un múltiplo y una infravaloración del otro. Es decir, la frecuencia con que el P/E de una compañía está por encima del de su sector, y el EV/EBITDA por debajo; o bien la frecuencia con que el P/E está por debajo y el EV/EBITDA por encima.

En segundo lugar, y de darse la anterior circunstancia, se tratará de entender si la contradicción entre múltiplos se origina en sus determinantes, entendiendo por tal el crecimiento, rentabilidad y riesgo. Así, se comprobará si la solidez o fortaleza de los determinantes justifican bien una sobrevaloración, bien una infravaloración, y la medida en que esta circunstancia aumenta la probabilidad de que medie contradicción entre los dos ratios.

Por último, se analizará si en la respuesta a la cuestión planteada interviene la pertenencia a una industria o a otra, plateándose como hipótesis si determinadas empresas, como consecuencia de su actividad principal y de los efectos financieros que de ella se derivan, tienen una cierta tendencia a recibir valoraciones distintas de cada múltiplo.

I. INTRODUCCIÓN TEÓRICA A LA VALORACIÓN POR MÚLTIPLOS

I.1. Valoración de empresas

Con anterioridad a adentrarnos de lleno en el estudio empírico en que consiste este trabajo, creo que es adecuado incidir resumidamente en la teoría que subyace al ejercicio que se va a hacer. De esta forma, comenzaré por exponer introductoriamente qué es aquello que se conoce como la valoración de una empresa, algunos de los muchos métodos que existen, y la importancia de esta actividad en las distintas disciplinas financieras.

La determinación del valor de una compañía no tiene un sentido distinto al que puede tener la valoración de cualquier bien que podamos imaginar, que no es otro que precisar un valor económico para, en función de él, tomar una decisión u otra. Como puede verse, es una proclama extremadamente intuitiva, y que tenemos todos totalmente asimilada en nuestro día a día. A la hora de comprar cualquier bien, desde un coche hasta un abrigo pasando por cualquier alimento que podamos imaginar, valoramos la medida en que las necesidades que tenemos quedan cubiertas por tal bien, y si el precio que pagamos por esa satisfacción es razonable o adecuado.

La valoración de una empresa no dista demasiado de esta idea, con la salvedad de que la finalidad no es la satisfacción de una necesidad -o, por lo menos, no es ésta la única finalidad- sino lograr una rentabilidad determinada a partir de nuestra inversión. No obstante, al invertir en una empresa, lo que finalmente conduce a tomar una u otra decisión es si el precio que se está pagando es adecuado a la viabilidad de la empresa; se valora, por tanto, si la compañía es “cara” o “barata” en relación tanto con su propio negocio como con el mercado en el que opera.

La siguiente pregunta a responder se refiere a cuáles son los métodos utilizados para valorar empresas. Lo cierto es que existen multitud de ellos, y estos difieren en función del sector de actividad, de si la empresa cotiza en algún índice o no, de la madurez de la compañía, y de muchos otros factores que habrán de ser considerados caso por caso. No obstante, tradicionalmente suelen distinguirse dos categorías principales: la valoración intrínseca y la valoración relativa (Damodaran, 2008).

La valoración intrínseca es aquella que viene definida por la capacidad de generación de flujos de caja de la empresa, dado un riesgo y unas perspectivas de crecimiento (Kaplan & Ruback, 1995). El método más conocido, y que pese a su antigüedad parece que nunca pasa de moda, es el descuento de flujos de caja o DCF (por sus siglas en inglés, *Discounted Cash Flows*). Este método consiste fundamentalmente en proyectar los flujos de caja que la empresa

generará en el futuro, y descontarlos al presente utilizando una tasa que refleje de manera adecuada el perfil de riesgo de la compañía (WACC, o *Weighted Average Cost of Capital*). Este es probablemente el método que, desde un punto de vista teórico, mejor refleja el valor que tiene una empresa, por centrarse en la capacidad de generación de caja a partir de su actividad operativa. Además del DCF, otros métodos de valoración intrínseca de empresas serían el DDM o *Dividend Discount Model*, utilizado para valorar, entre otras, instituciones financieras, y que consiste en el descuento de los dividendos proyectados a futuro; el NAV o *Net Asset Value*, utilizado en sectores en los que es más importante el tamaño del balance o para empresas en riesgo de insolvencia, y que otorga un valor a la empresa en función de la diferencia entre sus activo y su pasivo; o el descuento del beneficio operativo neto (FFO o *Funds From Operations*) en el caso de empresas inmobiliarias.

La valoración relativa, por su parte, es aquella que, en base a una serie de métricas, infiere el valor de la empresa a través de un ejercicio de comparación con otros activos de similares características (Damodaran, 2011). De esta manera, se contrasta el valor de la empresa analizada con el de aquellas que comparten una serie de atributos comunes, generalmente sector, modelo de negocio, tamaño y mercado. Pero, para realizar esta comparación, debemos ser cuidadosos a la hora de elegir las métricas que utilizamos como patrón comparativo. Pensemos, por ejemplo, en el precio de la acción como primer indicador para determinar el valor relativo de una empresa. Podría parecer que, dado que este número reflejaría el precio que estamos pagando por cada acción de la compañía, funcionaría bien como punto de comparación. No obstante, si nos detenemos brevemente a reflexionar acerca del significado de este número, que no es otro que la capitalización bursátil dividida entre el número de acciones, nos daremos cuenta de que el precio por acción no es la métrica más adecuada para realizar comparaciones, por el simple hecho de que viene determinado por el número de acciones que, a su vez, no representa valor alguno.

Podría pensarse, por otro lado, en tomar alguna de las partidas de la cuenta de Pérdidas y Ganancias, como por ejemplo el beneficio neto, y comparar a partir de ahí si una empresa vale más que otra. Sin embargo, una empresa puede generar más beneficio que otra por el simple hecho de ser más grande y tener un mayor volumen de ingresos y no por ello tener más valor.

La clave por tanto parece estar en combinar ambos elementos, dando con algún tipo de métrica que incluya información relativa tanto al precio como al desempeño financiero de la empresa; o, lo que es lo mismo, que relacione la cantidad de dinero que el mercado considera que vale la acción, con los números que genera la empresa a través de su actividad.

Consecuentemente, Suozzo et al. (2001) definen los múltiplos como la razón o relación que existe entre una expresión del valor de mercado de una empresa (numerador) y alguna partida de su cuenta de resultados que refleje los ingresos o beneficios obtenidos por ésta (denominador). Y serán precisamente estos múltiplos los que se utilicen para obtener información acerca del valor relativo que tiene una compañía con respecto a sus competidores, por ser una forma idónea de mitigar las diferencias relativas al tamaño, número de acciones o capitalización, y poder, por tanto, establecer una comparación coherente y en términos iguales.

I.2. Los distintos tipos de múltiplos de valoración

El universo de posibles múltiplos de valoración es muy extenso y depende del sector y modelo de negocio de la empresa: no es lo mismo valorar una empresa como Procter and Gamble, cuyas operaciones consisten fundamentalmente en la venta de productos de consumo, que un banco como Bank of America, cuyos principales ingresos provienen de la recolección de intereses y de las comisiones que cobran a sus clientes; igual que no es lo mismo valorar una empresa hotelera como Marriott, con enormes cifras de amortización o alquileres, que una farmacéutica como Pfizer, cuyo principal coste es la investigación y desarrollo.

No obstante, en general, suelen distinguirse dos grandes categorías (Suozzo et al., 2001): *Enterprise Value multiples* (múltiplos de valor de empresa, en adelante “EV Multiples”) y *Equity multiples* (múltiplos de capitalización bursátil).

Los EV multiples son aquellos que en el numerador contienen el EV y en el denominador cualquier métrica que represente ingresos, beneficios o flujos de caja con anterioridad al pago de intereses de deuda, principalmente ingresos, EBITDA, EBIT o UFCF (Suozzo et al., 2001). De esta manera, tanto numerador como denominador reflejan cantidades que corresponden a todos aquellos que tienen algún interés en la empresa; esto es, accionistas y acreedores. Para Koller et al. (2010), el EV se define como:

$$EV = MVE + MVD + MVPS + Participaciones minoritarias - Caja y equivalentes$$

Donde:

- MVE: “Market value of equity” o capitalización bursátil
- MVD: “Market value of debt” o valor de mercado de la deuda
- MVPS: “Market value of preferred stock” o valor de mercado de las acciones preferentes.

Por tanto, como se puede observar, el EV es una figura muy útil para representar el valor de mercado que tiene la empresa en su conjunto; que no está representada únicamente por la participación de los accionistas, sino también por las aportaciones de los tenedores de deuda, que no dejan de ser también “propietarios” de parte de los activos del balance.

Por otro lado, los Equity multiples son aquellos que en el numerador contienen un número que se corresponde únicamente con la participación que tienen los accionistas en la empresa (capitalización bursátil o precio por acción), y en el denominador alguna métrica que represente propiedad, beneficios o flujos de caja exclusivamente para este mismo colectivo (Suozzo et al., 2001), principalmente patrimonio neto, beneficio neto o LFCF (*Levered Free Cash Flow*, que se define como el flujo de caja generado para los accionistas de la sociedad, después de intereses y amortizaciones obligatorias de deuda).

Así, los múltiplos comúnmente utilizados serían:

EV Múltiples	Equity Multiples
EV/REVENUES	P/B
EV/EBITDA	P/E
EV/EBIT	P/LFCF
EV/UFCF	

Además de estos, existen infinidad de múltiplos que se utilizan concretamente para determinados sectores e industrias, como podría ser EV/EBITDAX para empresas petroleras, donde la “X” representa los gastos de exploración, ya que algunas empresas los capitalizan y otras no (Kumar & Sukumaran, 2017); EV/EBITDAR para empresas de venta minorista, donde la “R” representa los gastos de alquiler, puesto que algunas compañías tienen los locales comerciales en propiedad y otros los alquilan (Matthews, 2012); u otros más creativos, como EV/Usuarios, para los casos en que las empresas están en fases prematuras y todavía no generan flujos de caja ni beneficios.

Así, por ejemplo, una empresa que tuviese un $EV/EBITDA = 7x$, estaría siendo valorada, términos de EV, a siete veces su EBITDA; o, lo que es lo mismo, un inversor que quisiese comprar esta empresa estaría dispuesto a pagar por ella siete euros por cada euro de EBITDA que ésta generase. Y es precisamente este significado de los múltiplos, el de razón o relación entre dos métricas estandarizadas, el que nos permite hacer comparaciones entre distintas empresas y valorar cómo de barato o caro es el precio de mercado de unas frente a otras.

A modo de ejemplo, veamos la siguiente tabla, en la que se representa el valor relativo de una serie de marcas estadounidense de refrescos y bebidas no alcohólicas:

Compañías estadounidenses de refrescos y bebidas no alcohólicas, en \$ mn

Company Name	Ticker	Fiscal Period	Mkt Cap	EV	Net Income	EBIT	EBITDA	Sales	P/E	EV/EBITDA	EV/EBIT
Coca-Cola	KO-US	04/01/2022	280.086	313.427	10.307	10.697	12.149	38.726	27,2x	25,8x	29,3x
PepsiCo	PEP-US	03/19/2022	237.407	270.670	10.165	11.447	14.157	79.468	23,4x	19,1x	23,6x
Keurig Dr Pepper	KDP-US	03/31/2022	52.685	65.342	2.406	2.868	3.576	12.683	21,9x	18,3x	22,8x
Monster Beverage	MNST-US	03/31/2022	44.239	41.550	1.356	1.803	1.857	5.541	32,6x	22,4x	23,0x
National Beverage	FIZZ-US	01/29/2022	4.282	4.325	163	225	256	1.072	26,2x	16,9x	19,2x
Celsius	CELH-US	12/31/2021	3.120	3.105	4	-4	-3	314	792,3x	-	-
Primo Water	PRMW-CA	12/31/2021	2.266	3.879	-3	126	345	2.073	-	11,3x	30,9x
Average									26,3x	18,9x	24,8x
Median									26,2x	18,3x	23,0x

Fuente: FactSet

A la izquierda podemos observar los distintos elementos necesarios para el cálculo de los múltiplos mientras que, a la derecha, como resultado de las correspondientes divisiones, se encuentran el P/E, EV/EBITDA y EV/EBIT. Dicho esto, es preciso realizar una serie de observaciones.

La primera de ellas es que, para aquellos casos en los que la figura del denominador es negativa, el cálculo de múltiplos carece de significado. Conceptualmente, no tiene sentido decir que un inversor está dispuesto a pagar por una empresa, por ejemplo, menos 5 veces su EBITDA. Es por ello por lo que en estos casos se prescinde de los múltiplos correspondientes a estas empresas para cualquier tipo de análisis.

En segundo lugar, los múltiplos, por su propia naturaleza de razón o relación, son sensibles a los valores que tomen el numerador y el denominador, pudiendo obtenerse resultados poco razonables o que, por lo menos, debemos tomar como *outliers* en cualquier análisis que hagamos. En este caso Celsius, con un resultado del ejercicio de 4 millones de dólares y una capitalización bursátil de unos 3.000 millones de dólares, está valorada a un P/E de 729x. Viendo los números de sus comparables, parece razonable pensar que se ha dado algún acontecimiento o circunstancia durante este ejercicio que ha motivado a la empresa a anotarse una serie de gastos extraordinarios, resultando en una reducción inusual de sus beneficios. Es por eso por lo que, en estos casos, no tiene demasiado sentido decir que Celsius está sobrevalorada con respecto a su sector, sino que, más bien, lo ideal sería normalizar la cuenta de pérdidas y ganancias para eliminar elementos extraordinarios que difuminan el análisis, y realizar ciertos ajustes sobre a la hora de calcular los múltiplos. En cualquier caso,

y porque el ejercicio que aquí estamos haciendo tiene una finalidad ejemplificativa, lo tomaremos como un *outlier* y lo ignoraremos.

Otra circunstancia merecedora de reflexión, a la que de hecho dedicaré el resto del trabajo, es la aparente contradicción o anomalía que existe en aquellos casos en los que una empresa, en términos de un múltiplo concreto, como puede ser el EV/EBITDA, está valorada por encima de sus comparables, pero, sin embargo, tomando como referencia otro múltiplo distinto, como podría ser el P/E, se valora por debajo. Este es el caso de Pepsi, que en EV/EBITDA está por encima de la media de sus competidores, pero en P/E, no obstante, está por debajo. El resultado de esta cuestión es que, por un lado, se recibe una señal de compra por parte del P/E, que está transmitiendo que Pepsi está infravalorada con respecto a sus comparables y que, por lo tanto, la tendencia sería un aumento de precio que la igualase al resto de empresas; pero, por el otro lado, el EV/EBITDA transmite el mensaje contrario, poniendo de manifiesto una oportunidad de venta porque, de nuevo, el múltiplo debería reducir su valor hasta situarse a la altura de la de sus comparables.

Existe una relación matemática entre ambos múltiplos que, a través de una serie de ajustes, permite obtener uno a partir del otro. Así, Jullens (2012) define los siguientes factores de ajuste que permiten reconciliar el valor del EV/EBITDA y el del P/E: depreciación, amortización, activos “secundarios” como inversiones en empresas asociadas, deuda e intereses, impuestos y participaciones minoritarias.

Ahora bien, pese a que ambos múltiplos estén matemáticamente relacionados, y exista un forma de explicar las diferencias que median entre ellos, lo verdaderamente interesante es explorar los motivos sustantivos por los que ambos múltiplos nos dan señales contradictorias, y tratar de dilucidar cuáles son las consecuencias de esta contradicción. Buscar, por tanto, la forma de convertir una aparente incoherencia -entre un múltiplo que sugiere comprar frente a otro que sugiere vender- en una oportunidad para anticiparse y tomar la decisión apropiada. Por tanto, a lo largo de este trabajo, exploraremos los siguientes puntos:

1. En primer lugar, tomando como referencia empírica un registro histórico de empresas del S&P 500, se estudiará si realmente ocurre, con mayor o menor frecuencia, que la valoración relativa de una empresa difiera en función de si se utiliza un múltiplo u otro.
2. En segundo lugar, en caso de producirse lo anterior, se analizará si esta circunstancia guarda algún tipo de relación con los determinantes de los múltiplos.
3. Y, en tercer lugar, se examinará si la cuestión planteada viene determinada, o sucede con mayor o menor frecuencia, en unas industrias frente a otras.

II. DESCRIPCIÓN Y TRATAMIENTO DE DATOS

II.1. Base de datos de COMPUSTAT para información sobre fundamentales

Para la realización del estudio que tenemos por delante debemos, en primer lugar, detenernos en la descripción y análisis de los datos empleados. Así, dado que los múltiplos de valoración constan de una parte directamente relacionada con el mercado en el numerador, y otra basada en los fundamentales de la empresa en el denominador, utilizaremos dos bases de datos: Bloomberg, para todo aquello que se refiere a la información de mercado; y COMPUSTAT, para el análisis fundamental de las distintas empresas.

Nos enfocaremos, en un primer lugar, en COMPUSTAT, base de datos perteneciente a la empresa Standard & Poor's y que contiene información histórica sobre los estados financieros de miles de empresas alrededor de todo el mundo.

Una característica singular del conjunto de datos utilizado es su extenso volumen. El archivo trimestral, que será al que recurriremos principalmente, consta de 585.784 observaciones distribuidas a lo largo de 55 variables, número razonable si se tiene en cuenta que refleja información de los años 1984 a 2019. En términos matriciales, estaríamos hablando de 32.218.120 elementos. No obstante, para el propósito de nuestro estudio, únicamente tenemos interés en aquellas empresas que forman parte del índice S&P 500, por lo que podemos reducir en gran medida el número de observaciones. Si utilizamos los tickers como variable identificativa de cada empresa veremos que, en un primer momento, la base de datos contaba con 10.267 empresas distintas, dado que incluía compañías que quizás no cotizasen en el S&P 500 pero sí en el S&P MidCap 400, en el S&P SmallCap 600, en algún otro índice creado por Standard & Poor's, o que simplemente no tuviese presencia en ninguno de los anteriores.

Se procedió a la extracción de una lista de las empresas que constituyen el S&P 500, y se obtuvo un subconjunto de datos a partir de ella. Para este proceso, y para otros tantos que iremos tratando, se utilizó un paquete denominado "sqldf", que permite agrupar, filtrar e intersecar datos de distintas tablas utilizando las mismas *queries* que se emplearían en el lenguaje MySQL. De esta forma, ejecutando el siguiente código:

```
sqldf('SELECT tic FROM compustat_quarterly INTERSECT SELECT tic
FROM sp_500_stocks'),
```

se retuvieron únicamente aquellas empresas del primer conjunto de datos cuyo ticker aparecía en la lista de empresas del S&P 500. Una vez hecho esto, se comprobó que algunas

de las empresas de esta lista no aparecían en aquel conjunto de datos (concretamente cinco), y se estudió el motivo. En algunos casos, el vacío venía dado por el hecho de que Compustat únicamente incluye un tipo de acción por compañía, por lo que Google o Under Armour, que cotizan a través de diferentes clases de acciones (Class A y Class C, por ejemplo), solamente llevan aparejadas un ticker, en lugar de dos; y, en los demás casos, el ticker era ligeramente distinto, por lo que se localizaron las diferencias y se enmendaron. Así, se consiguieron dos conjuntos de datos, “sp500_quarterly” y “sp500_yearly”, que contenían las mismas columnas que los archivos iniciales, pero con un número de observaciones mucho más reducido (56.488 y 16.492, respectivamente).

La siguiente cuestión con la que se lidió fue la relativa a la concreción del sector e industria al que pertenece cada empresa. Nótese que ésta es una cuestión fundamental para el desarrollo del trabajo dado que la industria constituyó el patrón de referencia para la definición del universo de comparables de las distintas empresas.

Para la asignación de un sector e industria a cada ticker, se utilizaron dos procedimientos distintos para, a continuación, elegir aquél que ofreciese resultados más completos y satisfactorios.

El primero de los métodos empleados trata de hacer *Web Scraping* de la información pública disponible en Yahoo Finance. La técnica del Web Scraping, cuya traducción literal sería “raspado de web”, consiste en la obtención automatizada de datos disponibles en la web, ya sea en formato numérico, texto, imagen o vídeo (Horton, N. J., & Kleinman, K., 2015). Fundamentalmente, lo que se hace es, a través de un código, descargar y almacenar datos públicamente disponibles en la web. Los paquetes necesarios para la realización de esta labor difieren en función del lenguaje que se utilice; pero, en nuestro caso, RStudio requiere principalmente los paquetes “rvest” y “httr”. El código consiste principalmente en una función que tiene como variable el ticker y que, una vez introducido este, rastrea y almacena los datos que se encuentran en la siguiente url:

```
url =  
paste0('https://finance.yahoo.com/quote/',x,'/profile?p=',x,'')
```

donde la “x” se correspondería con el ticker que hemos introducido.

Una vez la función rastrea la página de Yahoo Finance en busca del sector (“Sector”) y la industria (“Industry”) que se corresponde con cada ticker, los extrae, y los almacena en una base de datos que tiene cuatro variables: Ticker, Nombre de la compañía, Sector e Industria.

El código descrito tiene una gran utilidad, no sólo en términos de obtención de información sectorial, sino también relativa a los estados financieros, la situación de mercado, o cualquier otro dato que se encuentre en Yahoo Finance; bastaría con realizar unas pequeñas modificaciones para que, a partir del ticker, se pudiese extraer información de todo tipo. No obstante, tiene algunos inconvenientes. El primero de ellos no se refiere tanto a la función en sí sino más bien al Web Scraping en general. Ésta es una tarea pesada, que requiere bastante esfuerzo computacional para su ejecución, por lo que su utilidad se optimiza si el volumen de datos a extraer no es demasiado grande; sin embargo, cuando tratamos de hacerlo con quinientos tickers distintos, podemos encontrarnos con que requiere una cantidad de tiempo considerables.

El segundo inconveniente viene dado por la diferencia que puede existir entre la información disponible en Yahoo Finance y el conjunto de datos que estamos utilizando. Efectivamente, si se introducen manualmente uno por uno los tickers en la función, comprobando su nombre exacto en Yahoo Finance, este no debería ser un problema; no obstante, cuando tratamos de automatizar este proceso creando un vector con todos los tickers que necesitamos y que obtenemos a partir de nuestra base de datos, nos podemos encontrar con que no se corresponden exactamente con la forma en que se presentan en Yahoo Finance. Concretamente, en nuestro caso, esto ocurriría con 39 compañías.

En este otro caso lo que se hizo fue, a partir de dos de las variables disponibles en el conjunto de datos, “gsector” y “gsubind”, hacer Web Scraping para obtener los datos de sector y sub-industria de cada compañía de acuerdo con los estándares GICS. Los estándares GICS (Global Industry Classification Standard) son unos criterios taxonómicos promovidos por MSCI y Standard & Poor’s que tienen por finalidad aportar una clasificación de los sectores de actividad a los que se dedican las empresas. Así, dividen las actividades por sector, grupo de industria, industria y sub-industria, siendo cada una de las categorías más concreta que la anterior:

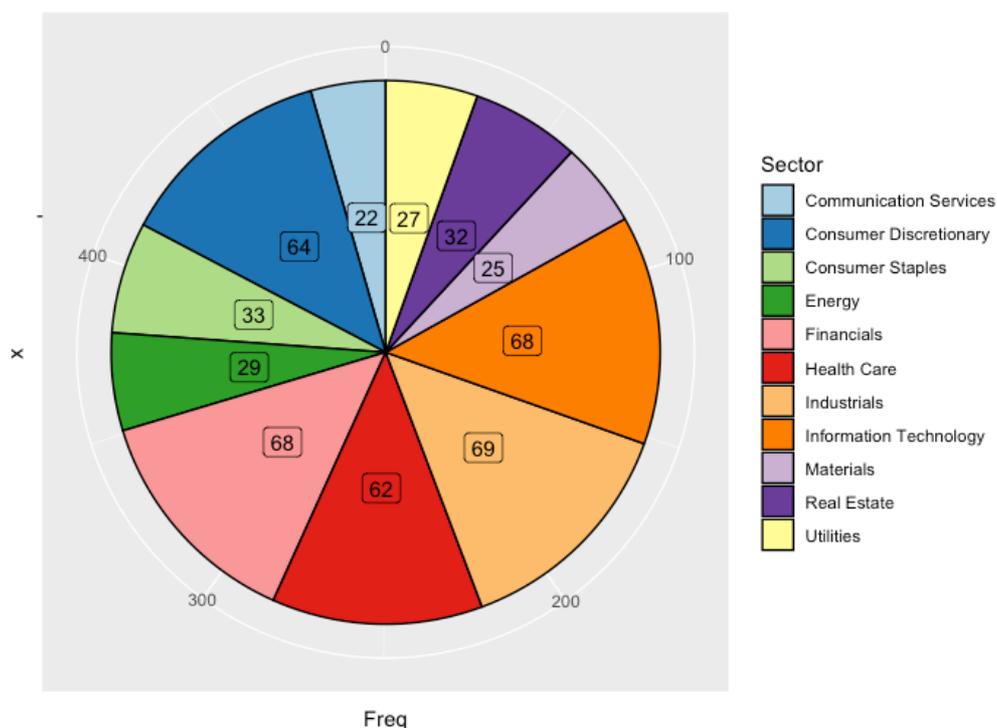


Cada uno de los sectores, grupos de industria, industrias y sub-industrias tienen asignado un código, y este código es, precisamente, el que se encuentra en las dos variables mencionadas anteriormente. Por lo tanto, la labor a realizar aquí se reduce a hacer Web Scraping de Wikipedia, donde encontramos una lista con el nombre que corresponde a cada código numérico para, a continuación, almacenarlos todos ellos en una tabla:

```
url <-
"https://en.wikipedia.org/wiki/Global_Industry_Classification_Standard
"

GIC_subindustries = url %>%
  read_html() %>%
  html_node(xpath = '//*[@id="mw-content-
text"]/div[1]/table/tbody') %>%
  html_table(fill = TRUE)
```

Y, una vez tenemos esta tabla, simplemente utilizamos la función “inner_join” incluida en el paquete “sqldf” para asignar a cada empresa, en función de los códigos de sector y sub-industria que lleven asignados, el nombre del sector, grupo de industria, industria y sub-industria que les corresponden.



S&P 500 por sector, visualización creada con ggplot2

Este gráfico refleja de manera muy sencilla la composición sectorial del índice. Se puede observar que existe una distribución poco homogénea, al existir una serie de sectores, como el industrial, el financiero o el tecnológico, que concentran un gran número de empresas; mientras que otros, como los servicios de comunicación, las infraestructuras o los materiales, recogen un número mucho más reducido de compañías.

Una vez añadida la información sectorial al conjunto de datos, se procedió a valorar la importancia de las variables en él contenidas, de manera que se eliminaron aquellas que no aportaban información relevante al estudio. Así, nos desprendimos de aquéllas que eran totalmente irrelevantes, por tener un mismo valor para todas las observaciones, así como aquéllas que únicamente llevaban aparejados valores desconocidos (NAs o *missing values*). Igualmente, se eliminaron aquéllas que se correspondían con el código de sector, industria y sub-industria, por tener ya los nombres correspondientes a cada una y evitar redundancias.

A continuación, se procedió al estudio de los valores desconocidos, calculando la proporción de NAs con respecto al total de observaciones utilizando el siguiente código:

```
apply (sp500_quarterly, 2,
function(col) sum(is.na(col))/length(col))
```

Las que se muestran a continuación son las variables que, en un principio, deberían haber sido eliminadas por tener un número de NAs superior al 10%, estableciéndose este límite como un número razonable a partir del cual la proporción de *missing values* es excesiva.

Variable	%NAs
Total Current Assets	21%
Total Non-Curret Assets	28%
Debt in current liabilities	10%
DD&A	16%
Goodwill	47%
Intangible Assets	39%
Total Current Liabilities	21%
Total LT liabilities	28%
Minorities Interest in income	21%
EBITDA	14%
Equity	59%
Interest expense	15%
SG&A	27%

Variables con un porcentaje de NAs > 10 %

Marcadas en rojo encontramos aquellas que, por representar el total de otras partidas del balance, fueron eliminadas. Más adelante, se sumaron las distintas partidas que corresponden a cada total para obtener estas variables.

Las marcadas en verde son aquellas que, pese al elevado número de NAs, se consideró que tenían una relevancia suficiente como para no ser eliminadas. El caso más notorio fue el del EBITDA, que tiene una importancia fundamental para el cálculo de los múltiplos. Se descartó la sustitución de este valor por un cero dada la fuerte distorsión que implicaría sobre los datos y sobre nuestro estudio; recuérdese que el EBITDA es el denominador de uno de los múltiplos y que, por lo tanto, un cero o un valor nulo nos imposibilitaría el cálculo de esta métrica. No obstante, estudiando más de cerca de los datos, se vio que, en la inmensa mayoría de los casos, el EBITDA venía calculado como la diferencia entre los ingresos y los costes, tanto directos como indirectos (Revenues – COGS – SG&A). Así, excluyendo los NAs, se observó que, sobre un total de 35.135 observaciones, la diferencia entre el EBITDA calculado como diferencia entre ingresos y gastos, y el EBITDA tal y como venía en el conjunto de datos, era mayor que 1 en tres casos, y menor que -1 en ciento noventa y nueve. Por lo tanto, se tomó este método de cálculo para obtener el EBITDA.

Y, viendo la tabla anterior, parece que el principal problema que existía aquí, y que motivaba el elevado número de NAs en el EBITDA, venía dado por la cantidad de *missing values* en la variable SG&A. No obstante, esto no debería extrañarnos, habidas las diversas formas de imputación de costes directos e indirectos que tienen las empresas y que justifica que, en bastantes ocasiones, la partida SG&A no esté presente en las cuentas trimestrales publicadas, ya sea porque se consideran todos los costes como directos, ya sea porque directamente no tienen costes indirectos. Por tanto, se creyó apropiado sustituir todos los NAs de la columna de SG&A por ceros, así como los de las columnas de ingresos y COGS, y volver a calcular el EBITDA con la fórmula anteriormente mencionada. De esta forma conseguimos no tener *missing values* en el EBITDA.

Por otra parte, se prestó atención detallada a los *missing values* correspondientes a las variables EPS (*Earning per share* o beneficio por acción) dado que, al igual que el EBITDA, constituye el denominador de uno de los múltiplos a utilizar -en este caso, P/E- y, por tanto, fue de gran importancia asegurarse que no hubiese NAs en esta variable.

Compustat calcula esta variable como el resultado de dividir el resultado del ejercicio (*Net Income* o NI) entre el número de acciones. Por tanto, la concurrencia de un NA en cualquiera de estos valores motivaría que el EPS tampoco tuviese un valor asignado. Así, en un primer momento, se trataron los *missing values* del resultado del ejercicio. Estos ascendían

únicamente un 0,8% de las observaciones, por lo que no hubo demasiado problema en eliminarlas. No obstante, como la idea era tener cifras para cada uno de los trimestres del año (y calcular, como más adelante se explicará, los números correspondientes a los últimos doce meses en cada caso), se decidió que, para todo ticker que en un determinado año y trimestre tuviese un valor desconocido, se eliminarían todas las observaciones correspondientes a esa empresa en ese año. Además, esta aproximación vino avalada por el hecho de que, en la mayoría de los casos, faltaban las cifras de resultado del ejercicio para los cuatro trimestres del año, por lo que apenas se redujo el conjunto de datos.

La segunda de las variables que hubo que tratar fue el número de acciones. Tras estudiar las distintas posibilidades, se decidió implementar la misma solución que con la variable anterior; esto es, se identificaron todos los tickers que llevaban asociados un *missing value* en el número de acciones, y se procedió a estudiar en cuántos casos esto se daba durante los cuatro trimestres del año. Resultó ser la mayoría de ellos. Además, también se llegó a la conclusión de que, mayoritariamente, los NAs se daban durante el primer o los primeros años en los que ese ticker aparecía en la base de datos. Por tanto, se optó por eliminar todas las observaciones correspondientes a trimestres en los que, para un determinado año, la empresa tuviese al menos un NA en la variable número de acciones (tal y como hicimos para el resultado del ejercicio).

Y, por último, nos restaba únicamente calcular el nuevo EPS tras las modificaciones que hemos hecho. Y así se hizo, dividiendo la columna de beneficio neto entre las del número de acciones, obteniendo un resultado muy coherente con el que ya venía por defecto calculado en la base de datos.

El último de los puntos que nos toca tratar en referencia a la base de datos de Compustat es el cálculo de las métricas de la cuenta de pérdidas y ganancias en el formato LTM (*Last Twelve Months*, últimos doce meses). Como se indicó, la base de datos trimestral de Compustat incluye información únicamente de los tres meses a los que se refiere. No obstante, el cálculo de los múltiplos de valoración requiere que las figuras del denominador se refieran a un año entero, dado que las del numerador, sea la capitalización bursátil, sea el EV, están expresadas en términos anuales.

Por tanto, para el cálculo del LTM del EBITDA, EPS, y demás partidas de la cuenta de resultados, se recurrió al uso de un bucle “for”, dentro del cual se construyeron una serie de instrucciones en formato condicional. Así, el código funcionaba de la siguiente manera: para todas las observaciones de la base de datos que tuviesen un mismo valor en la columna correspondiente a los tickers, se almacenó en una columna nueva el resultado de sumar al valor de la propia observación los tres valores anteriores. Al mismo tiempo, si esa observación se

correspondía con la primera, la segunda o la tercera vez que aparecía el ticker en la base de datos, se asignaba un NA en la nueva variable, dado que para los tres primeros trimestres registrados no existe un LTM. Este proceso se hizo para todas las partidas de la cuenta de resultados, obteniendo en consecuencia 14 observaciones, todas ellas con 1.492 *missing values* correspondientes a los tres primeros trimestres en que aparece cada ticker.

Es cierto que los bucles no son la forma más eficiente de proceder en RStudio. No obstante, dado el número de condiciones que había que introducir en el bucle, probablemente ésta fuese la forma más intuitiva de proceder.

II.2. Extracción de información de mercado

Una vez se ha descrito el procedimiento de obtención y tratamiento de los datos relativos a los fundamentales, lo que supondría disponer ya del denominador de los múltiplos, procede a continuación abordar la extracción de datos de mercado.

Antes de ello, fue conveniente realizar un pequeño ajuste sobre los datos tal y como quedaron expresados tras el tratamiento anterior. La realidad es que los múltiplos, si se calculan dividiendo el precio de la acción que corresponde a cada una de las fechas sobre las que tenemos información financiera, no están del todo bien expresados. Y el motivo es, fundamentalmente, el momento en el que el mercado conoce los fundamentales de una empresa y puede reflejarlos en el precio de sus acciones. De esta forma, el precio por acción de una empresa a, por ejemplo, 31-12-2018, no está fundamentado en los resultados obtenidos por la compañía durante los cuatro trimestres de ese año, sino en la última información que es pública y conocida, que sería la correspondiente a los primeros tres trimestres de 2018 y el último de 2017. Es decir, las compañías no publican sus resultados el último día del periodo al que se refieren, sino que lo hacen con un diferimiento temporal máximo de entre 45 días (en el caso del 10-Q o resultados trimestrales) o de 60 días (para el 10-K o resultados anuales). En consecuencia, el dato fundamental que debemos tomar no es el del propio periodo, sino el correspondiente al periodo anterior, que es el que el mercado conocía en el momento en que otorgó un precio a una acción.

Ahora bien, las compañías tienen calendarios de resultados diversos, que no necesariamente deben coincidir con el año natural. No obstante, mi compañero Álvaro de Artiñano, logró construir un código en Python que hiciese Web Scraping a través de la página web de la SEC (Securities and Exchange Commission, organismo regulador de los mercados financieros en Estados Unidos), de manera que, a partir del código identificativo CIK, se

obtuviese la fecha en que esa empresa publicó sus resultados trimestrales o anuales. Esta información se almacena en un conjunto de datos que contiene todas las fechas de publicación que constan en la página web para cada empresa, junto con su ticker, su código CIK, y si se trata del 10-Q o el 10-K.

Una vez se obtuvieron las fechas en las que las empresas publicaron sus resultados entre 1996 y 2018, se guardó el archivo en formato Excel y, a partir de ahí, se utilizó Bloomberg para obtener el EV y la capitalización bursátil de cada empresa en el momento de publicación de los resultados. Seguidamente, y dado que, como se dijo, el periodo máximo de publicación de resultados es de 60 días, se asignó a cada fecha de publicación la fecha correspondiente al último día del trimestre anterior. De esta forma, al unir este conjunto de datos al que ya teníamos con las cifras LTM calculadas, se obtuvo una base de datos en la que a cada observación se le asignaron los fundamentales que eran conocidos en el momento en el que el mercado otorgaba un determinado precio la acción. Por tanto, esta modificación nos permitió obtener un conjunto de datos que reflejase la información que tenía el mercado a la hora otorgarle un determinado valor a una compañía.

Pongamos un ejemplo. Una empresa publica sus resultados anuales de 2016 el 18-02-2017. Por tanto, lo que se haría en este caso es obtener de Bloomberg la capitalización bursátil y el EV a fecha 18-02-2017 para, a continuación, utilizar el último día del trimestre anterior (31-12-2016) como conducto para poder unir esta información a la relativa a los fundamentales. Así, para el día 31-12-2016 tendríamos la capitalización bursátil y el EV que reflejan la reacción del mercado a las cifras publicadas el 18-02-2017, pero que corresponden al ejercicio finalizado el 31-12-2016.

II.3. Cálculo de múltiplos y valoración relativa

Hecho todo lo anterior, se estuvo en disposición de calcular los múltiplos de valoración:

$$P/E = \frac{\text{Capitalización bursátil}}{NILTM}$$

$$EV/EBITDA = \frac{(\text{Capitalización bursátil} + \text{Deuda Total} - \text{efectivo})}{EBITDALTM}$$

A continuación, se calcularon los múltiplos medianos para cada industria y cada año. Idealmente, el ejercicio de establecimiento de un universo de comparables, como ya se comentó, debe ser mucho más exhaustivo, y debería tener en cuenta también otras circunstancias como el crecimiento, la rentabilidad y el riesgo (Koller et al., 2010). No obstante, para los efectos que aquí interesan, que son más bien un estudio estadístico general, la industria sirve como aproximación razonablemente válida.

Por último, se crearon dos nuevas variables, “expev” y “exppe”, de carácter dicotómico y que tomaron valor 1 para aquellas observaciones en las que el EV/EBITDA fuese superior a la mediana y al mismo tiempo el P/E inferior, y 0 cuando ambos fuesen superiores o inferiores (“expev”); o bien 1 cuando sucediese que el P/E fuese superior a la mediana y el EV/EBITDA inferior, y 0 para los casos en que ambos fuesen superiores o inferiores (“exppe”). De esta forma, hubo observaciones que tuvieron 0 en ambas variables, pero en ningún caso una empresa en un mismo momento pudo tomar el valor 1 en ambas variables.

Así, se obtuvo que, para el caso del EV/EBITDA, 3.373 observaciones cumplían con los requisitos para tener asignadas un 1; y, en el caso del P/E, 3.430. En total, por tanto, teniendo en cuenta el número total de observaciones, 26.989, se observó una sobrevaloración en términos de EV/EBITDA en el 12,5 % de las observaciones; y, en el P/E, en el 12,7 %. En total, un 25,2 % de las observaciones presentaron una discrepancia valorativa en términos de EV/EBITDA y de P/E; esto es, en aproximadamente un cuarto de



Distribución de las observaciones sobrevaloradas en EV/EBITDA (rojo) y en P/E (azul)

las ocurrencias, se observó que el EV/EBITDA y el P/E se contradijeron a la hora de reflejar si una empresa vale más o menos que sus comparables.

III. HIPÓTESIS QUE JUSTIFICAN LA DIFERENCIA VALORATIVA ENTRE MÚLTIPLOS

III.1. Hipótesis 1: Por determinantes de los múltiplos

A continuación, se va a tratar de analizar si las diferencias de valoración entre EV/EBITDA y P/E vienen determinadas por las características intrínsecas de la actividad y negocio que lleva a cabo una empresa. Así, se recurrirá al estudio de una serie de métricas conocidas como los determinantes de los múltiplos, que son, en principio, los actores que justifican que una empresa esté peor o mejor valorada que sus comparables.

Siguiendo a Koller et al. (2010) podríamos decir que una empresa crea valor invirtiendo el dinero que genera de tal manera que, en un futuro, pueda llegar a generar aún más dinero. Consecuentemente, el valor de una empresa se define como la expectativa de generación futura de flujos de caja a partir de los que actualmente se generan. Pueden establecerse, por tanto, tres factores que definen el valor de una empresa: rentabilidad, crecimiento y riesgo.

Según Damodaran (2007), existe una dualidad entre rentabilidad y crecimiento, de manera que ambos factores son directamente proporcionales a la creación de valor. El crecimiento al que aquí nos referimos viene determinado por la fortaleza o idoneidad con que las empresas localizan el resultado de sus operaciones. Piénsese, por ejemplo, es una compañía que lo único que hace es distribuir entre sus accionistas todo el dinero generado a partir de su actividad. Es probable que esta solución fuese ideal en el corto plazo para aquéllos que únicamente tienen interés en rentabilizar sus inversiones cuanto antes, pero una empresa que funciona de esta forma está condenada a estancarse, no crecer y, probablemente, acabar desapareciendo. Es por tanto fundamental para una empresa reinvertir una parte de sus ganancias en ella misma, porque esta es la única forma que tiene de crecer de manera orgánica y, en un futuro, alcanzar volúmenes de ingresos superiores a los actuales. Y es precisamente este factor, el de la reinversión, el que actúa como palanca de unión entre rentabilidad y crecimiento.

Damodaran (2007) define el crecimiento sostenible a largo de plazo de una empresa podría definirse de la siguiente manera:

Crecimiento en beneficio operativo (g) = % de Reinversión x ROIC

$$= \frac{Capex - DD\&A + \Delta NWC}{EBITDA (1 - t)(1 - D)} \times ROIC = \frac{Capex - DD\&A + \Delta NWC}{NOPAT} \times ROIC$$

Donde $Capex - DD\&A$ representan las inversiones netas en activo fijo, ΔNWC (Incremento en *Net Working Capital* o Capital Circulante) refleja las inversiones realizadas en activo corriente, t el tipo impositivo aplicable, D el porcentaje de depreciación, y el $ROIC$ la rentabilidad que obtiene a partir del capital invertido (*Return on Invested Capital*).

Análogamente:

Crecimiento en beneficio (g) = % de Reinversión en Equity x ROE

$$= \frac{Capex - DD\&A + \Delta NWC - \Delta Deuda}{NI} \times ROE$$

Donde NI es el beneficio neto (*Net Income*) y el ROE representa la rentabilidad generada por las inversiones en el *Equity*

Así, crecimiento y rentabilidad interaccionan entre sí para la creación de mayores flujos de caja en el futuro. Y es que tiene sentido pensar que aquellas empresas que asignen una mayor parte del efectivo que generan a su propio crecimiento a través de la reinversión en sus activos, y que lo hagan además de una manera eficiente, tengan unas mayores expectativas de crecimiento y el mercado, por tanto, las valore mejor que a otras que no satisfagan estas circunstancias.

El contrapeso del crecimiento y la rentabilidad es, por supuesto, el riesgo asociado a la empresa en cuestión, que puede ser tanto un riesgo de negocio, como un riesgo financiero, como un riesgo del entorno con el que interacciona (Lai & Shad, 2017). En primer lugar, se entiende por riesgo de negocio todo riesgo asociado a la actividad desarrollada por la compañía, y que puede venir dado por la industria y modelo de negocio, por los clientes o proveedores, por los competidores, por la aparición de productos sustitutivos, o por el equipo directivo, entre otros. En segundo lugar, riesgo financiero se refiere a aquél relacionado con la propia estructura de capital, y que por norma general aumenta a medida que crece el porcentaje de deuda frente a *equity* en el balance. Por último, los riesgos de entorno sería una categoría heterogénea que incluiría circunstancias políticas, económicas o sociales que se escapan totalmente del margen de actuación de la compañía.

La combinación de todos estos riesgos sobre la actividad y futuro de una empresa se traduce en un determinado retorno exigido por los agentes que invierten en la misma (Berry et

al., 2014). Evidentemente, cuanto mayor sea el riesgo asumido por el inversor, mayor será la rentabilidad mínima que exija. A los efectos que aquí nos interesan, el riesgo viene determinado en términos porcentuales por el WACC, para el caso de inversores en el conjunto de la empresa; y por el *cost of equity* (K_e) para los accionistas.

Y, en la medida en que los múltiplos son indicadores del valor de una compañía, al menos en términos relativos, parece razonable pensar que sus determinantes sean, igualmente, la rentabilidad, el crecimiento y el riesgo.

En primer lugar, en referencia al EV/EBITDA, definiremos los siguientes indicadores:

$$\begin{aligned} UFCF &= NOPAT - Capex + DD\&A - \Delta NWC \\ &= EBITDA \times (1 - t) \times (1 - D) - Capex + DD\&A - \Delta NWC \end{aligned}$$

$$RR (\% \text{ de reinversión}) = \frac{Capex - DD\&A + \Delta NWC}{NOPAT}$$

$$g = ROIC \times RR$$

Y, por lo tanto, podemos definir el UFCF de la siguiente manera:

$$UFCF = NOPAT - (NOPAT \times RR) = NOPAT \times (1 - RR)$$

Siguiendo a Jullens (2012), definiremos el EV como un DCF en su forma más elemental, que no es más que asumir una generación perpetua de flujos de caja que crecen anualmente a un ritmo constante g (fundamentalmente, considerar que el DCF se presenta como *growing peruity*):

$$EV = \frac{UFCF}{WACC - g}$$

Si definimos:

$$RR = \frac{g}{ROIC}$$

Podemos sustituir y definir el EV como:

$$EV = \frac{NOPAT \times \left(1 - \frac{g}{ROIC}\right)}{WACC - g} = \frac{NOPAT \times (ROIC - g)}{ROIC \times (WACC - g)}$$

$$= \frac{EBITDA \times (ROIC - g)}{ROIC \times (WACC - g)} \times (1 - t) \times (1 - D)$$

$$EV/EBITDA = \frac{ROIC - g}{ROIC \times (WACC - g)} \times (1 - t) \times (1 - D)$$

Así, llegamos a la expresión propuesta por Suozzo et al. (2001), que permite concluir que los determinantes del EV/EBITDA son, por tanto, los mismo que lo son para el valor de empresa: rentabilidad, crecimiento y riesgo.

Por otra parte, poniendo el enfoque esta vez en el P/E, la demostración es similar a la anterior, pero con ciertas peculiaridades. Piénsese que, en este caso, la medida que estamos utilizando para determinar la rentabilidad no es el ROIC sino el ROE; y el riesgo viene representado no por el WACC sino por el Ke. Además, aquí la caja se genera únicamente para los accionistas, no para todos los interesados en la empresa, por lo que en lugar del UFCF debemos utilizar el LFCF.

$$LFCF = NI - Capex + DD\&A - \Delta NWC + \Delta Deuda$$

$$RRe (\% \text{ de reinversión}) = \frac{Capex - DD\&A + \Delta NWC - \Delta Deuda}{NI}$$

$$ge = ROE \times RRe$$

Análogamente al caso anterior, podemos definir el LFCF como:

$$LFCF = NI - (NI \times RRe) = NI \times (1 - RRe)$$

Y expresamos el *Equity Value* como flujos de caja permanentes que crecen a un cierto *g*, y los descontamos con el Ke al momento presente:

$$Eq = \frac{LFCF}{Ke - ge}$$

Así:

$$RRe = \frac{ge}{ROE}$$

$$Eq = \frac{NI \times \left(1 - \frac{ge}{ROE}\right)}{Ke - g} = \frac{NI \times (ROE - ge)}{ROE \times (Ke - ge)}$$

$$P/E = \frac{ROE - ge}{ROE \times (Ke - ge)}$$

De nuevo, logramos probar de esta manera que los determinantes del P/E son los mismos que pasa el EV/EBITDA y que para el valor de empresa: rentabilidad, crecimiento y riesgo (Suozzo et al., 2001).

En cuanto a las medidas de rentabilidad, nos detendremos brevemente a reflexionar acerca de sus componentes. De esta forma, Damodaran (2007) define el ROIC como:

$$\begin{aligned} ROIC &= \frac{NOPAT}{Capital\ Invertido_{t-1}} \\ &= \frac{NOPAT}{Total\ Activos_{t-1} - Pasivo\ Corriente_{t-1} - Caja_{t-1}} \\ &= \frac{NOPAT}{Ingresos} \times \frac{Ingresos}{Capital\ Invertido_{t-1}} \end{aligned}$$

Realmente, la segunda expresión es una descomposición de la primera, y refleja las dos componentes que forman el ROIC. En primer lugar, está el margen operativo, que refleja la capacidad de generación de beneficios operativos (después de impuestos y antes de intereses). La segunda componente, por su parte, es un indicador de eficiencia, y representa la medida en que el capital invertido durante el periodo anterior se concretó en ingresos en el actual. En su conjunto, por tanto, representan la rentabilidad operativa producida por las inversiones en capital.

En cuanto al ROE, de nuevo Damodaran (2007) propone:

$$ROE = \frac{NI}{Patrimonio\ Neto_{t-1}} = \frac{NI}{Ingresos} \times \frac{Ingresos}{Total\ Activo_{t-1}} \times \frac{Total\ Activo_{t-1}}{Patrimonio\ Neto_{t-1}}$$

La segunda expresión es la que se conoce como la descomposición de DuPont, que tiene una gran utilidad para diferenciar los principales motores del ROE: margen de beneficios, rotación de activos y apalancamiento. En primer lugar, el margen de beneficios parece que no necesita demasiada explicación. En segundo lugar, la rotación de activos, comúnmente denominada *asset turnover*, que constituye un medidor de eficiencia y representa la capacidad que tiene una empresa para “activar” su balance; esto es, lograr que los activos de su balance se concreten en ingresos. En tercer lugar, el *Equity multiplier*, que es un indicador que muestra la idoneidad del apalancamiento como forma de aumentar la rentabilidad de los accionistas. En este sentido, un incremento del balance de la compañía que no requiere un mayor desembolso por parte de los accionistas, por aumentarse el capital a base de deuda, se concreta en un crecimiento del volumen de activos (y en teoría, de ingresos y beneficios) del que participan éstos sin haber tenido que aumentar su aportación. Por lo tanto, el apalancamiento, como norma general y en la medida en que no sea excesivo, “multiplica” la rentabilidad del accionista.

La importancia de la discusión del ROIC y el ROE viene motivada porque, en lo que resta de capítulo, se aplicará el marco teórico expuesto a la práctica, y para ello se utilizarán como veremos las respectivas componentes como variables independientes. El riesgo, ya sea en forma de WACC ya sea en forma de K_e , se ha asumido como constante, dada la dificultad de acceso a información histórica en este sentido y la posible similitud entre empresas de la misma industria

Así, a continuación, se estudiará la relación que podría pensarse que existe entre los respectivos determinantes de los múltiplos y el hecho de que una compañía esté sobrevalorada conforme a uno e infravalorada al mismo tiempo de acuerdo con el otro. Nótese, antes de empezar con las distintas propuestas exploradas, que las variables dependientes son, en este caso, variables dicotómicas, y que por lo tanto no podremos acudir a la clásica regresión lineal múltiple, sino que tendremos que sustituirla por la regresión logística, propia de aquellos estudios en los que el dato que se quiere predecir no tiene carácter numérico. La regresión logística es un método que permite cuantificar la asociación que existe entre un resultado y una serie de variables; pero no modela directamente un resultado, sino la probabilidad de pertenecer a una clase o a otra (Walsh, 1987).

Un concepto importante cuando realizamos ejercicios de clasificación por medio de regresión logística es el de *odds*, o la relación de probabilidad que existe entre pertenecer a una clase o a otra. De esta forma, siendo p la probabilidad de pertenecer a la clase 1, definiremos los *odds* de pertenecer a esta clase como:

$$Odds(Y = 1) = \frac{p}{1 - p}$$

Así, el modelo logit se define como:

$$Logit = Log(Odds) = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$$

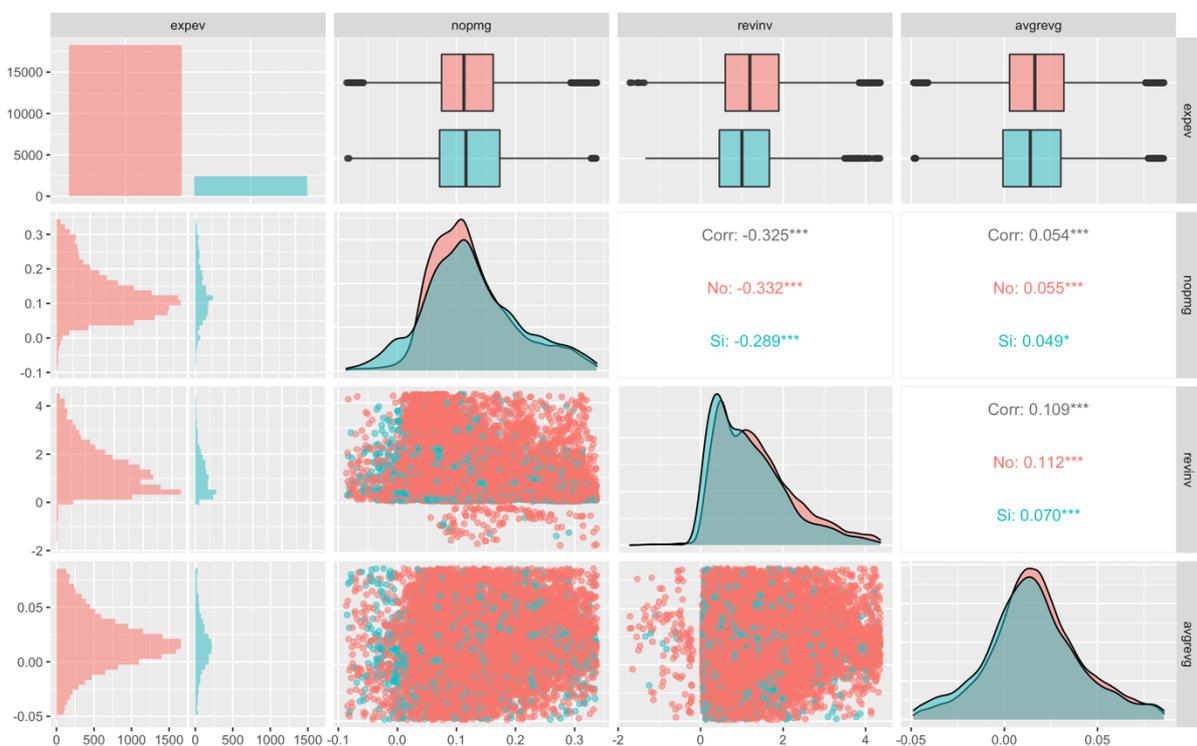
Es importante destacar que los β de la regresión logística no representan efectos marginales de la misma forma lo hacen en la lineal. Por tanto, si tenemos, por ejemplo, que $\beta_1 = 2$, en una regresión lineal ello significaría que cada incremento unitario de la primera variable implicaría un incremento de dos unidades en la variable *target*. No obstante, para el caso de la regresión logística, un $\beta_1 = 2$ significaría que el incremento unitario de la primera variable se traduce en un incremento de dos unidades en el log (odds) de pertenecer a la clase, no en la variable *target* -pues, recordemos, que esta es una variable dicotómica y el problema que abordamos es de clasificación, por lo que no tendría demasiado sentido decir que “y” aumenta en dos unidades cuando realmente solo puede ser 1 o 0-.

En primer lugar, estudiaremos la relación que guardan los determinantes del EV/EBITDA con el hecho de que este múltiplo dé una valoración relativa por encima de la mediana de su sector, mientras que el P/E dé una por debajo. Para ello, del conjunto de datos original, tomaremos únicamente aquellos que son relevantes a estos efectos. Esto es, el margen operativo, los ingresos partido por el capital invertido, y el crecimiento medio de los últimos cuatro trimestres. A continuación, y dado que el objetivo es obtener un modelo que no sirva sólo para explicar nuestro conjunto de datos, sino que tenga un afán universal, deberemos eliminar los valores atípicos. De mantenerlos, correríamos el riesgo de sobreajuste y que nuestro modelo explicase extremadamente bien nuestros datos, pero sin ser extrapolable a otros distintos. Tras todo esto, obtenemos una base de datos con el siguiente aspecto:

expev	nopmg	revinv	avgrevg
No:18294	Min. : -0.08631	Min. : -1.7133	Min. : -0.048727
Si: 2476	1st Qu.: 0.07446	1st Qu.: 0.5782	1st Qu.: 0.002335
	Median : 0.11305	Median : 1.1700	Median : 0.016043
	Mean : 0.12533	Mean : 1.3448	Mean : 0.017472
	3rd Qu.: 0.16323	3rd Qu.: 1.8708	3rd Qu.: 0.031766
	Max. : 0.33746	Max. : 4.3571	Max. : 0.085551

Donde la proporción de observaciones cuyo EV/EBITDA está por encima de la mediana y el P/E por debajo representa el 13,53% de las observaciones.

De manera más visual, podemos estudiar la relación que guardan las variables entre sí a través de la siguiente manera:



Representación visual de la relación que guardan las variables entre sí. Visualización creada con la función "ggpairs" del paquete "GGally".

Aquí podemos realizar alguna observación que puede llegar a cobrar sentido más adelante cuando ahondemos en el modelo. La primera de ellas resulta de fijarse en los *boxplot* de la parte superior, en los que se parece que la distribución de los valores que toman el margen operativo, los ingresos/capital invertido y el crecimiento medio para cada una de las categorías no es demasiado distinto. De hecho, en contra de lo que podría decirnos la intuición, la mediana de estas dos últimas variables es mayor para los casos en que no existe una valoración distinta entre EV/EBITDA y P/E. Por otra parte, también se puede deducir de esta visualización que la

variable ingresos/capital invertido correlaciona negativamente con el margen operativo, puesto que aquellas empresas pertenecientes a industrias que operan con amplios márgenes no tienen necesidad de ser eficientes en los términos aquí expresados. Por último, es interesante que la correlación que existe entre margen operativo-crecimiento e ingresos/capital invertido-crecimiento es más fuerte para aquellos casos en los que tanto EV/EBITDA como P/E coinciden en la posición relativa que tiene la empresa frente a sus comparables.

Tras dividir el conjunto de datos en dos, conjunto de entrenamiento o *training set* (70 %) y conjunto de validación o *test set* (30 %), ejecutamos el siguiente código sobre los datos de entrenamiento:

```
modelo<-glm(formula=expev~.,family=binomial(logit), data=train)
summary(modelo)
```

Y el resultado es el siguiente:

```
Call:
glm(formula = expev ~ ., family = binomial(logit), data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7162  -0.5339  -0.4950  -0.4431   2.3984

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.51710     0.07324  -20.715  < 2e-16 ***
nopmg       -1.02244     0.38084   -2.685  0.007260 **
revinv      -0.22916     0.03070   -7.463  8.44e-14 ***
avgrevg     -3.63578     1.04277   -3.487  0.000489 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 10635  on 14537  degrees of freedom
Residual deviance: 10555  on 14534  degrees of freedom
AIC: 10563

Number of Fisher Scoring iterations: 5
```

Y observamos como el resultado no es el esperado. Es más, es contrario a lo que en un principio podría esperarse: los *log odds* de que una empresa sea cara en EV/EBITDA pero barata en P/E disminuyen a medida que aumenta el margen operativo, la variable ingresos/capital invertido, y el crecimiento. Y es que, además, las tres variables son estadísticamente significativas, en virtud de lo reflejado por el p-valor ($\Pr(>|z|)$). Esto es, si

tomamos la hipótesis nula como cierta, que en este caso consistiría en admitir que no existe relación alguna entre las tres variables predictoras y aquella que nos indica que existe una discrepancia entre los múltiplos en favor del EV/EBITDA, el p-valor nos indica la probabilidad de obtener los resultados que hemos obtenido. Por tanto, observando el resultado obtenido, es más que razonable rechazar la hipótesis nula y concluir que las tres variables son estadísticamente significativas.

Lo que aquí queda reflejado es que, cuanto mayores sean los determinantes de un múltiplo, mayor probabilidad existe de que la valoración de una compañía esté alineada en términos de EV/EBITDA y en términos de P/E. Parece por tanto que lo que viene a expresar este modelo es que lo que realmente es consistente con el análisis fundamental del negocio de una empresa es que su posición de mercado relativa a sus competidores sea la misma sin importar el múltiplo en el que nos fijamos; o, lo que es lo mismo, que el fundamental motor de discrepancia entre múltiplos (por lo menos en aquellos casos en que una empresa está cara en términos de EV/EBITDA) no se encuentra en su crecimiento, en su eficiencia o en su rentabilidad operativa, sino más bien en factores ajenos, probablemente relacionados con circunstancias específicas del mercado en cada momento.

Una forma de corroborar que lo anteriormente expuesto es consistente, es ver qué sucede si, en lugar de tomar el valor de las variables para cada empresa, se comparasen estos con los correspondientes de su industria. Así, se calculó el crecimiento medio del margen operativo, de la variable ingresos/capital invertido y de crecimiento medio, para cada industria en cada año, y se asignó un 1 o un 0 en función de si el valor correspondiente para cada empresa en cada momento era mayor o menor que la media de la industria. Y los resultados fueron similares:

```

Call:
glm(formula = expev ~ avgnopmg + avgg + avgreinv, family = binomial(logit),
    data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.6421 -0.5149 -0.4900 -0.4651  2.3184

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.43934    0.11478  -12.540 < 2e-16 ***
avgnopmg    -1.77925    0.61351   -2.900  0.00373 **
avgg        -0.04758    1.80730   -0.026  0.97900
avgreinv    -0.26467    0.04451  -5.947  2.73e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 10599  on 14602  degrees of freedom
Residual deviance: 10561  on 14599  degrees of freedom
AIC: 10569

Number of Fisher Scoring iterations: 4

```

De nuevo, si bien el grado de significación estadística es distinto, se observa como el hecho de crecer más que el resto de la industria, ser más eficiente u ofrecer una mayor rentabilidad operativa, no justifica las discrepancias entre múltiplos. Parece por tanto que lo que aquí se prueba es que existe un paralelismo entre los fundamentales y la coherencia entre múltiplos, de manera que aquellos que registran unos fundamentales mejores que los de su industria, es esperable que estén por encima tanto en términos de P/E como en términos de EV/EBITDA; y, los que tienen un peor desempeño, estén por debajo en ambas métricas.

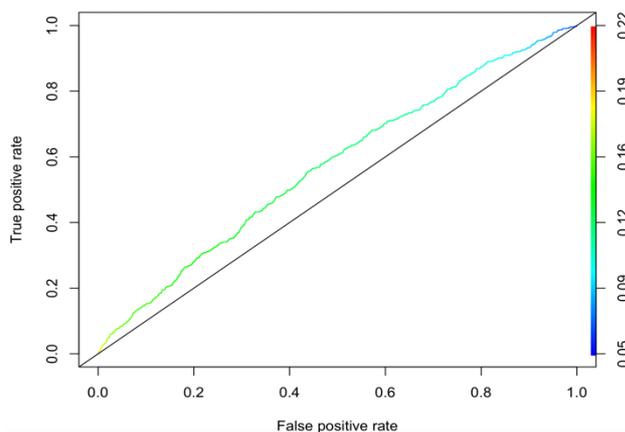
Volviendo a los datos originales -esto es, tomando los valores absolutos en lugar de las medias- estudiaremos a continuación la idoneidad de nuestro modelo, y su capacidad predictiva, sobre *test set*. Para ello, tenemos que fijar un valor de corte. Ello se debe a que las predicciones que hace el modelo son en términos probabilísticos, por lo que habrá que establecer un límite a partir del cual se asigne un grupo a cada predicción. Y en este punto es en el que tenemos que valorar, en términos de sensibilidad y especificidad, si preferimos que nuestro diagnóstico hierre por exceso de unos o por exceso de ceros.

Teniendo una matriz de confusión, que relaciona los valores reales con los predichos (en nuestro caso, una matriz de 2x2, en la que las filas son “Sí” o “No” y se refieren a los valores predichos; y las columnas son “Sí” o “No” y se refieren a los valores reales) Loong (2003) define sensibilidad y especificidad de la siguiente manera:

$$\text{Sensibilidad} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos negativos}} \rightarrow \text{Fracción de verdaderos positivos}$$

$$\text{Especificidad} = \frac{\text{Verdaderos negativos}}{\text{Verdaderos negativos} + \text{Falsos positivos}} \rightarrow \text{Fracción de verdaderos negativos}$$

Por tanto, las medidas de sensibilidad y especificidad nos indican cómo de preciso es el modelo a la hora de predecir aquellos que pertenecen a la categoría 1, en la que los múltiples dan señales contradictorias (sensibilidad); frente a los que pertenecen a la categoría 0, en la que ambos están alineados (especificidad). En consecuencia, existe una contraposición entre sensibilidad y especificidad, de tal forma que en función del valor de corte elegido se estará sacrificando una categoría en favor de la otra. Para visualizar esta relación, es útil acudir a la curva ROC, que para nuestro concreto caso tiene el siguiente aspecto:



Curva ROC del modelo glm (expev ~)

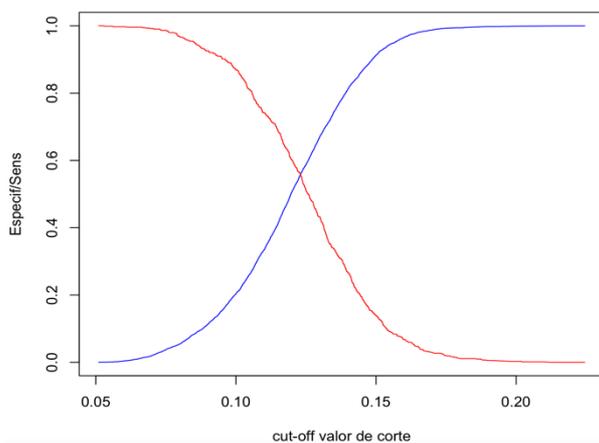
Viendo esta curva, podemos concluir fácilmente que el modelo logrado no es especialmente bueno. En el gráfico se representa, en el eje de ordenadas, la sensibilidad, tal y como se explicó *supra*. En el de abscisas se recoge, por su parte, el número de falsos positivos, que representa la probabilidad inversa a la correspondiente a la especificidad (siendo p la proporción de

verdaderos negativos, denominada especificidad, $1-p$ será la que representemos en la curva ROC). De esta forma, podríamos definir un punto ideal (0,1), que se correspondería con un modelo que predice correctamente el 100 % de los verdaderos positivos y verdaderos negativos; mientras que la bisectriz trazada del punto (0,0) al punto (1,1) reflejaría un modelo que simplemente asignase de manera aleatoria observaciones a una categoría o a otra. Cualquier curva que esté entre medias de ambas sería mejor que la predicción aleatoria; pero, desde luego, cuanto menos se parezca al *random guessing*, mejor será el modelo. Y, en este caso, vemos que se parece bastante, lo que quiere decir que es complicado encontrar un punto de corte en el que tanto especificidad como sensibilidad representen simultáneamente una proporción elevada.

Se puede afinar más en el análisis mediante la introducción del *AUC* (*Area Under the Curve*, o superficie bajo la curva), que es un concepto íntimamente relacionado con la curva ROC por representar el área encerrada por ésta. Nótese que para modelos con gran capacidad

predictiva el AUC estará próximo a 1, por ser el área similar a la del cuadrado formado por los ejes. Por el contrario, la recta de predicción aleatoria encierra un área de 0,5 (Bradley, 1997). En nuestro caso, el AUC es 0,572, lo que significa que no estamos prediciendo mucho mejor que si optásemos por asignar aleatoriamente cada observación a una categoría o a otra. Es cierto, en cualquier caso, que la regresión logit funciona mejor cuando existe una proporción de 0 y 1 más próxima a la paridad que la que observamos en nuestro caso, que como dijimos está en torno al 13,5% y que, por lo tanto, afecta negativamente a las predicciones que se pudiesen hacer de acuerdo con este modelo.

Retomando la cuestión relativa a la determinación del punto de corte como referencia clasificatoria, se puede trazar una curva que represente la relación especificidad/sensibilidad y puntos cortes.



Relación Especificidad/Sensibilidad – Valores de corte

Así, la línea azul representa la especificidad, de manera que el aumento del valor de corte se traduce en un aumento de la proporción de verdaderos negativos propiamente calculados; y, contrariamente, la roja representa la sensibilidad, que guarda una relación inversa con el *cut-off*. A medida que aumentamos el valor de corte, lo que realmente estamos haciendo es clasificar

más observaciones como 0 que como 1, por lo que la probabilidad de predecir correctamente los verdaderos negativos aumenta.

El criterio que se utilice en la fijación del punto de corte dependerá de cuál sea el objetivo del modelo. En el caso paradigmático del modelo que trate de predecir si una persona padece una enfermedad que supone un riesgo grave para su vida, parece razonable que se busque premiar la sensibilidad frente a la especificidad, de manera que se trate de garantizar que los pacientes enfermos son propiamente identificados, y se asuma la negativa de clasificar a un mayor número de personas sanas como enfermas. No obstante, en nuestro caso, parece que no está tan clara cuál es nuestra preferencia. Sería interesante realizar un estudio en el que se incluyese la dimensión temporal, de manera que se analizase qué le pasa el precio de las acciones que están caras en EV/EBITDA, pero baratas en P/E, y tratar de ver si existe alguna posibilidad de arbitraje. De ser así, y resultar que para aquellas empresas en las que se da esta circunstancia existe una variación significativa en el precio de sus acciones, probablemente

también querríamos identificar preferentemente los verdaderos positivos frente a los verdaderos negativos, toda vez que, además, aquéllos representan una proporción reducida respecto a éstos.

En nuestro caso el objetivo es más bien relacionar los fundamentales con la coherencia entre múltiplos, por lo que idealmente estableceríamos un punto de corte que lograra una mayor *performance* clasificatoria en su conjunto. Sin embargo, no se encontró ningún punto que lograra un desempeño mejor que la regla de la mayoría, que consiste en asignar todos los puntos a la categoría mayoritaria (en este caso 0) y que en nuestro modelo asciende al 88,13 %. Fundamentalmente esto lo que nos dice es que, si estableciésemos un punto de corte que motivara la clasificación de todos los valores predichos como 0, estaríamos prediciendo correctamente el 88,13 % de las observaciones. El motivo evidente de ello es la elevada proporción de 0 con respecto a 1, y que permite concluir que por lo general EV/EBITDA y P/E estarán alineados y que, en aquellos casos en que no lo estén, el motivo no radica en los determinantes de los múltiplos.

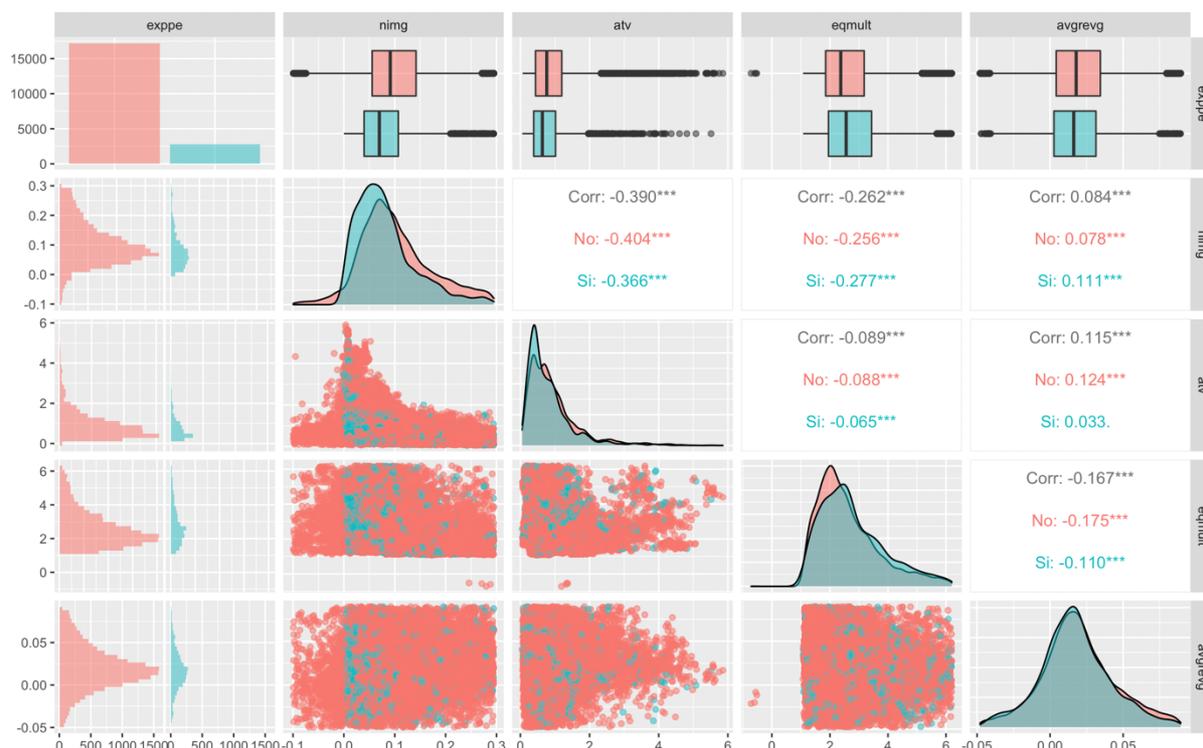
Si fijamos, por ejemplo, 0,145 como valor de corte, la sensibilidad resultante sería de 19,32 % y la especificidad de 86,54 %. La *performance* clasificatoria en su conjunto equivale a predecir correctamente el 78,6 % de los valores, y la matriz de confusión queda de la siguiente manera:

		REAL	
		No	Si
PREDICHO	0	3.688	420
	1	1.804	320

Matriz de confusión

Se procederá a continuación a realizar el análisis análogo, pero esta vez con los determinantes del P/E en lugar de los del EV/EBITDA. Así, se tomó como variable *target* aquella que asigna un 1 a las observaciones cuyo P/E está por encima del de sus competidores, pero su EV/EBITDA por debajo; y, como variables independientes, el margen de beneficios, el *asset turnover* y el *equity multiplier*. En este caso, las observaciones de la variable *target* que tenían asignado un 1 por estar sobrevaloradas en términos de P/E, toda vez se eliminaron los valores atípicos para cubrirnos del riesgo de sobreajuste, representaron el 15,96 % de las observaciones, lo que supone una diferencia de casi 2,5 puntos porcentuales con respecto al caso anterior.

De nuevo se representa visualmente la relación que existe entre variables:



Representación visual de la relación que guardan las variables entre sí. Visualización creada con la función "ggpairs" del paquete "GGally".

En este sentido, podemos realizar una serie de observaciones. La primera de ellas ya nos da ciertas nociones acerca del modelo que se obtendrá, al apreciarse que tanto el margen de beneficios, como el *asset turnover*, como el crecimiento en ingresos, agrupados en torno a la variable *target*, tienen mayor mediana para los casos en que los múltiplos son coherentes entre sí que para aquéllos en los que no lo son. Razonamiento parecido al expresado anteriormente: el impacto de variaciones en estas variables se traduce en el P/E y en el EV/EBITDA de igual manera, sin que tenga un efecto preferente en el P/E sobre el EV/EBITDA. No obstante, sí que parece que el *equity multiplier* afecta de manera diferente. Lo que aquí se aprecia es que la mediana es mayor para aquellos casos en que sí existe una sobrevaloración del P/E. Parece por tanto que un aumento del apalancamiento financiero de la empresa sí que afecta positivamente al P/E frente al EV/EBITDA, de manera que los datos infieren que el endeudamiento, por implicar un mayor volumen de activos a repartir entre el mismo número de accionistas, conduce a una mejor valoración en términos de P/E que en términos de EV/EBITDA.

Otra de las circunstancias que se infiere de esta visualización es, en primer lugar y de manera similar a lo que nos ocurría para el caso anterior, que el margen de beneficios y el *asset turnover* correlacionan de manera de negativa; y, de nuevo, la justificación se halla en el hecho

de que aquellos modelos de negocio generadores grandes márgenes tienen unos requerimientos de eficiencia menores que aquéllos de márgenes más estrechos. Por otra parte, es llamativo observar también que el margen de beneficios está inversamente correlacionado con el *equity multiplier*, algo que tampoco sorprende si pensamos en las consecuencias que tiene sobre el resultado del ejercicio un aumento de la deuda (mayor pago de intereses y, por tanto, menor beneficio). Así mismo, el crecimiento en ingresos se traduce en un mayor crecimiento en beneficios y por tanto mejores márgenes; y, al mismo tiempo, parece que la mayor eficiencia operacional (*asset turnover*) tiene un impacto positivo en el crecimiento en ingresos.

Por último, se puede apreciar que el *asset turnover* tiene una marcada distribución asimétrica a la derecha, que parece que puede venir motivada por el hecho de que, pese a que la mayoría de observaciones oscilan entre el cero y el uno, encontramos un número considerable de empresas que son tremendamente eficientes en la conversión de sus activos en ingresos. La explicación más razonable es que estos puntos se corresponden con empresas de servicios o de nueva economía que, en términos contables, tienen un volumen de activos reducido.

Dicho esto, dividimos el conjunto de datos en datos de entrenamiento y datos de validación, de la misma manera que se hizo anteriormente, y ejecutamos el siguiente modelo:

```
modelo<-glm(formula=expe~.,family=binomial(logit), data=train)
summary(modelo)
```

Obteniendo el siguiente resultado:

```

Call:
glm(formula = exppe ~ ., family = binomial(logit), data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0889  -0.5845  -0.5046  -0.4031   2.7214

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.767899   0.105891  -7.252 4.11e-13 ***
nim          -6.163083   0.417071 -14.777 < 2e-16 ***
atv          -0.571477   0.045287 -12.619 < 2e-16 ***
eqmult       0.008546   0.023618   0.362  0.717
avgrevg     -0.416156   0.989813  -0.420  0.674
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 11255  on 13963  degrees of freedom
Residual deviance: 10902  on 13959  degrees of freedom
AIC: 10912

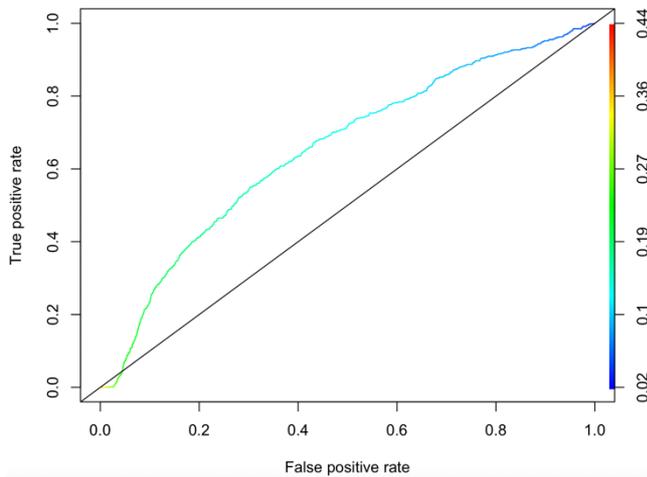
Number of Fisher Scoring iterations: 5

```

Tal y como se anticipó anteriormente, al analizar la visualización, lo que aquí observamos es que tanto el margen de beneficios, como el *asset turnover*, como el crecimiento en ingresos, operan en este modelo de manera opuesta a cómo cabría esperar. Intuitivamente, parece que lo razonable sería que estas variables, por ser los determinantes de los múltiplos, tuviesen un impacto positivo en la probabilidad de que una observación esté sobrevalorada en términos de P/E e infravalorada en términos de EV/EBITDA. No obstante, el modelo resuelve lo contrario: cuanto mayor sea el valor de estas variables, menor es la probabilidad de que una observación sea cara de acuerdo con su P/E. Por tanto, el margen operativo, el *assets turnover* y el crecimiento en ingresos opera en el mismo sentido tanto para el P/E como para el EV/EBITDA, de manera que el aumento de estas variables resulta en una mayor probabilidad de que una observación tenga una misma posición relativa a sus comparables en P/E y en EV/EBITDA.

En este caso, serían significativas únicamente el margen de beneficios y el *assets turnover*. El incremento en *equity multiplier*, por su parte, aumentaría las probabilidades de que una observación estuviese sobrevalorada en términos de P/E. No obstante, el p-valor no nos permite considerarla estadísticamente significativa.

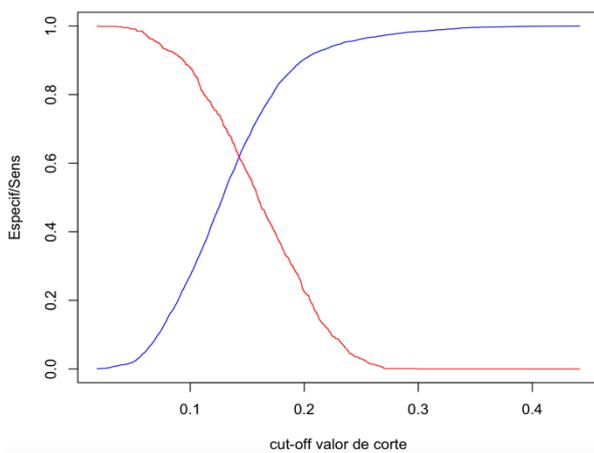
En este caso, la curva ROC sería la siguiente:



Curva ROC del modelo $glm(\text{exppe-})$

Parece que la capacidad predictiva del modelo en este caso es algo mejor que en el caso anterior. Así, el AUC que se obtiene es de 0,653, frente al 0,572 del caso anterior. La mejor *performance* predictiva parece que viene determinada, en este caso, por un mayor número de empresas en las que se da una sobrevaloración del P/E frente al EV/EBITDA, pues como dijimos, el algoritmo logit se ve influenciado por la distribución de las observaciones en cada categoría. Es llamativo, por otra parte, que para niveles de sensibilidad bajos y de especificidad altos, el algoritmo es peor predictor que el *random guess*.

Por otra parte, la regla de mayoría concluye en este caso que, si predecimos que en todas las observaciones el P/E y el EV/EBITDA van de la mano en relación con el resto de las empresas de su industria, estaríamos prediciendo adecuadamente el 86,53 %.



Relación Especificidad/Sensibilidad – Valores de corte

En este caso, sensibilidad y especificidad serían iguales si establecemos como punto de corte alguno que se situase más próximo al 0,1. En cualquier caso, situaremos en este caso el punto de corte en 0,19, de manera que obtenemos una sensibilidad del 29,90 % y una especificidad del 87,39 %.

En conjunto, habremos predicho correctamente el 79,6 % de las observaciones, todavía por debajo de la regla de la mayoría, pero algo mejor que en el caso del EV/EBITDA.

La matriz de confusión queda de la siguiente manera:

		REAL	
		No	Si
PREDICHO	No	4.526	565
	Si	653	241

Matriz de confusión

III.2. Hipótesis 2: Por industria

La segunda de las hipótesis que se podría formular es que esta diferencia de valoración tiene un contenido sectorial; esto es, que para determinados sectores -o, siendo más específicos, industrias- se da una recurrencia en la sobrevaloración de un múltiplo sobre otro, por considerarse que uno de ellos refleja de manera más adecuada el valor que tiene una empresa. Tendría sentido pensar, en cierta medida, que determinadas industrias, en las que el EBITDA o el beneficio neto tienen una mayor importancia a efectos de determinar el valor de una empresa, deberían estar sujetas a una sobrevaloración de aquel múltiplo que contenga la métrica más relevante. No obstante, parece que esta hipótesis no acaba de realizarse. Sin embargo, lo que sí ocurre, como a continuación veremos, es que hay determinadas industrias en las que existe una mayor probabilidad de encontrar sobrevaloraciones tanto por el lado del EV/EBITDA como por el lado del P/E.

Para la realización de este estudio se extrajeron, en primer lugar, aquellas variables dotadas de relevancia: la industria y las dos variables dicotómicas que toman valores de 0 o 1 en función de si la observación está sobrevalorada en EV/EBITDA frente a P/E, o en P/E frente a EV/EBITDA. A continuación, se convirtieron ambas variables a factor, y se ejecutaron los siguientes códigos:

Para EV/EBITDA:

```
modelo<-glm(formula=expev~.,family=binomial(logit), data=industry)
```

Y para P/E:

```
modelo2<-glm(formula=exppe~.,family=binomial(logit), data=industry2)
```

Una vez se obtuvieron los correspondientes *log odds* para cada industria en cada caso, se almacenó la información conjunta en formato HTML con el fin de poder comparar de manera rápida los resultados. Parece que lo más interesante de este ejercicio es centrarse en aquellas industrias consideradas estadísticamente significativas, de manera que la pertenencia a alguna de ellas aumenta o reduce la probabilidad de que exista una discrepancia entre múltiplos. Así:

<i>Predictors</i>	Expensive EV/EBITDA			Expensive P/E		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p-value</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p-value</i>
(Intercept)	0,12	0,09 - 0,17	<0,001	0,13	0,10 - 0,17	<0,001
Industry [Biotechnology]	0,52	0,32 - 0,83	0,007	0,46	0,28 - 0,74	0,002
Industry [Containers & Packaging]	1,78	1,21 - 2,63	0,004	1,64	1,12 - 2,43	0,012
Industry [Distributors]	0,25	0,07 - 0,62	0,008	0,18	0,04 - 0,49	0,004
Industry [Diversified Telecommunication Services]	1,66	0,98 - 2,76	0,052	1,67	1,00 - 2,74	0,047
Industry [Electric Utilities]	1,53	1,10 - 2,16	0,014	1,5	1,08 - 2,11	0,017
Industry [Equity Real Estate Investment Trust (REITs)]	1,88	1,38 - 2,59	<0,001	1,77	1,31 - 2,43	<0,001
Industry [Interactive Media & Services]	2,74	1,29 - 5,50	0,006	1,81	0,79 - 3,81	0,135
Industry [Media]	1,78	1,20 - 2,65	0,004	1,58	1,07 - 2,35	0,023
Industry [Multi - Utilities]	1,66	1,18 - 2,37	0,004	1,52	1,08 - 2,16	0,017
Industry [Trading Companies & Distributors]	0,39	0,19 - 0,74	0,006	0,3	0,14 - 0,60	0,001

Las columnas de la tabla son, en primer lugar, el *Odds Ratio*, que determina la razón que existe entre los *Odds* de que la variable que se trata de estimar sea 1, y los *Odds* de que la misma sea 0. De esta manera, de acuerdo con Szumilas (2010), podríamos decir que:

Odds Ratio = 1 → La industria no produce ningún efecto a la existencia de una sobrevaloración de un múltiplo frente al otro (aleatorio).

Odds Ratio > 1 → Pertenencia a una industria incrementa la probabilidad de sobrevaloración de un múltiplo frente al otro.

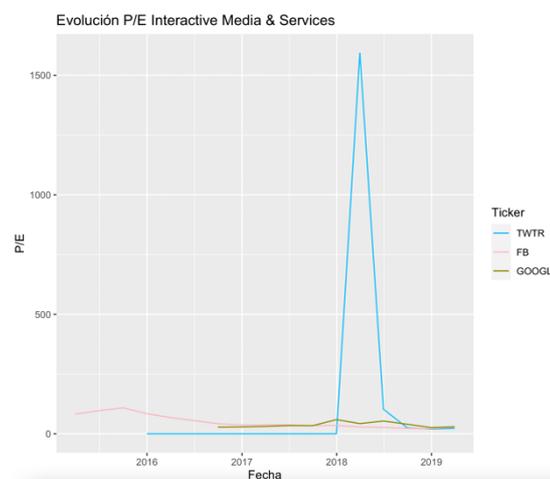
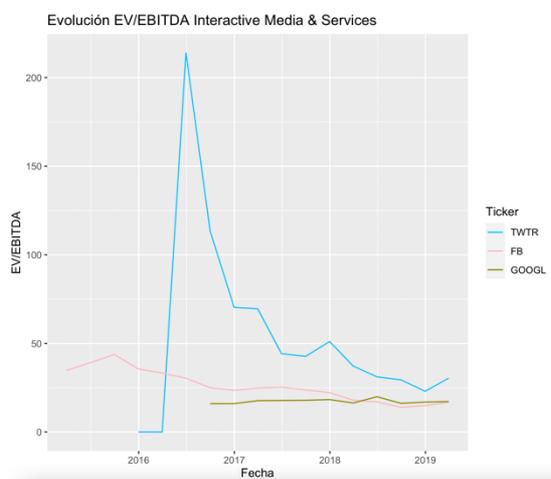
Odds Ratio < 1 → Pertenencia a una industria reduce la probabilidad de sobrevaloración de un múltiplo frente a otro.

En virtud de ello, se observa que para la industrias *Biotechnology*, *Distributors* y *Trading Companies & Distributors*, el hecho de pertenencia a estas industrias conlleva una menor probabilidad de que exista una diferencia valorativa entre los múltiplos, tanto para los casos en que el EV/EBITDA está por encima del de sus competidores, como para los casos en que P/E hace lo propio. En el resto de industrias la relación es la opuesta. Además, es interesante anotar que, en ningún caso, se produce que el *Odds Ratio* esté por encima de 1 en un múltiplo, pero por debajo en el otro. Por tanto, la incremento o reducción de las probabilidades de hallarse sobrevalorado en términos de EV/EBITDA o en términos de P/E se mueve en el mismo sentido para ambas variables en todas las industrias, si bien lo hacen con distinta intensidad.

Por otra parte, la columna *CI* nos indica el intervalo de confianza en el que es probable que se encuentre el *Odds Ratio*, y el p-valor, tal y como se explicó anteriormente, nos da información relativa a la significación estadística.

La primera cuestión que parece interesante abordar se refiere a las industrias relacionadas con el sector de los medios de comunicación. Podemos observar, por un lado, que tanto la industria *Interactive Media & Services* como la industria *Media* muestran una importante diferencia en términos de relevancia estadística en lo que se refiere al EV/EBITDA y al P/E. En el caso de la primera, vemos que, si bien en ambos casos el *Odds Ratio* es mayor que 1, su relevancia estadística únicamente es considerada en términos de EV/EBITDA; y que, además, este ratio es de 2,74 para el primer múltiplo frente a 1,81 para el segundo. Por su parte, para el caso de *Media*, existe también un orden de diferencia entre la relevancia estadística del EV/EBITDA frente al P/E, y el ratio es mayor en aquel múltiplo que en éste. Por lo tanto, lo que parece deducirse aquí es que, para estas compañías, existe una probabilidad mayor de encontrar sobrevaloraciones en EV/EBITDA que en P/E.

En lo que se refiere a *Interactive Media & Services*, el motivo de esta diferencia parece radicar fundamentalmente en dos factores. El primero, el reducido número de empresas pertenecientes a esta industria, pues solo se dispone de datos de Twitter, Facebook y Google. El segundo, las altas valoraciones que reciben algunas de estas compañías, por descontarse un enorme crecimiento futuro incluso antes de que la empresa sea generadora de flujos de caja o de beneficios.

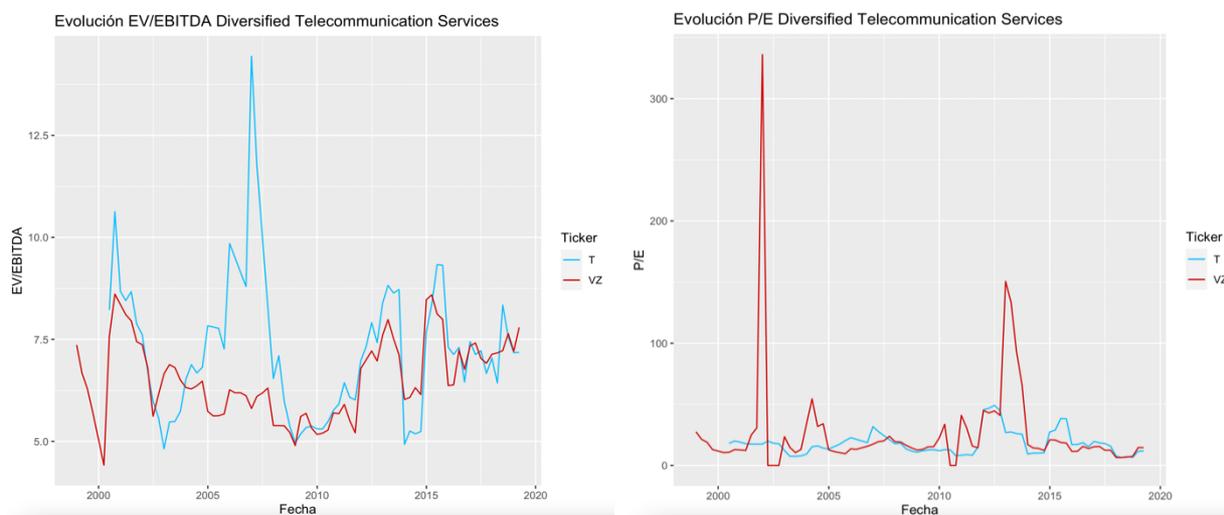


Evolución de los múltiplos de la industria Interactive Media & Services

Parece que el principal motor distorsionador de las coherencia entre valoraciones es, a estos efectos, Twitter. Claramente, sea aprecia que el mercado otorgaba a esta compañía una altísima valoración, que en términos de EV/EBITDA se impulsa en marzo de 2016, cuando el EBITDA entra en terreno positivo; y en términos de P/E, a principios de 2018, cuando el NI hace lo propio. Sea como fuere, lo que parece que aquí se representa es que, pese al salto inicial

que recibe Twitter cuando comienza a registrar unos fundamentales positivos, los trimestres próximos se normaliza y se pone a la altura de sus competidores. No obstante, parece que esto sucede de manera más pronunciada en P/E que en EV/EBITDA, dado que, mientras que en aquel caso llega a estar por debajo de sus comparables, en éste siempre está por encima. Y esta circunstancia, sumada al hecho de que únicamente tenemos tres observaciones, precipita que nuestro modelo reconozca que existe una mayor probabilidad de que, perteneciendo a la industria de *Interactive Media & Services*, el EV/EBITDA esté por encima de la mediana y el P/E por debajo.

En el caso de *Diversified Telecommunication Services* la situación es la contraria, pues la significancia estadística de esta industria es mayor en el caso del P/E que en el del EV/EBITDA. Lo cierto es que la diferencia entre los *Odds ratios* no es demasiada, ni tampoco lo es el p-valor. No obstante, merece la pena detenerse momentáneamente en esta industria. Únicamente encontramos dos empresas, AT&T (T) y Verizon Communications (VZ), por lo que el estudio se reduce a analizar aquellos casos en que la posición relativa de ambas empresas en EV/EBITDA y en P/E sea contraria.

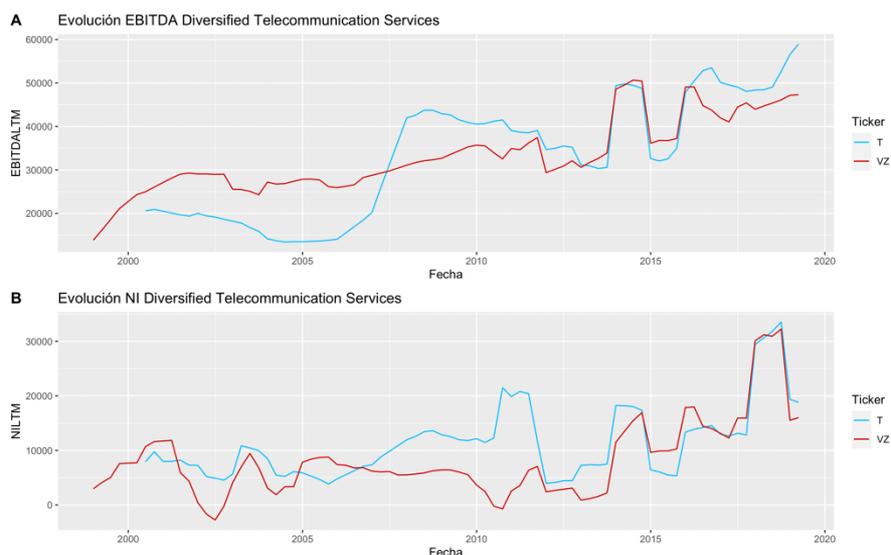


Evolución de los múltiplos de la industria Diversified Telecommunication Services

Las rectas verticales negras delimitan los momentos en que se produce una inversión en la valoración relativa de cada empresa, de manera que la que estaba por encima en, por ejemplo, EV/EBITDA, pasa a partir de este punto a estar por debajo. Se aprecia que, si bien el número de inversiones es el mismo, el momento en el que ocurren es distinto. Además, es llamativo que AT&T parece estar mejor valorada que Verizon la mayor parte del tiempo en términos de EV/EBITDA; pero, sin embargo, en P/E, se da la circunstancia opuesta. Lo lógico,

sin embargo, sería que ambos múltiplos estuviesen solapados, de manera que condujesen a similares conclusiones en torno a qué empresa vale más.

Parece, además, que la evolución de las valoraciones otorgadas por el P/E presenta una mayor volatilidad, apreciándose picos muy marcados y un mayor número de pequeñas inversiones en la valoración relativa de ambas empresas. Una explicación podría estar contenida en la siguiente visualización:



Evolución EBITDA y NI en la industria Diversified Telecommunication Services

Así, la volatilidad en el P/E parece que viene justificada por la propia volatilidad en el resultado del ejercicio. No obstante, viendo la evolución de tanto el EBITDA como el beneficio neto, parece que esta industria tiende a valorar mucho más la obtención de beneficio que la generación de flujos de caja. Esto es, la industria de las telecomunicaciones, especialmente en lo que se refiere a los operadores móviles y de red, se caracteriza por encontrarse en una fase madura con alta competencia, teniendo por tanto un estrecho margen de crecimiento. Por supuesto que hay capacidad de innovación, como en cualquier otro sector, pero quizás en este sentido se parezca más a empresas de suministro, centradas en la captación y mantenimiento de clientes, que en crear nuevos productos o servicios disruptivos.

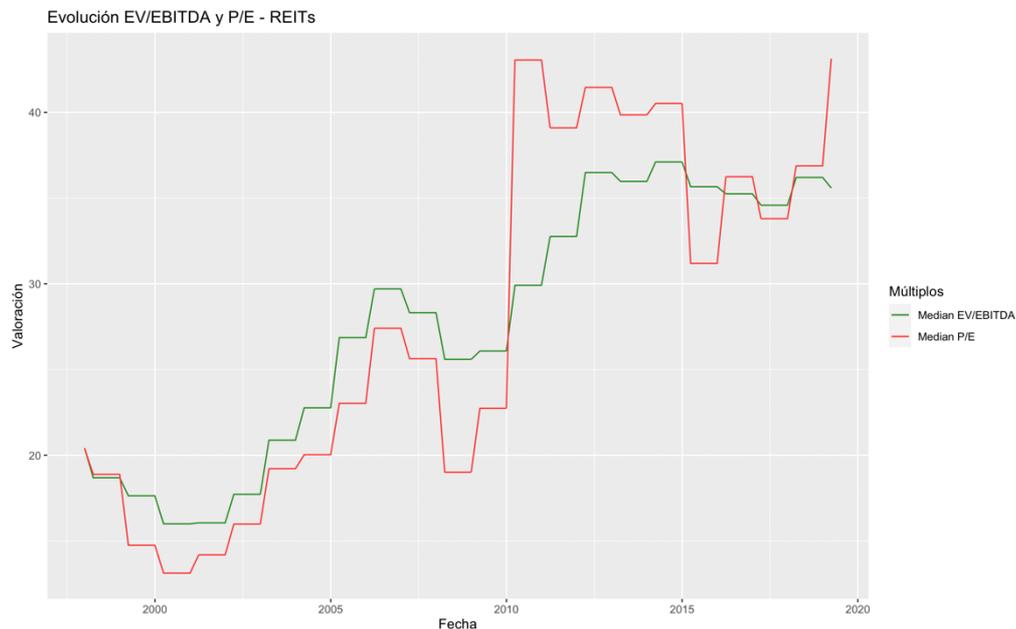
Por tanto, tiene sentido que el inversor de estas empresas sea más averso al riesgo y tenga un mayor interés en recibir dividendos que en ver la empresa crecer. Y esto queda reflejado en las visualizaciones que se muestran. Para el caso de Verizon, el EBITDA muestra una tendencia creciente sostenida durante gran parte del periodo mostrado; no obstante, en el beneficio neto, se aprecian pérdidas en varios momentos y periodos de más o menos prolongados de progresiva reducción del resultado del ejercicio. Y el mercado, sin duda, penaliza este tipo circunstancias.

Y, curiosamente, esta penalización se ve reflejada con más claridad en el gráfico que representa la evolución del EV/EBITDA, dado que éste no está distorsionado por cambios bruscos provocados por fundamentales negativos o de muy reducida dimensión. Así, en este caso, podríamos decir que, si bien la métrica importante de esta industria es el resultado del ejercicio, el EV/EBITDA parece reflejar mejor el valor de la compañía.

No obstante, debe remarcar que este análisis es difícilmente extrapolable al universo de todas las empresas que operan en el sector de las telecomunicaciones, pues el estudio aquí realizado se fundamenta únicamente en dos compañías.

La siguiente de las industrias que vamos a tratar es la de *Equity Real Estate Investment Trust*, en adelante REITs. La pertenencia a esta industria hace que se incrementen los *odds* de estar tanto sobrevalorado en P/E e infravalorado en EV/EBITDA, como sobrevalorados en EV/EBITDA e infravalorados en P/E. No obstante, no deben sacarse conclusiones precipitadas en este respecto.

La representación de los múltiplos medianos a lo largo del tiempo pone de manifiesto una anomalía que, con alta certeza, explica la enorme significación estadística y explica por qué, para esta industria, la elección de los múltiplos no es la adecuada.



Evolución de los múltiplos en REITs

Es extraño, dado que matemáticamente no suele sostenerse, y además es muy poco frecuente observarlo en el mercado, que el EV/EBITDA sea mayor que el P/E. Por lo general, la adición que supone la deuda neta a la capitalización bursátil es bastante menor en términos relativos que todo lo que hay que sumar para llegar del beneficio neto al EBITDA. Por lo tanto,

lo observado en este gráfico viene provocado por una circunstancia que es bastante característica de este tipo de empresas, que es el alto porcentaje que representa la deuda en su estructura de capital. Así, en numerosas ocasiones, los REITs se apalancan para financiar la compra de activos, dado el reducido coste de la deuda y el fácil acceso a la misma, pues el propio activo que se adquiere opera como colateral y por tanto garantía de pago. Además, sucede algo curioso con los datos, y es que un elevado número de observaciones presenta beneficios netos negativos, por lo que el P/E automáticamente se va a cero y, en consecuencia, a efectos prácticos, se estrecha la diferencia entre EBITDA y beneficio. Es por estos motivos por los que se explica que el EV/EBITDA sea mayor que el P/E.

Sea como fuere, esta explicación, por razonable que parezca, no es relevante, ya que, la práctica habitual de valoración de los REITs no suele incluir un análisis relativo en términos de P/E y en términos de EV/EBITDA como en el resto de industrias. Dada la naturaleza de este modelo de negocio, suelen aplicarse, o bien fórmulas valorativas centradas en los activos de la empresa, como el NAV (*Net Asset Value*), por considerarse que el valor en libros de los activos representa de manera precisa su valor de mercado; o métodos de valoración intrínseca, como el DCF o el DDM. Además, el NI suele ajustarse y pasa a denominarse FFO (*Funds From Operations*), que fundamentalmente añade la depreciación y la amortización, así como las pérdidas o ganancias por la venta de activos, al beneficio neto.

Por todo ello, no parece razonable sacar conclusiones acerca de si la valoración de los REITs viene determinada en mayor medida por el P/E o por el EV/EBITDA, porque la realidad es que el precio de sus acciones refleja un determinado valor que surge de métodos alternativos a los múltiplos por comparables.

Por último, y sin ánimo de extender más este apartado del trabajo, destacaré que aquellas industrias cuyo *Odds Ratio* es menor que uno (*Biotechnology, Distributors y Trading Companies and Distributors*) serán las que presenten una mayor coherencia entre múltiplos, de tal manera que la probabilidad de toparnos con una observación cara en EV/EBITDA y barata en P/E (o viceversa) se reduce si dicha observaciones pertenece a estas industrias. A este respecto, por lo tanto, podríamos decir que ambas métricas reflejan propiamente el valor que tiene una empresa en términos relativos, y que por lo tanto sería más complicado encontrar alguna forma de hacer arbitraje entre ellos.

Concluiremos este apartado con una última reflexión. En general, para la mayor parte de industrias, no existe un grado de significación estadística suficiente como para decir que, en función de la industria a la que pertenezca la empresa, existe una mayor o menos probabilidad de que encontrar una divergencia entre EV/EBITDA y P/E, sea en el sentido que sea. No

obstante, hay algunas en las que sí se da esta circunstancia, y se da en distinta medida para cada múltiplo, pero siempre en el mismo sentido (o bien mayor probabilidad o bien menor). Por tanto, para los casos mostrados en la tabla *supra*, la combinación entre los *Odds Ratio* y el p-valor podrá dar una noción de la tendencia que tiene esa industria a sobrevalorar o infravalorar un múltiplo frente al otro; y, por tanto, se podrá utilizar esta información para examinar la medida en que cada uno de los múltiplos refleja el valor relativo de una empresa, y tomar decisiones en consecuencia.

CONCLUSIONES

A lo largo del trabajo, se ha ofrecido una explicación detallada de los múltiplos de valoración y del papel que juegan a efectos de obtener el valor relativo de una empresa. Igualmente, se ha desarrollado el trasfondo matemático que subyace tanto al P/E como al EV/EBITDA, y se ha demostrado la relación fundamental que existe entre ellos. No obstante, pese a todo, se ha comprobado cómo, inevitablemente, los múltiplos se ven afectados por factores que trascienden el análisis puramente fundamental, y que por tanto la coherencia teórica que media entre ellos no se cumple en la totalidad de los casos. Además, se ha podido observar que la conexión que existe entre estas incoherencias y los determinantes no es, precisamente, la que inicialmente cabría haber esperado; y, finalmente, se ha visto que la industria juega un papel considerable en esta cuestión.

Así, la primera de las conclusiones que aquí se exponen es que la realidad nos muestra que el EV/EBITDA y el P/E ofrecen información contradictoria en lo que se refiere a la valoración de las empresas en un número importante de ocasiones. Como se expuso, existe una sobrevaloración del EV/EBITDA y una infravaloración del P/E en un 12,5 % de las ocasiones; y una sobrevaloración en P/E y una infravaloración en EV/EBITDA en un 12,7 % de las ocasiones. En total, en el 25,2 % de las observaciones el EV/EBITDA y el P/E dan señales contradictorias. Por tanto, dado que en un cuarto de las observaciones sucede que nos hallamos ante una contradicción entre vender y comprar, parece que este hecho que se ha venido estudiando es un supuesto frecuente y merecedor de atención.

La segunda de las conclusiones se refiere a que los determinantes, si bien adquieren relevancia a la hora de estimar la probabilidad de que se dé un supuesto de contradicción entre múltiplos, la forma en que lo hacen es distinta a la que inicialmente cabría esperar. Así, aquellas empresas de fuerte crecimiento y con altos niveles de rentabilidad, tienen una tendencia a que sus múltiplos de valoración sean congruentes entre sí. En este sentido, se demostró que, a medida que aumenta el crecimiento y la rentabilidad, se reduce la probabilidad de desacuerdo entre valoraciones relativas. Parece por tanto que la solidez de los determinantes no justifica una sobrevaloración por parte de un múltiplo. Más bien, el modelo elaborado sugiere que aquellas empresas que no crecen tanto o que no son tan rentables son las que, en realidad, presentan una mayor probabilidad de contradicción. De esta conclusión se deriva además otra, y es que los determinantes de ambos múltiplos, bien los del P/E bien los del EV/EBITDA, están alineados y realmente explican la valoración obtenida por una empresa, por ser los casos de

mayor crecimiento y de mayor rentabilidad los que presentan mayor coherencia, y por tanto sugieren valoraciones relativas más sólidas.

Finalmente, se ha probado que la actividad desarrollada por las empresas adquiere relevancia a la hora de recibir señales distintas por parte de P/E y de EV/EBITDA. Y lo hace en ambos sentidos, dado que en algunos casos la pertenencia a una industria reduce la probabilidad de presentar esta anomalía, como es el caso de la biotecnología o de la distribución; y en otros la aumenta, como sucede es las empresas de suministros o en las de media y comunicación. Existe, además, diferencia entre la intensidad con que la pertenencia a una industria afecta a la sobrevaloración en P/E y en EV/EBITDA, pero en ningún caso sucede que la industria aumente la probabilidad de una y disminuya la de la otra. El aprendizaje por tanto que de aquí se puede extraer es que, dependiendo de la industria a la que pertenezca la empresa, tiene sentido darle más importancia a los supuestos de sobrevaloración del EV/EBITDA frente a los del P/E, o viceversa. Además, para aquellas industrias en las que se reduce la probabilidad de contradicción, merece más la pena analizar en profundidad el origen concreto de esa incongruencia, por ser mucho menos habitual y por tener, en consecuencia, potencial de ofrecer información valiosa acerca de la mejor decisión a adoptar.

Por último, se propone como área de estudio futuro el análisis de la evolución temporal de la incoherencia valorativa entre múltiplos. Así, si bien hemos presentado a lo largo de todo este trabajo la diferencia entre múltiplos como una anomalía, lo razonable es que, dada esa anomalía en un momento dado, exista una corrección posterior encaminada a que ambos múltiplos se alineen en sus mensajes. De esta forma, el estudio que viene a continuación debería ocuparse del tiempo durante el que se extiende la contrariedad, así como el origen de la corrección posterior, de manera que si llegado un punto determinadas diferencias entre múltiplos desencadenasen una ordenación en términos valorativos promovida por una subida en el precio de la acción, se estaría dando con una oportunidad de arbitraje en la que merecería la pena profundizar. De la misma forma, si las incongruencias se reparasen con una reducción del valor de mercado hasta que ambos múltiplos tuviesen la misma posición relativa a sus competidores, se estaría descubriendo una voz de alarma que permitiría deshacer posiciones y vender activos antes de que perdiesen su valor.

BIBLIOGRAFÍA

1. Berry, S.G., Betterton, C.E., & Karagiannidis, I. (2014). “Understanding Weighted Average Cost of Capital: A Pedagogical Application”, *Journal of Financial Education*, 40, 115-136.
2. Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7), 1145-1159.
3. Damodaran, A. (2007). Return on capital (ROC), return on invested capital (ROIC) and return on equity (ROE): Measurement and implications. *Return on Invested Capital (ROIC) and Return on Equity (ROE): Measurement and Implications (July 2007)*.
4. Damodaran, A. (2008), Introduction to Valuation. En Fabozzi, F.J. (Ed.), *Handbook of Finance* (1ª ed., Vol. 3, pp. 303 - 309). John Wiley & Sons.
5. Damodaran, A. (2011). *The little book of valuation: how to value a company, pick a stock and profit* (Vol. 34). John Wiley & Sons.
6. De Artiñano, A. (2022). *Fecha de publicación de resultados de las empresas del S&P 500*.
7. Horton, N. J., & Kleinman, K. (2015). *Using R and RStudio for data management, statistical analysis, and graphics*. CRC Press.
8. Jullens, D. (2012). Expensive on P/E; cheap on EV/EBITDA: European Distillers. *UBS Investment Research. The footnotes Compendium*, 60-65.
9. Jullens, D. (2012). Reconciling EV/EBITDA and DCF. *UBS Investment Research. The footnotes Compendium*, 55-57.
10. Kaplan, S. N., & Ruback, R. S. (1995). The valuation of cash flow forecasts: An empirical analysis. *The journal of Finance*, 50 (4), 1059-1093.
11. Koller, T., Goedhart, M., & Wessels, D. (2010). *Valuation: measuring and managing the value of companies* (5ª ed.,). John Wiley & Sons.
12. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5), 1 - 26. doi:<http://dx.doi.org/10.18637/jss.v028.i05>.
13. Kumar, R., & Sukumaran, S. K. (2017). Value drivers in Oil Companies: An Application of Variance Based Structure Equation Model. *Contemporary Management Research*, 13(1), 31-52.
14. Lai, F. W., & Shad, M. K. (2017). Economic Value Added Analysis for Enterprise Risk Management. *Global Business & Management Research*, 9, 338 – 347.

15. Loong, T. W. (2003). Understanding sensitivity and specificity with the right side of the brain. *Bmj*, 327(7417), 716-719.
16. Matthews, G. E. (2012). Valuation Methods in Fairness Opinions: An Empirical Study of Cash Transactions. *Business Valuation Review*, 31(2-3), 55-74.
17. Standard & Poor's (2003). *Standard & Poor's Compustat User's Guide*, The McGraw-Hill Companies.
18. Suozzo, P., Cooper, S., Sutherland, G., & Deng, Z. (2001). Valuation multiples: A primer. *UBS Warburg: Valuation and Accounting*, 1, 1-47.
19. Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian academy of child and adolescent psychiatry*, 19(3), 227.
20. Walsh, A. (1987). Teaching Understanding and Interpretation of Logit Regression. *Teaching Sociology*, 15(2), 178-183.
21. Wickham H, François R, Henry L, Müller K (2022). *dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>
22. Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from <https://ggplot2.tidyverse.org>.
23. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.

ANEXOS

ANEXO 1: WEB SCRAPING DE YAHOO FINANCE

```
#web scrapping from yahoo finance

#install libraries

install.packages('data.table')
install.packages('tidyr')
install.packages('tidyverse')
install.packages('XML')
install.packages('xml2')
install.packages('rvest')
install.packages('httr')
install.packages('sqldf')

library(data.table)
library(tidyr)
library(tidyverse)
library(XML)
library(xml2)
library(rvest)
library(httr)

library(readxl)
sp_500_stocks <- read_excel("Documents/TFG Analytics/CRSP/sp_500_stocks.xlsx")
View(sp_500_stocks)

#function

getSecInd = function(x)
{
  url = paste0('https://finance.yahoo.com/quote/',x,'/profile?p=',x,"")
  a <- suppressWarnings(read_html(url))
  Industry <- a %>% html_nodes('div p span') %>% .[4] %>% html_text()
  Industry <- as.data.frame(Industry)
  Sector <- a %>% html_nodes('div p span') %>% .[2] %>% html_text()
  Sector <- as.data.frame(Sector)
  Ticker <- as.data.frame(x)
  colnames(Ticker) <- "Ticker"
  cbind(Ticker,Sector,Industry)
}

tickers <- data.frame(x = sp_500_stocks$Ticker)
tickers <- distinct(tickers)
tickers <- tickers$x
```

tickers

```
ALL <- lapply(as.list(tickers),function(x){
```

```
  tmp<-try(getSecInd(x))
  if(!inherits(tmp,'try-error'))
    tmp
```

```
})
```

```
ALL <- ALL[lapply(ALL,length) > 0 ]
```

```
ALL <- do.call(rbind,ALL)
```

```
ticker_industry <- ALL
```

```
#see the missing elements in ALL
```

```
require(sqldf)
```

```
sp500dbtickers <- data.frame(sp_500_stocks$Ticker)
```

```
scrapticker <- data.frame(ticker_industry$Ticker)
```

```
missing_tickers <- sqldf('SELECT * FROM sp500dbtickers EXCEPT SELECT * FROM  
scrapticker')
```

```
view(missing_tickers)
```

```
missing_tickers_clean<-data.frame(missing_tickers[-c(1,16,38),])
```

```
view(missing_tickers_clean)
```

```
nrow(missing_tickers_clean) #missing 39 rows
```

ANEXO 2: REGRESIÓN LOGÍSTICA PARA EV/EBITDA

```
#create new database with key variables
drv<-data.frame(dt[,c('expev','nopmg','revinv','avgrevg')])
summary(drv)

#remove outliers
outlier_revinv <- boxplot.stats(drv$revinv)$out
boxplot(drv$revinv, main="Revenue/Invested Capital", boxwex=0.1)
print(outlier_revinv)
drv <- drv[-which(drv$revinv %in% outlier_revinv),]
summary(drv$revinv)

outlier_nopmg <- boxplot.stats(drv$nopmg)$out
boxplot(drv$nopmg, main="Operating Marign", boxwex=0.1)
print(outlier_nopmg)
drv <- drv[-which(drv$nopmg %in% outlier_nopmg),]
summary(drv$nopmg)

outlier_avgrevg <- boxplot.stats(drv$avgrevg)$out
boxplot(drv$avgrevg, main="Avg. Revenue g", boxwex=0.1)
print(outlier_avgrevg)
drv <- drv[-which(drv$avgrevg %in% outlier_avgrevg),]
summary(drv$avgrevg)

drv$expev<-factor(ifelse(drv$expev==1,"Si","No"))
summary(drv)

ggpairs(drv,aes(color=expev,alpha=0.3))

set.seed(1)
trainIndex<-sample(1:nrow(drv), nrow(drv)*0.7, replace=FALSE )

train<-drv[trainIndex,]
test<-drv[-trainIndex,]

#### full model
modelo<-glm(formula=expev~.,family=binomial(logit), data=train)
summary(modelo)

##### predictions
test$pred<-predict.glm(modelo, newdata=test, type="response")

### cut off
test$predclas<-factor(ifelse(test$pred>=0.145,"Si","No"))
summary(test)

matrizconfusion<-table(PREDICHO=test$predclas,REAL=test$expev)
matrizconfusion
```

```

### performance measures
total<-round( (matrizconfusion[1,1]+matrizconfusion[2,2])/sum(matrizconfusion),3)
paste("casos correctos", total)

###Majority - Benchmark
table(test$expev)/nrow(test)*100

##### Sensitivity
sensibilidad1<-matrizconfusion[2,2]/(matrizconfusion[1,2]+matrizconfusion[2,2])*100
sensibilidad1

##### Specificity
especificidad<-matrizconfusion[1,1]/(matrizconfusion[1,1]+matrizconfusion[2,1])*100
especificidad

##### ROC Curve

install.packages("ROCR")

library(ROCR)

pred1<-prediction(test$pred,test$expev)

perf1<-performance(pred1,"tpr", "fpr")

#ROC
plot(perf1, colorize=TRUE)
abline(a=0,b=1)

#AUC
perf2<-performance(pred1,"auc")
perf2@y.values[[1]]

#Sensitivity-Specificity-cut off curve
perf3<-performance(pred1,"sens", "spec")
plot(perf3@alpha.values[[1]],perf3@x.values[[1]], xlab="cut-off valor de corte",
ylab="Especif/Sens", type="l", col="blue")
lines(perf3@alpha.values[[1]],perf3@y.values[[1]], col="red")

mosaic(matrizconfusion,shade=T,colorize=T,
gp=gpar(fill=matrix(c("green3", "red2", "red2", "green3"),2,2)))

```