

## LA ELECCIÓN DE BASE DE DATOS Y SUS EFECTOS SOBRE LA INVESTIGACIÓN CONTABLE EMPÍRICA EN ESPAÑA

*Database choice and its effects over empirical accounting research in Spain*

**Juan Manuel García Lara**<sup>1</sup>  
Universidad Carlos III de Madrid

**Beatriz García Osma**  
Universidad Autónoma de Madrid

**Belén Gill de Albornoz Noguera**  
Universidad Jaime I de Castellón

### RESUMEN

La escasa disponibilidad de datos con los que llevar a cabo estudios empíricos es uno de los problemas que tradicionalmente han sufrido los investigadores en el área de contabilidad y finanzas corporativas. Sin embargo, en los últimos años han aparecido diversas bases de datos que ofrecen información contable y financiera sobre empresas de todo el mundo. En este trabajo se analiza si la elección de la base de datos por parte del investigador afecta a los resultados de la investigación contable relacionada con el mercado de capitales. Los resultados obtenidos de una simple adaptación del modelo de Ohlson (1995) para empresas españolas muestran diferencias significativas dependiendo de la base de datos elegida para extraer la muestra analizada. Si se utilizan únicamente las observaciones comunes entre las bases de datos, las diferencias desaparecen, poniendo de manifiesto que las diferencias se deben a la distinta cobertura de empresas de cada base de datos.

### Palabras Clave

Bases de datos, Investigación empírica, España, Datos contables

---

<sup>1</sup> Dirección para correspondencia [jmglara@emp.uc3m.es](mailto:jmglara@emp.uc3m.es)

Este trabajo ha sido presentado en el XIII congreso de AECA (Oviedo, 2005). Los autores agradecen las sugerencias de dos revisores anónimos que han contribuido a mejorar el trabajo. Juan Manuel García Lara y Beatriz García Osma agradecen la financiación de la Cátedra Carlos Cubillo de AECA así como del proyecto de investigación del Ministerio de Ciencia y Tecnología "La convergencia contable como factor de competitividad empresarial y de desarrollo de los mercados bursátiles" (SEC2002-04608-C02-02) y del proyecto de investigación INCENNOR (SEJ2005-08644-C02-01) financiado por el Ministerio de Educación y Ciencia y por fondos FEDER. Juan Manuel García Lara agradece igualmente la financiación del proyecto SEJ2004-08176-C02-02/ECO.

### **ABSTRACT**

*One of the problems faced by researchers to undertake empirical accounting studies in Spain has been that data were difficult to obtain. However, in recent years several databases, offering what they claim to be comprehensive worldwide accounting and financial data, have appeared. In this study we analyse whether database choice affects the results of market based accounting research in Spain. We empirically analyse the relation between accounting and market data using a simplification of the Ohlson (1995) model, retrieving Spanish data from seven databases. The results we obtain change significantly across the different databases. If we use the common observations across all databases then the differences in the results disappear, highlighting that the differences across databases are mainly attributable to a different firm coverage.*

### **Key Words**

*Databases, Empirical research, Spain, Accounting data*

## **1. INTRODUCCIÓN**

El creciente interés por la investigación empírica en contabilidad y finanzas corporativas observado en España durante la última década ha ido paralelo a la creación de diversas bases de datos que, desde comienzos de la década de los 90, ofrecen lo que las compañías que las comercializan denominan una “cobertura completa de información contable y financiera para empresas de todo el mundo”. Normalmente, los trabajos de referencia en cada línea de investigación sientan precedente sobre cuáles son las bases de datos a utilizar, facilitando las comparaciones entre los resultados obtenidos. Este es el caso de los estudios realizados en muestras estadounidenses y del Reino Unido, en los que *Compustat / CRSP* y *DataStream* son, respectivamente, las fuentes de datos utilizadas por excelencia. No ocurre lo mismo en la investigación en el ámbito internacional o centrada en países concretos, como España, donde la aún incipiente tradición investigadora y la existencia de un amplio abanico de bases de datos no han propiciado que la literatura empírica se decante por una única fuente de información. Las diferencias potenciales existentes entre las bases de datos disponibles dificultan la extrapolación de los resultados, lo que sin duda tiene implicaciones para el desarrollo futuro de la investigación.

En el contexto internacional, algunos trabajos previos han expresado cierta preocupación por los posibles efectos que las diferencias existentes entre las distintas bases de datos tienen sobre los resultados de los estudios empíricos. Por ejemplo, Ball *et al.* (2000) sugieren que las diferencias entre sus resultados y los obtenidos por Pope y Walker (1999) para empresas del Reino Unido podrían deberse al uso de distintas fuentes de información<sup>2</sup>. Sin embargo, ningún trabajo previo ha estudiado directamente cómo la elección de la base de datos de la que se obtiene la muestra analizada afecta a los resultados de la investigación empírica. Este trabajo

---

<sup>2</sup> Ball *et al.* (2000) utilizan *Global Vantage* y Pope y Walker (1999) *DataStream*. Otros trabajos previos investigan más en concreto las diferencias entre los datos contenidos en diversas bases de datos para EEUU (Courtenay y Keller, 1994, Guenther y Rosean, 1994, Kern y Morris, 1994 o Elton *et al.*, 2001).

cubre, al menos en parte, la laguna existente en la literatura, aportando evidencia sobre algunas de las diferencias existentes entre las distintas bases de datos utilizadas en la investigación empírica en contabilidad, y apuntando algunas consecuencias derivadas de dichas diferencias. En particular, este estudio se centra en el contexto español, donde, a diferencia de otros países con mayor tradición investigadora, no se observa en los trabajos publicados una homogeneidad en cuanto a las fuentes de información utilizadas para llevar a cabo trabajos de carácter empírico.<sup>3</sup> Para ello, se plantea un ejercicio de investigación habitual en la literatura contable utilizando alternativamente las muestras proporcionadas por las siete fuentes de información empresarial, contable y financiera, con cobertura internacional más conocidas: *Company Analysis*, *DataStream*, *Extel Financials*, *Global Vantage (Compustat Global)*, *BvD Osiris*, *Thomson Financials* y *WorldScope*. En particular, se plantea la estimación de una sencilla adaptación del modelo de Ohlson (1995), en el que la capitalización bursátil se explica en función del resultado contable y del valor contable de los fondos propios, utilizando las muestras de empresas españolas que, para el periodo 1990-1999, ofrecen estas siete bases de datos alternativas.

En un primer análisis descriptivo se pone de manifiesto que las siete muestras difieren sustancialmente, lo que explica que, en la estimación del modelo de Ohlson con las distintas muestras, los coeficientes y medidas de bondad del ajuste sean también significativamente distintos, asimismo, se observa una relación inversa entre el número de observaciones y el tamaño medio de las empresas cubiertas por las bases de datos, lo que sugiere que el uso de las bases de datos con menor número de observaciones puede introducir sesgos de tamaño y supervivencia en las muestras obtenidas. En este sentido, los resultados del trabajo ponen de manifiesto que la elección de la base de datos afecta significativamente a los resultados de la investigación planteada. Por otro lado, en un análisis adicional en el que las muestras se limitan a las observaciones comunes a todas las bases de datos las diferencias en los resultados de la estimación del modelo desaparecen, lo que nos permite concluir que la dispar cobertura ofrecida por las bases de datos consideradas es la causa de las divergencias en los resultados, y que la información disponible para las empresas comunes es similar, no viéndose afectada por errores o diferencias en la definición de las variables.

Este estudio contribuye a la literatura empírica en contabilidad y finanzas corporativas al poner de manifiesto que la elección de la base de datos a utilizar no es una cuestión baladí. La evidencia empírica aportada sugiere que las distintas bases de datos que actualmente se encuentran a disposición de los investigadores proporcionan visiones distintas de la población empresarial española. Esta circunstancia tiene dos implicaciones fundamentales: (1) impide una comparación directa de los resultados de distintos estudios si éstos utilizan fuentes de información distintas para obtener las muestras analizadas; y (2) invalida en cierto modo la generalización de las conclusiones de la investigación, que únicamente sería posible si la muestra objeto de estudio es representativa de la población que pretende estudiarse.

El resto del trabajo se ha estructurado como sigue. En el apartado siguiente se expone el problema de la elección de base de datos y se describen las principales bases de datos empresariales que en el ámbito internacional están a disposición de los investigadores

---

<sup>3</sup> El trabajo de Arce (2001) comenta de forma descriptiva algunos aspectos de las bases de datos *Amadeus*, *Global Vantage* y *Extel Company Analysis*.

interesados en llevar a cabo estudios empíricos en la línea de investigación contable relacionada con el mercado de capitales. Seguidamente, en el apartado tercero se presenta el diseño metodológico del análisis empírico desarrollado. En el apartado cuarto se muestran los resultados y, finalmente, en el quinto apartado se exponen las principales conclusiones derivadas de los mismos.

## 2. LA ELECCIÓN DE BASE DE DATOS

Los investigadores interesados en llevar a cabo trabajos empíricos en la línea de investigación contable relacionada con el mercado de capitales disponen de varias opciones para recopilar la información contable y de mercado que necesitan. Entre las bases de datos más conocidas en el ámbito internacional se encuentran *Company Analysis*, *DataStream*, *Extel Financials*, *Global Vantage (Compustat Global)*, *BvD Osiris*, *Thomson Financials* y *WorldScope*. La elección entre estas fuentes alternativas de información se basa en varios criterios: (1) disponibilidad de la base de datos en la institución a la que pertenece el investigador; (2) disponibilidad de las variables necesarias para llevar a cabo el estudio planteado; (3) práctica común en la investigación previa; y (4) cobertura de las bases de datos alternativas, tanto en cuanto al número de empresas como al periodo de tiempo para el que ofrecen información.

El primero de los criterios señalados, que viene marcado por las limitaciones presupuestarias, es especialmente importante. Debido a sus restricciones presupuestarias, las instituciones académicas a menudo adquieren sólo una base de datos, limitando con ello incluso el tipo de investigación que puede llevarse a cabo. El resto de criterios cobra importancia si se dispone de más de una base de datos, pero en realidad, la propia decisión institucional de qué base de datos suscribir dadas sus restricciones presupuestarias, debería considerar dichos criterios. Cuando se analizan muestras de empresas de algunos países, como Estados Unidos y Reino Unido, la base de datos a utilizar viene marcada por la propia tradición investigadora. Así, la gran mayoría de estudios empíricos publicados en las revistas científicas de prestigio que analizan muestras norteamericanas y del Reino Unido utilizan respectivamente *Compustat* y *DataStream* como fuentes de información contable. Este no es el caso de la investigación basada en muestras internacionales o de países concretos con menor tradición investigadora, como España.

Entre los estudios internacionales más relevantes que incluyen a España en la muestra que analizan se encuentran Leuz *et al.* (2003), que usan *WorldScope*, Raonic *et al.* (2004), que usan *DataStream*, y Bushman *et al.* (2005), que usan *Global Vantage*. En cuanto a trabajos que únicamente se centran en muestras de empresas españolas podemos citar a Arcas y Rees (1999), Giner y Reverte (2001), Arce y Mora (2002) o García Lara y Mora (2004), que utilizan diversas versiones de *Company Analysis / Extel*, y Gallizo *et al.* (2003), que utilizan la base de datos *Business Accounts Harmonized Data Bank (BACH)*, menos conocida internacionalmente. Estos ejemplos ponen de manifiesto la existencia de cierta heterogeneidad en las bases de datos utilizadas para llevar a cabo estudios empíricos sobre contabilidad y mercado de capitales con datos españoles.

A continuación describimos las principales características de las distintas bases de datos que ofrecen información contable y financiera de empresas cotizadas de todo el mundo, también españolas. Todas ellas dicen ofrecer la misma información: lo que denominan una “completa cobertura de empresas a escala mundial”.

Debido a que, como se ha señalado previamente, la tradición en investigación empírica es mucho mayor en el Reino Unido que en otros países europeos, la base de datos comúnmente utilizada en los trabajos desarrollados con muestras de este país, *DataStream*, es una de las más conocidas en la literatura internacional. Esta base de datos ofrece información contable no sólo de empresas del Reino Unido sino también de compañías de otros países, tanto desarrollados como emergentes, así como cotizaciones de acciones y valores de renta fija, índices, tipos de cambio, indicadores macroeconómicos y derivados. Una cuestión a tener en cuenta por parte de los investigadores en años sucesivos respecto a esta base de datos es que, como consecuencia de la adquisición de *Primark*, antiguo propietario de *DataStream*, por la Corporación *Thomson*, ésta ha dejado de actualizar la información contable de *DataStream* y, a partir de 2004, dicha información proviene de *WorldScope*.

*Global Vantage* es el nombre con el que se conocen los archivos internacionales de *Compustat*, la base de datos de compañías norteamericanas y canadienses comercializada por *Standard & Poor's*. La información de *Global Vantage* puede ser recopilada a través de *Research Insight*<sup>4</sup> o de otras plataformas independientes como *WRDS*<sup>5</sup>. Una característica destacable de esta base de datos es que la información contable que proporciona está estandarizada de acuerdo con las definiciones establecidas por *Standards & Poor's*, compatibles con la normativa contable estadounidense. Por tanto, la información no se puede obtener tal y como las empresas la reportan en sus Cuentas Anuales. La base de datos está formada por cuatro archivos: (1) *Industrial / Commercial File*, que contiene más de 200 variables con información de las Cuentas Anuales de compañías industriales y comerciales; (2) *Financial Services File*, que contiene aproximadamente 500 variables de las Cuentas Anuales de bancos, compañías de seguros, inmobiliarias y otros sectores financieros. (3) *Issue File*, que proporciona información de mercado, incluyendo cotizaciones mensuales, tanto de compañías individuales como de índices, dividendos o beneficios por acción; y (4) *Currency File*, que incluye información sobre tipos de cambio mensuales de diversas monedas. Los dos primeros archivos (*Industrial / Commercial* y *Financial Files*) cubren más de 13.000 compañías cotizadas en 80 países. Según *Standard & Poor's* la cobertura de *Global Vantage* representa más del 90% de la capitalización bursátil mundial, incluyendo el 96% de los mercados europeos. La cobertura histórica alcanza los 12 años. Además, la información de mercado cubre más de 90 índices de mercados locales y 110 monedas.

*WorldScope*, *Thomson Financials*, *Extel (Financials)* y *Company Analysis* son productos actualmente controlados, al igual que *DataStream*, por la división financiera de *Thomson Corporation*.

<sup>4</sup> *Research Insight* es un programa diseñado específicamente por *Standard & Poor's* para realizar búsquedas de información de sus bases de datos. También es posible elaborar diversos tipos de informes utilizando este software.

<sup>5</sup> *Wharton Research Data Services*, *WRDS*, es una plataforma desarrollada en la Universidad de Pensilvania que permite suscribir el acceso a diversas bases de datos, y que facilita de gran manera la obtención de la información.

La compañía inglesa *Barron and Smithers* es quien diseña y elabora la base de datos *Extel Financials*, que actualmente comercializa *Thomson* básicamente a través de dos plataformas: *Company Analysis* y *Thomson One Banker*. La base de datos *Extel Financials* contiene información contable de 26.500 compañías de más de 55 países distintos desde 1985. Al contrario que otras bases de datos, como *WorldScope* o *Global Vantage*, *Extel Financials* proporciona la información contable tal y como ésta es reportada por las compañías. Entre las distintas bases de datos analizadas ésta es la única que describe detalladamente el proceso de recopilación e introducción de la información<sup>6</sup>. En principio, existen aproximadamente 1.450 variables contables, aunque en realidad sólo una proporción de las mismas está disponible para todas las compañías. Por ejemplo, para el Reino Unido, donde la normativa requiere mayor grado de detalle, el número de variables disponible es mayor que para otros países, como España. La información se encuentra organizada en forma piramidal, de modo que se muestran las relaciones de unas variables con otras.

*Company Analysis* es un producto creado en 1995 por Mike Barron y Jeremy Smithers y fue específicamente pensado como un potente instrumento para manipular la información contenida en la base de datos *Extel Financials*. Se trata de un producto atractivo para el investigador por cuanto que hace muy asequible, en términos de tiempo y esfuerzo, la recopilación de grandes cantidades de información. La base de datos ligada a este producto se actualiza semanalmente mediante CD.

La información de *Extel* era anteriormente comercializada por el *Financial Times* bajo la denominación “*Financial Times Extel Company Analysis*”, y posteriormente por *Primark* como “*Primark Extel Company Analysis*”, siendo accesible únicamente a través de *Company Analysis*. Después de que *Primark* fuese adquirida por *Thomson*, la información de *Extel* se ofrece también a través de la plataforma *Thomson One Banker*. En teoría, por tanto, la información obtenida de *Extel Financials* utilizando *Thomson One Banker* y *Company Analysis* debería ser la misma. Sin embargo, como se pone de manifiesto más adelante en el trabajo, las muestras obtenidas para España a partir de estas dos fuentes difieren sustancialmente, y el caso español no es un caso aislado. Las muestras disponibles a través de *Company Analysis* y a través de *Extel Financials* en *Thomson One Banker* difieren también para otros países.

Por su parte, los responsables de la base de datos *WorldScope* también la ofertan como única fuente necesaria para obtener información contable y de mercado detallada (más de 1.500 variables) de más de 40.000 compañías cotizadas (31.000 activas y 9.000 inactivas) en más de 50 mercados mundiales, tanto desarrollados como emergentes<sup>7</sup>, con hasta 20 años de cobertura histórica, incluyendo información tanto anual como trimestral. Los datos contables de *Worldscope*, que pueden ser actualizados diariamente, se pueden obtener a través de *Thomson One Banker*; *Global Access*, *DataStream*, FTP (en formato ASCII), archivos personalizados y

<sup>6</sup> En este sentido, se señala que “...los profesionales que recopilan la información son normalmente graduados con background en contabilidad...siendo de aproximadamente siete horas el tiempo invertido en introducir la información de una compañía, mucho más cuando se trata de un caso complejo. Además, después de la introducción inicial de los datos, se llevan a cabo más de 150 rutinas de validación para asegurar que el editor no ha cometido ningún error. Si los tests anteriores se superan con éxito, la información pasa a un editor senior que la vuelve a comprobar. Además, existe un equipo encargado de realizar pruebas constantemente sobre la calidad de la información contenida en la base de datos (*Quality Assurance team*) para asegurar que todos los posibles errores son identificados”.

<sup>7</sup> Esta cobertura representa, según la información ofrecida por la propia base de datos, más del 96% de la capitalización bursátil mundial.

otras plataformas independientes, como *WRDS*. Esta base de datos obtiene la información directamente a partir de las Cuentas Anuales y los comunicados de prensa de las compañías, de los archivos oficiales de los organismos reguladores de los distintos mercados, o de otros canales de comunicación. Al igual que *Global Vantage*, la información se proporciona en un formato estandarizado.

La base *Thomson Financials* combina, de un modo poco transparente, información de varias bases de datos (*Compustat* y *Worldscope* entre otras). La información de esta fuente puede obtenerse a través de *Thomson One Banker*. *Thomson One Banker* permite descargar directamente la información a través de su página Web y también proporciona una herramienta para descargar información a Microsoft Excel.

Finalmente, *BvD Osiris* es propiedad de *Bureau van Dijk Electronic Publishing, SA*, y ofrece información contable y financiera de empresas de más de 120 países. También *Osiris* dice cubrir todas las compañías cotizadas del mundo<sup>8</sup>. La información de *Osiris* puede recopilarse *on-line*, mediante la versión en DVD o a través de plataformas como *WRDS*. Además de la información contable, que obtiene fundamentalmente a través de la empresa *WorldVest Base*, esta base de datos proporciona información sobre la estructura de propiedad de las compañías, noticias, *ratings*, cotizaciones o predicciones de analistas. Igual que *Extel Financials*, *Osiris* proporciona la información contable tal y como es reportada por la compañía, si bien algunas variables se ofrecen también en un formato estandarizado, al objeto de permitir comparaciones entre países. *BvD Osiris* cubre 24.700 compañías cotizadas y cerca de 3.600 no cotizadas. La base de datos es actualizada semanalmente a través de Internet y diez veces al año en la versión de DVD. Aunque la información comienza a estar disponible para algunas empresas en 1986, como en la mayoría de las bases analizadas, el número de variables disponible en los primeros años es más reducido.

En definitiva, existe una considerable variedad de bases de datos que, en principio, parecen ofrecer un mismo producto: información contable y financiera de empresas a escala mundial.

### 3. DISEÑO DE LA INVESTIGACIÓN

El objetivo fundamental de este trabajo es contrastar si la información ofrecida durante el periodo 1990-1999 para empresas españolas en las siete bases de datos alternativas mencionadas en el epígrafe anterior es de hecho comparable y, en su caso, si las diferencias observadas tienen un impacto significativo sobre los resultados de la investigación. Para ello, se replica el comportamiento de un investigador que, teniendo únicamente una base de datos disponible, estima un modelo comúnmente utilizado en la línea de investigación contable relacionada con el mercado de capitales. En particular, se eligió una sencilla adaptación del modelo de Ohlson (1995) consistente en explicar la cotización bursátil en función del beneficio y los fondos propios, como indica la expresión [1].

---

<sup>8</sup> *Bureau van Dijk* oferta otras bases de datos que proporcionan información de pequeñas y medianas compañías no cotizadas en diversos países, básicamente europeos. *Amadeus* es la base de datos a nivel europeo con empresas de estas características de 34 países. Versiones locales de esta base de datos son *FAME* en el Reino Unido o *SABI* para compañías españolas y portuguesas.

$$P_t = \alpha + \beta BV_t + \gamma X_t + \mu_t \quad [1]$$

donde  $P$  es la cotización de la acción,  $BV$  es el valor en libros de los fondos propios por acción,  $X$  es el beneficio neto por acción y  $t$  es el indicador del periodo.

La elección de este modelo se basa en dos razones fundamentalmente: (1) se trata de una especificación comúnmente utilizada en la literatura. El modelo planteado utiliza las tres variables más frecuentemente empleadas en la investigación contable relacionada con el mercado de capitales: cotización, fondos propios y beneficios. Además, éstas son seguramente tres variables estándar que están disponibles en todas las bases de datos para la gran mayoría de compañías; y (2) si se encuentran diferencias en la estimación del modelo más sencillo posible, la hipótesis de que la elección de la base de datos a utilizar en la investigación tiene implicaciones importantes para los resultados de la misma sería claramente corroborada.

Para contrastar las diferencias entre bases de datos, recopilamos la información necesaria para estimar el modelo [1] durante el periodo 1990-1999 de empresas españolas cotizadas en las siete bases de datos previamente descritas (*DataStream*, *Global Vantage*, *WorldScope*, *Company Analysis*, *Thomson Financials*, *Extel Financials* y *BvD Osiris*). La información fue recopilada en enero de 2004. Dadas sus características específicas, las compañías del sector financiero y asegurador fueron excluidas del análisis. La información se recopiló en euros. De cada una de las siete muestras generadas eliminamos las observaciones para las que no existía alguna de las cuatro variables necesarias en la estimación del modelo [1] (cotización, fondos propios, beneficio neto y número de acciones). Como es habitual en la literatura, se eliminaron las observaciones con fondos propios negativos. También se eliminó el 1% de las observaciones extremas en ambas colas de la distribución de las variables fondos propios por acción, resultado por acción y precio.

## 4. RESULTADOS

### 4.1. Muestras alternativas obtenidas: estadísticos descriptivos

Como se aprecia en la Tabla 1, la muestra final obtenida varía sustancialmente dependiendo de la base de datos de que se trate. BvD Osiris es claramente la que ofrece el menor número de observaciones, lo cual se debe a la inexistencia de datos de mercado con anterioridad a 1996. Pese a que Osiris ofrece información contable desde 1986, la información sobre la capitalización de mercado (ítem #35907) comienza a estar disponible sólo a partir de 1996. Debido a esta limitación, aunque también se presentan los resultados obtenidos para esta base de datos, se evitará realizar comentarios o inferencias basadas en los mismos.

Analizando el resto de bases de datos, la muestra obtenida oscila entre las 813 observaciones año-empresa de *Thomson Financials*, y las 343 de *DataStream*. La muestra de *Thomson Financials* es más del doble de la que se obtiene a partir de *DataStream* -2,37 veces mayor-. Como se comentó al describir las bases de datos, *Thomson Financials* combina de manera relativamente oscura observaciones de otras bases de datos. Para el caso español, parece que la mayor parte de la información de *Thomson Financials* viene directamente de *WorldScope*, otra de las bases de datos disponibles a través de *Thomson One Banker*. *WorldScope* ofrece 810 observaciones para el período considerado y el resto de sus estadísticos descriptivos



coincide con los ofrecidos por *Thomson Financials*. Por su parte *Company Analysis* ofrece 454 observaciones, *Extel Financials* 681 y *Global Vantage* 455. Se observa por tanto que, la información que *Thomson* ofrece a través de *Thomson One Banker* como *Extel Financials* es distinta de la obtenida a través de *Company Analysis*, a pesar de que, como se ha mencionado en el apartado anterior, deberían ser coincidentes.

Puesto que el tamaño de las muestras varía sensiblemente, parece inevitable que los resultados del análisis empírico propuesto reflejen estas diferencias, a menos que todas y cada una de las muestras obtenidas sean una representación perfecta y fidedigna de la población, algo poco probable.

Una de las posibles consecuencias de las diferencias en la cobertura ofrecida por las bases de datos estudiadas, es la existencia de un sesgo de tamaño en las muestras seleccionadas, y potencialmente, un sesgo de supervivencia, siendo más probable que las empresas grandes persistan en el tiempo. Es de esperar que todas las bases de datos ofrezcan información sobre las empresas más grandes y difieran en su cobertura de las más pequeñas. La Tabla 1 muestra el valor mediano de la capitalización bursátil (aproximación del tamaño) de las empresas de las siete muestras. Se observa que las muestras generadas por las bases de datos con más observaciones (*Thomson Financials*, *WorldScope*) tienen un tamaño medio menor, mientras que la base de datos que ofrece el menor número de observaciones (*DataStream*) tiene las empresas de mayor tamaño medio. Por su parte, el Gráfico 1 aporta evidencia adicional sobre el sesgo de tamaño en algunas bases de datos. Se aprecia claramente la asociación inversa existente entre el número de observaciones y el tamaño medio de las empresas contenidas en las bases de datos (medido empleando la mediana de la capitalización bursátil). Esta evidencia es consistente con la hipótesis de que las bases de datos que ofertan un menor número de observaciones se centran en las empresas de mayor tamaño, y que probablemente, las diferencias entre bases de datos se deben a divergencias en la cobertura de empresas pequeñas. Por tanto, parece que, en principio, cuanto mayor sea la muestra obtenida menor será el sesgo tamaño hacia las empresas grandes.

Si analizamos el resto de estadísticos descriptivos, que incluyen las variables necesarias para estimar la ecuación [1], éstos varían sustancialmente de una base de datos a otra, lo que no es de extrañar a juzgar por las diferencias en la cobertura ofrecida por las bases de datos estudiadas.

#### **4.2. Análisis de regresión: réplica del comportamiento de un investigador tipo**

En un primer análisis se replica el comportamiento de un investigador tipo, que sólo tiene una base de datos disponible en su institución. Por ello, se utiliza la muestra total obtenida de cada base de datos para estimar el modelo [1].

Los resultados de este análisis se recogen en el Panel A de la Tabla 2. Como puede observarse tanto los coeficientes de las variables como las medidas de ajuste varían considerablemente al utilizar las muestras alternativas para estimar el modelo. El término independiente toma valores que van desde 15,25 en *DataStream* a 3,24 en *Extel Financials*; el coeficiente de fondos propios (b) toma valores entre 0,58 en *Extel Financials* y 0,09 en *DataStream*; el coeficiente del resultado contable (c) oscila entre 8,36 en *Company Analysis* y 4,11 en *WorldScope*; y finalmente, el coeficiente de determinación corregido oscila entre 0,72 en *Extel Financials* y 0,62 en *Company Análisis* y *Global Vantage*. Por su parte, el Panel B de la Tabla

2 muestra la significatividad estadística de estas diferencias. Los resultados de este Panel ponen de manifiesto que las bases de datos analizadas podrían formar tres grupos: (1) el primer grupo estaría formado por *Extel Financials*, *Thomson Financials*, y *WorldScope* entre las que no se aprecian diferencias significativas, pero sí con el resto; (2) un segundo grupo lo componen *Company Analysis* y *Global Vantage*, que tampoco muestran diferencias significativas entre ellas pero sí con las demás; y (3) *DataStream* sería el único componente del tercer grupo ya que ofrece resultados diferentes del resto.

En definitiva, estos resultados confirman la hipótesis de que la elección de base de datos afecta a los resultados de la investigación empírica incluso cuando se emplea un mismo desarrollo metodológico.

#### **4.3. Análisis adicional: muestra común a todas las bases de datos**

Los resultados del análisis anterior se explican, al menos en parte, por la existencia de diferencias en las muestras obtenidas de cada una de las bases de datos. Sin embargo, no queda aún claro si estas diferencias en las muestras se deben (1) exclusivamente a diferencias en la cobertura ofrecida por las bases de datos (esto es, las muestras contienen empresas diferentes), o (2) a diferencias en los datos contenidos en las distintas fuentes incluso para las mismas empresas, bien por la comisión de errores en el tratamiento de la información por parte de las bases de datos o bien porque las definiciones de las variables que cada base de datos utiliza son distintas<sup>9</sup>. Para identificar hasta qué punto las divergencias en los resultados se deben a esta segunda cuestión, se planteó un análisis adicional utilizando las observaciones comunes a todas las bases de datos, con excepción de *BvD Osiris*<sup>10</sup>.

Este análisis adicional no refleja la realidad cotidiana a la que se enfrenta el investigador, particularmente en España, ya que éste difícilmente se tiene acceso a más de una base de datos e, incluso en ese caso, cuando se selecciona una no suele plantearse ningún ajuste adicional para emplear sólo aquellos datos que aparecen en todas las demás fuentes de información disponibles. Como ejemplo, Land y Lang (2002) argumentan que se decantan en favor de *Global Vantage* en lugar de *DataStream* o *WorldScope* debido a la mayor cobertura ofrecida por la primera de estas bases de datos. Sin embargo, una vez que seleccionan *Global Vantage*, emplean todas las observaciones disponibles en la base de datos, sin limitar la muestra a aquellas observaciones que son comunes a las tres bases de datos. Esta razón justifica que en el análisis principal del trabajo no se restringiese a la muestra seleccionada en cada de datos a las observaciones comunes de todas ellas. Sin embargo, limitar la muestra a las observaciones comunes permite aportar evidencia adicional sobre las diferencias entre las bases de datos, y por ello, sobre los resultados inicialmente expuestos.

Para seleccionar las observaciones comunes a todas las bases de datos empleamos el código SEDOL. Como se aprecia en la Tabla 3, la muestra de observaciones comunes (137 observaciones año-empresa) es mucho menor que incluso la más pequeña de las anteriores, lo que confirma que la cobertura de las distintas bases de datos analizadas es bastante heterogénea. Se aprecia que tanto los valores absolutos de los fondos propios, la capitalización

<sup>9</sup> Como se aprecia en la Tabla 5, es poco probable que las diferencias encontradas se deban a definiciones dispares de las variables, ya que cada variable se ha definido en las distintas bases utilizadas ajustándolas para que representen exactamente el mismo ítem.

<sup>10</sup> No incluimos *BvD Osiris* en este análisis adicional porque ello limitaría la muestra al periodo 1996-1999, lo que supondría perder observaciones de 6 años.

y el resultado neto como sus valores por acción son muy similares entre las bases de datos y las diferencias son sensiblemente menores que las mostradas en la Tabla 1. En lo referente a los datos por acción, existen dos fuentes de datos: por una parte, *Extel Financials*, *Thomson Financials* y *WorldScope* ofrecen datos por acción prácticamente idénticos para las tres variables analizadas; por otra, *Company Analysis*, *DataStream* y *Global Vantage* también ofrecen valores por acción muy similares. Por su parte, en los valores absolutos se aprecia que las diferencias entre las bases de datos son mínimas. Continúan, sin embargo, siendo significativas las diferencias entre los valores por acción. Estas diferencias son debidas a disparidades en la variable "acciones comunes desembolsadas" que cada base de datos proporciona.

Cuando se replica el análisis de regresión referido en el epígrafe 4.2 para la muestra de observaciones comunes (véase Tabla 4), se observa que las diferencias desaparecen por completo, no quedando ninguna diferencia significativa entre los coeficientes obtenidos utilizando las distintas bases de datos. Como vemos en el Panel A de la Tabla 4, el término independiente oscila entre 3,23 en *Extel Financials* y 4,67 en *Global Vantage*, el múltiplo de los fondos propios oscila entre 0,48 en *Extel Financials* y 0,55 en *DataStream*, el múltiplo del resultado contable va desde 5,59 en *WorldScope* y *Thomson Financials* a 8,09 en *Company Analysis*. Finalmente, el coeficiente de determinación corregido va desde 0,58 en *Company Analysis* y *Global Vantage* a 0,66 en *Extel Financials*. En este caso, ninguna de las diferencias observadas, como indican los resultados del Panel B de la Tabla 4, es estadísticamente significativa.

## 5. RESUMEN Y CONCLUSIONES

En este trabajo se pone de manifiesto que la elección de base de datos en la investigación contable empírica relacionada con el mercado de capitales afecta de forma significativa a los resultados obtenidos. Usando como fuente de información siete bases de datos comúnmente empleadas por los investigadores internacionales en esta línea, se replica el comportamiento de un investigador que sólo tiene disponible una base de datos y se lleva a cabo un estudio empírico muy sencillo consistente en la estimación de un modelo de valoración en el que la capitalización bursátil se expresa en función del valor contable de los fondos propios y del resultado.

Los resultados muestran que tanto los parámetros como las medidas de bondad de ajuste del modelo difieren significativamente dependiendo de la base de datos empleada. Dos conclusiones fundamentales pueden obtenerse: en primer lugar, existen diferencias entre las bases de datos, y estas diferencias son lo suficientemente significativas para alterar los resultados de un sencillo estudio empírico, que emplea variables contables y de mercado básicas; y, en segundo lugar, el análisis realizado sugiere que las diferencias son atribuibles a la distinta cobertura empresarial de cada base de datos. De hecho, cuando el análisis se centra en las observaciones comunes a todas las bases de datos, las diferencias desaparecen por completo.

Este estudio tiene implicaciones tanto para los investigadores interesados en llevar a cabo estudios empíricos como para las instituciones académicas, en tanto que potenciales adquirentes de las bases de datos. La evidencia empírica suministrada pone en duda la generalización y comparabilidad de los resultados de estudios empíricos previos. Asumiendo que la información ofrecida por todas las bases de datos tiene credibilidad similar, puesto que los resultados sugieren que las diferencias observadas se deben casi exclusivamente a la dispar cobertura ofrecida por las mismas, la elección del investigador debería favorecer a aquella/s que diera/n lugar a muestras más representativas de la población objetivo de su trabajo. Trabajos futuros son necesarios para abordar la cuestión de qué base de datos contiene la información más generalizable, para lo que probablemente sea preciso emplear modelos más sofisticados que el empleado en este trabajo. Por otro lado, aunque los resultados obtenidos sugieren que las diferencias de información reportada por las distintas bases de datos para las observaciones comunes son mínimas, es interesante un análisis más detallado de las diferencias en las definiciones de determinadas variables ofrecidas por bases de datos alternativas.

## REFERENCIAS

- Arcas, M.J., y W.P. Rees. [1999]: "Regularities in the equity price response to earnings announcements in Spain". *The European Accounting Review*, 8 (4): 585-607.
- Arce, M. [2001]: "Bases de datos de información contable en el ámbito internacional", *Revista Valenciana de Economía y Hacienda*, 3 (III): 225-237.
- Arce, M., y A. Mora. [2002]: "Empirical evidence of the effects of European accounting differences on the stock market valuation of earnings and book value": *The European Accounting Review*, 11 (3): 573-599.
- Ball, R., S.P. Kothari, y A. Robin. [2000]: "The effect of international institutional factors on properties of accounting earnings". *Journal of Accounting and Economics*, 29: 1-51.
- Bushman, R.M., y J.D. Piotroski [2005]: "Financial reporting incentives for conservative accounting: The influence of legal and political institutions". *Journal of Accounting and Economics*, forthcoming.
- Courtenay, S.M., y S.B. Keller. [1994]: "Errors in databases revisited: An examination of the CRSP shares-outstanding data". *The Accounting Review*, 69 (1): 283-291.
- Elton, E.J., M.J. Gruber, y C.R. Blake. [2001]: "A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases". *The Journal of Finance*, 56 (6): 2415-2430.
- Gallizo, J.L., F. Jiménez, y M. Salvador. [2003]: "Evaluating the effects of financial ratio adjustments in European financial statements". *The European Accounting Review*, 12 (2): 357-377.
- García Lara, J.M., y A. Mora. [2004]: "Balance sheet versus earnings conservatism in Europe". *The European Accounting Review*, 13 (2): 261-292.
- Giner, B., y C. Reverte. [2001]: "Valuation implications of capital structure: a contextual approach". *The European Accounting Review*, 10 (2): 291-314.
- Guenther, D.A., y A.J. Rosman. [1994]: "Differences between COMPUSTAT and CRSP SIC codes and related effects on research". *Journal of Accounting and Economics*, 18 (1): 115-128.

Kern, B.B., y M.H. Morris. [1994]: “Differences in the COMPUSTAT and expanded *Value Line* databases and the potential impact on empirical research”. *The Accounting Review*, 69 (1): 274-284.

Land, J., y M.H. Lang. [2002]: “Empirical evidence on the evolution of international earnings”. *The Accounting Review*, 77 (Supplement): 115-133.

Leuz, C., D.J. Nanda y P. Wysocki. [2003]: “Earnings management and investor protection: An international comparison”. *Journal of Financial Economics*, 69: 505-527.

Ohlson, J.A. [1995]: “Earnings, Book Values, and Dividends in Equity Valuation”. *Contemporary Accounting Research*, 11 (2): 661-687.

Pope, P.F., y M. Walker. [1999]: “International differences in the timeliness, conservatism, and classification of earnings”. *Journal of Accounting Research*, 37 (Supplement): 53-86.

Raonic, I., McLeay, S.J., y I. Asimakopoulos. [2004]: “The timeliness of income recognition by European companies: An analysis of institutional and market complexity”. *Journal of Business, Finance and Accounting*, 31: 115-148.

White, H. [1980]: “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”. *Econometrica*, 48 (May): 817-838.

Tabla 1: Estadísticos descriptivos

|            | Company Analysis |            | DataStream |            | Extel Financials |            | Global Vantage |            | BvD Osiris |            | Thomson Financials |            | WorldScope |            |
|------------|------------------|------------|------------|------------|------------------|------------|----------------|------------|------------|------------|--------------------|------------|------------|------------|
| N. obs.    | 454              | 343        | 681        | 455        | 140              | 813        | 810            |            |            |            |                    |            |            |            |
|            | Mediana          | Media      | Mediana    | Media      | Mediana          | Media      | Mediana        | Media      | Mediana    | Media      | Mediana            | Media      | Mediana    | Media      |
| BVPS       | 10.2             | 13.8       | 9.1        | 15.2       | 5.8              | 9.6        | 10.5           | 14.3       | 4.2        | 5.9        | 5.3                | 8.7        | 5.2        | 8.7        |
| Precio     | 15.1             | 21.4       | 15.0       | 25.6       | 7.4              | 12.0       | 15.3           | 22.5       | 9.4        | 11.5       | 7.2                | 10.9       | 7.2        | 11.0       |
| EPS        | 1.0              | 1.4        | 0.9        | 1.8        | 0.5              | 0.7        | 0.9            | 1.4        | 0.5        | 0.7        | 0.4                | 0.7        | 0.4        | 0.7        |
| Capitaliz. | 229              | 1629       | 471        | 2206       | 219              | 1197       | 314            | 1295       | 534        | 3118       | 183                | 1048       | 187        | 1052       |
|            | Máximo           | Mínimo     | Máximo     | Mínimo     | Máximo           | Mínimo     | Máximo         | Mínimo     | Máximo     | Mínimo     | Máximo             | Mínimo     | Máximo     | Mínimo     |
| BVPS       | 58.7             | 0.3        | 380.5      | 0.4        | 83.8             | 0.1        | 71.6           | 0.3        | 30.1       | 0.3        | 79.5               | 0.1        | 79.5       | 0.1        |
| Precio     | 113.6            | 0.3        | 320.9      | 0.8        | 92.6             | 0.3        | 113.5          | 0.5        | 38.1       | 0.4        | 84.7               | 0.4        | 84.7       | 0.4        |
| EPS        | 8.1              | -3.4       | 62.3       | -2.9       | 8.9              | -3.3       | 7.8            | -4.4       | 6.5        | -0.1       | 6.5                | -3.2       | 6.5        | -3.2       |
| Capitaliz. | 79328            | 1.3        | 80918      | 10.8       | 79328            | 1.3        | 39177          | 6.3        | 80918      | 16.9       | 80915              | 1.6        | 80915      | 1.6        |
|            | Desv. Tip.       | Desv. Tip. | Desv. Tip. | Desv. Tip. | Desv. Tip.       | Desv. Tip. | Desv. Tip.     | Desv. Tip. | Desv. Tip. | Desv. Tip. | Desv. Tip.         | Desv. Tip. | Desv. Tip. | Desv. Tip. |
| BVPS       | 10.9             | 28.4       | 10.3       | 12.1       | 5.6              | 9.6        | 9.6            | 9.6        | 9.6        | 9.6        | 9.6                | 9.6        | 9.6        | 9.6        |
| Precio     | 20.6             | 31.4       | 13.0       | 20.6       | 10.1             | 11.1       | 11.1           | 11.1       | 11.1       | 11.1       | 11.1               | 11.1       | 11.1       | 11.1       |
| EPS        | 1.7              | 4.6        | 1.4        | 1.7        | 0.8              | 1.2        | 1.2            | 1.2        | 1.2        | 1.2        | 1.2                | 1.2        | 1.2        | 1.2        |
| Capitaliz. | 5236             | 6004       | 4326       | 3351       | 8743             | 4017       | 4024           | 4024       | 4024       | 4024       | 4024               | 4024       | 4024       | 4024       |

Notas:

N. obs es el número de observaciones empresa-año. BVPS es el valor contable de los fondos propios por acción (incluyendo únicamente las acciones ordinarias), Precio es el precio bursátil de una acción en la fecha del balance, EPS es el resultado contable por acción después de partidas extraordinarias y Capitaliz. es la capitalización bursátil total. Ver definición exacta de las variables en la Tabla 5. BVPS, Precio y EPS expresados en euros. Capitaliz. expresado en millones de euros. Ninguna conversión de pesetas a euros se ha realizado de manera manual, sino obteniendo la información directamente en euros de la base de datos.

Tabla 2: Resultados de la regresión usando todas las observaciones disponibles por base de datos

| Panel A: Resultados de la regresión |               |               |               |                          |
|-------------------------------------|---------------|---------------|---------------|--------------------------|
|                                     | estadístico t | estadístico t | estadístico t | R <sup>2</sup> corregido |
| <i>Company Analysis</i>             | <b>5.77</b>   | <b>0.28</b>   | <b>8.36</b>   | 0.62                     |
|                                     | 5.79          | 2.44          | 9.41          |                          |
| <i>DataStream</i>                   | <b>15.25</b>  | <b>0.09</b>   | <b>4.95</b>   | 0.64                     |
|                                     | 7.31          | 0.35          | 4.05          |                          |
| <i>Extel Financials</i>             | <b>3.24</b>   | <b>0.58</b>   | <b>4.31</b>   | 0.72                     |
|                                     | 5.68          | 6.75          | 9.63          |                          |
| <i>Global Vantage</i>               | <b>6.30</b>   | <b>0.48</b>   | <b>6.76</b>   | 0.62                     |
|                                     | 6.92          | 4.41          | 7.67          |                          |
| <i>BvD Osiris</i>                   | <b>4.96</b>   | <b>0.44</b>   | <b>5.48</b>   | 0.59                     |
|                                     | 8.08          | 2.65          | 3.73          |                          |
| <i>Thomson Financials</i>           | <b>3.55</b>   | <b>0.53</b>   | <b>4.12</b>   | 0.68                     |
|                                     | 7.86          | 7.49          | 10.45         |                          |
| <i>WorldScope</i>                   | <b>3.36</b>   | <b>0.57</b>   | <b>4.11</b>   | 0.67                     |
|                                     | 9.66          | 9.94          | 10.53         |                          |

Panel B: Significatividad de la diferencia en los coeficientes entre las bases de datos

|           | <i>Company Analysis</i> | <i>DataStream</i> | <i>Extel Financials</i> | <i>Global Vantage</i> | <i>BvD Osiris</i> | <i>Thomson Financials</i> |
|-----------|-------------------------|-------------------|-------------------------|-----------------------|-------------------|---------------------------|
| <i>DS</i> | a***,g**                |                   |                         |                       |                   |                           |
| <i>EF</i> | a**,b**,g***            | a***,b*           |                         |                       |                   |                           |
| <i>GV</i> | -                       | a***              | a***,g**                |                       |                   |                           |
| <i>OS</i> | g*                      | a***              | a**                     | -                     |                   |                           |
| <i>TF</i> | a**,b*,g***             | a***              | -                       | a***,g***             | a*                |                           |
| <i>WS</i> | a**,b**,g***            | a***,b*           | -                       | a***,g***             | a**               | -                         |

Notas:

$$\text{Modelo: } P_t = \alpha + \beta BV_t + \gamma X_t + u_t$$

donde  $P$  es el precio bursátil por acción,  $BV$  es el valor contable de los fondos propios por acción,  $X$  es el resultado por acción después de partidas extraordinarias,  $u$  es el residuo y  $t$  es el indicador temporal.

Los estadísticos  $t$  han sido corregidos por heteroscedasticidad utilizando la matriz de varianzas-covarianzas de White (1980).

Se estima la significatividad estadística de la diferencia entre los coeficientes de las regresiones usando diferentes bases de datos como  $\frac{\beta_1 - \beta_2}{\sqrt{\frac{2}{n_1} + \frac{2}{n_2}}}$  dividido por  $\sqrt{\frac{2}{n_1} + \frac{2}{n_2}}$  donde  $\beta_i$  es el coeficiente estimado y  $\beta_i$  es el error estándar de la variable  $i$ . Este estadístico se distribuye aproximadamente como una  $t$  de student. \*\*\*, \*\* y \* indica que la diferencia es significativa al 1%, 5% y 10% respectivamente.

Se han usado las siguientes abreviaturas: CA = Company Analysis; DS = DataStream; EF = Extel Financials; GV = Global Vantage; OS = BvD Osiris; TF: Thomson Financials.

**Tabla 3: Estadísticos descriptivos (medianas) de la muestra común**

|  | <i>Company Analysis</i> |                   | <i>DataStream</i>         |                   | <i>Extel Financials</i> |                   |
|--|-------------------------|-------------------|---------------------------|-------------------|-------------------------|-------------------|
|  | <i>Absoluto</i>         | <i>Por acción</i> | <i>Absoluto</i>           | <i>Por acción</i> | <i>Absoluto</i>         | <i>Por acción</i> |
| <i>Fondos propios</i>                  | 197.8                   | 9.0               | 202.5                     | 9.3               | 197.8                   | 4.4               |
| <i>Cap. Mercado</i>                    | 517.9                   | 15.3              | 517.9                     | 15.3              | 517.9                   | 7.2               |
| <i>Resultado cont.</i>                 | 28.2                    | 0.9               | 28.2                      | 0.9               | 28.2                    | 0.4               |
|  | <i>Global Vantage</i>   |                   | <i>Thomson Financials</i> |                   | <i>WorldScope</i>       |                   |
|  | <i>Absoluto</i>         | <i>Por acción</i> | <i>Absoluto</i>           | <i>Por acción</i> | <i>Absoluto</i>         | <i>Por acción</i> |
| <i>Fondos propios</i>                  | 199.4                   | 9.5               | 197.8                     | 4.3               | 197.8                   | 4.3               |
| <i>Cap. Mercado</i>                    | 522.1                   | 15.7              | 517.9                     | 7.2               | 517.9                   | 7.2               |
| <i>Resultado cont.</i>                 | 28.5                    | 0.9               | 28.2                      | 0.4               | 28.2                    | 0.4               |
| <i>Número de observaciones comunes</i> |                         |                   | 137                       |                   |                         |                   |

*Notas:*

*Fondos propios es el valor contable de los fondos propios (excluyendo acciones preferentes). Cap. mercado es la capitalización bursátil de la empresa. Resultado cont. es el resultado contable después de partidas extraordinarias.*



Tabla 4: Resultados de la regresión usando la muestra común

| <i>Panel A: Resultados de la regresión</i> |                      |                      |                      |                                |
|--|----------------------|----------------------|----------------------|--------------------------------|
|  | <i>estadístico t</i> | <i>estadístico t</i> | <i>estadístico t</i> | <i>R<sup>2</sup> corregido</i> |
| <i>Company Analysis</i>                    | <b>4.28</b><br>3.89  | <b>0.53</b><br>2.79  | <b>8.09</b><br>3.96  | 0.58                           |
| <i>DataStream</i>                          | <b>4.04</b><br>3.72  | <b>0.55</b><br>2.91  | <b>8.02</b><br>3.93  | 0.59                           |
| <i>Extel Financials</i>                    | <b>3.23</b><br>5.72  | <b>0.48</b><br>2.22  | <b>6.33</b><br>3.32  | 0.66                           |
| <i>Global Vantage</i>                      | <b>4.67</b><br>4.16  | <b>0.51</b><br>2.72  | <b>7.97</b><br>4.08  | 0.58                           |
| <i>Thomson Financials</i>                  | <b>3.28</b><br>5.91  | <b>0.51</b><br>2.66  | <b>5.59</b><br>3.17  | 0.65                           |
| <i>WorldScope</i>                          | <b>3.28</b><br>5.91  | <b>0.51</b><br>2.66  | <b>5.59</b><br>3.17  | 0.65                           |

*Panel B: Significatividad de la diferencia en los coeficientes entre las bases de datos*

|           | <i>Company Analysis</i> | <i>DataStream</i> | <i>Extel Financials</i> | <i>Global Vantage</i> | <i>Thomson Financials</i> |
|-----------|-------------------------|-------------------|-------------------------|-----------------------|---------------------------|
| <i>DS</i> | -                       |                   |                         |                       |                           |
| <i>EF</i> | -                       | -                 |                         |                       |                           |
| <i>GV</i> | -                       | -                 | -                       |                       |                           |
| <i>TF</i> | -                       | -                 | -                       | -                     |                           |
| <i>WS</i> | -                       | -                 | -                       | -                     | -                         |

Notas:

$$\text{Modelo: } P_t = \alpha + \beta BV_t + \gamma X_t + u_t$$

donde  $P$  es el precio bursátil por acción,  $BV$  es el valor contable de los fondos propios por acción,  $X$  es el resultado por acción después de partidas extraordinarias,  $u$  es el residuo y  $t$  es el indicador temporal.

Los estadísticos  $t$  han sido corregidos por heteroscedasticidad utilizando la matriz de varianzas-covarianzas de White (1980).

Se estima la significatividad estadística de la diferencia entre los coeficientes de las regresiones usando diferentes bases de datos como  $\frac{\beta_1 - \beta_2}{\sqrt{\frac{2}{n_1} + \frac{2}{n_2}}}$  dividido por  $\sqrt{\frac{2}{n_1} + \frac{2}{n_2}}$  donde  $\beta_i$  es el coeficiente estimado y  $\beta_i$  es el error estándar de la variable  $i$ . Este estadístico se distribuye aproximadamente como una  $t$  de student. \*\*\*, \*\* y \* indica que la diferencia es significativa al 1%, 5% y 10% respectivamente.

Se han usado las siguientes abreviaturas: CA = Company Analysis; DS = DataStream; EF = Extel Financials; GV = Global Vantage; TF: Thomson Financials.

Tabla 5: Referencia de las variables utilizadas por base de datos

|               | Company Analysis                               | DataStream            | Extel Financials  | Global Vantage       | BvD Osiris                                    | Thomson Financials                               | WorldScope   |
|---------------|--|-----------------------|---|----------------------|---|--|--|
| <i>B/PS</i>   | $\frac{[ \{eq\} - \{eq.s.ps\} ]}{\{mkt.nsh\}}$ | DS305/DSIC            | (EX.Shareholders Equity- EX.ShareholdersEquityPreferShare)/EX.CommonSharesOutstanding | (G193-G268)/CSHO     | Shareholders Funds / No of outstanding shares | TF.TotalCommon Equity/TF.CommonSharesOutstanding | WS.TotalCommon Equity (03501) / WS.CommonSharesOutstanding (05301) |
| <i>Precio</i> | {mkt.shpe}                                     | DSMV/DSIC             | EX.PriceClose   | prcem                | Mark Cap / No of outstanding shares           | TF.YrEndMarketCap/TF.CommonSharesOutstanding     | WS.PriceClose (05001*)   |
| <i>EPS</i>    | $\frac{\{ni\}}{\{mkt.nsh\}}$                   | (DS625 +DS193) / DSIC | EX.NetIncome/EX.CommonSharesOutstanding   | (G378 + G381) / CSHO | Net Income / No of outstanding shares         | TF.NetIncome/TF.CommonSharesOutstanding          | WS.NetIncome (01751) / WS.CommonSharesOutstanding (05301)          |

Notas:

También definimos EPS en *Thomson Financials* como:  $(TF.NetIncome + TF.ExtraItemsAndGnLsSaleofAsset) / TF.CommonSharesOutstanding$   
 También definimos EPS en *WorldScope* como:  $(WS.NetIncome + WS.ExtraItemsAndGnLsSaleofAsset) / WS.CommonSharesOutstanding$

Usamos estas definiciones alternativas para el resultado contable en estas dos bases de datos ya que no parece completamente claro en los manuales de las bases de datos de Thomson Financials y WorldScope si TF.NetIncome y WS.NetIncome incluyen o no partidas extraordinarias. En cualquier caso, los resultados no cambian cualitativamente independientemente de la definición utilizada.

Gráfico 1: Relación entre cobertura y tamaño de las empresas incluidas en cada base de datos

