Pacific-Basin Finance Journal 53 (2019) 410–431

Contents lists available at ScienceDirect

Pacific-Basin Finance Journal

journal homepage: www.elsevier.com/locate/pacfin

Ball and Brown (1968) after fifty years

Ray Ball[a], Philip Brown[b,*]

[a] The University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637-1610, USA
[b] The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

ARTICLE INFO

Keywords:
Anomaly
Ball and Brown (1968)
Earnings
EPS
Event study
Income numbers
Market efficiency
PEAD
Replication
Research design
Stock returns
Timeliness
Value relevance

ABSTRACT

The Editor commissioned this replication of Ball and Brown (1968) for a special issue of the *Pacific-Basin Finance Journal* commemorating the 50th anniversary of its publication. We also describe the background to the original paper and its research design, and offer observations on its interpretation, its impact on the literature and practice, some negative consequences, and its relation to papers published around the same time. One of the pleasing attributes of our original paper is that its results consistently replicate, the implication being that the research design and its implementation uncovered a universal relation between accounting earnings and changes in firm values. The current replication is in two dimensions: time; and geography, with an emphasis on Pacific Basin countries. In the USA and in a selection of 16 other countries, annual accounting earnings continue to contain a large proportion of the information that investors trade into share prices over the year, though not in a timely fashion. Post earnings announcement drift, the first acknowledged share market anomaly, continues today despite being reported five decades ago. One change is that reporting lags internationally have shortened on average and their range has narrowed. A notable change in USA data is that the proportion of the information incorporated in share prices that is contemporaneously incorporated in annual accounting earnings has declined, though the conclusions we draw from this decline are more cautious than in Lev and Gu (2016).

1. Introduction

A little more than half a century ago, we started a project together that succeeded beyond our wildest dreams. We did believe from the outset that it would have a substantial impact on the accounting literature of the day, for we knew it provided the first confrontation of that literature's central ideas with systematic – as distinct from anecdotal – evidence. And we were confident those ideas would fail the test. But we had no way of knowing how widespread and enduring the impact of the paper would be on the subsequent literature. In retrospect, that is the accomplishment of Ball and Brown (1968) – hereafter, BB68 – that gives us the most pride: it paved the way for an entirely new, robust and still-expanding evidence-based accounting literature.

As now is well known, BB68 studied the relation between revision in stock price and change in accounting earnings. By using revision in price as a measure of the flow of value-relevant information about a firm, the research design allowed us to estimate two fundamental properties of accounting earnings. Does accounting earnings incorporate information that investors consider value-relevant? (Yes). Does accounting earnings incorporate value-relevant information in a timely fashion? (No).

In addition to the voluminous evidence-based accounting literature that has ensued, two features of BB68 of which we are very

* Corresponding author.
E-mail addresses: ray.ball@chicagobooth.edu (R. Ball), philip.brown@uwa.edu.au (P. Brown).

https://doi.org/10.1016/j.pacfin.2018.12.008
Received 19 December 2018; Accepted 22 December 2018
Available online 24 December 2018
0927-538X/

proud provide the motivation for the present study. First, the results continue to replicate well.[1] Along with its impact on the academic literature, on classroom teaching and on the wider world, replicability possibly is the most important criterion for scientific work. Indeed, results that do not replicate are unlikely to have any enduring impact, so in that sense replicability is the more fundamental property. Consistent replication implies that the BB68 research design, and careful attention to its implementation, was successful in revealing a fundamental relation between two important indicators of corporate performance – accounting earnings and stock returns – that is universal. Second, the paper has received widespread international attention. It is on doctoral syllabi and has been cited by thousands of scholars worldwide. Consequently, it is fitting that the Editor commissioned us to revisit BB68 and replicate it, in the process utilizing more recent data and addressing financial reporting in a range of countries.

Data availability confined BB68 to studying a sample of 261 USA firms over a period of nine years, 1957 to 1965, a sample that is tiny by today's standards. Times have changed, so it is natural to ask how different the USA results will be with over 50 years of more recent data and with a larger and possibly wider set of listed firms. For example, has the increased importance of intellectual property made accounting earnings (as distinct from balance sheets) less relevant? Has the precipitous decline in information processing costs increased the amount of forward-looking information incorporated in stock prices, rendering fundamentally backward-looking accounting earnings even less timely?

Similarly, despite increased globalization, countries continue to differ in their economic, legal, social and political structures. Do the results differ materially by countries or regions? In view of the audience of this journal, the Editor asked us to replicate BB68 focusing on Pacific Basin countries, and to replicate the original USA study for comparison. We have added a few additional countries provided we were comfortable with their data quality, for further comparison.

We are pleased to report that the principal results in BB68 replicate well. One financial statement variable alone (change in annual accounting earnings) continues to encapsulate a large fraction of the information that investors trade into share prices over the year (in modern parlance, earnings continue to be "value relevant"). Almost all of that information is incorporated into prices before earnings are announced (accounting earnings are not particularly timely; prices lead earnings). Abnormal returns after earnings announcements still can be predicted to a degree from the announcement (subsequently known as "post earnings announcement drift", or PEAD, which is the first acknowledged share market "anomaly").

There are, however, some major differences. For example, the reporting lag has shortened in most countries and the lag's interquartile range typically has narrowed. In the USA data, changes in annual accounting earnings have become less associated with contemporaneous share price changes, consistent with Lev and Gu (2016); we discuss various interpretations of this result.

2. Background to the 1968 study

This section briefly discusses some of the context from which BB68 emerged. Readers interested in a more comprehensive coverage of the background can refer to our joint Presidential Scholar Address at the 2012 American Accounting Association meetings (Ball and Brown, 2014).

2.1. The reigning research paradigm

When we commenced work on the project, the research paradigm of the day comprised largely a priori reasoning, occasionally backed by anecdotal evidence. A central conclusion from this school of thought was that financial statements prepared under existing accounting standards (then called "principles" and "techniques," depending on their granularity) are completely devoid of meaning. There was much debate on what radical change to the accounting system was necessary to restore meaningfulness. But – strange as it may seem in hindsight – there was no important dissent from the central conclusion that accounting information is meaningless.

To our knowledge, the first explication of this thesis was in Canning's (1929) classic based on his doctoral thesis in economics at the University of Chicago. In a section entitled "Net Income Has No Qualitative Attributes," Canning argued (Canning, 1929, p. 126) that "No propositions that assign a qualitative nature to net income can be maintained. It is a wholly quantitative thing. This is, of course, to be expected from the fact that it expresses the magnitude of a difference between two summations of non-homogeneous things." Versions of this argument proliferated over the decades, and continued well into the 1960s and 1970s. Echoing Canning's thesis, Ray Chambers, the preeminent accounting theorist when we commenced our project, concluded that balance sheet and income statement aggregates based on historical cost accounting are "of very doubtful utility" because "it is pointless to add unlike things" (Chambers, 1966, p. 4).

2.2. Challenging the paradigm with economic reasoning and systematic evidence

We both had been thoroughly schooled in the reigning paradigm by W. J. McK. (Bill) Stewart at the University of NSW when we arrived to study at Chicago. We quickly found the culture was entirely different from anything we had experienced. All ideas were viewed as candidates for challenge, with alternative arguments and with data. A frequently heard question was: "what's the evidence for that?" The school crackled with energy. They were exciting times.

[1] As described below, a successful replication on a different sample had been conducted even before BB68 was published and since then there have been far too many to single them out. For years it was not uncommon for first-year doctoral students in accounting to be assigned the task of replicating BB68's Fig. 1; failure to do so was viewed as failure of the student to correctly execute the research.

There was a prevailing ethos – based on the work of Coase, Knight, Hayek and Friedman, among others – that an evolved institutional structure (firms, markets, and the forms they take) is more rational than is obvious to the individual observer. At the same time, there was a strong belief that economic analysis could provide the key to unlocking puzzles as to why institutions take the shape they do. That background culture must have influenced our thinking, because we began asking questions like "why do firms and auditors put so many resources into calculating net income and preparing balance sheets, if they're meaningless?"[2] Perhaps because investors actually find them useful? We decided to find out.

2.3. Resources available to us

By today's standards, the available data and computational resources circa 1967–68 were miniscule. But by the standards of the day, they seemed abundant. The Center for Research in Security Prices (CRSP) was founded at Chicago in 1960, and in 1964 it completed constructing the CRSP file, the first comprehensive machine-readable archive of stock price and return data. Data were monthly in frequency, going back to 1926. Standard & Poor's released the Compustat file in 1962, providing annual earnings-related information. Earnings announcement dates were hand collected, mostly by us, from the archives of the Wall Street Journal.

Computing was primitive and expensive. The University of Chicago computer (yes, there was only one suitable computer on campus to our knowledge) was an IBM 7094 mainframe that cost tens of millions of US dollars (USD) in the early 1960s yet had only 32 kilobytes of addressable memory. The 7094 was housed in a computer centre and operated in batch processing mode: when our job was running, no-one else on campus could use it. Our computer code was written in Fortran IV; and the code, together with all data not on the CRSP or Compustat magnetic tapes, was loaded via punch cards. The only available statistical package had been developed recently for biomedical research. It was slow, clunky, occupied a lot of precious memory, and it was not designed for a task like ours. Consequently, we wrote our own statistical routines. To give some idea of the changes in computers over the last 50 years, consider this. While not strictly comparable, the IBM 7094's addressable memory was a mere sixteen-millionth of the memory of a present-generation 512 gigabyte Apple iPhone XS. The computer code for the current replication, which is orders of magnitude larger in scope than the original study, has been written by us entirely in Gawk (a derivative of Awk), all graphs were initially generated with the R statistical package, and all processing has been done in a Linux environment on a small-footprint desktop computer costing about USD1,000. The comparison gets even starker. In principle, all processing and analysis could be done on a Raspberry Pi computer used by high school students that costs around USD50 and measures 200 cm^3 including the plastic case!

An important resource available to us was Fama, Fisher, Jensen and Roll (1969; hereafter FFJR). Fama (1965) earlier had made the conceptual breakthrough of framing stock price behaviour over time as a function of information flows to the market. While obvious in retrospect, this way of thinking was completely novel at the time. It now permeates financial economics. An event study turns out to be a natural way of studying price behaviour as a function of information. The event study design addresses a single relatively homogeneous information event (in their study, stock splits) that occurs on different calendar dates for different firms. Time then is defined relative to the event date and is referred to as "event time." For example, in BB68 event time 0 is the month in which the individual firm announced its annual earnings (the calendar time of which varies by firm and year), event time −1 is one month before the earnings announcement, and event time +1 is one month after it. With all firm/year abnormal return observations thus aligned in event time, it is a simple matter of aggregating them in cross-section to depict the typical firm's price response to the event. By studying price behaviour before, at and after the event, FFJR provided us with a template we could adapt to study the value relevance and timeliness of annual accounting earnings, as well as post-announcement drift.

3. BB68

Against the background described above, we set out to study the available USA data. A brief description of the study follows.

3.1. Sample construction

We decided to limit the sample to firms that ended their financial years on 31 December. We had two reasons for that decision. First, we fitted a firm-specific regression of the *change* in each firm's earnings on the *change* in the earnings of the other firms in our sample; including firms with financial years ending on months other than December could have induced serial dependence in our proxy for the earnings of other firms (Working, 1960). The second reason was that more USA firms end their financial years in December than in any other month. The most recent CRSP price data ended in June 1966, so 1965 was the final year of the earnings sample. Earnings data were available on Compustat back to 1946, but (as described below) we lost 1946 from the sample because one prior observation was required to first-difference the earnings variable and we lost 1947–56 because we judged at least 10 observations would be required for the "market model" earnings regression. Consequently, the sample period was 1957–65. After imposing data availability constraints, the sample comprised 261 companies over the nine years. A further nine observations were lost because we detected data errors, so the final sample was reduced to 2340 company-years.

[2] In retrospect, we were encouraged by the results of our forerunner study to BB68, Brown and Ball (1967), which had identified market and industry factors in accounting earnings. The implication was that earnings reflect "real" economic phenomena, and hence they cannot be totally meaningless to users. In that sense, BB68 was a natural extension of the forerunner.

We noted (p. 168) that the sample selection criteria limited the generality of the results, but took heart from the fact that – even before our paper was published (publication was slowed down by it first being rejected by *The Accounting Review*) – the basic results had been replicated for a smaller sample over a longer period by fellow students. That replication is described at p. 168 of BB68, including fn. 24.

3.2. Checking the data

We went to considerable lengths to clean the data and to fill in missing observations. Monthly stock returns were filtered for outliers, which were verified from microfilm copies of daily newspapers and hand-corrected if necessary. Outliers that appeared erroneous and could not be verified were reported to CRSP and treated as missing. Similarly, earnings announcement dates, initially gathered from the *Wall Street Journal Index*, were compared with announcements by the same firm in adjacent years and major differences were verified by reference to the *Wall Street Journal*. Dates missing from the *Wall Street Journal Index* were obtained by searching the newspaper on microfilm. It was a painful and lengthy exercise but a benefit was that we came to understand our data in detail.

This approach is rarely used today, due at least in part to the size of modern databases. Instead, many researchers assume outliers are bad data and either pull them back to some truncation point, or toss them out entirely. Polite words, such as "winsorizing," "trimming" and "censoring", are used to describe what is done. Economic variables such as accounting income, book value, stock return and market value are naturally skewed, so one consequence of outlier modification procedures is that the researcher is not studying the true economic variable, but a mechanical transformation of its distribution. For example, the relation between stock price and a symmetrically trimmed version of accounting income cannot be said to represent the true relation between stock price and naturally-skewed accounting income.[3] The objective of BB68 was to determine important properties of accounting income, as actually reported, so it is appropriate that the variable was not trimmed and consequently that a non-parametric Good news – Bad news design was used.[4]

3.3. Defining the announcement event

In FFJR the event date was the date a stock's price was first quoted on the stock market on a post-split basis. For a small subsample of observations, they alternatively defined the event as the announcement of the split. Unlike splits, which occur with relatively low frequency, earnings announcements for all public companies are compulsory and routine. The fact that a firm's earnings are announced is not per se a surprise to the market. Therefore the first issue we addressed was how to measure the newsworthy component of an earnings announcement.

We fell back on some fundamental economics, reasoning that the information in the earnings announcement would lie in its unexpected component. Implementing the notion of an unexpected component required an earnings expectation model, as described below. We then defined the event as the announcement of the sign of unexpected earnings, namely the difference between reported and expected earnings (i.e., "good news" if unexpected earnings is positive or "bad news" if it is negative).

3.4. Earnings variables

Three earnings variables were investigated: the prediction error from a "market model" regression using changes in net income; the equivalent prediction error using changes in earnings per share (EPS); and simple changes in EPS, the implicit prediction model being a random walk. We noted (p. 172) that, since there was a predictable upward drift in actual eps over time, the random walk assumption meant the third variable misclassified a substantial number of bad news observations as good news.

The market model was motivated by the evidence we had earlier reported (Brown and Ball, 1967) of a "market" factor in earnings, and the likelihood that share prices would have incorporated information about the market factor into firms' earnings expectations by the time the average firm announced. It also was consistent with the procedure (described below) of removing an equivalent market factor from stock returns that was devised by FFJR (1969) – the first factor model. Our earnings model was a simple OLS linear relation, estimated separately for each company, regressing its change in net income (or in EPS) on the change in the average income of all other sample members (the earnings-based "market" index). Because it formed the basis of an expectations model, the regression was estimated using only past data.

3.5. Estimating abnormal returns

Following the precedent in FFJR, a "market model" was used to separate expected and unexpected ("abnormal") returns. An OLS regression was estimated separately for each company, using all available stock return data for that company within the period January 1946 to 1966 (the sample selection criteria guaranteed a minimum of 100 months of data). The market index was Fisher's (1966) "Combination Investment Performance Index." Stock returns had been adjusted by CRSP for changes in the basis of quotation,

[3] Kothari et al. (2004) draw attention to the fact that trimming a skewed variable changes its mean and its conditional mean (e.g., the slope in a regression model).

[4] The final footnote of the paper (BB68, fn. 40) cautioned against using linear parametric models.

such as stock splits and dividends, and Fisher's index was calculated on the same basis. Abnormal returns were measured by the residuals from fitting the regression models.

The event period studied began twelve months before the end of the earnings announcement month and ended six months after it. Abnormal returns were averaged across companies in event time, and then accumulated over the event period in what we labelled an “Abnormal Performance Index,” or API.

3.6. Principal BB68 results

BB68's now classic Fig. 1 depicted the overall pattern in the form of the API from months −12 through +6. The most fundamental result was that accounting earnings and stock returns were correlated (in current terminology, they were “value relevant”). In other words, accounting earnings appeared to have incorporated at least some of the information about firm value that investors were building into prices that same year. More specifically, the information incorporated in annual accounting earnings (“AI”) was approximately one half of the net information incorporated in annual stock price revisions (“NI”; see BB68 p. 175 for their definitions of AI and NI).

Nevertheless, annual accounting earnings lacked timeliness. Approximately 85% of the growth in the API over the 12 months ending with the announcement month had been priced by the end of month −1. Equivalently, only about 15% was priced in the announcement month itself. Even in that month, the value of the information contained in annual accounting earnings constituted, on average, only a fifth of the value of all “abnormal” information coming to the market. Access to higher-frequency stock returns would have revealed further anticipation of the earnings outcome within the month, as subsequently reported by Foster et al. (1984) and Bernard and Thomas (1989). This interpretation is confirmed in the replication reported below, which uses daily data.

After the earnings announcement month, the API (which cumulates abnormal returns) continued to drift in the same direction. The implication was that post-announcement stock returns were predictable using publicly-available earnings information. If anything, BB68 understated the PEAD's short-term effect because any price movements after the announcement day until the end of the announcement month were treated as “announcement effects”. The PEAD, as it came to be known in the accounting literature, was the first acknowledged systematic stock market anomaly of any kind in accounting or finance.

BB68's Table 3 revealed a shortening of the reporting lag over the study period 1957–1965, as evidenced by trends in the first, second and third quartile values of the lag.

4. Current replication

4.1. Initial remarks

Our replication is undertaken using data for three groups of countries: the USA, site of BB68; 11 “regional” countries, which the Editor considered are likely to be of special interest to readers of this journal; and five “other” countries we have added because they may also be of some interest. The 11 regional countries, which we identify by their three-character ISO mnemonics, are Australia (AUS), China (CHN), Hong Kong (HKG), Indonesia (IDN), Japan (JPN), Malaysia (MYS), the Philippines (PHL), Singapore (SGP), Republic of Korea (KOR), Taiwan (TWN) and Thailand (THA); the five others are Canada (CAN), France (FRA), Germany (DEU), New Zealand (NZL) and the United Kingdom (GBR).[5] Because the data sources for the second and third groups are the same, we treat them as a single group when discussing data issues and the results.

The data requirements to replicate BB68 comprise a set of listed companies' annual earnings signals, each signal being interpreted either as unambiguously “Good” or as “Bad” news; the announcement date of each signal; and the company's periodic stock market returns in the 12 months leading up to the announcement and in the 6 months following it.

4.2. Replicating BB68 for the USA

4.2.1. Data

For this replication we continue to focus on complete financial years (accounting periods of 12 months) but employ daily stock returns sourced from the CRSP Daily Stock dataset (ret for stocks, ewret for the market), to allow us to explore issues arising from the coarseness of monthly returns data. As noted above, in BB68 we went to considerable lengths to clean the data and to fill in missing observations. In this replication, we take the data “as read”; that is, we assume any data errors are not influential enough to affect our conclusions.

Compustat reports four measures of annual eps: primary and fully-diluted EPS, including and excluding extraordinary items (epspi, epspx, epsfi, epsfx; the fourth and fifth characters identify the specific measure). We chose Compustat's epspi as the base case for our replication because it appears to more closely correspond to BB68's choice of Compustat's net income including nonrecurring items as their primary earnings variable (p. 173). The fourth quarter's report date (rdq4) is taken as the annual EPS announcement date and is sourced from the CRSP/Compustat Merged/Fundamentals Quarterly dataset. All USA financial statement data, comprising annual eps, financial year-end date and stock price at the end of the financial year, are sourced from the CRSP/Compustat Merged/

[5] For convenience only, and in particular without taking a position on the status of Taiwan, we refer to all as “countries.” The 3-character mnemonics are published in the International Organization for Standardization's ISO 3166.

Table 1
Distribution of the reporting lag in calendar days and the median change made by firms from the previous year's lag, all USA firms.

Fyr	N	Median	Centile 1	Quartile 1	Quartile 3	Centile 99	N_Chs	Medn Ch
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1971	1132	48	18	35	66	119		
1972	1815	51	17	37	66	118	1098	−1.5
1973	2387	50	11	36	65	120	1774	0
1974	2488	51	15	36	69	121	2310	+1
1975	2513	50	15	35	68	121	2419	−1
1976	2529	49	13	33	64	114	2441	−1
1977	2507	48	13	33	65	116	2438	0
1978	2454	47	11	32	63	114	2388	−1
1979	2393	47	10	32	63	116	2314	−1
1980	2381	48	14	33	62	117	2286	−1
1981	2369	48	13	33	62	123	2247	−1
1982	2811	48	13	33	66	121	2207	−1
1983	3595	49	13	35	68	116	2638	−1
1984	3941	51	15	37	71	122	3243	0
1985	3966	52	15	37	73	121	3499	0
1986	4091	52	14	37	73	119	3497	−1
1987	4435	53	14	36	76	127	3683	−1
1988	4355	52	16	36	75	122	3850	−1
1989	4225	51	12	36	75	122	3838	−1
1990	4253	51	15	36	76	128	3850	−1
1991	4445	50	14	35	73	121	3937	−1
1992	4825	49	14	34	71	125	4172	−1
1993	5430	48	14	34	70	118	4546	−1
1994	6168	47	13	32	67	118	4955	−1
1995	6561	46	13	31	66	118	5614	−1
1996	7017	44	15	30	64	110	5932	−1
1997	7256	44	14	30	64	112	6309	−1
1998	7157	44	13	29	62	113	6348	−1
1999	7384	46	13	30	68	125	6332	−1
2000	7398	45	14	30	69	137	6618	−1
2001	6778	44	14	30	67	135	6583	−1
2002	6295	44	14	30	67	126	6191	−1
2003	5946	43	14	29	64	121	5838	−1
2004	5908	47	15	32	69	118	5575	0
2005	5816	48	17	33	68	110	5530	−1
2006	5718	51	16	35	69	118	5419	−1
2007	5605	52	16	36	71	113	5254	−1
2008	5309	52	18	36	70	110	5205	−1
2009	5039	53	19	35	68	105	4962	−1
2010	4934	50	19	35	67	105	4761	−1
2011	4832	52	18	37	67	107	4680	−1
2012	4781	51	16	37	66	110	4642	−1
2013	4867	51	16	36	65	115	4612	−1
2014	5051	52	20	36	66	110	4693	−1
2015	5020	54	19	36	68	115	4824	−1
2016	4887	53	18	38	67	117	4764	−1
2017	3853	47	17	36	59	91	3800	−1

Fundamentals Annual dataset. Because we measure earnings surprise by the change in eps from one year to the next, our database is restricted to financial periods of 12 months. Like BB68, we ignore noise introduced by subtle differences in the length of successive financial years due to leap years, or differences in the number of weeks or of trading days in the financial year. USA data for all listed companies, including "dead" companies (those that were delisted at some point during the full period) were downloaded in 2018 via the WRDS (Wharton Research Data Services) web portal.

The latest practicable cut-off date for our USA replication is financial years ending 31 December 2017. As of 30 September 2018, CRSP's time-series of daily stock returns available via WRDS terminated on Friday 29 June. The time-series of earnings announcement dates begins in 1971, although relatively few observations are available in the first year.

Table 1 contains information about the distribution, by the calendar year in which the company's financial year ended, of the number of companies for which we could obtain data and their reporting lags (the reporting lag is the number of calendar days from the end of the company's financial year until its EPS announcement date). These statistics relate to all available companies, before we applied filters that restricted the sample to companies with EPS data for two successive financial years with accounting periods of 12 months, and matching daily stock returns sourced from CRSP. Table 1 shows that the median reporting lag, according to Compustat, remained around 45–50 days throughout the almost 50 years in the replication period. However, changes to the composition of the dataset underlying Table 1 mask a steady shortening of the reporting lag from year-to-year, which is revealed when we use the same set of firms in each pair of adjacent years and extract the median change. The result is shown in the last column of Table 1. For the 46 years 1972–2017, the median change from one year to the next was to shorten the lag by 1 day in 40 years, to make no change in 5 years, and to lengthen it by 1–2 days in just one year (1974). These changes are half the average change in BB68's Table 3, which shows (for their constant sample of 261 firms) that the median reporting lag shortened by 17 days between 1957 and 1965.

4.2.2. Method

Given the data on earnings and stock returns, there are two more steps: (1) classifying the earnings signal as conveying "Good" or "Bad" news to stock market investors; and (2) conditional on the earnings news, measuring abnormal stock price movements over the year leading up to the earnings announcement and in the following six months.

The second step, measuring abnormal stock price movements, can be more challenging than we may like. For the primary analysis in BB68 we settled on a market index calculated and published by CRSP, and fitted a version of the Market Model to identify abnormal returns; but as a robustness test, we also collected data on USA bond yields and fitted an ex post version of the CAPM. We approach things differently in this paper, in that we study both unadjusted and market-adjusted returns. The full window is 541 days (about 18 months), beginning 360 days before the presumed announcement date and ending 180 days after it. We divide this window into three sub-periods: the pre-event window, which is from the beginning of day −360 to the close of the last trading day before the announcement; the event window, from the close of the last trading day before the announcement to the close of the first trading day after it; and the post-event window, from the close of the first trading day after the announcement to the close of day +180. Our event day zero differs from the announcement date recorded by Compustat for two related reasons. The first is that the price effect of announcements made after the closing price has been determined are not observable in USA domestic stock exchanges' traded prices until the next trading day.[6] The second reason is some announcements are made during weekends and holidays, although the number is relatively small (announcements on weekends account for about 0.25% of our USA replication sample). The trade-off is between weakening the experiment by adding noise – by extending the event period an extra day when the announcement is made before the closing price is determined – and strengthening the experiment by more closely matching the event period to on-market trading opportunities, thereby avoiding overstating the PEAD. Because BB68 were the first to note the PEAD, we chose to shift the announcement date from that reported by Compustat (i.e., rdq4) to the next available trading day as recorded in the CRSP daily returns file.

Stock returns, sourced from CRSP, are expressed in both unadjusted and market-adjusted forms. A stock's unadjusted daily stock return is the natural log of (1 + ret); correspondingly, the daily "market" return is the natural log of (1 + ewret); and the market-adjusted stock return is the stock's return less the market return. These daily unadjusted and market-adjusted cumulative stock returns are each set to zero at the beginning of event day −360, which we assume is the end of day −361. They are then cumulated progressively over the next 541 days.

It may seem odd that we consider returns that are, and are not, market-adjusted. The appropriateness of the chosen market index is always subject to doubt. Consider a hypothetical event study in which all events are on the same date and the only question is whether one set of observations on average performs differently from the complementary set — in BB68's setting, whether returns to the foreknowledge of "Good" earnings news on average outperform the returns to "Bad" news. Then if the same common market factor were to apply equally to all observations, market adjustment would be irrelevant when answering the question. The extent to which choice of an index is relevant depends on the strength of the market factor, its pervasiveness among observations, the extent to which there is diversity in event dates, and the reliability of the proxy used to measure the market factor. In yet another demonstration of the robustness of BB68's study, it turns out that the conclusions we draw from our replication of BB68 Fig. 1 do not turn on this question. For convenience we focus on cumulative log returns that are market-adjusted. We do not deny that the choice, whether or not to adjust for market conditions and if so how, can matter a great deal in other settings.

Many event studies are conducted in transaction rather than calendar time. This replication is international in character. Because there are differences across countries in the number of days in a year when the domestic stock market is closed, and closures occur on different calendar days, we work in calendar time. The upshot is that the width of the complete window (comprising returns from day −360 to day +180 inclusive) is identical across countries in terms of the number of calendar days. To work in calendar days we assume the return is zero on days when the stock market is closed. Shifting the announcement date to the "effective date", namely the first date on which the market is open for trading following Compustat's published date, will result in event windows of unequal width: for any given earnings announcement, the cumulative return in the event period is assumed to be the return from buying at the closing price on the last trading day before the published date and selling at the closing price on the first trading day after it. For the great majority of announcements in most weeks, when the USA stock market is open Monday-Friday and closed on weekends, the event period comprises either 2 calendar days (69% of observations in our replication sample) or 4 (25%); in the other 5–6% of observations it was almost always 5 calendar days. Although we have not investigated whether our results turn on this specific design choice, we cannot imagine that they do given the robustness so characteristic of BB68. That said, we note that BB68's study was itself conducted in calendar time. Their complete window was from the end of month −12 to the end of month +6, where month zero was the calendar month in which the earnings announcement was made.

[6] "More than 40% of the earnings announcements of Russell 3000 firms in the period 2000 to 2004 were made after the close of trading." (Berkman and Truong, 2009, p. 71). Beaver et al., 2017) estimate that around a third of USA quarterly earnings announcements were made after market closing in 1999, rising to more than half by 2010. Their time-series ends with 2012. We calculated, in mid-2018, the proportion of annual EPS announcement times (ANNTIMS_ACT) reported by I/B/E/S to have been 4 pm or later. That proportion for USA stocks was two-thirds for financial years ending in 1998, rising to four-in-five by 2003 and exceeding four-in-five thereafter. It was 84% in 2017.

4.2.3. Results

To recap, the base case for this replication is the year-to-year change in a company's primary eps including extraordinary items (Compustat's epspi), the sign of the change being used to classify the earnings signal as good or bad news, and the company's cumulative market-adjusted log return on the stock market to measure abnormal stock price movements beginning 360 days before the effective earnings announcement date (being the first stock exchange trading date after Compustat's rdq4) and ending 180 days after it.

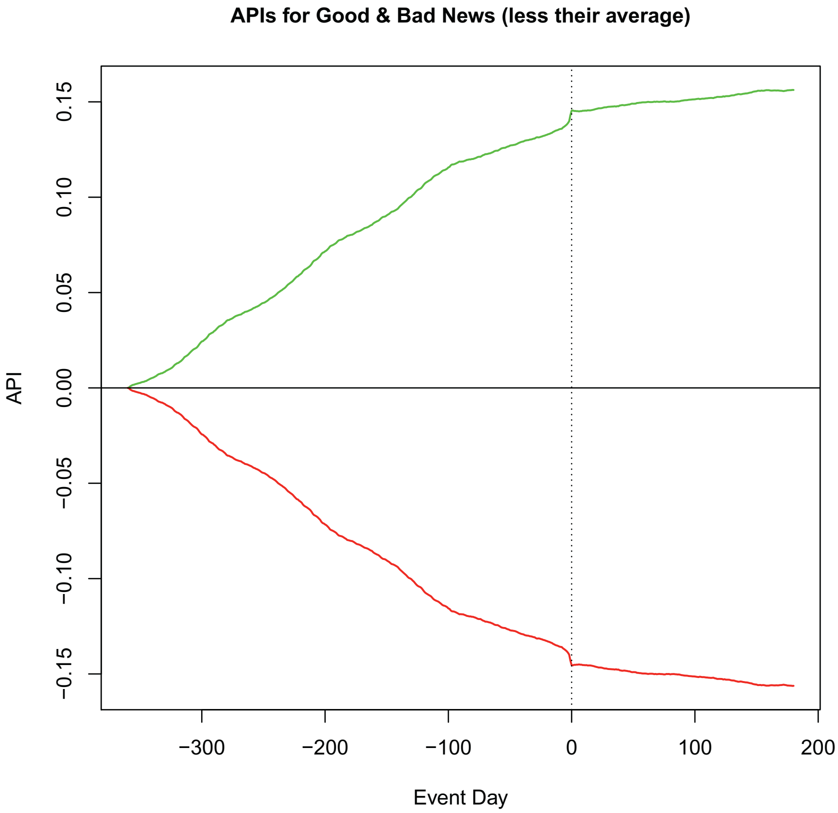

Fig. 1. Replication of BB68 Fig. 1 for USA, all years (1971–2017) pooled.

One of the most recognisable features of BB68 is Fig. 1. It is replicated here, also as Fig. 1, which relates to all observations (firm-years) in our sample and to the base case. The Good and Bad news time-series in Fig. 1 are transformed in a manner designed to compensate for any bias induced by the simple market-adjustment process or by our choice of the CRSP equally-weighted return index. The transformation is as follows. For the Good news and separately for the Bad news portfolios, we subtract the average of the two. Obviously the transformed “API” time-series are symmetric about the event-time axis, i.e., when API = zero.[7] The cost of this transformation is that we give up the ability (in Fig. 1) to identify the separate contributions of the Good and Bad news portfolios to patterns in their time-series. Instead, the sizes of changes in the time-series are averaged across the two. However, since many of BB68's results relate to differences in the APIs for Good and Bad news, Fig. 1 is a useful starting point.

Fig. 1 is strikingly clear. The transformed APIs separate from the start of the time-series, the difference between them increasing on 85% of days and reaching almost 30% by the end of the presumed announcement date, which is indicated by the vertical dotted line at event day 0. A relatively small jump in the separation between the APIs occurs in the announcement period, although the jumps on the Compustat rdq4 day (0.6%) and on the effective announcement day itself (0.5%), both days being in our announcement period, are larger than on any other single day within the 541-day window. Thus we replicate BB68's Fig. 1. In doing so, we point out that the magnitude of the separation in the current Fig. 1 is almost twice the size of the separation in BB68's Fig. 1; differences in the characteristics of the respective datasets and in the measurement methods would be part of the explanation for the relative sizes of the separations. Fig. 1 also exhibits evidence of the PEAD: the separation continues to grow in the post-announcement period, albeit at a smaller average rate than prior to the announcement date. Overall, the consistency of these results with those reported fifty years ago in BB68 indicates the strengths of the original research and the universality of the relation between accounting earnings and stock returns.

[7] We retain the mnemonic “API” since that is the term used in BB68. “CAR”, or cumulative abnormal return, is commonly used today.

Fig. 1 depicts price behaviour for Good news and Bad news that in each case has been averaged across a large number of firm/years. We caution that the pattern for individual firms or years is substantially more volatile. The seemingly smooth behaviour of the averages is a manifestation of the "Working effect" alluded to earlier, that "[i]n the study of serial correlations in prices series it is important to bear in mind that the use of averages can introduce correlations not present in the original series" (Working, 1960, p. 916).

BB68 applied separate Chi-square tests for each event month, to establish the statistical significance of the relationship between the earnings signal and cumulative stock returns that was apparent in their Fig. 1. Despite the Chi-square test's weakness, the strength of the relationship was such that BB could convincingly reject chance as an explanation for the pattern observed. Not only was the association statistically significant in months prior to the announcement month, it was significant in the announcement month and for their Naïve Model in the first two months after it (BB68, Table 5).

In this replication we address two issues with the BB68 significance tests. The first is testing significance over longer periods: there is no reason to single out any particular month in the pre-event or post-event sub-periods as a suitable basis for rejecting the null hypothesis of independence. Instead, we test the significance of the relationship between the earnings signal and the API for the three sub-periods already identified, which we label the pre-event, event, and post-event sub-periods. The second difference is that we establish statistical significance by applying a standard, non-parametric resampling procedure to generate the empirical distribution of the statistic were the null hypothesis of independence to be true. In particular, we estimate the relative frequency with which a random assignment of observations to the Good and Bad news portfolios results in a separation between their APIs that is less than the separation observed for the two experimental portfolios.

Table 2 contains the results, by year, for the base case. Columns (2) and (3) show the number of observations allocated to the Good and Bad news portfolios. The remaining six columns constitute three pairs, each pair comprising, for the indicated sub-period, (1) the difference between the APIs for the Good and Bad news portfolios and (2) the resampling test results. The sample sizes in Table 2 are much smaller than those in Table 1 for the reasons given when we discussed Table 1.

For the pre-event sub-period, the maximum separation between the Good and Bad news experimental portfolios' APIs is 42.8% in 2000; the minimum is 14.2% in 2014. 1971 is the only year for which the resampling frequency of the separation between Good and Bad news, via the eps signal, is < 0.99 for the pre-event period. Equivalently, we estimate for the pre-event period that the probability of incorrectly rejecting the null hypothesis of independence is < 0.01 for every year except the first. So strong is the separation that, when we apply a sign test to just the $N = 18$ earnings signals for the 1971 financial year, we can be confident the principal BB68 finding still holds: earnings signals and stock price movements are indeed positively correlated. For the event sub-period, despite its brevity (2 to 4 calendar days in 95% of observations) the separation between Good and Bad news portfolio APIs is positive in every year, ranging from $< 0.1\%$ to 2.9%, while the difference between the APIs is significant at 0.05 or better for 43 of the 47 years. For the 180-day post-event sub-period, the corresponding statistics for the difference in APIs is 38 out of 47 being positive, ranging from a loss of $> 9\%$ to a gain of the same magnitude, while the difference between the APIs is significantly positive at 0.05 or better for 20 of the 47 years (including 10 of the last 20.) Thus PEAD, which was first observed and commented upon by BB68, is still with us, consistent with the thesis in Ball (1978).

The persistence of PEAD over time is evidence it does not constitute market inefficiency. Grossman and Stiglitz (1980) develop the notion of an adaptively efficient market, in which inefficiencies are traded out of the market over time. McLean and Pontiff (2016) test a specific variant of this thesis, concluding that academic publication of an anomalous result is one channel that brings the anomaly to the attention of traders and causes it to disappear. It now is five decades since BB68 reported the earnings anomaly, and it shows no signs of being traded out of the market. The implication is that PEAD does not appear to be due to systematic market mispricing of earnings information, again consistent with the thesis in Ball (1978).

An innovation in BB68 was to estimate the extent of the abnormal return from a hypothetical, perfect foresight investment strategy implemented by equal-weighted buying and holding on future Good news, and selling short and maintaining the short position based on future Bad news. BB68 contemplated two scenarios. In the first scenario, the investor was endowed with perfect foresight of stock prices in 12 months' time; in the second the investor was endowed with perfect foresight of the earnings signal alone, to the exclusion of all other information about the future. BB68 referred to the average return to the first strategy as NI (Net Information) and to the second as AI (Accounting Information). NI was calculated as the return on the long position held in the first scenario minus the return on the short position, while AI was the result of the same procedure applied to the long and short positions in the second scenario. BB68 then calculated the ratio of AI to NI, which they interpreted as the percentage of value-relevant information from all sources that typically was captured in the annual income number. Around the 1960s the percentage averaged approximately 50% (BB68, p. 175).

We replicate the calculations on a year-by-year basis, from 1972 to 2017 (1971 is excluded because there were only 18 observations for that year.) The graph, for the base case, is included as Fig. 2. An unweighted least squares linear regression of the ratio of AI to NI on calendar time discloses a downward trend in the *proportional* relevance of EPS. The simple r^2 of the regression is 0.14. The change is driven primarily but not exclusively by an increase in stock price volatility that NI reflects.[8] From the first half of the

[8] BB68's AI/NI ratio is bounded: its maximum value is $+1$ (when all observations have the same sign for both the change in EPS and the cumulative abnormal return) and minimum value is -1 (when all observations have the opposite sign; i.e., when they invariably disagree). Where the ratio, for a particular set of observations, lies between these two limiting values depends upon the relative frequency with which the two signs disagree and the average absolute value of the abnormal returns when they agree relative to when they disagree. While the APIs in Fig. 1 vary little according to whether or not we adjust stock returns for market conditions (largely because it plots the difference between the APIs and their average values), the same statement cannot be made about the AI/NI ratio. This is because the frequency of disagreement (between the signs) can be affected greatly by whether the returns are market-adjusted or not.

Table 2
Difference between the average investment performance of Good and Bad news stocks and resampling test statistics in the pre-event, event and post-event sub-periods, all USA firms, by year.

Fyr	N(G)	N(B)	Pre G-B	Pre C < E	Event G-B	Event C < E	Post G-B	Post C < E
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1971	7	11	0.357	0.975	0.010	0.687	0.007	0.495
1972	56	42	0.210	0.993	0.021	0.986	−0.051	0.154
1973	259	122	0.315	1	0.022	1	0.015	0.687
1974	282	232	0.324	1	0.029	1	−0.017	0.228
1975	303	256	0.349	1	0.002	0.607	−0.039	0.040
1976	408	174	0.232	1	0.011	0.996	0.019	0.835
1977	387	211	0.268	1	0.010	0.994	0.054	0.995
1978	413	189	0.239	1	0.005	0.867	0.005	0.588
1979	399	224	0.234	1	< 0.001	0.521	0.023	0.876
1980	315	316	0.241	1	0.009	0.991	0.076	1
1981	325	337	0.227	1	0.010	0.995	0.019	0.820
1982	443	618	0.346	1	0.008	0.992	0.010	0.736
1983	781	792	0.245	1	0.004	0.956	0.019	0.913
1984	1051	766	0.293	1	0.011	1	0.018	0.884
1985	901	1066	0.279	1	0.006	0.994	0.034	0.991
1986	912	1105	0.268	1	0.012	1	0.034	0.992
1987	1186	988	0.296	1	0.007	0.995	0.032	0.992
1988	1357	953	0.278	1	0.011	1	0.059	1
1989	1263	1089	0.310	1	0.009	0.997	0.044	0.997
1990	1196	1206	0.373	1	0.024	1	0.018	0.867
1991	1195	1289	0.374	1	0.011	0.999	−0.021	0.058
1992	1482	1238	0.336	1	0.016	1	0.023	0.959
1993	1706	1320	0.291	1	0.006	0.990	0.013	0.866
1994	2005	1375	0.275	1	0.010	1	0.014	0.886
1995	2112	1546	0.317	1	0.011	1	0.018	0.919
1996	2035	1753	0.318	1	0.005	0.974	0.026	0.983
1997	2325	1819	0.293	1	0.010	1	0.032	0.993
1998	2087	2042	0.315	1	0.017	1	−0.019	0.067
1999	2168	1868	0.285	1	0.021	1	0.037	0.980
2000	2051	1852	0.428	1	0.018	1	0.090	1
2001	1640	2257	0.352	1	0.012	1	0.047	0.999
2002	1947	1743	0.252	1	0.013	1	0.048	1
2003	2234	1271	0.290	1	0.009	0.994	−0.028	0.017
2004	2247	1142	0.220	1	0.015	1	0.017	0.928
2005	1937	1410	0.197	1	0.012	1	0.022	0.986
2006	1881	1373	0.238	1	0.014	1	0.025	0.986
2007	1765	1436	0.226	1	0.010	0.997	0.051	1
2008	1487	1769	0.348	1	0.021	1	−0.093	0.000
2009	1574	1583	0.257	1	0.011	0.998	−0.019	0.037
2010	2050	999	0.176	1	0.009	0.997	0.009	0.756
2011	1789	1240	0.229	1	0.014	1	0.038	0.999
2012	1551	1441	0.238	1	0.011	1	0.031	0.993
2013	1637	1366	0.144	1	0.011	1	0.017	0.952
2014	1692	1380	0.142	1	0.011	1	0.014	0.859
2015	1632	1581	0.257	1	0.009	0.994	−0.034	0.005
2016	1727	1471	0.180	1	0.012	1	0.023	0.932
2017	1400	943	0.149	1	0.011	0.998	0.005	0.679

EPS is primary EPS including extraordinary items, as reported by Compustat. Good/Bad news is determined by whether EPS increased over the previous year. Investment performance is measured by the excess of the stock's cumulative log return over the corresponding return on CRSP's equally weighted daily return. Column 2 (3) contains the number of companies for which EPS increased (decreased). Columns 4, 6 and 8 contain the difference between the APIs for Good and Bad news firms in the three indicated sub-periods, and columns 5, 7 and 9 contain the associated resampling test statistics. The PEAD (column 8) for 2017 may be slightly understated due to the time-series of returns for USA stocks being forward-filled, where necessary, from 29 June until 180 days beyond their presumed announcement dates.

sample period (1971 to 1994) to the second half (1995 to 2017), NI increased by 16%, from 0.3518 to 0.4097. At the same time, AI declined by 5%, from 0.1265 to 0.1197. These changes combined to reduce the proportion AI/NI from 0.3607 to 0.2878. Thus, in the case of the USA, changes in annual accounting earnings have become less associated with contemporaneous share price changes.

We are cautious in drawing conclusions from this result because there are no immediate welfare-economic implications of a weaker link between accounting earnings and stock returns, and also because there are several possible explanations for it. Our caution here echoes the last sentence of BB68, that “a mechanism has been provided for an empirical approach to a restricted class of the controversial choices in external reporting” (Ball and Brown, 1968, p. 177).

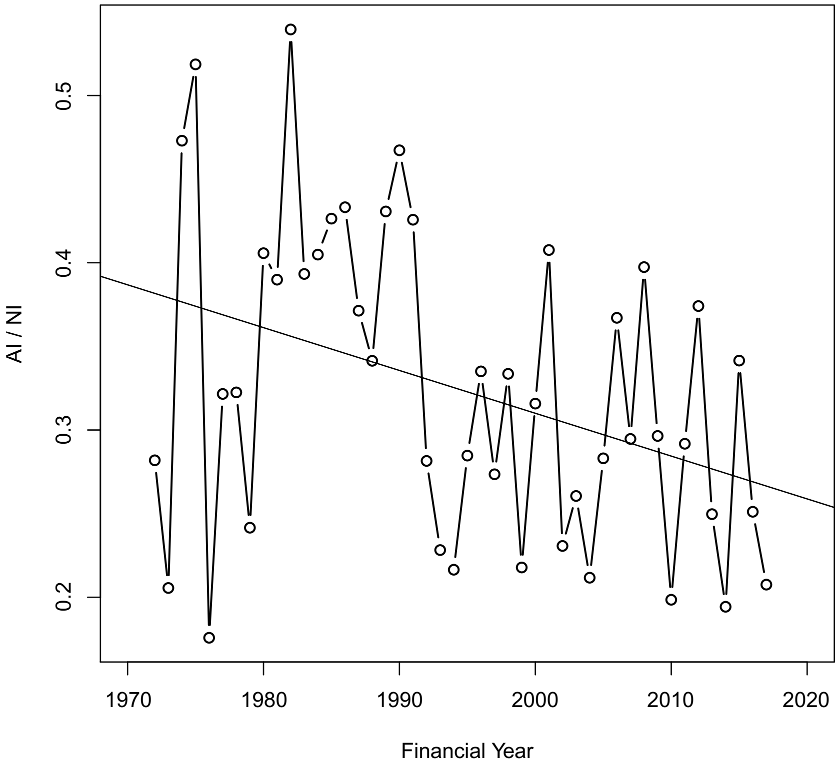

Fig. 2. Time-series graph of BB68's ratio of AI to NI for USA, 1972–2017.

The lack of immediate welfare-economic implications of a change in the earnings-returns association (value relevance) arises from the fact that association-study metrics take the amount of information in stock returns as given.[9] Our non-parametric ratio AI/NI controls for the denominator NI, a measure of the total information incorporated into share prices over the year, and hence it takes the welfare effects of information on shareholders as given. The parametric equivalent ratio is the r-squared between earnings and returns, whose denominator controls for the amount of information as measured by stock return variance. Thus, evidence of a decline in the earnings-returns association does not per se imply investors would be better informed under accounting standards that increase the association, because the evidence is based on investors being informed in the first place.[10] If an alternative accounting standard increased AI without affecting NI, or increased the earnings-returns r-squared without affecting return variance, there would be no immediate welfare-economic implications for shareholders. It might make accountants better off, but not equity investors. The appropriate welfare-economic criterion for financial reporting is not the degree of association with prices, but whether it leads to better prices, including equity prices.

An obvious accounting-related candidate for explaining a weaker link between accounting earnings and stock returns is an increase in the importance of intellectual property for the average listed firm, as reflected in heavier research and development expenditures. There is a timing mismatch between immediate expensing of research and development expenditures as they are incurred and the revenues they later produce. The mismatch makes accounting earnings a noisier measure of firm performance, thereby weakening the returns-earnings relation. This explanation is explored by Lev and Gu (2016). Another possible explanation is the increased adoption of “fair value” financial accounting standards that incorporate more changes in asset values into accounting earnings, though one would have thought that, if anything, this would have the opposite effect.

It would be a mistake to focus entirely on accounting-related explanations for the weakened link between accounting earnings and stock returns, ignoring market- and firm-related explanations. That is, the characteristics of returns can change also. In fact, as noted above, the fall in the proportion AI/NI is mostly due to an increase in NI, which reflects annual stock return volatility, not to a decrease in AI. This implies that something has changed in the real properties of listed firms, in the supply of information about them, or in the market's processing of information. For firms, substantial shifts in industry economics are likely to occur over periods as long as 50 years. Indeed, listed firms have become larger on average; while Table 1 reveals that in the second half of our sample period the number of publicly traded firms almost halved, their aggregate market value has been relatively constant. Shifts in composition between public and private firms also can occur. Gao et al. (2013) document a decline in the number of IPOs. For markets, prices

[9] This point is made in Ball et al. (2007), who also argue that association study metrics such as the earnings-returns r-squared do matter in contexts where contractual rights and payoffs are a function of reported financial outcomes, including debt contracting.

[10] For example, this comment applies to the evidence in Lev and Gu (2016, Figs. 3.1–3.4).

could have become noisier over time, due for example to index funds replacing active management, a decline in the number of analysts, or automated trading algorithms. Reduction in the cost of information production by firms, analysts and others, and reduction in the cost of processing information by investors, also could affect the earnings-returns relation.

One factor that cannot be ignored is the increase over time in the frequency of reported losses (Elliott and Hanna, 1996; Collins et al., 1997), including the increased frequency of "one time" charges that are predominantly negative and induce left-skew in accounting earnings (Basu, 1997). Gao et al. (2013) note that recent IPOs have been more likely to be loss-making. It is difficult to attribute these changes either to accounting factors or to real firm and market factors, especially since changes in accounting standards and practices are likely to be endogenous responses to real factors.

To summarise, while the data reveal a weakening over time of the contemporaneous association between annual accounting earnings and stock returns, the implications of the result are not obvious. There are no immediate welfare-economic implications of association metrics, and the feasible causes of the weakening are many and complex.

4.3. Other countries

4.3.1. Data

Thomson-Reuters datasets cover the 16 non-USA countries. Data sourced from I/B/E/S comprise EPS (ACTUAL, i.e., the "actual" value as reported by I/B/E/S), the financial year end date (FPEDATS), the announcement date (ANNDATS_ACT), and (split-adjusted) stock price (PRICE) on the I/B/E/S forecast cut-off date for the month in which the fiscal year ends. Stock return data were sourced from Datastream in the form of the daily RI series for observations we could match with the I/B/E/S data. Table 3 gives the count of the number of companies, by financial year and country, in our I/B/E/S-sourced dataset, which starts in a financial year that varies across countries and includes financial years up to 31 December 2017. Considerable variation occurs in the number of companies across the 16 countries and over the years for which we have data. Before 1998 more sample companies were domiciled in the UK than in any of the other 15 countries but starting with 1998 Japan has the largest representation.

There has been some discussion in the literature on the suitability of I/B/E/S earnings announcement dates for event studies (Livnat and Mendenhall, 2006; Acker and Duck, 2009; Berkman and Truong, 2009; Barber et al., 2013; WRDS, 2018). We investigated this issue using reliable data from two countries. For Australian companies, we compared the I/B/E/S announcement date with the official release date stated by the Australian Stock Exchange and found many instances, notably since 2009, where the I/B/E/S date was earlier, mostly by one day. Of the 5494 matched observations for financial years from 1993 to 2017 and for which the I/B/E/S and ASX announcement dates were within a day of each other, the dates were the same on 3646 occasions, the I/B/E/S date

Table 3
Number of companies (observations), by financial year and country, excludes USA. Countries are identified by their 3-character ISO mnemonics.

Fyr	AUS	CAN	CHN	DEU	FRA	GBR	HKG	IDN	JPN	KOR	MYS	NZL	PHL	SGP	THA	TWN
1989	75	72		33	89	423	56		158		33	2	2	21	2	
1990	97	207		53	132	538	70		242	14	28	5	10	25	18	
1991	56	238		50	182	502	84		279	62	66	19	5	33	16	
1992	90	234		58	182	568	107	19	296	25	94	34	12	39	28	6
1993	116	235		75	216	603	103	52	305	40	99	35	23	42	90	57
1994	137	231		117	228	673	135	53	324	81	122	38	28	47	115	91
1995	132	274	28	96	200	729	151	69	318	456	130	45	37	56	173	164
1996	204	270	69	85	293	920	209	86	330	581	155	66	44	80	281	211
1997	179	427	56	89	249	731	167	55	580	490	147	50	56	77	140	133
1998	180	363	54	111	242	696	171	51	700	357	158	54	59	69	68	144
1999	188	359	60	98	249	700	175	52	1356	300	172	51	47	69	63	150
2000	186	343	47	NA	NA	658	155	42	1459	309	147	41	44	64	56	188
2001	179	307	47	133	227	633	131	37	1520	297	125	41	41	63	52	180
2002	213	310	48	177	248	522	148	55	897	208	121	38	37	56	52	135
2003	244	301	6	174	249	484	139	49	820	155	114	39	31	64	50	109
2004	199	305	3	132	233	448	143	42	854	49	116	43	27	61	61	72
2005	204	396	56	134	232	438	173	39	813	43	115	44	26	61	97	83
2006	223	432	115	147	239	490	201	50	743	47	143	43	21	91	128	101
2007	240	443	141	153	236	534	207	50	710	59	149	42	27	83	131	102
2008	246	514	233	155	260	568	236	56	763	57	163	47	30	101	129	123
2009	268	492	339	163	272	566	199	53	781	53	185	51	28	95	92	101
2010	261	346	426	156	276	534	151	49	723	46	200	47	31	89	85	99
2011	246	476	544	152	258	532	159	50	715	50	179	47	29	94	89	105
2012	244	463	698	156	255	530	156	63	697	84	158	45	34	98	90	102
2013	246	522	779	159	243	526	167	59	648	148	167	41	37	95	82	124
2014	231	487	753	160	230	502	169	67	647	173	168	44	42	94	91	131
2015	221	483	725	169	231	478	175	82	687	245	174	43	44	99	102	160
2016	215	444	737	159	236	472	154	82	712	279	168	43	41	94	95	148
2017	213	419	575	140	174	293	187	71	701	335	182	44	27	92	113	154

Data are missing in 2000 for DEU and FRA because their adoption of the Euro on 1 January 2002 created a discontinuity in the time-series of year-to-year changes in EPS.

Table 4
Median reporting lag in calendar days, by financial year for 16 countries (see Table 1 for USA). Reporting lag is the number of calendar days from the end of the company's financial year to the first trading day after the announcement date reported by I/B/E/S.

Fyr	AUS	CAN	CHN	DEU	FRA	GBR	HKG	IDN	JPN	KOR	MYS	NZL	PHL	SGP	THA	TWN
1989	105	64		157	151	126	111		126		109	159	122	109	159	
1990	100	59		168	151	128	130		110	136	136	134	99	110	90	
1991	122	58		154	135	127	131		88	162	138	165	115	122	160	
1992	79	60		168	127	113	126	118	77	103	120	101	90	105	79	63
1993	104	55		137	104	97	112	137	86	70	97	137	133	89	161	117
1994	88	54		143	112	87	111	129	75	94	118	93	111	93	87	145
1995	73	54	121	132	102	92	111	128	73	145	103	77	123	104	96	151
1996	97	52	132	140	100	80	113	133	70	118	113	85	114	104	107	132
1997	80	54	125	140	86	77	115	124	63	84	98	71	131	91	68	132
1998	62	57	112	147	82	76	105	124	59	105	104	64	130	96	88	119
1999	61	56	117	98	90	76	103	125	69	80	66	62	137	84	84	122
2000	59	57	93	NA	NA	72	100	95	62	85	60	60	138	81	61	120
2001	60	54	102	144	142	70	100	115	58	109	62	61	118	74	66	120
2002	59	54	102	91	90	71	92	92	54	83	63	59	133	60	63	120
2003	60	53	94	84	79	71	88	92	50	82	58	56	135	56	61	149
2004	55	55	91	81	82	69	97	91	48	73	59	54	130	54	61	129
2005	55	57	90	83	76	68	97	90	46	68	59	96	128	53	61	132
2006	56	60	95	84	74	68	92	94	45	73	59	85	131	55	60	127
2007	55	63	87	79	73	69	93	80	44	77	59	61	107	57	59	107
2008	57	64	99	79	71	71	93	91	43	83	57	60	107	55	58	113
2009	57	68	92	77	70	69	88	91	43	34	56	57	106	55	56	102
2010	55	66	89	76	69	69	83	89	42	32	56	54	89	53	55	90
2011	54	68	87	75	69	68	82	88	42	37	58	53	83	55	55	87
2012	53	67	80	74	70	67	81	85	41	36	58	54	71	53	54	85
2013	53	65	72	78	65	66	79	79	40	38	57	52	105	54	55	85
2014	53	64	75	77	65	65	80	79	39	40	57	52	84	48	56	82
2015	55	64	81	74	60	68	82	79	42	36	57	51	76	50	55	84
2016	53	61	75	76	68	68	82	80	42	41	57	54	76	49	54	80
2017	53	61	59	74	61	65	81	79	42	40	59	53	73	49	54	78

was a day earlier on 1416 occasions and it was a day later on 432 occasions. We undertook a similar comparison of I/B/E/S announcement dates with those recorded in a Japanese dataset, made available to us for this purpose by Hiroyuki Aman, which included dates sourced from the Nomura Research Institute (NRI). We were able to compare the dates for 25,102 earnings announcements made between 1998 and 2015 and found: 19,303 were recorded within a day of each other; 15,963 (82%) of these dates were on the same day, while 894 I/B/E/S dates were a day earlier and 2446 were a day later. Also, only 95 of the I/B/E/S dates were earlier by more than a day. Of the remaining comparisons, the I/B/E/S date trailed by 2–10 days on 4155 occasions, by 11–31 days on 1189 occasions, and by > 31 days on 345 occasions (including 89 observations where the I/B/E/S date trailed by > 180 days.) Because we did not have ready access to any suitable, alternative country-specific databases, we proceeded with I/B/E/S as the source of announcement dates. These two comparisons again highlight the importance of "bumping" the announcement date to the next trading date if we wish to get a more accurate fix on the PEAD.

Table 4 gives the median reporting lag based on the "bumped" I/B/E/S earnings announcement date, again by financial year and country. The median lag in 2017 is shorter than in the first year with data in every country except Taiwan, and even that case seems due to its tiny sample size in 1992. In the last five years the shortest median lags were in Korea and Japan. We also calculated the interquartile range for each country-year. In untabulated results, the spread around the median decreased between 1995 and 2017 in all countries other than China, Korea, and Taiwan. In both China and Korea the interquartile range increased by 14 days and in Taiwan it increased by 10 days. Elsewhere, in Australia and in New Zealand the interquartile range decreased by 21 days; in the Philippines, Hong Kong, Malaysia and Japan the decreases were 41, 40, 27 and 14 days respectively; and in Germany, France and the UK the decreases were 27, 22 and 11 days. The spread decreased by 3 days in Thailand and 2 days in Singapore. The spread in the reporting lags of our USA sample also decreased over this 22-year period, by 12 days.

4.3.2. Method

We created our own daily "market-wide" index for each country by averaging the log return on every stock in our dataset with return data on that day. In all other respects the method used to analyse the data for the 16 "other" countries mirrors that described already for the USA replication.

4.3.3. Results

Fig. 3 contains the API graphs, corresponding to Fig. 1 for the USA, for a selection of four countries: Australia, Japan, Korea and Malaysia. Table 5 reports, for the 16 countries and the USA and for all years pooled (as in BB68), the following statistics: the numbers of Good and Bad news observations; and (separately for the pre-event, event and post-event sub-periods) the difference between the Good and Bad news APIs and the relative frequencies with which the observed API difference exceeded the difference that resulted

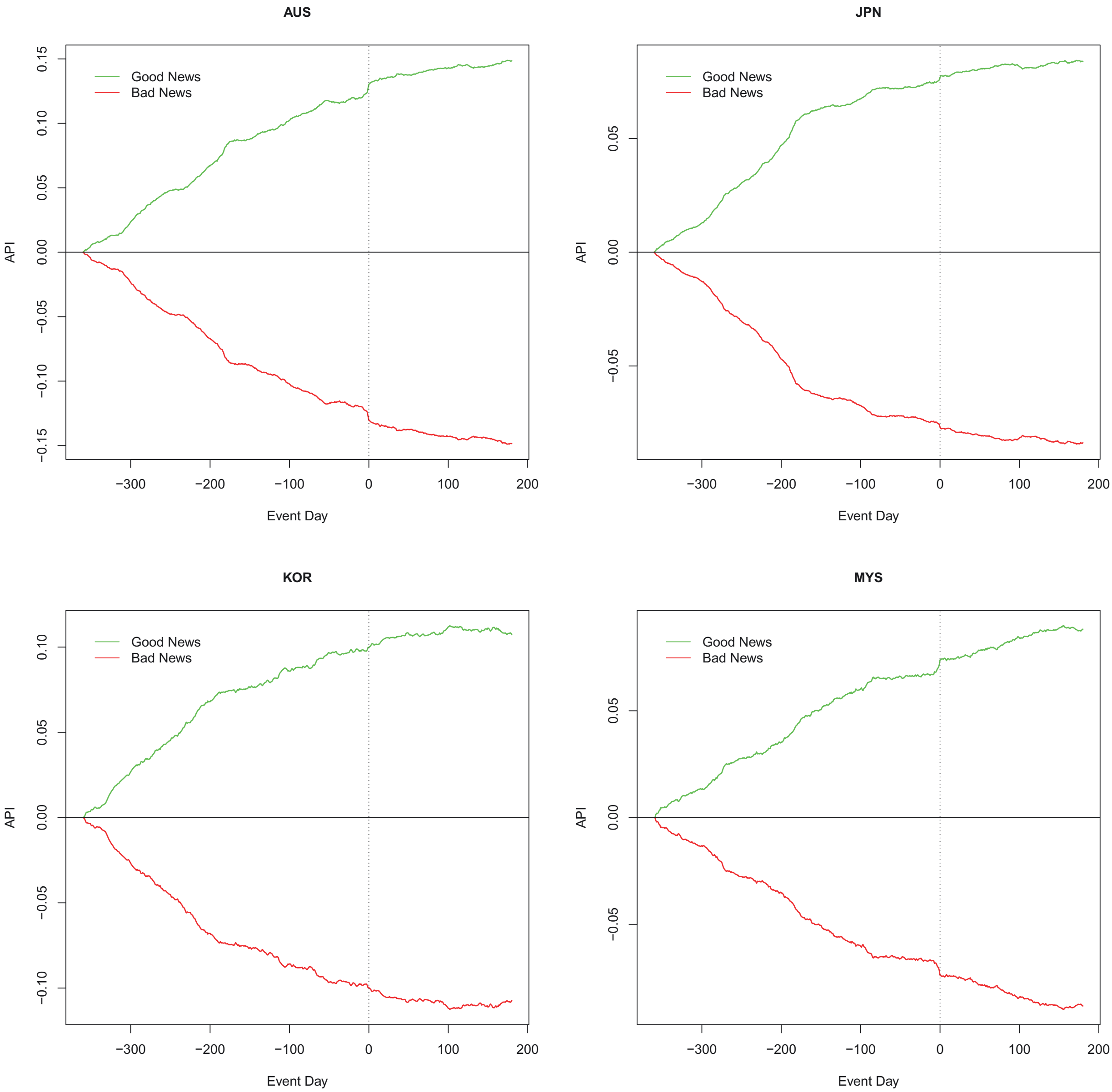

Fig. 3. Replication of BB68 Fig. 1 for Australia, Japan, Korea and Malaysia, all available years pooled.

from a random allocation of stocks to the Good and Bad news portfolios. The total number of observations by country in Table 3 differs from the sum of the numbers of Good and Bad news observations in Table 5 because the former table includes observations for which the change in EPS was zero, due to rounding in the data. Table 6 contains year-by-year information on BB68's AI/NI ratio for the 16 countries (excludes USA).

In brief, the four countries graphed in Fig. 3 exhibit three properties that characterised BB68's Fig. 1: an increasing separation between the APIs of the Good and Bad news portfolios as the event date approaches; a relatively small change in the APIs on the effective announcement date compared to the total change over the 360 days beforehand; and an increased separation over the 180 days following the announcement. Table 5 is compelling. The separation in the pre-event period is of the order of 15–25% for all countries other than USA, where it is 28%, and in every country the separation is extremely significant according to the resampling test. Similarly the separation increases in all countries in the event period (in 5 countries, including USA, by > 1%), although the increase is not statistically significant at the conventional 5% level in Indonesia and the Philippines. Interestingly, PEAD is apparent in all 17 countries. Accordingly, we should not be unduly surprised to find it is not statistically significant at the 5% level in one of 17 countries, namely Korea.

Finally, Table 6 replicates BB68's AI/NI ratio, by country-year. The ratio is in principle bounded by ± 1, although if the sample is large then a negative value would have to be considered pathological! We built a Monte Carlo simulation model to investigate this

Table 5
Miscellaneous statistics for 17 countries (includes USA), all year-by-year data pooled.

	Investment performance of Good *less* Bad news stocks								Regression of AI/NI on (Fyr-2000)			
Country	N(Good)	N(Bad)	Pre-event	Rel. freq.	Event	Rel. freq.	Post-event	Rel. freq.	Constant	Slope	*r*	N
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
AUS	3204	2190	0.2477	1	0.0131	1	0.0362	1	40.5	0.215	0.140	29
CAN	5661	4146	0.2294	1	0.0148	1	0.0241	0.998	37.1	−0.575	−0.348	29
CHN	3943	2442	0.1939	1	0.0047	1	0.0343	1	36.8	0.165	0.055	23
DEU	2044	1384	0.2195	1	0.0068	1	0.0177	0.956	39.5	−0.073	−0.042	28
FRA	3585	2546	0.1885	1	0.0091	1	0.0169	0.996	42.2	−0.110	−0.082	28
GBR	9870	5854	0.2657	1	0.0109	1	0.0330	1	44.6	−0.378	−0.347	29
HKG	2578	1793	0.2004	1	0.0187	1	0.0385	1	40.7	−0.046	−0.024	29
IDN	842	563	0.2040	1	0.0042	0.938	0.0675	0.999	30.2	0.376	0.134	26
JPN	10,984	8731	0.1503	1	0.0033	1	0.0137	1	29.9	0.881	0.574	29
KOR	2460	2469	0.1968	1	0.0026	0.964	0.0151	0.923	32.6	−0.054	−0.024	28
MYS	2239	1697	0.1407	1	0.0077	1	0.0289	1	26.8	1.547	0.686	29
NZL	651	483	0.2413	1	0.0064	0.999	0.0458	1	48.0	0.389	0.137	28
PHL	548	337	0.1711	1	0.0038	0.896	0.0461	0.998	30.0	−0.638	−0.192	28
SGP	1061	927	0.1533	1	0.0096	1	0.0393	0.999	38.0	0.148	0.065	29
THA	1315	1176	0.2510	1	0.0060	1	0.0356	0.991	35.6	0.982	0.400	28
TWN	1573	1504	0.1647	1	0.0044	1	0.0212	0.988	31.4	1.563	0.661	26
USA	61,600	51,204	0.2788	1	0.0120	1	0.0216	1	31.0	−0.003	−0.373	46

Contents of the columns are: (2, 3) the numbers of Good and Bad news observations; (4–9) the difference between the average investment performance of Good and Bad news stocks and the resampling test statistics for the (4–5) pre-event, (6–7) event and (8–9) post-event sub-periods; and (10–13), for the simple time-series regression of BB68's AI/N ratio on the financial year (minus 2000), the (10) intercept, (11) slope coefficient, (12) correlation coefficient and (13) number of observations. (See Fig. 2 for a plot of the ratio's time-series for USA).

Table 6
BB68's AI/NI ratio, expressed to the nearest percentage, by financial year for 16 countries (excludes USA).

Fyr	AUS	CAN	CHN	DEU	FRA	GBR	HKG	IDN	JPN	KOR	MYS	NZL	PHL	SGP	THA	TWN
1989	63	7		32	38	41	44		10		30			23		
1990	23	31		63	57	50	53		0	78	11	−19	5	53	20	
1991	19	39		34	40	42	51		34	35	11	83	100	52	−6	
1992	43	44		29	51	55	26	90	24	23	40	89	40	44	48	33
1993	19	45		21	40	52	52	54	15	41	10	57	4	48	58	18
1994	23	37		39	55	47	62	28	19	44	7	12	50	21	34	32
1995	40	39	0	71	54	58	49	12	24	26	26	27	36	45	26	0
1996	35	65	45	54	55	53	46	−8	21	21	17	66	34	44	35	25
1997	67	49	44	67	28	47	24	31	39	34	−5	59	52	21	40	27
1998	52	55	40	0	21	57	34	31	42	24	16	35	17	59	23	14
1999	55	25	24	34	43	41	−7	27	37	21	6	36	40	−5	13	34
2000	41	49	13			32	21	22	34	17	21	47	−4	16	1	40
2001	59	33	50	25	38	40	50	25	27	36	29	81	42	44	43	19
2002	32	43	41	43	24	45	26	12	35	29	21	58	20	−2	61	40
2003	46	41	60	43	39	36	26	27	40	42	4	33	7	30	26	25
2004	39	40	100	38	19	31	55	3	26	−9	58	30	41	83	67	43
2005	43	61	28	34	59	49	77	27	39	64	50	55	45	55	44	62
2006	47	37	17	38	46	44	52	21	49	8	57	43	17	34	46	9
2007	22	43	14	50	46	45	31	27	50	29	53	52	59	49	63	52
2008	47	37	41	38	56	48	44	17	42	5	38	59	−30	29	17	17
2009	36	9	32	26	19	18	19	19	27	54	48	65	3	24	38	23
2010	40	14	34	23	41	24	34	28	20	15	30	44	21	44	41	37
2011	44	27	52	41	44	48	64	30	46	30	53	62	−15	68	68	47
2012	34	28	63	51	52	46	40	74	50	28	45	44	17	45	50	51
2013	52	25	39	47	35	33	46	24	50	40	45	42	53	48	20	48
2014	42	22	26	59	46	46	41	59	26	52	39	73	59	32	25	56
2015	61	37	34	39	46	47	40	60	50	59	48	34	52	53	63	67
2016	44	26	21	31	34	39	41	44	41	30	62	80	3	20	47	57
2017	28	19	41	29	47	45	38	31	25	34	44	33	22	38	68	64

The results are based on Good and Bad news "portfolios" being constructed according to the sign of ΔEPS.

question: how large should the sample be to yield a statistically reliable value > 0 for BB68's AI/NI ratio? For a given country, say JPN, we assumed the pooled sample of individual stocks' cumulative market-adjusted log returns over the window [−360, 0], together with their eps codings as good or bad news, were representative of the population values for Japan. We then ran a series of 10,001 trials, in each trial first drawing a sample of predetermined size, sampling with replacement, from the presumed population and then calculating the AI/NI ratio for that trial. At the conclusion of the 10,001 trials the distribution of the randomly-generated AI/NI ratios was summarised. This procedure was repeated for sample sizes ranging from 10 to 200 observations, and for all 16 countries. Briefly, we noted that the median was, as expected, an unbiased estimate of the presumed "true" value of the ratio regardless of the sample size. However, the minimum sample size for which 95% of the 10,001 trial AI/NI ratios were positive was: AUS 26, CAN 51, CHN 35, DEU 38, FRA 30, GBR 27, HKG 26, IDN 41, JPN 31, KOR 49, MYS 31, NZL 18, PHL 38, SGP 37, THA 30 and TWN 34. For a 99% cut-off the corresponding sample sizes were AUS 54, CAN 103, CHN 68, DEU 77, FRA 60, GBR 52, HKG 53, IDN 82, JPN 61, KOR 97, MYS 64, NZL 35, PHL 84, SGP 74, THA 57 and TWN 67. The upshot is that much of the variation in the ratio over time is likely to be due to sampling error, as a result of the small sample size for many country-years.

5. Robustness tests

5.1. Evidence of robustness

BB68 undertook a number of tests to demonstrate their results were robust. Four "income numbers" reported by Compustat were compared: net income including non-recurring items, EPS, operating income (which BB68 interpreted as a rough proxy for cash flow from operations), and net income excluding non-recurring items. Abnormal returns were measured by the residuals from the Market Model fitted to monthly logarithmic rates of return, as an alternative to arithmetic returns, while an ex post variant of the Capital Asset Pricing Model was also employed as a substitute for the Market Model. BB68 reported relatively small differences in their APIs for these alternatives. While we did not attempt in this paper to replicate the specific robustness tests in BB68, we did investigate some alternative design choices available to us in order to demonstrate once again that BB68's principal results are robust.

For the USA, the base case is primary EPS including extraordinary items as reported by Compustat (epspi), with increases over the previous year coded as Good news and decreases as Bad (BB68's Naïve model) and daily abnormal returns measured by the stock's daily log return after subtracting the log of the market's equally-weighted daily return as reported by CRSP. Details of the results are available from the corresponding author, but in brief, the USA results in Fig. 1 are robust to the following alternatives: three other EPS measures reported by Compustat (epspx, epsfi and epsfx); deflating the change in EPS by the stock price at the end of the previous financial year and coding the deflated price change as Good/Bad news if it was above/below the median for that country-year[11]; not adjusting stock returns for the return on the market[12]; and mitigating a subtle bias that can be introduced by "double counting" overlapping event periods.[13]

For non-USA markets, the base case employs EPS as reported by I/B/E/S (in lieu of Compustat's epspi), stock return data are sourced from Datastream and our proxy for the market return is a self-constructed, unweighted average stock return time-series. There has been some discussion in the literature about the nature of the I/B/E/S EPS measure (Brown and Larocque, 2013). I/B/E/S began reporting cash flow per share (CPS) data for our international sample of countries in the mid-1990s, and GAAP EPS (GPS) about 10 years later. We considered calculating, for each country, the API and the ratio of AI/NI, using a common sample comprised of observations (company-years) where all three earnings measures are reported by I/B/E/S and pooling all years together. We decided not to generate these results because the sub-samples may not be sufficiently representative for our purposes.

5.2. An implication of robustness

There can be little doubt that the BB68 principal findings are robust across countries, with their different accounting methods and institutional arrangements, and over a long period of time. In particular, while year-to-year changes in "accounting income numbers" have continued, since 1968, to encapsulate a substantial fraction of the information that is reflected in abnormal stock price changes over the year leading up to their announcement dates, the bulk of the price changes typically occur well beforehand. One, sobering, implication of the extreme robustness we have observed is that almost any feasible "income number" would yield a statistically significant association between that measure and abnormal share price movements. So while BB68's experimental design was well-suited to its original purpose, it is probably too coarse to be much help in resolving more complex questions in financial accounting.

6. Some questions we have been asked over the years

It is not surprising for a paper that has been so influential and in the public domain for so long to attract many questions. In this section, we provide our views on some of the more important questions we have encountered.

[11] This "above/below the median" coding rule can be implemented in a way that guarantees equal numbers of Good and Bad news cases, thereby addressing the bias in the sign rule discussed by BB68 (pp. 171–172). However, since the median change in EPS is unknown until later in the announcement season, the median coding rule can introduce a look-ahead bias in coding an earlier announcement, leading to an appearance of reversal in the post-event period of some of the prior cumulative return.

[12] See the previous discussion (p. 416) on adjusting for market effects in settings where the focus is on the difference between average returns to complementary portfolios.

[13] Overlapping, as we use the term, occurs when the announcement date this year is < 360 calendar days after last year's announcement date. Bias can arise because Good news this year following Bad news last year tends to be announced earlier (60% in our sample) than where Bad news is reported again this year (48%). Brown (1970, fn. 11) discounted any effects from overlapping periods when replicating BB68 with Australian data; and we find in the current replication that the inferences are indistinguishable whether we do or do not set daily returns in any overlapping period this year to zero.

6.1. Why did BB68 have a larger impact on the literature than other seemingly contemporary studies?

An important feature of BB68 that affected its impact is that it was embedded in the accounting literature of the day. It challenged the fundamental way of thinking of accounting scholars worldwide. It put that thinking to a test, which the prevailing paradigm appeared to fail. The possibility there would be a paradigm shift attracted the attention of many participants in that literature, and while some reactions were vitriolic, they were reactions. In that sense, the paper drew almost immediate attention from the prior generation. Over time, a new generation of younger scholars saw related empirical research opportunities that seemed more promising than offering yet another opinion on the best alternative to prevailing accounting practice, and the impact of BB68 grew further.

A related feature of BB68 that most likely increased its impact is that it analysed the relation between annual accounting earnings and stock returns over the entire year before the announcement. This allowed us to address the relation between the information that investors trade into prices over the year and the information that is incorporated in the annual accounting earnings number (in current terminology, whether earnings are "value relevant"). This in turn allowed us to comment on the apparently low timeliness of earnings, which is another important accounting property. In other words, BB68 addressed some fundamental properties of *accounting*.[14]

In comparison, Benston (1967) fitted several cross-sectional models to a sample of announcements made by 483 firms in the third quarter of 1964. The models were fitted by regressing returns, in the announcement month and in pre-announcement periods as long as three months, on growth rates in earnings and sales. The two primary conclusions were that earnings and returns were only weakly related (even with test statistics that assumed independence in a single cross-section), and that sales growth rates had approximately ten times the effect on returns as earnings (Benston, 1967, p. 27). The impact of Benston's paper was adversely affected by limitations of the research design, the small sample and the non-result for earnings.

Beaver (1968) subsequently introduced a highly influential announcement effect design, addressing abnormal price volatility and trading volume over a 17-week period centred on the announcement week. Major innovations were focusing more closely on the announcement period using weekly rather than monthly data, and the volume metric. The selected sample was 5 years of announcements for 143 non-December 31 firms with few competing non-earnings announcements: no dividend announcements in the earnings announcement week; no stock splits announced during 17 weeks centred on the earnings announcement week; and fewer than 20 total news announcements per year. Interpretation of the results was somewhat blunted by the sample selection criteria, which tended to exaggerate the timeliness and surprise content of accounting income.[15] Interpretation also was complicated by the ambiguity of volume as a dependent variable (Dopuch, 1983). Beaver's (1968, p. 69) interpretation of volume was that it reflects changes in expectations of individual investors, as distinct from the market generally. However, volume also could reflect: portfolio rebalancing in response to event-time price changes that unbalance investors' portfolio weights; portfolio rebalancing when a publicized earnings announcement brings the stock and its past price changes to the fore of investors' limited attention; unobservable noise in the reported earnings number causing investor confusion; or tax rebalancing because investors believe earnings changes increase the likelihood of dividend changes. Nevertheless, the research design was very innovative and the accounting literature was enthused by the appearance of substantial new information in earnings announcements. The paper has had a substantial impact, is widely cited, and studies of changes in trading volume remain popular.

Another feature of BB68 that added to its impact is that it reported returns in the six months after earnings announcements, and devoted almost a full page attempting unsuccessfully to explain why there was a "drift" (Ball and Brown, 1968, pp. 173–174). The paper thereby presented the first evidence in the accounting or finance literature of a market "anomaly", namely the PEAD (Ball, 1978). The seeming inconsistency of this result with market efficiency, together with the apparent promise of trading profits, naturally generated a substantial literature of its own (e.g., Foster et al., 1984; Bernard and Thomas, 1989, 1990).

Three final features that perhaps added to the impact of BB68 are the robustness of the research design, the attention we gave to magnitudes rather than just statistical significance, and the care we took in writing it. We have already commented on the research design above. While we reported simple non-parametric chi-squared significance tests, our focus was on how much of the contemporaneous value-relevant information is reflected in accounting earnings, and when. Consequently, magnitudes were emphasized. In writing the paper, we put considerable effort into explaining the research choices we made as exactly and clearly as possible. It goes without saying that clear and complete exposition is desirable in published research generally, in part because it permits replication. Back in 1967 we were aware that the paper's audience had at best a limited background in empirical research, and no experience with the data and event study technique we were using, which made careful writing even more important than usual.

[14] While the BB68 evidence involves only accounting earnings, it indirectly informs us about *accounting*. Net Income obviously is an important accounting variable in its own right. In addition, the logic of double entry accounting implies that earnings automatically flow on to revise Balance Sheet accounts. Other things equal (notably, dividends and stock issuances and purchases), profits strengthen balance sheets and losses weaken them. For example, timely recognition of gains and losses in the Income Statement implies timely revision of Balance Sheet asset and liability amounts.

[15] Bamber et al. (2000) conclude that return volatility and volume in this sample also were exaggerated by an over-representation of small stocks. They note that despite Beaver (1968, p. 72) clearly detailing the sample selection criteria, and inviting others to try different criteria, the subsequent literature over-generalized the results. Ball and Shivakumar (2008) argue that comparing event-day or event-week volume and volatility with their average daily or weekly equivalents creates an exaggerated impression of the announcement effect, because earnings announcements occur only infrequently (quarterly, half-yearly, or annually) and hence event-period abnormal volume or volatility is a small proportion of the total.

6.2. Was there an impact on practice?

We noted in the Introduction that the single accomplishment that gives us the most pride is the robust accounting literature that ensued. But was there a wider impact? Was accounting standard setting affected? Financial market or disclosure regulation? Investment practice?

These are difficult questions to answer. As is the case with the accounting literature, impact is difficult to assess because one does not know the counter-factual. What would have been the sequence of events if we had *not* published? Further, BB68 was published around the same time as other Chicago-originated empirical studies, the most notable being FFJR (1969), so its incremental contribution is muddied. Impact on practice has no equivalent to the literature citation counts that provide imperfect but useful measures. We can say that one and perhaps two generations of accountants, standard setters, regulators and asset managers now have been educated by faculty with a different view of the world than the accounting professors who taught us, which is bound to have had an effect.

6.3. Does "useful" disprove "meaningless"?

There was much subsequent discussion about whether demonstrating usefulness was the same as demonstrating meaningfulness. As noted above, our intuition was that if accounting earnings are correlated with the information investors appear to find useful by incorporating it into price revisions, then investors must find earnings to be meaningful.

We now realize that the connection can be made more formally, based on the work of the pioneering linguist Charles Ogden. Ogden noted that for a word to have meaning "requires that it form a context with further experiences (Ogden and Richards, 1923, p. 270)". Thus, what earnings mean to investors arises from them using earnings along with other information in their decisions. This notion of meaning is different from definitional meaning, or correspondence with a precise analytical construct. In a legal context, the eminent American jurist Judge Learned Hand warned against placing too much weight on definitional meaning, stating: "it is one of the surest indexes of a mature and developed jurisprudence not to make a fortress out of the dictionary" Cabell v. Markham, 148 F.2d 737, 739 (2d Cir. 1945).[16]

6.4. Were there negative effects of BB68?

6.4.1. Effect on the prior literature

One negative effect was perhaps inevitable because it tends to occur with paradigm changes: some positive aspects of the prior literature were lost to the current generation. This phenomenon was identified in Kuhn's (1970) famous thesis on the development of a science. Kuhn distinguished between "normal" periods in which knowledge is accumulated over time by applying the paradigm to an increasingly wide set of applications, and "revolutionary" periods in which paradigm shifts occur and knowledge is not cumulative because some past achievements are not preserved (1970, p. 92). The phenomenon of discarding some past knowledge is known as "Kuhn-loss."[17]

The historical importance of the largely a priori literature that preceded BB68 cannot be over-emphasized. These were the scholars who established accounting as an area of serious intellectual inquiry, and solidified its place in scholastic institutions. Its struggles with issues such as the optimal marking of assets to market (at replacement costs, realizable values, discounted present values, or inflation-indexed historical costs) provide a background to contemporary debates on "fair value accounting." Some of the literature informs the debate on the role of accounting in stewardship relative to informing the capital markets.[18] While BB68 successfully challenged a basic tenet of that literature, it is a pity the current generation of students tends not to read it.

So why did the pre-BB68 literature fall out of favour, particularly among researchers and schools with a strong economics bent? After all, it had been described as "the golden age in the history of a priori research in accounting" (Nelson, 1973, p. 4). We believe there are two primary reasons. First, it failed to put its ideas to systematic empirical testing. The limited empirical work undertaken was selective, designed more to support ideas than to offer a genuine opportunity to refute them. Contrary to some perceptions, BB68 did not challenge the role of "normative" theory about optimal accounting rules or systems; rather, it made the case for confronting theory with data. Indeed, adopting usefulness (value relevance) and timeliness as important accounting properties implied a normative approach to financial reporting.

Second, the pre-BB68 literature did not fail in many places because it was normative/theoretical; it failed because its economic

[16] One way of illustrating the narrowness of a definitional interpretation of meaning is to consider some aggregates that educators commonly use, such as final grades in the courses we teach, GPAs, and SAT scores. For example, GPAs are aggregates of course grades that are based on widely different and not directly comparable techniques (assignments, cases, lab projects, essay questions, numerical questions, multiple choice questions, class participation, attendance, etc.), in much the same way that accounting earnings are aggregations of costs and revenues measured using many different and not directly comparable accounting methods. As is the case with GAAP calculations, some grade components such as answers to unambiguously worded true/false questions are comparatively objective; other grade components such as essays, case solutions, class participation scores and the weights attached to the components are comparatively subjective. Yet we aggregate these components to obtain individual course grades, and to obtain GPAs we aggregate grades in courses that weight the components differently, because the aggregates mean something to us in the contexts in which we use them. That is, we find them meaningful and useful.

[17] Stanford Encyclopaedia of Philosophy (2011).

[18] Notably, Chambers (1966).

theory was inadequate for the task in hand. The fundamental limitation of that literature was its failure to take seriously the concept of market frictions. Imperfections in market prices must be invoked to explain even the existence of firms (Coase, 1937). Indeed, in economic theory it is market imperfections that determine the existence and the optimal shape of every institutional characteristic, including financial reporting. Consequently, modern theory models economic frictions such as information asymmetry, moral hazard and adverse selection when pursuing the optimal shape of accounting and auditing.[19] Without diminishing the historical importance of the early accounting theory literature in earning a place for accounting in academic institutions, it floundered because it was mired in a pre-Coasian view of the world. Consequently, it was unable to explain even the existence of financial reporting or independent auditing, let alone provide a logical preference ordering of any accounting rules or systems, including those it advocated (e.g., historical cost accounting versus marking balance sheets to market using realizable values). While BB68 no doubt accelerated the demise of this literature, our conclusion is that it most likely would not have survived the advent of modern costly-contracting theory anyway.

6.4.2. Over-emphasis on the equity market

Another negative was an initial focus in the mainstream literature on the secondary equity market, at the exclusion of other contexts in which accounting information is used such as the debt, labour, corporate control, and goods and services markets, as well as internal "managerial" contexts. However, it would be unfair to attribute this imbalance to BB68. Much of the early emphasis on share markets was due to the CRSP and Compustat data being the only widely-accessible databases for some time. Indeed, the increased availability of other data has tended to redress the imbalance somewhat, one notable exception being large-scale data on internal uses of accounting within firms. Other contributing factors could be the importance of equity markets, particularly in North America and other common law countries, as well as developed Western economies, and the importance that key standard setters (FASB and IASB) place on informing investors.

6.5. Who came first?

There has been some discussion over the decades concerning provenance. For the historical record, we now offer some facts in relation to timing.

Our first comment concerns the publication dates of BB68 and FFJR (1969). The FFJR paper was well-known, particularly among Chicago PhD students, several years before it was published; our research method was greatly influenced by theirs in ways that we have always acknowledged. So whereas BB68 was the first published event study, FFJR (1969) provided a template for studying price behaviour over time and preceded us by a comfortable margin.

With respect to BB68's place in the accounting literature, an early version of BB68 was presented at a Chicago accounting research workshop on 1 June 1967 under the ambitious title: "A theoretical and empirical evaluation of accounting income numbers." The next circulated version was completed by early Autumn, 1967. That version was presented at the November 1967 Seminar on the Analysis of Security Prices organized by CRSP at the University of Chicago, a biannual conference attended by fee-paying asset managers and by Chicago faculty and doctoral students. The paper was included in the November 1967 Proceedings of the Seminar on the Analysis of Security Prices. It was titled: "The information value of the annual earnings report," in part because its intended audience was primarily the asset managers rather than the accounting researchers in attendance. It contained all the results that subsequently were published in BB68. Because of its historical importance, we have made this version available on SSRN (Ball and Brown, 1967).

For academic publication, the text and title were revised again to frame the paper as a challenge to the accounting literature of the day, but the results remained the same as in the circulated November 1967 version. This version was rejected by *The Accounting Review* in the winter of 1968 and then solicited by Nick Dopuch for the *Journal of Accounting Research*, where it was published with a slightly revised text in Autumn, 1968, again with the same results that were circulated in the November 1967 version. This required final proofs to be with the editor by the end of Spring 1968. So while BB68 was not published until a year later, a draft containing its complete results was circulated – at least among Chicago faculty and doctoral students, as well as some asset managers – in the October–November 1967 time frame. With the single exception of Ashley (1962), which we discuss below, we are not aware of any other earnings-returns association results having been circulated in any form prior to that time.

Subsequent to our workshop and CRSP Conference papers being circulated, two studies by graduates of the Chicago doctoral program circulated related results: Benston (1967) and Beaver (1968), which we have already mentioned in this paper. Both were presented at a Conference on Empirical Research in Accounting (later renamed the *JAR* Conference) and published in a conference proceedings issue. Pinning down the timing of these publications requires an understanding of *JAR*'s idiosyncratic dating of the proceedings issue at the time. The Conference was (and still is) held in Chicago each May. Following normal practice, authors were given ample opportunity to revise their manuscripts after the conference, and they were published in Summer of the following year. At the time, *JAR* subscribers paid for and received two issues per year, in Spring and Fall (Autumn). But to promote its new commitment to empirical research, when the conference was initiated in 1966 *JAR* began publishing the proceedings issue as a free extra "supplement" to the two annual issues, and here is where the dating became peculiar and downright misleading. The supplement was dated with the year the Conference was held, even though it was not published until the following year. *JAR* maintained this odd

[19] Holmstrom and Tirole (1989) and Tirole (2006) provide summaries. A particular friction – agency costs – was introduced to the accounting literature by Watts (1977) and Watts and Zimmerman (1986), based on Jensen and Meckling (1976).

practice until 2001, when it first became published by a commercial publisher.[20] The Conference issue no longer is described as a supplement and is published and dated correctly. For example, the 2017 Conference proceedings are in Volume 56, Issue 2, and dated 2018, not 2017.

Benston (1967) was presented at the May 1967 Conference but published in the Summer of 1968. In private correspondence, Nick Dopuch (Dopuch, 2008) stated: "George [Benston] did not have any empirical results when he presented his paper at the 1967 conference. He knew about the Ball-Brown paper because he attended the Crisp (sic) conference where your paper was presented. We allowed him to rewrite his paper [to include his results] after the conference." This was confirmed in two places by the proceedings editor (Davidson, 1967, iii and 14). If the proceedings issue had been dated according to the current and normal procedure, the publication date would have been 1968, not 1967.

Similarly, Beaver (1968) was presented at the May 1968 Conference but published in the Summer of 1969. If the proceedings issue had been dated according to normal procedure, the publication date would have been 1969, not 1968.

If there is a progenitor to BB68, it is Ashley (1962), a short four-page *Journal of Political Economy* paper on the earnings/price relation. We only learned of this study several years later, but there are some remarkable similarities. The expectation that earnings announcements are related to stock price changes was implicit in the research. The terms "good news" and "bad news" were used. News was defined as change in earnings. Earnings announcement data were taken from the *Wall Street Journal*. Price changes were tabulated over various intervals from the day of the announcement to 28 days after. There may be other precedents of which we are unaware, but to our knowledge Ashley (1962) is the first academic study of earnings announcements and stock price changes.

There also are important differences between the two papers. Ashley motivated his research in terms of classical ideas in Investments, whereas we saw our paper as a test of a paradigm in Accounting. Perhaps as a consequence, Ashley did not analyse how earnings related to price changes over any period before the announcements, whereas we studied the entire year of pre-announcement price behaviour. Consequently, Ashley did not address important accounting issues such as value relevance and timeliness. Rather, he studied a narrow time window, namely the day of and the month *after* the announcement, presumably to make sure he captured the full announcement effect (Fama had not yet introduced the notion of market efficiency). In current terminology, Ashley (1962) was the first study of the event-day announcement effect and post-announcement "drift."

A close reading of the statistics in Ashley (1962) reveals he most likely did not discover significant stock price reactions. The sample was limited to 65–93 stocks. Moreover the sample period totalled only eight months (August to November in two years, 1951 and 1955) and thus it probably would have consisted primarily of interim reports (of earnings for the first six months of the financial year). Relative to the size of the earnings changes, the reported magnitudes of the price effects were tiny. For example, earnings increases of 100–200% were associated with an approximately 1% average price increase on the announcement day and the day following the announcement; earnings decreases in excess of 50% were associated with an average 2% price decline. The standard errors appear to assume independence, despite the small number of independent time periods. Compounding the significance issue is the fact that the small sample was analysed in 64 overlapping categories as if the analyses were independent.

In sum, Ashley (1962) is the first known, published paper to investigate an association between earnings announcements and price changes. Ashley investigated price changes at and after, but not before, announcements. It is also fair to say that Ashley believed he had found a significant association, although a careful interpretation of his research design and reported numbers suggests that he most likely did not.

7. Importance of replicability

We noted in the Introduction that replicability possibly is the most important criterion for scientific work. In that regard, three insidious practices that undermine the credibility of research have drawn increased attention from concerned researchers worldwide: p-hacking, HARKing, and selective publication.

P-hacking or data mining involves altering research design choices such as model specification and data trimming until a statistically significant *p*-value is obtained. Modern software packages provide multiple design options that provide many combined choices. If enough combinations of models and sample selection criteria are estimated then, depending on the degree to which they are independent, the likelihood of at least one appearing to be significant by chance exceeds the chosen p-value, even when the true treatment effect is zero.

HARKing (Hypothesizing After the Results are Known) involves altering hypotheses after the fact to fit the observed results. In the social sciences, researchers usually can find some set of assumptions to rationalize any set of results. The effect of this practice is accentuated by focusing on coefficients' *p*-values rather than their magnitudes; in many contexts, modern sample sizes make it comparatively easy to obtain a statistically – if not "economically" – significant result.

Selective publication occurs when papers reporting insignificant results are discarded, either by researchers not submitting them for publication or by journal editors and referees rejecting submissions with insignificant results. When a large number of researchers test their ideas, some are likely to appear significant by chance when the true treatment effect is zero, and these are the tests that are more likely to be submitted and published.

In all such cases, the p-values in conventional tests statistics are over-stated, perhaps substantially so, because they assume that a

[20] Until then, *JAR* was published internally by the Chicago accounting group, which arranged refereeing and editing, copy-editing, typesetting, proof-reading, printing, mailing, subscriptions, etc. Over time that became unsustainable, hence the shift to commercial publication. The University of Chicago retained the editorial rights.

single specification has been tested. That is not to say that researchers consciously commit academic fraud (though there are notorious cases of that). Even well-intentioned researchers can be lulled by the ease with which one can "turn the dials" in statistical packages and obtain a result consistent with one's favourite hypothesis.

Some idea of the extent of the problem can be obtained from a large project undertaken by the Open Science Collaboration.[21] Research teams around the world were each assigned the task of independently replicating one of one hundred studies published in three top psychology journals. The widely-publicized results were alarming. Only 36% of the independent replications were statistically significant at the conventional level (i.e., with p-values $\leq .05$), whereas 97% of the original studies had passed that test. In other words, almost two out of every three studies failed to replicate. Problems such as p-hacking and HARKing (and, more insidiously, faking the data) only can be exposed by failure to replicate results using different data samples or different research specifications.

Replication does face some serious headwinds. Colleagues and deans tend to view mere replication as lacking substantial contribution. Academic journals tend to view it as of low interest to their readers and detrimental to their impact factors and ratings. So we were delighted when the Editor asked us to conduct and report the current replication.

8. Concluding observations

To commemorate the 50th anniversary of the publication of Ball and Brown (1968) we investigate how well its basic results replicate subsequently in the USA and in 16 other countries, successful replication being a hallmark of sound research. We also provide further explanation of the background to the original paper and its research design, as well as some comments on its interpretation, its impact on research and practice, and its relation to papers published around the same time.

We are pleased to report that the principal results replicate well. Changes in annual accounting earnings continue to encapsulate a large fraction of the information that investors trade into share prices over the year, so earnings remain "value relevant". Almost all of that information is incorporated into prices before earnings are announced, so annual earnings still are not particularly timely. Abnormal returns after earnings announcements can be predicted to a degree from the announcement, so anomalous "post earnings announcement drift", or PEAD, continues.

Nevertheless, there are notable differences across countries and over time. For the USA in particular, changes in accounting earnings have been associated with a declining fraction of share price changes over the year leading up to the earnings announcements.[22] There are several possible interpretations of this result. One possibility is that an increase in the importance of intellectual capital investments with long-term payoffs has reduced the proportion of value-relevant information that is captured in annual earnings. Another possibility is a database issue, involving a shift in mix between the properties of public and private companies, with mature firms being more likely to adopt private status and with intellectual capital firms being more likely to adopt public status. Perhaps stock prices have become noisier.

Another notable change is that the reporting lag has shortened in most countries over the years, and the lag's interquartile range typically has narrowed. This presumably reflects reduced information processing costs shortening the period required to compile, review and audit the numbers and, perhaps as a consequence, tightened regulatory filing deadlines.

It has been a marvellous half century. We hope, and expect, that the sheer vibrancy of the literature since BB68 will continue for many years to come.

Acknowledgments

We thank Robert Faff for encouraging us to undertake this project and for his and Marvin Wee's advice and support during the process. We also thank Hiroyuki Aman for assisting us with a sample of the dates and times of profit announcements made by Japanese companies, and the Securities Industry Research Centre of Asia-Pacific (SIRCA) for supplying data on similar announcements disseminated by the Australian Stock Exchange. Neither author has a financial interest in the outcome of this research. We received helpful comments from Bryan Howieson, S.P. Kothari, Christian Leuz, Douglas Skinner and Terry Walter. Ball gratefully acknowledges financial support from the University of Chicago, Booth School of Business and Brown from the UWA Business School.

References

Acker, D., Duck, N.W., 2009. On the Reliability of I/B/E/S Earnings Announcement Dates and Forecasts. Discussion Paper No. 09/611, Department of Economics, University of Bristol, 8 Woodland Road, Bristol UK.
Ashley, J.W., 1962. Stock prices and changes in earnings and dividends. J. Polit. Econ. 70, 82–85.
Ball, R., 1978. Anomalies in relationships between securities' yields and yield surrogates. J. Financ. Econ. 6, 103–126.
Ball, R., Brown P., 1967. The information value of the annual earnings report. Working Paper Presented at the November 1967 Seminar on the Analysis of Security Prices Organized by the Center for Research in Security Prices (CRSP) at the University of Chicago. SSRN https://ssrn.com/abstract=2128616 or https://doi.org/10.2139/ssrn.2128616 (accessed 7 December 2018).
Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. J. Account. Res. 6, 159–178.
Ball, R., Brown, P., 2014. Ball and Brown (1968): a retrospective. Account. Rev. 89, 1–26.

[21] Open Science Collaboration (2015).
[22] Outside the USA, the number of observations we had available to calculate the ratio on an annual basis is too small to provide more than an indication of whether this is a universal phenomenon.

Ball, R., Shivakumar, L., 2008. How much new information is there in earnings? J. Account. Res. 46, 975–1016.
Ball, R., Robin, A., Sadka, G., 2007. Is financial reporting shaped by equity markets or by debt markets? An international study of timeliness and conservatism. Rev. Acc. Stud. 13, 168–205.
Bamber, L.S., Christensen, T.E., Gaver, K.M., 2000. Do we really "know" what we think we know? A case study of seminal research and its subsequent overgeneralization. Acc. Organ. Soc. 25, 103–129.
Barber, B.M., DeGeorge, E.T., Lehavy, R., Trueman, B., 2013. The earnings announcement premium around the globe. J. Financ. Econ. 108, 118–138.
Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. J. Account. Econ. 24, 3–37.
Beaver, W.H., 1968. The information content of annual earnings announcements. J. Account. Res. 6 (Supplement), 67–92.
Beaver, W.H., McNichols, M.F., Wang, Z.Z., 2017. Increased information content of earnings announcements in the 21st century: An empirical investigation. SSRN. https://doi.org/10.2139/ssrn.3071317. https://ssrn.com/abstract=3071317 (accessed 27 December 2018).
Benston, G.J., 1967. Published corporate accounting data and stock prices. J. Account. Res. 5 (Supplement) (1–14 and 22–54).
Berkman, H., Truong, C., 2009. Event day 0? After-hours earnings announcements. J. Account. Res. 47 (1), 71–103.
Bernard, V.L., Thomas, J.K., 1989. Post-earnings-announcement drift: delayed price response or risk premium. J. Account. Res. 27 (Supplement), 1–36.
Bernard, V.L., Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. J. Account. Econ. 13, 305–340.
Brown, P., 1970. The impact of the annual net profit report on the stock market. Aust. Account. 40 (July), 277–283.
Brown, P., Ball, R., 1967. Some preliminary findings on the association between the earnings of a firm, its industry, and the economy. J. Account. Res. 5 (Supplement), 55–77.
Brown, L.D., Larocque, S., 2013. I/B/E/S reported actual eps and analysts' inferred actual eps. Account. Rev. 88, 853–880.
Canning, J.B., 1929. The Economics of Accountancy. The Ronald Press Co, New York.
Chambers, R.J., 1966. Accounting, Evaluation and Economic Behavior. Prentice-Hall, Inc., New Jersey.
Coase, R.H., 1937. The nature of the firm. Economica 4 (16), 386–405.
Collins, D.W., Maydew, E.L., Weiss, I.S., 1997. Changes in the value-relevance of earnings and book values over the past forty years. J. Account. Econ. 24, 39–67.
Davidson, S., 1967. Editor's preface. J. Account. Res. 5 (Supplement).
Dopuch, N., 1983. The impact of technology on accounting research - past, present, and future. In: Summers, E.L. (Ed.), Technological Change: Its Impact on Accounting. 1982 Proceedings of the Arthur Young Professors' Roundtable. The Council of Arthur Young Professors, Reston, VA, pp. 175–199.
Dopuch, N., 2008. Personal Email to Ray Ball. (February 15, 2008).
Elliott, J.A., Hanna, J.D., 1996. Repeated accounting write-offs and the information content of earnings. J. Account. Res. 34, 135–155.
Fama, E.F., 1965. The behavior of stock-market prices. J. Bus. 38, 34–105.
Fama, E.F., Fisher, L., Jensen, M.C., Roll, R., 1969. The adjustment of stock prices to new information. Int. Econ. Rev. 10 (1), 1–21.
Fisher, L., 1966. Some new stock-market indexes. J. Bus. 39, 191–225.
Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. Account. Rev. 59 (4), 574–603.
Gao, X., Ritter, J., Zhu, Z., 2013. Where have all the IPOs gone? J. Financ. Quant. Anal. 48, 1663–1692.
Grossman, S., Stiglitz, J., 1980. On the impossiblity of informationally efficient markets. Am. Econ. Rev. 70, 393–408.
Holmstrom, B.R., Tirole, J., 1989. The theory of the firm. In: Schmalensee, R., Willig, R. (Eds.), Handbook of Industrial Organization. Elsevier, Amsterdam, pp. 61–133.
Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. J. Financ. Econ. 3, 305–360.
Kothari, S.P., Sabino, J.S., Zach, T., 2004. Implications of survival and data trimming for tests of market efficiency. J. Account. Econ. 39, 129–161.
Kuhn, T.S., 1970. The Structure of Scientific Revolutions, 2nd ed. University of Chicago Press, Chicago.
Lev, B., Gu, F., 2016. The End of Accounting and the Path Forward for Investors and Managers. Wiley Finance, New York.
Livnat, J., Mendenhall, R.R., 2006. Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. J. Account. Res. 44 (1), 177–205.
McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? J. Financ. 71, 5–32.
Nelson, C., 1973. A priori research in accounting. In: Dopuch, N., Revsine, L. (Eds.), Accounting Research. A Critical Evaluation. Center for International Education and Research in Accounting University of Illinois, Urbana, Illinois, pp. 1960–1970.
Ogden, C.K., Richards, I.A., 1923. The Meaning of Meaning: A Study of the Influence of Language and Thought and of the Science of Symbolism. Harcourt Brace, Oxford, England.
Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. Science 349 (625) 28 August. http://science.sciencemag.org/content/349/6251/aac4716 (accessed 7 December 2018).
Stanford Encyclopaedia of Philosophy, 2011. Kuhn, T. https://plato.stanford.edu/entries/thomas-kuhn/ (accessed 13 February 2018).
Tirole, J., 2006. The Theory of Corporate Finance. Princeton University Press.
Watts, R.L., 1977. Corporate financial statements: A product of the market and political processes. Aust. J. Manag. 2, 52–75.
Watts, R.L., Zimmerman, J.L., 1986. Positive Accounting Theory. Prentice-Hall, Englewood Cliffs, N.J.
Working, H., 1960. Note on the correlation of first differences of averages in a random chain. Econometrica 28, 916–918.
WRDS, 2018. Difference Between I/B/E/S Earnings Announcement Date and Compustat Announcement Dates. https://wrds-web.wharton.upenn.edu/wrds/support/AdditionalSupport/WRDS KnowledgeBasewithFAQs.cfm?folder_id=653 (accessed 20 April 2018).