

イベントスタディで見た「買い推奨銘柄」の株価動向

辻幸民

2025 年 1 月 20 日『三田商学研究』第 67 巻 4 号再校校正後

概要

本稿ではイベントスタディの手法に依拠して、いわゆる「買い推奨銘柄」の情報がもたらす株価動向への影響を検討する。従来のイベントスタディと共通する特徴としては、イベントの与える株価への情報効果は確かに存在する。異常収益率 AR はイベント発出直後が最大の正値を記録し、その有意性は極めて高い。しかしイベント後の期間で他に正で有意な AR の時点は存在しない。他方、従来の研究と異なる点に、モーメンタム効果が捉えられることがある。イベント前の期間で AR が正で有意な時点がいくつか存在し、そのためイベント発出前から累積異常収益率 CAR が正の高い有意性を持つ。イベント前の期間を通じて、CAR はイベント当日に向け上昇基調をたどり、イベント当日直後に大きく跳ね上がって、イベント後は一進一退の動きとなる。この動向は、直近の株価が上昇基調にあった銘柄が「買い推奨銘柄」に選定されやすいといったことの反映ではないか。ただモーメンタムとはいっても、それほど長い期間にわたり効果が継続するわけではない。イベント後およそ 2 か月ほどすると、「買い推奨」という情報が株価に与えた正の効果は消滅してしまうことも分かる。

キーワード：イベントスタディ、異常収益率 AR、累積異常収益率 CAR、モーメンタム

1 はじめに

イベントスタディ event study とは、特定のイベントについて、その前後十数日間の株価動向を観察することで、そのイベントの株価に与える情報効果を検証するための手法である。イベントとは公開情報が発出されることをいう。ここの公開情報とは、本来は企業財務に関するニュースが主であったが、今日では企業業績との関連性の不明瞭な、合併買収 M&A や不祥事報道など、もっと広範囲の様々な種類のニュースがイベントとして取り上げられるようになっている。

イベントスタディの本来の目的は、次のような考え方の現実妥当性を検証することにあった。ファイナンス論 financial economics の見方によると、株価は、予想される将来配当金の現在価値合計であると考えられる。将来配当金はそのときそのときの業績次第である。従って、企業の将来業績に影響を与えそうな今のイベント（ニュース）は、その企業の将来配当金の予想に影響を与え、この影響は今の株価を左右することになる。それでは今発生したイベントは、どれぐらいのスピードで株価に影響を与えるのか。イベントの影響は即座に株価に反映されるか、それとも時間をかけてゆっくり株価に織り込まれるのか。この点を検証しようとするのがイベントスタディ本来の目的である。イベントが即座に株価に織り込まれることを情報効率的という。すなわち、イベントスタ

ディは歴史的には、株式市場が情報効率的な市場か否かを検証するための手段として利用されてきた。

株式市場の情報効率性に関する検証結果は、様々な解釈がなされているが、概ね肯定的に捉えてよからう。イベント発出に対し、その後の比較的短い期間で株価はニュースを織り込む。この織り込むさまが、イベントスタディにおける累積異常収益率 (cumulative abnormal return, 略して CAR) という尺度で見とれる。統計学の仮説検定により、イベント直後の数日間に関する CAR はゼロから有意に差がある。株価にプラスのイベントであれば正で有意な CAR が、また株価にマイナスのイベントであれば負で有意な CAR が観察されることでもって、ニュースが株価に織り込まれたと考える。ところがイベント発出の前は当然のこと、市場参加者はニュースを知らず、イベント前の株価動向はそのイベントとは無関係のはずであり、であるなら、イベント前の期間について CAR はゼロから有意に異ならない。以上が、情報効率的な場合の株価動向として観察されるはずの、イベント前後の CAR の特徴となる。

株価はあくまでも将来の企業業績を予想して値動きするから、イベントの中身が本当に企業の将来の業績を左右するものなのかどうか、まず最初に問われるべき点ではある。しかし今日のイベントスタディで取り上げられるイベントの中身は、あまりこの点を問題視することは少ないように思う。例えば企業の不祥事ニュースをイベントと考える場合、その不祥事が企業の将来業績にどう影響するのか。ケースバイケースで様々であろうことだけは明らかであるが、普遍的な因果関係についてはよく分からないのが率直なところではないか。イベントが買収合併であっても、多かれ少なかれ事情は同じであろう。買収合併が企業の将来にプラスかマイナスかはやはりケースバイケースであろう。イベントの中身が企業の将来業績にどう影響するか不明確であるにもかかわらず、イベントがその直後の株価動向にどう影響するかを調べることにどれほどの意義があるのか、筆者個人は昔からそう考えてきた。であるなら、むしろ逆説的に考えて、イベントの中身そのものが企業の将来業績に何の影響を持たないことが自明なものについて、イベントスタディを試みるとどのような結果を得るのか。そのような問題意識が根本にあり、本稿を執筆するに至った。

ここで筆者が取り上げる、イベント自体が企業の将来業績に影響しないイベントとは、第3者による株価の値上り予想である。株式銘柄企業とは何の関係もない第3者、例えば証券会社や金融情報会社等がいわゆる「買い推奨銘柄」に選定して公開発表すること、これはその第3者がその銘柄はこれから値上りしていくかもしれないという予想を広く表明していることを意味する。買い推奨銘柄に選定されたとしても、そのこと自体がその企業の将来業績に無関係なのは自明である。そもそも「買い推奨銘柄」が本当に将来の株価値上りを観察することになるのか。このように本稿では、当該企業とは無関係の第3者による、株価の値上り期待の表明が株価に影響を与えるか否かをイベントスタディの観点から検討したい。

ところで本稿の取り上げる「買い推奨銘柄」とは、東洋経済新報社がネットサイトで年4回の頻度で定期的に公開している「会社四季報オンライン・先取りサプライズ銘柄」である。2014年から2020年までの間に選定された1000銘柄超を対象として、これら銘柄にイベントスタディを適用した。通常のイベントスタディと同様の特徴としては、イベントの与える株価への情報効果は確かに存在する。「買い推奨銘柄」という情報は確かに株価に正の影響を与えるが、その効果は情報

の発出直後だけに限定される。他方、普通のイベントスタディとは若干異なる特徴も観察される。それはモーメンタム効果が捉えられる点である。ここのモーメンタム効果とは、過去上昇基調にあった株価は、これからもそのまま上昇基調を維持する傾向にあることをいう。本稿のイベントスタディの結果では、情報発出直後の株価の上昇のみならず、情報発出のかなり前から株価はゆっくりとした上昇基調を示す。これは、直近の株価が上昇基調にあった銘柄が「買い推奨銘柄」に選定されやすいといったことを反映しているのではないか。本稿のモーメンタム効果とは、過去上昇基調にあった株価が、情報発出でさらに跳ね上がり、その後も一定期間好調を維持するような様子であることをそう称している。ただモーメンタムとはいっても、それほど長い期間にわたって効果が継続するわけではない。本稿の計算の示すところでは、概ね 2 か月ほどで、「買い推奨」という情報の影響が一掃されてしまうのではないか。

本稿の構成は以下のとおりである。第 2 節でイベントスタディの手法について概説する。第 3 節では本稿の分析で用いたデータと設定方法を述べた後、計測結果を提示する。第 4 節は結びである。その後に付録として、イベントスタディの仮説検定の手法を詳細に導出しておく。

2 イベントスタディの手法

イベントスタディ event study とは、特定のイベントの前後数日間の株価動向を観察することで、そのイベントが株価にどう影響しているかを見るための手法である。特にイベントのもたらす情報が株価に織り込まれる様子を数量化して、株価への情報効果を検証しようというねらいがある。単独の株価動向で、イベントの情報効果を検証することも可能ではあるが、単独の株価動向が分析対象になることは稀で、普通は複数の株式銘柄を対象に、共通のイベントについて、その複数銘柄の平均的な株価動向から情報効果を検証する場合が多い。まずはここで典型的なイベントスタディを想定し、その手法に関する概説を与える。特に最重要な、仮説検定を行う際の統計量についてまとめておく。この導出についての詳細な議論は巻末の付録に記す。

イベントスタディで最も大切なのはイベント日の特定化である。イベントとは、典型的には増配(配当金増額)報道のような公開情報に関するニュースである。増配ニュースが発出された日がイベント日であるとする。本稿ではイベント日当日を「 $d \pm 0$ 」で表す。ただし、株価に直接影響を与え得るようなニュースは市場終了後の午後 3 時以降に発表されることも多いから、「 $d \pm 0$ 」が必ずしも本当のイベント初日というわけではない。イベントが株価に影響を与える情報効果の検証として、特にその株価動向が注目されるのはイベント日の翌日「 $d+1$ 」や翌々日「 $d+2$ 」である。ところで、影響があるか否かを観察するには、イベント日の前の期間についても株価動向を把握しておく必要があろうし、イベント日のしばらく後の期間についても注目しなければならない。そこで、イベント当日の 10 日前「 $d-10$ 」から、イベントの発出後 10 日間「 $d+10$ 」までの株価動向が問題にされる。イベント当日の前後 10 日間、合計でこの 21 日間のことをイベント期間と称する(英文では event window と称される)。また前後 10 日間ではなく、前後 20 日間を対象にした合計 41 日間をイベント期間とする場合もある。

今まで株価動向と称したが、イベントスタディの学術研究で株価そのものが対象となることはほ

とどなく、普通は株式収益率が株価動向を見るのに使われる。もう少し正確に述べると、OLS(最小 2 乗法) 推定量を用いた、次のような回帰モデルの予測誤差が株価動向を測る物差しとなる。

$$AR_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

ある個別のイベントを i という記号で表し、そのイベントの対象株式銘柄について、 t 時点における株式収益率を $R_{i,t}$ と表現する。 t 時点のマーケットポートフォリオ収益率は $R_{m,t}$ である。 $\hat{\alpha}_i$ と $\hat{\beta}_i$ は、次に定義される α_i と β_i の OLS 推定量である。今、イベントは複数あって N 個あることを想定しておく。つまり $i = 1, \dots, N$ である。この $AR_{i,t}$ は、イベントスタディの世界では異常収益率 abnormal return と称される (以下 AR と略称する) のが一般的であるが、これは統計学でいう回帰モデルの予測誤差に他ならない。

次の回帰モデルは、ファイナンス論ではマーケットモデルと称される。

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t}$$

任意の t 時点におけるイベント i の株式収益率 $R_{i,t}$ が被説明変数、マーケットポートフォリオ収益率 $R_{m,t}$ が説明変数で、 $\epsilon_{i,t}$ は攪乱項である。 $E(\epsilon_{i,t}) = 0$, $E(\epsilon_{i,t}^2) = \sigma_{\epsilon_i}^2$, $E(R_{m,t} \cdot \epsilon_{i,t}) = 0$ が仮定される。このモデルを推定するわけであるが、そのためのデータとして開始時点 t_0 、終了時点 t_1 とする標本があるとする。つまり、 $t = t_0, \dots, t_1$ という各時点における $R_{i,t}$ と $R_{m,t}$ の数字のペアが L 個あるものとしよう。データのサンプルサイズが L であり、 $L = t_1 - t_0 + 1$ である。

マーケットモデルの α_i と β_i は未知であるので、これに関するデータからの推定値 $\hat{\alpha}_i$ と $\hat{\beta}_i$ を得て、 $\hat{\alpha}_i + \hat{\beta}_i R_{m,t}$ と $R_{i,t}$ との差が AR であり、これを株価動向の尺度として分析対象とするのである。この手法の意味は次のような考え方が前提になっている。株式収益率のあるべき値が仮に存在するなら、そのあるべき値がマーケットモデルで決められ、そのあるべき値が $\hat{\alpha}_i + \hat{\beta}_i R_{m,t}$ であると考え。そのときの実際の値は $R_{i,t}$ である。両者は確率的な不特定要因により変動するので、両者の差額 AR は常にランダムに変動しているが、何かの特別な要因を $R_{i,t}$ だけが織り込んで、あるべき値の方に影響がないなら、AR の値はその特別な要因の分だけ従来とは違う変化をするはずである。従って AR は「異常収益率」という命名がなされ、AR の動向で、何か特別な要因に影響された株価動向と見なされる。^{*1}

t_0 時点から t_1 時点までのモデル推定のための標本となる期間を推定期間と称する。なお本稿では、推定期間はイベント期間の前の時点のデータを利用する。例えば、 t_0 時点が「d-220」、 t_1 時点が「d-21」として設定されるなら、 L は 200 という具合である。またイベント期間の開始時点と終了時点をそれぞれ t_2 と t_3 で表し、 $t_0 < t_1 < t_2 < t_3$ である。イベント期間を当日の前後 20 日を取った 41 日間とするなら、 t_2 は「d-20」、 t_3 は「d+20」である。ここで注意すべきは休日の

^{*1} もちろん、単純なマーケットモデルでもって株式収益率の「あるべき値」が算出できると考えるのはあまりにも非現実的であろう。しかし他に現実的かつ妥当なモデルが容易には判然としないという事情もあり、一つの方便としてマーケットモデルを用いているのが実用的慣例になっている。単純なマーケットモデルではなく、Fama-French に始まる複数ファクターモデルやそのバリエーションでもって AR を定義する場合もあるし、あるいはもっと単純に時系列の平均値との差額から AR を計算する場合も散見される。本稿では最も典型的と考えられる手法を採用している。

扱いである。当然のこと休日にはデータがない。従ってここの期間の長さとは、実日数の時間ではなく、標本を構成するデータの個数、つまりサンプルサイズを意味している。「d-20」は、イベント当日の20日前を表すが、これは標本の20要素前、イベント当日前の20番目の取引日を指す。通常1か月の取引日数は22日ほどであるから、実日数で見ると「d-20」はおおよそイベント1か月前の日ということになる。

推定される未知パラメタの推定量を定式化しておこう。これらは通常の統計学で学習する単純回帰の推定量そのままであるから、結果のみを記すに留める。標本は $t = t_0, \dots, t_1$ 時点のデータで、そのサンプルサイズは L である。

$$\begin{aligned}\hat{\alpha}_i &= \bar{R}_i - \hat{\beta}_i \bar{R}_m & \bar{R}_i &= \frac{1}{L} \sum_{t=t_0}^{t_1} R_{i,t}, & \bar{R}_m &= \frac{1}{L} \sum_{t=t_0}^{t_1} R_{m,t} \\ \hat{\beta}_i &= \frac{\sum_{t=t_0}^{t_1} (R_{i,t} - \bar{R}_i)(R_{m,t} - \bar{R}_m)}{\sum_{t=t_0}^{t_1} (R_{m,t} - \bar{R}_m)^2} \\ \hat{\sigma}_{\epsilon_i}^2 &= \frac{1}{L-2} \sum_{t=t_0}^{t_1} (R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t})^2\end{aligned}$$

この $\hat{\sigma}_{\epsilon_i}$ は回帰の標準誤差で、攪乱項の標準偏差 σ_{ϵ_i} の推定量となる。

次にこれらパラメタ推定量を用いて、 $t = t_1$ 以降の時点における予測がしたい。 t_2 時点から t_3 時点までのイベント期間について、そのときの $R_{m,t}$ の値からパラメタ推定量 $\hat{\alpha}_i$ と $\hat{\beta}_i$ を用いて、 $R_{i,t}$ の値を予測するのである。この予測誤差が $AR_{i,t}$ である。ただし対象の t は $t = t_2, \dots, t_3$ であって、以下では $K = t_3 - t_2 + 1$ とし、 K はイベント期間のサイズを表す。いうまでもないが、「d-20」から「d+20」のイベント期間なら $K = 41$ である。

イベントスタディでは、 $AR_{i,t}$ のみならず、これをイベント期間の中のある期間について足し上げた累積異常収益率 CAR という尺度もよく用いられる。足し上げる期間を τ_1 時点から τ_2 時点までとする。 $t_2 \leq \tau_1 < \tau_2 \leq t_3$ であり、

$$CAR_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}$$

と定義される。この $CAR_i(\tau_1, \tau_2)$ についても、 $AR_{i,t}$ と同様に、それぞれの平均と分散・標準偏差を計算できる。具体的な式は、もう少し後に検定のための統計量を明示する際に示す。

$AR_{i,t}$ あるいは $CAR_i(\tau_1, \tau_2)$ はどちらも単独の個別イベント、第 i 番目のイベントに関する変数である。イベントスタディでは、共通のイベントが複数個 (全部で N 個) あり、それぞれの i に関して $AR_{i,t}$ や $CAR_i(\tau_1, \tau_2)$ (ただし、 $i = 1, \dots, N$) を計算できるが、これらのクロスセクション平均を取って、株価の動向を捉えるというのが普通の手法になっている。ここで注意すべきことは、各イベント当日の日付はイベントにより異なり得るという点である。例えば「増配情報」の発出日はその銘柄によって様々であろう。イベント当日の日付が異なるから、そのイベント日からの相対的な日数で見た推定期間の初日 t_0 や終日 t_1 、イベント期間の初日 t_2 やその終日 t_3 、CAR の累積計算の期間 τ_1 や τ_2 、これらはすべてイベント毎に異なる日付となろう。しかしそうであった

としても、これら各日がイベント当日からの相対的な時間で決定されていて、その相対的時間の長さは複数イベントすべてで共通しているものとする。例で確認すると、イベント当日を「 $d \pm 0$ 」で表すとし、 t_0 は「 $d-220$ 」および t_1 は「 $d-21$ 」とする。つまり全イベントについて、推定期間の初日には「 $d-220$ 」を当て、推定期間の終日に「 $d-21$ 」を当てる。同様にイベント期間の初日 t_2 に「 $d-20$ 」を、イベント期間の終日 t_3 に「 $d+20$ 」を想定する。そして CAR の累積計算にはイベント期間の中の様々な期間を対象にする。一例として τ_1 が「 $d-5$ 」、 τ_2 が「 $d+5$ 」などとする。以上のような設定を全イベントの共通設定としておいて、複数イベントのクロスセクション集計を行うのである。

それでは、全イベントについてのクロスセクション平均を取ろう。 $i = 1, \dots, N$ として、各イベントの $AR_{i,t}$ および $CAR_i(\tau_1, \tau_2)$ から、

$$\overline{AR}_t = \frac{1}{N} \sum_{i=1}^N AR_{i,t}$$

$$\overline{CAR}(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^N CAR_i(\tau_1, \tau_2)$$

というように $\overline{AR}_t$ と $\overline{CAR}(\tau_1, \tau_2)$ を計算する。 $\overline{AR}_t$ は $AR_{i,t}$ の標本平均、 $\overline{CAR}(\tau_1, \tau_2)$ は $CAR_i(\tau_1, \tau_2)$ の標本平均である。イベントスタディにおける株価動向とは、このように計算された値がゼロから有意に異なっているかどうか問題にされる。そこでこれら標本平均がゼロに等しいという帰無仮説を検定するための統計量が必要になる。

以下、仮説検定で用いられる統計量をいくつか明示していく。

$$Z_0 = \frac{\overline{AR}_t}{\sqrt{\widehat{\text{var}}(\overline{AR}_t)}}$$

この分母に登場する分散 $\widehat{\text{var}}(\overline{AR}_t)$ は、

$$\widehat{\text{var}}(\overline{AR}_t) = \frac{1}{N^2} \sum_{i=1}^N s_{i,t}^2$$

とあるように、 N 個の $s_{i,t}^2$ の平均を取った値をさらに N で割るという式である。この $s_{i,t}^2$ は個別イベント $AR_{i,t}$ の分散推定量で、具体的な計算式は次のとおりである。

$$s_{i,t}^2 = \hat{\sigma}_{\epsilon_i}^2 \left[1 + \frac{1}{L} \left(1 + \frac{(R_{m,t} - \bar{R}_m)^2}{s(R_m)^2} \right) \right]$$

$$\text{ただし, } s(R_m)^2 = \frac{1}{L} \sum_{k=t_0}^{t_1} (R_{m,k} - \bar{R}_m)^2$$

これは昔からよく引用される有名な式である。^{*2} この式の導出は後で詳しく説明する。この分散が

^{*2} イベントスタディは 1980 年代から膨大な数の研究が存在するので、この式を最初に導出した研究者が誰なのか本当のところよく分からないのが率直なところであろうが、この式を明示している古い時代の研究には Mikkelsen-Partch(1986) や Schipper-Smith(1986) がある。

ら計算される AR に関する統計量を Z_0 で表す。

次に CAR を取り上げる。前と同様に標本平均をその標準偏差で除した比率を計算する。

$$Z_1 = \frac{\overline{CAR}(\tau_1, \tau_2)}{\sqrt{\widehat{\text{var}}(\overline{CAR}(\tau_1, \tau_2))}}$$

この分母に登場する分散は次のように求められる。

$$\widehat{\text{var}}(\overline{CAR}(\tau_1, \tau_2)) = \frac{1}{N^2} \sum_{i=1}^N \hat{\sigma}_i(\tau_1, \tau_2)^2$$

この $\hat{\sigma}_i(\tau_1, \tau_2)^2$ は、個別イベント i に関する CAR の分散推定値であり、この正確な形は

$$\hat{\sigma}_i(\tau_1, \tau_2)^2 = \hat{\sigma}_{\epsilon_i}^2 \left[T + \frac{(Ts(R_m))^2 + (G - T\bar{R}_m)^2}{Ls(R_m)^2} \right] \quad (1)$$

$$\text{ただし, } T = \tau_2 - \tau_1 + 1, \quad G = \sum_{k=\tau_1}^{\tau_2} R_{m,k}$$

のとおりに導出できるが、実はこの式が用いられることはあまりない。この正確な分散に依拠して計算される統計量をここでは Z_1 で表す。他方、よく用いられる分散の式は、

$$\hat{\sigma}_i(\tau_1, \tau_2)^2 = \hat{\sigma}_{\epsilon_i}^2 T \quad (2)$$

というものであって、正確な式に登場する複雑な第 2 項を無視したものである。^{*3} 分散をこの簡易式によって計算される統計量を以下では Z_2 とする。 Z_1 と Z_2 が実際どの程度の差異になるのか、これは具体的に確認する必要がある。

ところで $\widehat{SCAR}_i(\tau_1, \tau_2)$ という変数を次の比率で定義する。

$$\widehat{SCAR}_i(\tau_1, \tau_2) = \frac{CAR_i(\tau_1, \tau_2)}{\hat{\sigma}_i(\tau_1, \tau_2)}$$

この $\widehat{SCAR}_i(\tau_1, \tau_2)$ は、 $CAR_i(\tau_1, \tau_2)$ をその標準偏差で標準化したものと考えられる。イベントスタディの仮説検定では、これをクロスセクションについて標本平均した統計量をよく用いる。以下の形の比率を J という記号で表す。

$$J = \sqrt{\frac{N(L-4)}{L-2}} \cdot \overline{SCAR}(\tau_1, \tau_2)$$

この $\overline{SCAR}(\tau_1, \tau_2)$ は $\widehat{SCAR}_i(\tau_1, \tau_2)$ の標本平均

$$\overline{SCAR}(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^N \widehat{SCAR}_i(\tau_1, \tau_2)$$

^{*3} (2) 式を明示しているのは MacKinlay(1997) である。(1) 式を示した先行研究を筆者は知らない。が、筆者が知らないだけで、これを筆者の貢献と主張するつもりは毛頭ない。その導出は巻末の付録で詳しく述べられるが、ごくごくストレートな計算であるから、相当前の時代に多数の方々が既に導出しておられるであろう。

であり、また $\frac{L-2}{N(L-4)}$ はこの標本平均の分散に相当する。^{*4} 後で詳しく述べるが、このように計算される J は標準正規分布 $N(0, 1)$ に従うことが証明できる。

$\widehat{SCAR}_i(\tau_1, \tau_2)$ を計算する際の分母 $\hat{\sigma}_i(\tau_1, \tau_2)$ には計算方法が 2 種類存在した。(1) 式と (2) 式である。従って実際に J の値を求めたいなら、 $\widehat{SCAR}_i(\tau_1, \tau_2)$ を計算する方法が 2 種類あるので、どちらの方法を採用するかで、 J という比率の値も 2 種類計算できる。(1) 式に依拠して求められた J の値を以下では Z_3 で表現し、(2) 式から求められた J の値を Z_4 とする。

以下ではこのようにして計算される 5 つの統計量を求める。AR に関する Z_0 と、CAR に関する Z_1 から Z_4 という 4 種類の統計量である。これらはすべて標準正規分布 $N(0, 1)$ に従うことが導出できる。これら 5 つの統計量をまとめて称する場合、以下では Z 値ということにする。

3 データおよび計測結果

イベントスタディの一般的な手法を概説したので、次に本稿の分析対象を具体的に説明しよう。この節でデータおよび設定を詳しく述べた後に、その下で得られた計測結果を提示する。

3.1 データについて

本稿の対象とするイベントは、第 3 者による「買い推奨銘柄」の公表である。世の中のいわゆる「買い推奨」には様々な種類のおびただしい数があり、そのほとんどは証券会社関係者によるものが圧倒的に多いように思うが、証券会社とは異なる、経済ニュース報道など金融情報の提供を主とする会社によりなされる事実上の「買い推奨」も少なくない。具体的に本稿で取り上げたのが東洋経済新報社がホームページに掲載する「会社四季報オンライン・先取りサプライズ銘柄」(以下「サプライズ銘柄」と略称)である。その呼称こそ「買い推奨」となってないが、掲載された銘柄は、これから株価上昇が予想され、その購入は将来の値上り益獲得を惹起させるであろうところから、この掲載は東洋経済新報社という出版社による事実上の買い推奨と考えてよからう。

この「サプライズ銘柄」についてその経緯と背景をもう少し説明する。恐らく元々の発祥は、『会社四季報』の発刊に併せ発行されている有料別冊『会社四季報プロ 500』にかつてあった袋とじ付録にあるのではないかとはいわずとした、全上場企業の直近業績のデータ(実績値)およびその一部予想をコンパクトにまとめた冊子である。その予想とはあくまでも企業業績(売上や利益など)の予想であって、ストレートに今後の株価の上昇下落を予想しているわけではないことは注意しておく必要がある。筆者の個人的な印象かもしれないが、『会社四季報』そのものの記載情報が株価にどう影響しているかは思いの外曖昧な感じである。他方『会社四季報プロ 500』は、株式投資をより強く意識した内容になっていて、全上場銘柄ではなく一部銘柄(約 500 社)に掲載を厳選した上でその株価も予想されている。至るところ「厳選」という文字が目を引き、この厳選とは「この株価がこれから上がる可能性が大きいから、この(厳選された)銘柄を今購入しておけば、将来儲かりますよ!」という憶測を読者に想起させようとしていることは明白である。

^{*4} この点を利用している古い時代の研究に Eckbo(1983) がある。

『会社四季報』や『会社四季報プロ 500』は今でも一般書店で販売されるごくごく普通の書籍である。『会社四季報プロ 500』に収録されたかつての袋とじ付録とは、いうまでもないかもしれないが、いわゆる書店での立ち読み(ただ読み)で内容を見られないよう封で隠された一部ページのことである。『会社四季報プロ 500』自体が、3000 社ほどの全上場銘柄の中から 500 社ほどに「厳選」された銘柄群から成るが、その中でも厳選に厳選を重ねた特別に有望な十数銘柄を袋とじ付録として掲載したのである。この特別な情報を知りたいければ、是非書籍を購入した上で自分で封を破って見て下さいという趣旨である。ただ、印刷物からインターネットへと情報源の主役交代の時代的変遷に例外はなく、東洋経済新報社でも「会社四季報オンライン」という有料サイトを開設し、そこに『会社四季報』の果たしている役割を担わせようとしている。このような背景があって、『会社四季報』の毎季発刊に先立ち、「サプライズ銘柄」が有料サイトに発表されるようになった。ネットでの発表が始まったのは筆者の知る限り 2010 年過ぎ頃からであろうか。なお『会社四季報プロ 500』は現在でも刊行されているが、印刷物としての袋とじ付録は 2020 年の収録をもって最後となっている。

本稿の分析対象の期間では、『会社四季報プロ 500』袋とじ付録に収録されている銘柄と「会社四季報オンライン・先取りサプライズ銘柄」との 2 種類が混在し、これらは必ずしも一致しているわけではなく、営業上当然のこと不一致が多々見られよう。どちらがイベントスタディのイベントに相応しい情報源であろうか。あくまでも分析者の判断の問題になるが、本稿では会社四季報オンラインの方をイベントスタディのイベントとして取り上げる。どちらの情報源であったとしても、イベント日を完全に特定化できるという点が本稿のイベントスタディの特徴の一つであろうが、両者の間には、一度(発刊日)に多数発表される(印刷物の)袋とじ付録と、毎日 5 銘柄が発表される「サプライズ銘柄」という違いがある。『会社四季報』は年 4 回、3 月・6 月・9 月・12 月の各々中旬に発刊され、その同日に『会社四季報プロ 500』も発刊される。発刊と同時に袋とじ付録の公開である。これを情報源とするなら、この発刊日がイベント日ということになろう。他方「サプライズ銘柄」の方は、会社四季報の発刊に合わせ、概ね発刊日の 2~3 週間前ぐらいから始まって、連日 5 銘柄ずつ計 10 日間で 50 銘柄ほどが発表される。午後 5 時前後の時間に平日休日にかかわらず毎日である。「サプライズ銘柄」を情報源とするなら、適切なイベント日は各銘柄のネット上での公開日であり、イベント日は連日となる。

ところで東洋経済新報社の会社四季報オンラインは有料サイトであるから、そこで発表される情報は、厳密には公開情報ではない。しかしそのベーシック会員の会費は、いわゆる巷のサブスクの平均的レベルであって、書籍や新聞も当然のことそれぞれ入手コストがかかるわけであるから、ネットの有料サイトの発する情報であっても、(法外な利用料でなければ)それを公開情報と考えることに無理はなかろう。本稿では以上のような判断の下、サプライズ銘柄の公開日をイベント日と見なすことにする。取り上げた銘柄は、2014 年新春号から 2020 年秋号までの発刊に合わせる形で発表された銘柄群である。後で述べる推定のための株価の入手可能性を考慮し、分析対象となるのは全部で 1078 銘柄ある。2014 年新春号が発刊されるのは 2013 年 12 月であるが、「サプライズ銘柄」はそれに先立つ 2~3 週間前に発表される。従ってイベント日はその公開日の 11 月中旬の当日となる。2020 年秋号の発刊は 2020 年 9 月であり、サプライズ銘柄の公開は 2020 年 8 月当日で

ある。

イベント日が特定化されたので、その前の期間の株価から必要なパラメタ推定を行う。本稿では推定期間としてイベント前の最大 400 日 ($L = 400$) を想定する。それ以外にも 200 日と 100 日の場合も計算する。イベント期間にはイベント当日の前後 20 日間ずつと前後 40 日間ずつの 2 つを想定する。前後 20 日間ずつの場合は (当日もカウントして) $K = 41$ 、前後 40 日間ずつの場合は $K = 81$ である。本稿では推定期間をイベント期間の前日までとしている。すなわち、推定期間が 400 日で、イベント期間が 40 日ずつの場合なら、イベント当日の 41 日前から 440 日前までが推定期間となり、その期間の株式収益率を使って必要なパラメタ推定を行うことになる。

日々の株価から日次の株式収益率を計算する。 i 番目のイベントに関する銘柄の、 t 時点の株価を $P_{i,t}$ とすると、

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}}$$

という式から株式収益率 $R_{i,t}$ を計算する。この式から明らかなとおり、本稿では配当金を無視している。推定期間 400 日とは実日数なら 2 年弱の期間であり、期中に配当金は 3 回支払われるであろうが、サンプルサイズ 400 の中のわずか 3 つのデータが影響を受けるだけであるから、本稿の議論の範囲内なら、配当金を考慮するか否かの差異は軽微であろうと考えられる。

3.2 計測結果

ここで提示する計測結果は表 1 から表 3 で、これらは推定期間のサンプルサイズ L を 200、イベント期間をイベント当日の前後 20 日間ずつ (「d-20」から「d+20」までの $K = 41$) とする場合についての結果である。これら 3 つの表の違いは CAR を計算する際の累積期間の違いである。以下順に説明していくが、他の推定期間やイベント期間についてもほぼ同様な結果が得られていて、紙幅の都合で掲載は省略する。

表 1 は前後 20 日間ずつのイベント期間全体を俯瞰するためのものである。表の AR という列は $\overline{AR}_t$ の値で、 Z_0 の列は t 時点の $\overline{AR}_t$ に関する検定統計量である。なお有意性を示すために普通付される星印は、見映えが雑多になるだけなのでここでは省略している。統計量の Z_0 から Z_4 はすべて標準正規分布に従い、絶対値で見ても (両側検定で) 1.96 が 5% 有意、2.58 が 1% 有意であるから、 $Z_0 \sim Z_4$ の値を見れば概ねの有意性は自明であろう。またこの表 1 では、イベント期間の初日「d-20」を累積の初日 τ_1 時点に設定している。 CAR という列は $\overline{CAR}(\tau_1, \tau_2)$ の値であり、例えば表の t 時点が「d-7」であるなら、その行の CAR は $\tau_1 = d-20$ 、 $\tau_2 = d-7$ とする場合の $\overline{CAR}(\tau_1, \tau_2)$ である。 $\overline{CAR}(\tau_1, \tau_2)$ に関して、分散の計算式を (1) 式に依拠する場合の統計量が Z_1 、(2) 式に依拠する場合の統計量が Z_2 である。表の $SCAR_1$ と $SCAR_2$ という 2 つの列はどちらも $\overline{SCAR}(\tau_1, \tau_2)$ であるが、分散の計算式を (1) 式に依拠する場合が $SCAR_1$ で、これに関する統計量が Z_3 、分散を (2) 式に依拠する場合が $SCAR_2$ で、その統計量が Z_4 である。

表 1 の AR を見ると、イベントの事実上の初日である「d+1」時点で正で有意というだけでなく、 Z_0 の値と AR の値どちらも最大となっていて他の時点と比べ非常に大きい。イベント発出日

表1 AR と CAR の計算結果：イベント期間「d-20」から「d+20」（その1）

t	AR	Z_0	$\tau_1 = d-20, \tau_2 = t$						
			CAR	Z_1	Z_2	$SCAR_1$	Z_3	$SCAR_2$	Z_4
d-20	0.001941	2.454	0.001941	2.454	2.467	0.14906	4.869	0.14889	4.864
d-19	-0.000211	-0.267	0.001730	1.541	1.555	0.14127	4.615	0.14106	4.608
d-18	-0.000539	-0.682	0.001190	0.864	0.874	0.12062	3.940	0.12037	3.932
d-17	0.000822	1.039	0.002012	1.262	1.279	0.14172	4.630	0.14262	4.659
d-16	0.002077	2.624	0.004090	2.288	2.325	0.16025	5.235	0.16237	5.304
d-15	-0.000425	-0.537	0.003665	1.867	1.902	0.14796	4.833	0.15001	4.900
d-14	0.001310	1.656	0.004975	2.341	2.390	0.16775	5.480	0.17046	5.568
d-13	0.002615	3.305	0.007589	3.333	3.411	0.20130	6.576	0.20507	6.699
d-12	0.001449	1.832	0.009038	3.733	3.829	0.20557	6.715	0.20960	6.847
d-11	-0.000734	-0.928	0.008304	3.246	3.338	0.19323	6.312	0.19710	6.439
d-10	0.000627	0.793	0.008931	3.321	3.423	0.20012	6.537	0.20501	6.697
d-9	-0.000281	-0.356	0.008650	3.073	3.174	0.19218	6.278	0.19755	6.453
d-8	-0.000038	-0.049	0.008612	2.933	3.036	0.18435	6.022	0.18994	6.205
d-7	0.000579	0.733	0.009191	3.009	3.122	0.18982	6.201	0.19674	6.427
d-6	-0.000956	-1.210	0.008235	2.599	2.703	0.17314	5.656	0.17983	5.874
d-5	0.000506	0.641	0.008741	2.665	2.778	0.18098	5.912	0.18832	6.152
d-4	0.001150	1.455	0.009891	2.919	3.049	0.18470	6.033	0.19256	6.290
d-3	-0.000139	-0.177	0.009751	2.790	2.922	0.18061	5.900	0.18870	6.164
d-2	0.000502	0.635	0.010253	2.849	2.990	0.18045	5.895	0.18884	6.169
d-1	0.001494	1.890	0.011747	3.173	3.339	0.19400	6.337	0.20340	6.645
d±0	0.001292	1.633	0.013039	3.429	3.617	0.20553	6.714	0.21607	7.058
d+1	0.009857	12.470	0.022896	5.868	6.205	0.30002	9.801	0.31632	10.333
d+2	-0.000146	-0.185	0.022750	5.690	6.030	0.29409	9.607	0.31030	10.136
d+3	-0.000374	-0.473	0.022375	5.464	5.806	0.29000	9.473	0.30631	10.006
d+4	0.000310	0.391	0.022685	5.415	5.767	0.28920	9.447	0.30582	9.990
d+5	0.000444	0.562	0.023129	5.400	5.766	0.29083	9.501	0.30789	10.058
d+6	-0.000870	-1.099	0.022260	5.089	5.445	0.28059	9.166	0.29774	9.726
d+7	-0.002348	-2.968	0.019912	4.461	4.783	0.26054	8.511	0.27671	9.039
d+8	0.000827	1.045	0.020738	4.555	4.895	0.26328	8.601	0.28017	9.152
d+9	-0.001021	-1.291	0.019717	4.248	4.576	0.25399	8.297	0.27029	8.830
d+10	0.000045	0.057	0.019762	4.180	4.512	0.25264	8.253	0.26959	8.807
d+11	0.000605	0.764	0.020367	4.232	4.577	0.25666	8.384	0.27455	8.969
d+12	0.000477	0.602	0.020844	4.255	4.612	0.26041	8.507	0.27943	9.128
d+13	0.000380	0.481	0.021224	4.260	4.627	0.26308	8.594	0.28291	9.242
d+14	0.000765	0.966	0.021989	4.341	4.724	0.27054	8.838	0.29177	9.531
d+15	0.000525	0.664	0.022514	4.375	4.770	0.27277	8.910	0.29523	9.644
d+16	-0.000867	-1.096	0.021647	4.140	4.524	0.26614	8.694	0.28883	9.435
d+17	0.000061	0.077	0.021708	4.089	4.476	0.26611	8.693	0.28936	9.452
d+18	-0.000277	-0.350	0.021430	3.977	4.362	0.25965	8.482	0.28300	9.245
d+19	-0.000547	-0.690	0.020884	3.820	4.197	0.25309	8.268	0.27646	9.031
d+20	0.000612	0.772	0.021495	3.876	4.267	0.25389	8.294	0.27788	9.077

(注) t は時点、イベント当日(ニュース発出日)を「d±0」とし、その当日からの日数を表す。ただしこの日数とは時間的な実日数ではなく、構成サンプルの要素個数を表す。例えば「d-3」とはイベント当日から3つ前のデータを意味する。 AR は t 時点の $\overline{AR}_t$ を表し、 Z_0 はそれに関する検定統計量である。 CAR は $\overline{CAR}(\tau_1, \tau_2)$ であり、その検定統計量が Z_1 と Z_2 である。分散の計算式に(1)式を想定する場合が Z_1 、分散に(2)式を想定するのが Z_2 である。なおこの表の CAR 等の累積期間には、 $\tau_1 = d-20$ および $\tau_2 = t$ を想定している。すなわち、初期時点の「d-20」から各 t 時点までを累積期間とした $\overline{CAR}(\tau_1, \tau_2)$ を計算している。 $SCAR_1$ と $SCAR_2$ は両方とも $\overline{SCAR}(\tau_1, \tau_2)$ であり、分散の計算式を(1)式に依拠しているのが $SCAR_1$ で、(2)式に依拠したのが $SCAR_2$ である。 $SCAR_1$ に関する検定統計量が Z_3 、 $SCAR_2$ に関するものが Z_4 である。サンプルサイズは $L = 200$ 、 $N = 1078$ である。

以降の時点で、他に正の有意性を持つ日は存在しない。「d+7」時点だけ AR が負で有意となっているという点を除き、他はすべて AR がゼロから有意に差があるとはいえない。イベント後よりもイベント前の期間において、ゼロから正の有意差がある時点がいくつか観察され、「d-20」「d-16」「d-13」という3時点で正の有意差が認められる。この点は本稿のイベントスタディの特徴ではないかと考えられるが、これについては後述する。

次は表1の CAR の動向である。イベント前の期間で正で有意の AR がいくつか出現することを反映し、イベント後の期間は無しのこと、イベント前の期間であっても CAR の有意性は高い。 Z_3 と Z_4 という統計量で見ると、イベント期間全体で一貫して CAR の値は正で有意といえる。 Z_1 と Z_2 という統計量でも、「d-13」時点より後は一貫して CAR はゼロよりも有意に大きい。イベント発出後の初日「d+1」で、 AR の値そのものとその有意性が高いため、 CAR についてもその値は跳ね上がり、いずれの Z 値で見てもその有意性は急上昇している。

通常よくあるイベントスタディの結果では、市場が情報効率的であるなら、イベント前の CAR はゼロからの有意差がほとんどなく、イベント当日直後の AR の大きさを反映させて、イベント後の CAR はゼロから有意に異なるようになるというのが典型であろう。しかし本稿の結果では、イベント前の期間でも CAR は正の有意性が十分に高い。これは一体なぜなのか。この理由を考察すると、本稿のイベントスタディの特徴的な点が浮き上がる。

通常のイベントスタディのイベントとは、企業財務に関する公開情報の発出の場合であることが多いが、本稿のイベントは、企業とは関係ない第三者による株価値上り予想の発出である。イベント内容が企業財務の公開情報であるなら、それが公開時点より前に外部漏出した可能性は、一昔前ならいざ知らず今日においてはほとんどあり得ないことであろう。そのため、イベント前にイベントの中身が株価に影響を与えたとは考えにくく、 CAR の値もそれと整合的な動きを見せる。しかし、本稿のイベント内容は少なからず影響力ある出版社による、株価値上り予想の表明である。たとえその予想が事前に外部に漏出しなかったとしても、 CAR が正で有意になっている可能性は十分あり得る。というのは、予測者(=今の場合は出版社記者)はどうやって値上り予想銘柄を選別しているのかを考えてみよう。これから値上りしそうな銘柄を予測者はいろいろな角度から慎重に検討しているのかもしれないが、手取り早く有望銘柄を選別する方法として、モーメンタムという手法があるように、直近から今現在に至るまでに株価が値上り基調にあった銘柄をこれからの値上り有望銘柄として選んでいる可能性は少なからずあろう。この点を反映して、値上り予想が発出される前の期間に、 CAR は正で有意な値となっているのではないか。

表2は、累積期間の初日をイベント当日に近い「d-5」に設定して t 時点までの CAR を計算したものである。表1と比べ、 CAR の値そのものは小さくなり、 Z 値も全体的に小さくなるので有意性は下がるかの印象である。ただし Z_3 や Z_4 で見ると、ほとんどの時点でやはり正で有意であり、 Z_1 や Z_2 で見ても「d-1」や「d±0」では完全に正の有意差がある。このように累積期間をイベント当日間近に絞って計算したとしても、イベント前の時期に正の CAR は根強く観察されそうである。この点も前に述べたモーメンタム効果と整合的であろう。

表3では、イベント当日から少々の時間を経た後を累積期間の初日として CAR を計算したものである。「d+5」を初日としている。翌「d+6」から一貫して CAR は負値である。 Z_1 と Z_2 で見る

表 2 CAR の計算結果：イベント期間「d-20」から「d+20」(その 2)

t	$\tau_1 = d-5, \tau_2 = t$						
	CAR	Z_1	Z_2	$SCAR_1$	Z_3	$SCAR_2$	Z_4
d-5	0.000506	0.641	0.643	0.05689	1.858	0.05681	1.856
d-4	0.001656	1.478	1.488	0.06949	2.270	0.06892	2.251
d-3	0.001516	1.102	1.113	0.06065	1.981	0.06009	1.963
d-2	0.002018	1.267	1.283	0.06369	2.080	0.06333	2.069
d-1	0.003512	1.968	1.997	0.09558	3.122	0.09533	3.114
d±0	0.004804	2.450	2.493	0.11906	3.889	0.11990	3.917
d+1	0.014661	6.905	7.044	0.29305	9.573	0.29754	9.719
d+2	0.014515	6.378	6.523	0.27546	8.998	0.27989	9.143
d+3	0.014140	5.841	5.991	0.26377	8.616	0.26805	8.756
d+4	0.014450	5.647	5.808	0.25911	8.464	0.26330	8.601
d+5	0.014894	5.535	5.708	0.25958	8.480	0.26335	8.603

(注) 表 1 の脚注を参照せよ。ただしこの表では、 $\tau_1 = d-5$ および $\tau_2 = t$ を想定している。すなわち、累積期間の初期時点「d-5」とし、それ以降の各 t 時点までを累積期間とする CAR 等を計算している。

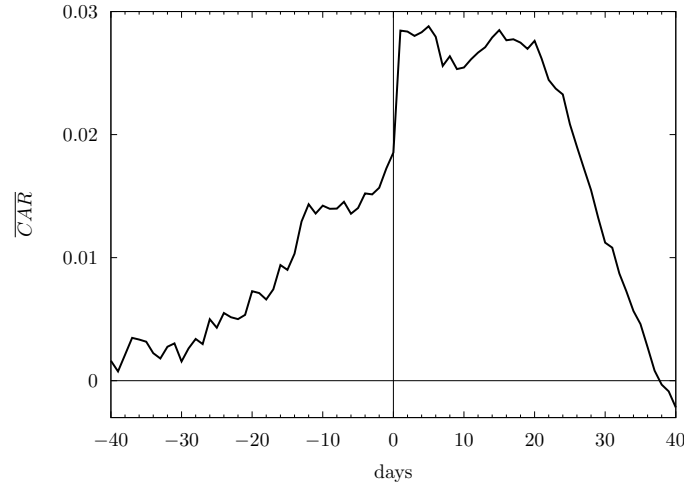
表 3 CAR の計算結果：イベント期間「d-20」から「d+20」(その 3)

t	$\tau_1 = d+5, \tau_2 = t$						
	CAR	Z_1	Z_2	$SCAR_1$	Z_3	$SCAR_2$	Z_4
d+5	0.000444	0.562	0.565	0.04260	1.392	0.04082	1.333
d+6	-0.000425	-0.379	-0.382	0.01371	0.448	0.01271	0.415
d+7	-0.002773	-2.014	-2.035	-0.03634	-1.187	-0.03748	-1.224
d+8	-0.001946	-1.221	-1.237	-0.00869	-0.284	-0.01018	-0.332
d+9	-0.002968	-1.660	-1.687	-0.01808	-0.591	-0.02176	-0.711
d+10	-0.002923	-1.489	-1.517	-0.00814	-0.266	-0.01147	-0.375
d+11	-0.002318	-1.090	-1.114	0.01090	0.356	0.00907	0.296
d+12	-0.001841	-0.808	-0.827	0.02769	0.905	0.02690	0.879
d+13	-0.001461	-0.603	-0.619	0.04131	1.349	0.04018	1.312
d+14	-0.000696	-0.272	-0.280	0.06242	2.039	0.06231	2.036
d+15	-0.000171	-0.064	-0.065	0.07202	2.353	0.07305	2.386
d+16	-0.001038	-0.369	-0.381	0.06486	2.119	0.06575	2.148
d+17	-0.000977	-0.333	-0.345	0.06959	2.273	0.07062	2.307
d+18	-0.001254	-0.411	-0.426	0.06252	2.042	0.06368	2.080
d+19	-0.001801	-0.569	-0.591	0.05543	1.811	0.05665	1.850
d+20	-0.001189	-0.363	-0.378	0.06095	1.991	0.06254	2.043

(注) 表 1 の脚注を参照せよ。ただしこの表では、 $\tau_1 = d+5$ および $\tau_2 = t$ を想定している。すなわち、累積期間の初期時点「d+5」とし、それ以降の各 t 時点までを累積期間とする CAR 等を計算している。

と、 CAR の値はすべての時点で有意ではない。この表 3 では、イベントスタディの手法上の興味深い点も出現する。 CAR が負値であるにもかかわらず、それを標準化した $SCAR_1$ や $SCAR_2$ は正の値となる時点が多々ある。 Z_3 や Z_4 で、後半の時点では正で有意な時点が続いて、 CAR とは異なる結果を示す。これが何を意味するのか。変数を定義する数式を検討すれば明らかであるが、 $\overline{CAR}(\tau_1, \tau_2)$ が負値であっても $\overline{SCAR}(\tau_1, \tau_2)$ が必ず負値であるとは限らず、正值になったとしても構わない。 CAR の標準偏差は 1 より相当に小さいので、標準偏差で除して標準化した値は元の CAR より (絶対値で見て) 大きな値となる。このクロスセクション平均が $\overline{SCAR}(\tau_1, \tau_2)$ である。他方、個々の $CAR_i(\tau_1, \tau_2)$ からクロスセクションの標本平均を計算したのが $\overline{CAR}(\tau_1, \tau_2)$ で

図1 $\overline{CAR}$ の推移： $\tau_1 = d-40$ の場合



(注) 縦軸の $\overline{CAR}$ は $\overline{CAR}(\tau_1, \tau_2)$ の値で、 $\tau_1 = d-40$ および $\tau_2 = t$ を累積期間としている。横軸 days は t 時点で、0 がイベント当日「 $d \pm 0$ 」を表している。サンプルサイズは $L = 200$ 、 $N = 1078$ である。

ある。従って、 $\overline{CAR}(\tau_1, \tau_2)$ が負値であるときに、たまたま正值の $CAR_i(\tau_1, \tau_2)$ で標準偏差が小さく、負値の $CAR_i(\tau_1, \tau_2)$ で標準偏差が大きい傾向にあるなら、標準偏差の逆数でウェイト付けされたクロスセクション平均 $\overline{SCAR}(\tau_1, \tau_2)$ は、正負逆転して正值になる可能性は十分考え得る。

確かに数式上はそうなのであるが、実際問題として、正負共存というのは統計量の解釈が厄介になりかねない。表1と表2と併せて考察すると、 Z_3 と Z_4 は Z_1 と Z_2 よりも(絶対値で見ても)全般的に大きな値になっている。つまり「有意差あり」との結論になりやすいといえるかもしれない。他方で、(1)式と(2)式という2種類の分散であるが、 Z_1 対 Z_2 、あるいは Z_3 対 Z_4 という統計量の数値上の差異は比較的小さい。この程度の差であれば、分散の計算式には簡易な(2)式を用いれば十分という主張の根拠になろう。

最後に図1は、CARをプロットしたものである。ただし今までの計算では、イベント期間を「 $d-20$ 」から「 $d+20$ 」としていたが、この図のイベント期間はまだ少し長くてイベント当日の前後40日間ずつ、 $K = 81$ の場合である点が異なっている。イベント当日の40日前から見ても、CARは多少ジグザグしながらもイベント前の期間全体で上昇基調にある。イベント発出直後の「 $d+1$ 」時点でCARは跳ね上がり、その後しばらくの間、「 $d+20$ 」ぐらいまでは一進一退しながら同水準のCARを維持しているが、「 $d+20$ 」以降から「 $d+40$ 」ほどまでにCARは一直線に下がっている。「 $d+40$ 」の直前にCARは正值から負値になる。つまり、情報発出の株価に与える効果はほとんど消滅するといえる。「サプライズ銘柄」という買い推奨の情報効果は、発出時にのみ明確に出現するが、約2か月ほどの時間を経てその効果はほぼ完全に消滅してしまう。この図1に記されたCARの動向がこの点を端的に表現しているように思う。

4 結びに代えて

本稿ではイベントスタディの手法に依拠して、いわゆる「買い推奨銘柄」の情報がもたらす株価動向への影響について検討を加えた。通常のイベントスタディと共通する特徴もあれば、そうでない特徴も観察された。共通する特徴としては、イベントの与える株価への情報効果は確かに存在する。イベント発出直後の AR は最大の大きさに、その有意性は非常に高い。しかし発出直後を除き、イベント後の期間で正で有意な AR は他に存在しない。「買い推奨銘柄」という情報効果は確かに存在するが、それは情報の発出直後だけのものに限定される。他方で普通とは異なる特徴としては、モーメンタム効果がイベントスタディで捉えられる点である。イベント前の期間で AR が正で有意な時点がいくつか存在し、そのためにイベント発出前から CAR が正の高い有意性を持つ。イベント前の期間を通じて、CAR はイベント当日に向け上昇基調をたどる。そして、イベント当日直後に大きく跳ね上がり、イベント後は一進一退の動きとなる。イベント前に CAR が上昇基調となる理由は、筆者の推測の域を出ない話であるが、「買い推奨銘柄」という情報がイベント前に漏出したことが原因とは考え難く、そもそものところ、直近の株価が上昇基調にあった銘柄が「買い推奨銘柄」に選定されやすいといったことに原因があるのではないかと推測される。ただモーメンタムとはいっても、それほど長い期間にわたって効果が継続するわけではない。イベント期間をもう少し長く取ってみると、イベント後 20 時点過ぎあたりから CAR は一直線に低下し始め、40 時点経過ちょっと前に CAR はゼロを乗り越え負値に転じる。これは「買い推奨」という情報が株価に与えた正の影響が、そのときまでに一掃されてしまうことを意味する。これが本稿の調べた「買い推奨銘柄」という情報の株価に与える効果である。

付録 A イベントスタディにおける検定方法の導出

ここではイベントスタディの計量経済学的な解析を与える。特に仮説検定の根拠となる統計量の確率分布をできる限り詳しく導出したい。イベントスタディでは、異常収益率 AR と称される回帰モデルの予測誤差が株価動向を測る物差しとなっている。OLS(最小 2 乗法) 推定量から計算される次のような回帰モデルの予測誤差が AR である。

$$AR_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

t 時点における i 銘柄の株式収益率を $R_{i,t}$ で、マーケットポートフォリオ収益率を $R_{m,t}$ で表す。 $\hat{\alpha}_i$ と $\hat{\beta}_i$ は、回帰係数の α_i と β_i に関する OLS 推定量である。

株式収益率 $R_{i,t}$ がファイナンス論のマーケットモデルに従うことを想定する。

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t}$$

マーケットモデルとは、株式収益率 $R_{i,t}$ を被説明変数、マーケットポートフォリオ収益率 $R_{m,t}$ を説明変数、 $\epsilon_{i,t}$ を攪乱項とする単純回帰を想定したモデルである。 $E(\epsilon_{i,t}) = 0$ 、 $E(\epsilon_{i,t}^2) = \sigma_{\epsilon_i}^2$ 、 $E(R_{m,t} \cdot \epsilon_{i,t}) = 0$ が仮定される。以下ではこのモデルを推定するところから話を始めたい。

A.1 パラメタ推定

上記のマーケットモデルを推定するためのデータとして、開始時点を t_0 、終了時点を t_1 とする標本があるとしよう。データのサンプルサイズが L であり、 $L = t_1 - t_0 + 1$ である。最小 2 乗法 OLS を使って、パラメーター α_i と β_i を推定するが、OLS 推定量の行列表記を利用したいので、その作法に従って、銘柄 i の時系列データとして t について数値を並べた次のような行列を定義する。

$$\mathbf{R}_i = \begin{pmatrix} R_{i,t_0} \\ \vdots \\ R_{i,t_1} \end{pmatrix}, \quad \mathbf{X}_i = (\mathbf{1}_L \quad \mathbf{R}_m), \quad \mathbf{R}_m = \begin{pmatrix} R_{m,t_0} \\ \vdots \\ R_{m,t_1} \end{pmatrix}, \quad \mathbf{1}_L = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \boldsymbol{\varepsilon}_i = \begin{pmatrix} \epsilon_{i,t_0} \\ \vdots \\ \epsilon_{i,t_1} \end{pmatrix}$$

ただし $\mathbf{R}_i$, $\mathbf{R}_m$, $\mathbf{1}_L$, $\boldsymbol{\varepsilon}_i$ は $L \times 1$ の列ベクトルで、 $\mathbf{X}_i$ は $L \times 2$ の行列である。回帰係数を 2×1 のベクトル $\boldsymbol{\theta}_i = (\alpha_i \ \beta_i)'$ で表すと (' は転置を表す)、先のマーケットモデルは

$$\mathbf{R}_i = \mathbf{X}_i \boldsymbol{\theta}_i + \boldsymbol{\varepsilon}_i \quad (3)$$

として書き換えることができる。ここの攪乱項 $\boldsymbol{\varepsilon}_i$ には、計量経済学の一般的な仮定として、

$$\boldsymbol{\varepsilon}_i | \mathbf{X}_i \sim N(\mathbf{0}, \sigma_{\epsilon_i}^2 \mathbf{I}_L)$$

を仮定する。つまり、 $\mathbf{X}_i$ を条件とする $\boldsymbol{\varepsilon}_i$ は正規分布に従い、平均で $E(\boldsymbol{\varepsilon}_i | \mathbf{X}_i) = \mathbf{0}$ 、分散では $E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' | \mathbf{X}_i) = \sigma_{\epsilon_i}^2 \mathbf{I}_L$ という conditional homoskedasticity が仮定される。ところで $\boldsymbol{\theta}_i$ の OLS 推定量 $\hat{\boldsymbol{\theta}}_i = (\hat{\alpha}_i \ \hat{\beta}_i)'$ は、

$$\hat{\boldsymbol{\theta}}_i = (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' \mathbf{R}_i \quad (4)$$

であり、(4) 式に (3) 式を代入すると、

$$\hat{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i + (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' \boldsymbol{\varepsilon}_i \quad (5)$$

という関係が導出される。これを使うと $\mathbf{X}_i$ を条件とする $\hat{\boldsymbol{\theta}}_i$ の分散共分散行列は、

$$\text{Var}(\hat{\boldsymbol{\theta}}_i | \mathbf{X}_i) = E\left((\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i)(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i)' | \mathbf{X}_i\right) = \sigma_{\epsilon_i}^2 (\mathbf{X}_i' \mathbf{X}_i)^{-1}$$

として定式化できるが、ここの σ_{ϵ_i} は未知であるから、この値を推定する必要がある。 σ_{ϵ_i} を推定するには、回帰の残差

$$\hat{\epsilon}_i = \mathbf{R}_i - \mathbf{X}_i \hat{\boldsymbol{\theta}}_i$$

から計算される回帰の標準誤差 $\hat{\sigma}_{\epsilon_i}$ を用いる。なお $\hat{\sigma}_{\epsilon_i}$ は次のように計算する。

$$\hat{\sigma}_{\epsilon_i}^2 = \frac{\hat{\boldsymbol{\varepsilon}}_i' \hat{\boldsymbol{\varepsilon}}_i}{L - 2} \quad (6)$$

この $\hat{\sigma}_{\epsilon_i}^2$ でもって $\sigma_{\epsilon_i}^2$ を置き換えた

$$\widehat{\text{Var}}(\hat{\boldsymbol{\theta}}_i | \mathbf{X}_i) = \hat{\sigma}_{\epsilon_i}^2 (\mathbf{X}_i' \mathbf{X}_i)^{-1}$$

という行列の対角要素が $\hat{\theta}_i$ の標準誤差の 2 乗値である。ここまでの議論は、計量経済学のあらゆる中級テキストで紹介される、回帰分析に関するごくごく一般的な記述である。イベントスタディならではの独特の議論は次の小節から展開される。^{*5}

A.2 AR の確率分布と検定方法

このようにして、 $t = t_0, \dots, t_1$ 時点のデータであるサンプルサイズ L の標本から、パラメタの推定値を得た後、これを用いて $t = t_1$ 以降の時点における予想をしたい。 $t_2(> t_1)$ 時点から $t_3(> t_2)$ 時点までについて、そのときの $R_{m,t}$ の値からパラメタ推定量 $\hat{\theta}_i = (\hat{\alpha}_i \ \hat{\beta}_i)'$ を用いて、 $R_{i,t}$ の値を予想するのである。この OLS 推定量に依拠した予測値として $\hat{\alpha}_i + \hat{\beta}_i R_{m,t}$ を計算する。ただし $t = t_2, \dots, t_3$ で、 $K = t_3 - t_2 + 1$ である。

この予測を取り扱うのに次のような行列を定義する。 t_2 時点から t_3 時点までのデータを

$$\mathbf{R}_i^* = \begin{pmatrix} R_{i,t_2} \\ \vdots \\ R_{i,t_3} \end{pmatrix}, \quad \mathbf{R}_m^* = \begin{pmatrix} R_{m,t_2} \\ \vdots \\ R_{m,t_3} \end{pmatrix}, \quad \mathbf{1}_K = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{X}_i^* = (\mathbf{1}_K \ \mathbf{R}_m^*)$$

の形のベクトルおよび行列で表現する。この $\mathbf{R}_i^*$, $\mathbf{R}_m^*$, $\mathbf{1}_K$ は $K \times 1$ の列ベクトルで、 $\mathbf{X}_i^*$ は $K \times 2$ の行列である。これらの表記を用いると、OLS 推定量を用いた予測値 $\hat{\alpha}_i + \hat{\beta}_i R_{m,t}$ は、 $\mathbf{X}_i^* \hat{\theta}_i$ と書ける。 $R_{i,t}$ の実現値ベクトルが $\mathbf{R}_i^*$ であるから、予測誤差を $\hat{\epsilon}_i^*$ で表すと、

$$\hat{\epsilon}_i^* = \mathbf{R}_i^* - \mathbf{X}_i^* \hat{\theta}_i$$

である。このベクトル $\hat{\epsilon}_i^*$ を構成する要素で、 t 時点に相当するものが $AR_{i,t}$ に他ならない ($t = t_2, \dots, t_3$)。

他方で、パラメタの真の値 θ_i と関係する攪乱項に相当する項 ϵ_i^* は

$$\epsilon_i^* = \mathbf{R}_i^* - \mathbf{X}_i^* \theta_i$$

であるから、これらの記号を使うと次の式を得る。

$$\hat{\epsilon}_i^* = \epsilon_i^* - \mathbf{X}_i^* (\hat{\theta}_i - \theta_i)$$

この式を用いて、 $\mathbf{X}_i^*$ を条件とする $\hat{\epsilon}_i^*$ の平均および分散共分散を以下のように計算する。

$$\begin{aligned} E(\hat{\epsilon}_i^* | \mathbf{X}_i^*) &= E(\epsilon_i^* | \mathbf{X}_i^*) - E(\mathbf{X}_i^* (\hat{\theta}_i - \theta_i) | \mathbf{X}_i^*) = \mathbf{0} \\ \mathbf{V}_i &= E(\hat{\epsilon}_i^* \hat{\epsilon}_i^{*'} | \mathbf{X}_i^*) = E(\epsilon_i^* \epsilon_i^{*'} | \mathbf{X}_i^*) - 2 E(\mathbf{X}_i^* (\hat{\theta}_i - \theta_i) \epsilon_i^{*'} | \mathbf{X}_i^*) \\ &\quad + E(\mathbf{X}_i^* (\hat{\theta}_i - \theta_i) (\hat{\theta}_i - \theta_i)' \mathbf{X}_i^{*'} | \mathbf{X}_i^*) \end{aligned}$$

^{*5} 計量経済学の一般的議論に関する教科書は非常に多数あるが、例えば代表的なものとして Hayashi(2000) をあげておく。イベントスタディは膨大な数が研究されたが、これに関する計量経済学的にキチンとした議論は実はほとんどなくて、むしろ例外的に MacKinlay(1997) と Campbell-Lo-MacKinlay(1997) 第 4 章の 2 つが存在するだけではないか。本稿の以下の議論は決して、これら 2 つの文献を超えるものではなく、これら 2 つの文献に依拠しながら、これらよりも若干詳しく導出過程を示している。この点が本稿の本当の存在意義なのではなからうか。

これら平均と分散共分散の項を一つ一つ確認していこう。まず平均の方である。最初の項 $E(\boldsymbol{\varepsilon}_i^* | \mathbf{X}_i^*)$ はゼロと考えて構わないであろう。次の項は、(5) 式を使って、さらに $\mathbf{X}_i$ と $\mathbf{X}_i^*$ とが独立であることに依拠して、

$$E(\mathbf{X}_i^*(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i) | \mathbf{X}_i^*) = \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' E(\boldsymbol{\varepsilon}_i | \mathbf{X}_i^*)$$

であるから、これもやはりゼロと見なせる。次に分散共分散の方を検討する。 $\mathbf{V}_i$ の第 1 項は、 $E(\boldsymbol{\varepsilon}_i^* \boldsymbol{\varepsilon}_i^{*'} | \mathbf{X}_i^*) = \sigma_{\varepsilon_i}^2 \mathbf{I}_K$ とする。 $\mathbf{V}_i$ の第 2 項は、やはり (5) 式および $\mathbf{X}_i$ と $\mathbf{X}_i^*$ の独立性を用いて、

$$E(\mathbf{X}_i^*(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i) \boldsymbol{\varepsilon}_i^{*'} | \mathbf{X}_i^*) = \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' | \mathbf{X}_i^*)$$

のように式展開できるが、この式の中で $E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' | \mathbf{X}_i^*)$ はゼロと考えて構わないだろう。これは異なる期間についての攪乱項 $\boldsymbol{\varepsilon}_i$ と $\boldsymbol{\varepsilon}_i^*$ とが無相関と仮定することを意味している。最後に $\mathbf{V}_i$ の第 3 項についても

$$\begin{aligned} \mathbf{X}_i^* E((\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i)(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i)' | \mathbf{X}_i^*) \mathbf{X}_i^{*'} &= \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' | \mathbf{X}_i^*) \mathbf{X}_i(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'} \\ &= \sigma_{\varepsilon_i}^2 \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'} \end{aligned}$$

というように展開できる。途中 $E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i' | \mathbf{X}_i^*) = \sigma_{\varepsilon_i}^2 \mathbf{I}_L$ を用いている。以上のことから、 $\mathbf{V}_i$ は次のような式となる。

$$\mathbf{V}_i = \sigma_{\varepsilon_i}^2 [\mathbf{I}_K + \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'}]$$

さて、この $\mathbf{V}_i$ の対角要素のスカラー表記はどうなるか。上の式第 2 項の行列の中身を具体的に展開して行けばよい。今、

$$(\mathbf{X}_i' \mathbf{X}_i)^{-1} = \frac{1}{(\mathbf{1}_L' \mathbf{1}_L)(\mathbf{R}_m' \mathbf{R}_m) - (\mathbf{1}_L' \mathbf{R}_m)^2} \begin{pmatrix} \mathbf{R}_m' \mathbf{R}_m & -\mathbf{1}_L' \mathbf{R}_m \\ -\mathbf{1}_L' \mathbf{R}_m & \mathbf{1}_L' \mathbf{1}_L \end{pmatrix}$$

は 2×2 の行列であり、 $\mathbf{1}_L' \mathbf{1}_L = L$ 、 $\mathbf{1}_L' \mathbf{R}_m = \sum_{k=t_0}^{t_1} R_{m,k} = L \bar{R}_m$ 、 $\mathbf{R}_m' \mathbf{R}_m = \sum_{k=t_0}^{t_1} R_{m,k}^2 (\equiv C)$ であるから、

$$(\mathbf{1}_L' \mathbf{1}_L)(\mathbf{R}_m' \mathbf{R}_m) - (\mathbf{1}_L' \mathbf{R}_m)^2 = L^2 \left(\frac{1}{L} \sum_{k=t_0}^{t_1} R_{m,k}^2 - \bar{R}_m^2 \right) = L^2 s(R_m)^2$$

というように書き換えることができる。この $s(R_m)^2$ は、推定期間における $R_{m,k}$ (ただし $k = t_0, \dots, t_1$) の分散推定量で、次のように定義されている。

$$s(R_m)^2 = \frac{1}{L} \sum_{k=t_0}^{t_1} (R_{m,k} - \bar{R}_m)^2 = \frac{1}{L} \sum_{k=t_0}^{t_1} R_{m,k}^2 - \bar{R}_m^2 \quad (7)$$

以上の逆行列の結果を使うと、次の式を得る。

$$\begin{aligned} \mathbf{X}_i^*(\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'} &= \frac{1}{L^2 s(R_m)^2} (\mathbf{1}_K \quad \mathbf{R}_m^*) \begin{pmatrix} \mathbf{R}_m' \mathbf{R}_m & -\mathbf{1}_L' \mathbf{R}_m \\ -\mathbf{1}_L' \mathbf{R}_m & \mathbf{1}_L' \mathbf{1}_L \end{pmatrix} \begin{pmatrix} \mathbf{1}_K' \\ \mathbf{R}_m^{*'} \end{pmatrix} \\ &= \frac{C(\mathbf{1}_K \mathbf{1}_K') - L \bar{R}_m(\mathbf{R}_m^* \mathbf{1}_K' + \mathbf{1}_K \mathbf{R}_m^{*'}) + L(\mathbf{R}_m^* \mathbf{R}_m^{*'})}{L^2 s(R_m)^2} \end{aligned} \quad (8)$$

この分子にある $K \times K$ の 4 つの行列は次のような形をしている。

$$\begin{aligned} \mathbf{R}_m^* \mathbf{1}'_K &= \begin{pmatrix} R_{m,t_2} & \cdots & R_{m,t_2} \\ \vdots & \ddots & \vdots \\ R_{m,t_3} & \cdots & R_{m,t_3} \end{pmatrix}, \quad \mathbf{1}_K \mathbf{R}_m^{*'} = \begin{pmatrix} R_{m,t_2} & \cdots & R_{m,t_3} \\ \vdots & \ddots & \vdots \\ R_{m,t_2} & \cdots & R_{m,t_3} \end{pmatrix}, \\ \mathbf{1}_K \mathbf{1}'_K &= \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}, \quad \mathbf{R}_m^* \mathbf{R}_m^{*'} = \begin{pmatrix} R_{m,t_2}^2 & \cdots & R_{m,t_2} R_{m,t_3} \\ \vdots & \ddots & \vdots \\ R_{m,t_3} R_{m,t_2} & \cdots & R_{m,t_3}^2 \end{pmatrix} \end{aligned}$$

これらの行列を用いて、 $\mathbf{X}_i^* (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'} の t 番目の対角要素を取り出すと次のとおりである。$

$$\begin{aligned} [\mathbf{X}_i^* (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'}]_t &= \frac{C - 2L\bar{R}_m R_{m,t} + LR_{m,t}^2}{L^2 s(R_m)^2} \\ &= \frac{s(R_m)^2 + \bar{R}_m^2 - 2\bar{R}_m R_{m,t} + R_{m,t}^2}{L s(R_m)^2} = \frac{1}{L} \left(1 + \frac{(R_{m,t} - \bar{R}_m)^2}{s(R_m)^2} \right) \end{aligned}$$

なおこの 1 行目から 2 行目の展開には、(7) 式の $C = \sum_{k=t_0}^{t_1} R_{m,k}^2 = L[s(R_m)^2 + \bar{R}_m^2]$ という関係を用いている。この計算結果を用いて、 $\mathbf{V}_i$ の任意の t 番目の対角要素をスカラー表記すると、

$$[\mathbf{V}_i]_t = \text{var}(AR_{i,t} | \mathbf{X}_i^*) = \sigma_{\epsilon_i}^2 \left[1 + \frac{1}{L} \left(1 + \frac{(R_{m,t} - \bar{R}_m)^2}{s(R_m)^2} \right) \right]$$

という形を導出することができ、これが第 i イベントの任意の t 時点 (ただし $t = t_2, \dots, t_3$) における $AR_{i,t}$ の分散 $\text{var}(AR_{i,t} | \mathbf{X}_i^*)$ である。

ここで議論を単純化するため、 ϵ_i^* は $\mathbf{X}_i^*$ を条件とする正規分布に従うことを仮定しよう。

$$\epsilon_i^* | \mathbf{X}_i^* \sim N(\mathbf{0}, \mathbf{V}_i)$$

であるから、その t 番目の要素である $AR_{i,t}$ については、

$$\frac{AR_{i,t} | \mathbf{X}_i^*}{\sqrt{[\mathbf{V}_i]_t}} \equiv y_{i(AR)} \sim N(0, 1)$$

が成立している。ところで $\mathbf{V}_i$ の中にある $\sigma_{\epsilon_i}^2$ は未知であるから、その代りに (6) 式で定義される回帰の標準誤差 $\hat{\sigma}_{\epsilon_i}^2$ を用いる。 $\sigma_{\epsilon_i}^2$ と $\hat{\sigma}_{\epsilon_i}^2$ の間には、

$$\frac{(L-2)\hat{\sigma}_{\epsilon_i}^2}{\sigma_{\epsilon_i}^2} \equiv q_i \sim \chi^2(L-2)$$

という関係のあることが分かっている。 q_i という変数が自由度 $L-2$ のカイ 2 乗分布に従うことを $\chi^2(L-2)$ でもって表現している。^{*6} この $\hat{\sigma}_{\epsilon_i}^2$ に依拠して

$$\hat{\mathbf{V}}_i = \hat{\sigma}_{\epsilon_i}^2 [\mathbf{I}_K + \mathbf{X}_i^* (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'}] \quad (9)$$

^{*6} この関係およびカイ 2 乗分布やステューデント t 分布の説明については、どの教科書にも紹介されるごくごく一般的議論であるから、ここでの具体的な証明は省略する。例えば Hayashi(2000) 第 1 章を参照願いたい。

というように $\widehat{\mathbf{V}}_i$ を定義すると、この $\widehat{\mathbf{V}}_i$ は、 $\widehat{\mathbf{V}}_i = \mathbf{V}_i q_i / (L - 2)$ ということでもある。さらに $\widehat{\mathbf{V}}_i$ の t 番目の対角要素を $[\widehat{\mathbf{V}}_i]_t$ と記す。以上のことから、次のように定義される $\widehat{SAR}_{i,t}$ は、

$$\widehat{SAR}_{i,t} = \frac{AR_{i,t} | \mathbf{X}_i^*}{\sqrt{[\widehat{\mathbf{V}}_i]_t}} = \frac{y_{i(AR)}}{\sqrt{q_i / (L - 2)}} \sim t(L - 2) \quad (10)$$

となって、自由度 $L - 2$ のステューデント t 分布に従うことが分かる。実際には L が数百となる大標本を想定していることが多く、その場合、ステューデント t 分布は標準正規分布に収束しているので、その場合は $N(0, 1)$ を用いて仮説検定することになろう。

A.3 CAR の確率分布と検定方法

次に CAR である。 $CAR_i(\tau_1, \tau_2)$ は、イベント期間 (t_2 時間から t_3 時点まで) の中のある期間、 τ_1 時点から τ_2 時点までの $AR_{i,t}$ を足し上げたもので、

$$CAR_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}$$

と定義される。ただし $t_2 \leq \tau_1 < \tau_2 \leq t_3$ である。 $CAR_i(\tau_1, \tau_2)$ の平均と分散を計算するには、次のような $K \times 1$ のベクトル γ を使う。

$$\gamma' = (0 \quad \cdots \quad 0 \quad 1 \quad \cdots \quad 1 \quad 0 \quad \cdots \quad 0)$$

これは K 個の要素で、 $\tau_1 - t_2 + 1$ 番目から $\tau_2 - t_2 + 1$ 番目までの要素は 1 が並び、それ以外の要素はすべてゼロというベクトルである。 $AR_{i,t}$ のベクトル $\hat{\epsilon}_i^*$ とこの γ を用いれば、

$$CAR_i(\tau_1, \tau_2) = \gamma' \hat{\epsilon}_i^*$$

であるから、前で求めた $\mathbf{X}_i^*$ を条件とする $\hat{\epsilon}_i^*$ の平均および分散共分散の計算結果から、 $CAR_i(\tau_1, \tau_2)$ の条件付き平均と分散は、 $E(CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*) = 0$ と $\text{var}(CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*) = \gamma' \mathbf{V}_i \gamma$ と表現できる。なお以下ではこの分散 $\text{var}(CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*)$ を $\sigma_i(\tau_1, \tau_2)^2$ として表すことにする。

この $CAR_i(\tau_1, \tau_2)$ を仮説検定するのに必要な確率分布を、前と同様に導出しよう。

$$\frac{CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*}{\sigma_i(\tau_1, \tau_2)} \equiv y_{i(CAR)} \sim N(0, 1)$$

であるが、 $\sigma_i(\tau_1, \tau_2)$ は $\mathbf{V}_i$ の中にある未知の $\sigma_{\epsilon_i}^2$ に依存しているので、 $\sigma_{\epsilon_i}^2$ を (6) 式の $\hat{\sigma}_{\epsilon_i}^2$ に置き換えた $\widehat{\mathbf{V}}_i$ を用いて、

$$\widehat{\text{var}}(CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*) = \gamma' \widehat{\mathbf{V}}_i \gamma = \hat{\sigma}_i(\tau_1, \tau_2)^2$$

を定義すると、これに依拠して展開される統計量を $\widehat{SCAR}_i(\tau_1, \tau_2)$ で表そう。

$$\widehat{SCAR}_i(\tau_1, \tau_2) = \frac{CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^*}{\hat{\sigma}_i(\tau_1, \tau_2)} = \frac{y_{i(CAR)}}{\sqrt{q_i / (L - 2)}} \sim t(L - 2) \quad (11)$$

このように $\widehat{SCAR}_i(\tau_1, \tau_2)$ は、自由度 $L - 2$ のステューデント t 分布に従うことが導出できる。

それでは $\hat{\sigma}_i(\tau_1, \tau_2)^2$ をどうやって計算したらよいか。これは、0 と 1 の並びである γ というベクトルを用いて、 $\gamma' \hat{\mathbf{V}}_i \gamma$ を実際に計算するしかない。この $\hat{\mathbf{V}}_i$ は (9) 式のように導出されているから、これを用いる。

$$\hat{\sigma}_i(\tau_1, \tau_2)^2 = \gamma' \hat{\mathbf{V}}_i \gamma = \hat{\sigma}_{\epsilon_i}^2 [\gamma' \gamma + \gamma' \mathbf{X}_i^* (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i^{*'} \gamma]$$

この式第 2 項の行列の中身は既に (8) 式で具体的に検討されている。また次の結果を確認するのは容易であろう。ここで $T \equiv \tau_2 - \tau_1 + 1$ と $G \equiv \sum_{k=\tau_1}^{\tau_2} R_{m,k}$ という変数を定義する。

$$\begin{aligned} \gamma' \gamma &= T \\ \gamma' (\mathbf{1}_K \mathbf{1}_K') \gamma &= T^2 \\ \gamma' (\mathbf{R}_m^* \mathbf{1}_K') \gamma &= \gamma' (\mathbf{1}_K \mathbf{R}_m^{*'}) \gamma = TG \\ \gamma' (\mathbf{R}_m^* \mathbf{R}_m^{*'}) \gamma &= G^2 \end{aligned}$$

(8) 式に従ってこれら結果を用いて式を整理すれば、

$$\hat{\sigma}_i(\tau_1, \tau_2)^2 = \hat{\sigma}_{\epsilon_i}^2 \left[T + \frac{(Ts(R_m))^2 + (G - T\bar{R}_m)^2}{Ls(R_m)^2} \right]$$

という結果を得ることができる。この右辺に登場する T と G が τ_1 と τ_2 の関数になっていて、 $\hat{\sigma}_i(\tau_1, \tau_2)$ は τ_1 と τ_2 に依存する。

A.4 クロスセクション集計

以上は単独のイベント、つまり i 番目のイベントに関する $AR_{i,t}$ あるいは $CAR_i(\tau_1, \tau_2)$ の確率分布の導出であった。次にここでは、同種の複数イベントが全部で N 個あって、それぞれに関する $AR_{i,t}$ や $CAR_i(\tau_1, \tau_2)$ (ただし、 $i = 1, \dots, N$) を求め、これらをクロスセクション集計する場合、その標本平均の確率分布を問題にしよう。複数イベントそれぞれについて、推定期間やイベント期間、累積期間がイベント当日からの相対的な時点で決定されていて、その相対時間の長さは複数イベントすべてで共通していると考ええる。

このような共通の設定の下、全イベントについてのクロスセクション平均を取ろう。 $i = 1, \dots, N$ として、各イベントの $AR_{i,t}$ および $CAR_i(\tau_1, \tau_2)$ から、

$$\begin{aligned} \overline{AR}_t &= \frac{1}{N} \sum_{i=1}^N AR_{i,t} \\ \overline{CAR}(\tau_1, \tau_2) &= \frac{1}{N} \sum_{i=1}^N CAR_i(\tau_1, \tau_2) \end{aligned}$$

というように $\overline{AR}_t$ と $\overline{CAR}(\tau_1, \tau_2)$ を計算できる。 $\overline{AR}_t$ は $AR_{i,t}$ の標本平均であり、 $\overline{CAR}(\tau_1, \tau_2)$ は $CAR_i(\tau_1, \tau_2)$ の標本平均である。それではこのように計算される標本平均の確率分布はどうか。

上記のクロスセクション平均について問題になる点は、個別イベントの場合と異なってクロスセクション集計の場合は、標本平均を標本標準偏差で除した分数がステューデント t 分布に従うことを証明できない。あるイベント i について、 $\sigma_{\epsilon_i}^2$ を $\hat{\sigma}_{\epsilon_i}^2$ に置き換える場合、個別イベントの場合のように、カイ 2 乗分布に従う q_i という変数を単純な形で乗じれば済むわけではないからである。^{*7} しかし以下で定義される統計量はいずれも、 N や L が大きい大標本においては、標準正規分布 $N(0, 1)$ に従うと考えることができる。統計学の初歩知識も含めて若干冗長な議論かもしれないが、大事な点でありかつ、導出の議論が意外と存在しない現状ゆえ確認のため端折らずに話を展開する。

N 個の確率変数 x_k (ただし、 $k = 1, \dots, N$) が独立で同一の確率分布に従うとする。独立で同一の確率分布を i.i.d. (independent and identical distribution) と称することが多い。同一の確率分布に従うから、 N 個の確率変数共通にその平均 $E(x_k)$ はゼロ、分散 $\text{var}(x_k)$ は σ^2 という一つの定数とする。今、 N 個の確率変数を使って標本平均 $\bar{x}$ を $\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k$ のように定義する。この $\bar{x}$ の確率分布はどうなるか。平均 $E(\bar{x})$ はゼロ、分散 $\text{var}(\bar{x})$ は σ^2/N となる。なぜなら N 個の確率変数は相互に独立であるから、

$$\text{var}(\bar{x}) = \text{var}\left(\frac{1}{N} \sum_{k=1}^N x_k\right) = \frac{1}{N^2} \sum_{k=1}^N \text{var}(x_k) = \frac{\sigma^2}{N}$$

というように式展開できるからである。さてこの段階ではまだ正規分布は仮定されていないが、ここで N が非常に大きな値であるような場合を考える。各 x_k が i.i.d. であるなら、Lindberg-Levy の中心極限定理が適用でき、これは $\sqrt{N} \cdot \bar{x} \sim N(0, \sigma^2)$ であることを教えてくれる。各 x_k の確率分布の形にかかわらず、 N が非常に大きくなると $\sqrt{N} \cdot \bar{x}$ は正規分布に従うようになる。正確には「分布収束する」というが、その形式的な定義は省略する。であるなら、標準偏差 σ で標準化した確率変数は、正規分布の特徴から、

$$\frac{\sqrt{N} \cdot \bar{x}}{\sigma} \sim N(0, 1)$$

となる。この左辺の分数は $\frac{\bar{x}}{\sigma/\sqrt{N}}$ であり、 N が大きいとこれが標準正規分布 $N(0, 1)$ に従うということである。つまり標本平均 $\bar{x}$ の平均はゼロ、分散は σ^2/N であったから、これを標準化した $\frac{\bar{x}}{\sigma/\sqrt{N}}$ は $N(0, 1)$ という確率分布でもって仮説検定すればよいという根拠になる。

ところで一般には確率分布の分散 σ^2 の値は未知である。そこで次の問題は σ の値をどうするかという点である。普通は未知パラメタ σ を標本統計量から推定するということになるが、ごく限られたケースで分散の値が既知と考えられる場合もある。以下でそれぞれ詳しく見てみる。

A.4.1 分散を既知と見なす場合

まずは分散が既知の場合である。前で $\widehat{SCAR}_i(\tau_1, \tau_2)$ は、自由度 $L - 2$ のステューデント t 分布に従うことを導いた。ステューデント t 分布の平均はゼロであり、その分散は自由度のみに依

^{*7} 例えば、ステューデント t 分布に従う複数個の確率変数を合計した確率変数はステューデント t 分布にはならない。このことを「(分布の) 再現性がない」という。正規分布やカイ 2 乗分布には再現性が備わっている。

存して $\frac{L-2}{L-4}$ という値となることが分かっている (ただし $L > 4$)。従ってすべての i について、 $\widehat{SCAR}_i(\tau_1, \tau_2)$ は共通の平均・分散を持った確率分布に従うと考えることができる。であるなら、 $\widehat{SCAR}_i(\tau_1, \tau_2)$ のすべての i に関する標本平均は

$$\overline{SCAR}(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^N \widehat{SCAR}_i(\tau_1, \tau_2)$$

のように計算され、 $\sqrt{N} \cdot \overline{SCAR}(\tau_1, \tau_2)$ は N が大きいとき、中心極限定理から平均ゼロ、分散 $\frac{L-2}{L-4}$ の正規分布に従う。以上のことから

$$(J \equiv) \frac{\overline{SCAR}(\tau_1, \tau_2)}{\sqrt{\frac{L-2}{N(L-4)}}} \sim N(0, 1)$$

が成立する。ただし本文で見たように、この左辺の形をした比率を J と表す。 J から作られる統計量は今日、イベントスタディのクロスセクション集計で最も一般的に用いられる統計量であろう。^{*8}

A.4.2 分散を推定する場合

次に、未知の分散を標本統計量から推定する場合についてまとめる。

まずいうまでもないことであろうが、確率変数 x_k と定数 v_k から成る比率 $\frac{x_k}{v_k}$ が標準正規分布 $N(0, 1)$ に従うなら、正規分布の特徴として $x_k \sim N(0, v_k^2)$ である。ここで v_k の値が k により異なる確率変数 x_k が N 個あるとする (ただし $k = 1, \dots, N$)。ということは、 N 個の確率変数 x_k の確率分布は同一ではなく、その分散 v_k^2 は異なる値となる。このとき、 $\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k$ で定義される $\bar{x}$ の確率分布は $N(0, \text{var}(\bar{x}))$ であり、この $\text{var}(\bar{x})$ は、 x_k が独立であるなら、

$$\text{var}(\bar{x}) = \frac{1}{N^2} \sum_{k=1}^N v_k^2$$

として定式化できる。当然のこと、 $\frac{\bar{x}}{\sqrt{\text{var}(\bar{x})}}$ は標準正規分布 $N(0, 1)$ に従う。

注意すべきは、この段階で v_k は任意の定数という点である。そこで次に、 $\text{var}(\bar{x})$ の値を具体的に計算するのに、この N 個の定数 v_k に各 x_k の標本標準偏差 $\hat{v}_k$ の値を当てるとする ($k = 1, \dots, N$)。^{*9} この点を強調して $\text{var}(\bar{x})$ は $\widehat{\text{var}}(\bar{x})$ のように表記される。その定義式は、前の式に $v_k = \hat{v}_k$ を代入して、

$$\widehat{\text{var}}(\bar{x}) = \frac{1}{N^2} \sum_{k=1}^N \hat{v}_k^2$$

^{*8} この本文では、累積異常収益率 $CAR_i(\tau_1, \tau_2)$ に関するクロスセクション集計の統計量を説明したが、実はまったく同様の議論を累積を取らない方の $AR_{i,t}$ でも展開することができる。 $AR_{i,t}$ をその標準偏差で標準化した値 (前の $\widehat{SAR}_{i,t}$) も同様に自由度 $L-2$ のステューデント t 分布に従うからである。が、どういうわけか、これをすべての i でクロスセクション標本平均した統計量が、やはり平均ゼロ、分散 $\frac{L-2}{L-4}$ の正規分布に従うという言説を、筆者は寡聞にして見たことがない。ここでは学界一般の言説に従って累積異常収益率の方の確率分布を取り上げた。

^{*9} これもいうまでもないことかもしれないが、標本標準偏差は次のように求められる。例えば確率変数 x_k の実現値が M 個あるとする。これら数値を 1 つの標本と見なし、 $\{x_{k,1}, x_{k,2}, \dots, x_{k,M}\}$ というように表現するとしよう。1

であるが、この式を使って N 個の $\hat{v}_k$ の値から $\widehat{\text{var}}(\bar{x})$ を得ることができる。 $\widehat{\text{var}}(\bar{x})$ はあくまでも $\text{var}(\bar{x})$ の 1 つの実現値 (具体的な値) であるから、 $\frac{\bar{x}}{\sqrt{\widehat{\text{var}}(\bar{x})}}$ はやはり標準正規分布 $N(0, 1)$ に従うとしてよい。

さて以上が統計学一般の議論であるが、これをこの議論に形式的に当てはめよう。個別イベント i について、 $AR_{i,t}$ は (10) 式で見たように、また $CAR_i(\tau_1, \tau_2)$ は (11) 式で見たように、それぞれの標準偏差で標準化した比率は自由度 $L - 2$ のステューデント t 分布に従う。ステューデント t 分布は自由度が大きくなると、標準正規分布 $N(0, 1)$ と同じになるという特徴がある。^{*10} 従って、 L が大きな値の大標本において、(10) 式の比率や (11) 式の比率は $N(0, 1)$ に従うと考えてよいので、これら比率を変形すれば、すべての i について

$$AR_{i,t} | \mathbf{X}_i^* \sim N\left(0, \left[\widehat{\mathbf{V}}_i\right]_t\right)$$

$$CAR_i(\tau_1, \tau_2) | \mathbf{X}_i^* \sim N\left(0, \hat{\sigma}_i(\tau_1, \tau_2)^2\right)$$

と書くこともできる。これらの i に関する標本平均が $\overline{AR}_t$ と $\overline{CAR}(\tau_1, \tau_2)$ であり、個々の $AR_{i,t}$ がイベント相互に (i について) 独立であるなら、 $\overline{AR}_t$ と $\overline{CAR}(\tau_1, \tau_2)$ の分散は次のように推定値を計算すればよい。

$$\widehat{\text{var}}(\overline{AR}_t) = \frac{1}{N^2} \sum_{i=1}^N \left[\widehat{\mathbf{V}}_i\right]_t$$

$$\widehat{\text{var}}(\overline{CAR}(\tau_1, \tau_2)) = \frac{1}{N^2} \sum_{i=1}^N \hat{\sigma}_i(\tau_1, \tau_2)^2$$

この分散の推定値に平方根を取った値が標本標準偏差であり、これで標本平均を除した比率は標準正規分布 $N(0, 1)$ に従う。

$$\frac{\overline{AR}_t}{\sqrt{\widehat{\text{var}}(\overline{AR}_t)}} \sim N(0, 1)$$

$$\frac{\overline{CAR}(\tau_1, \tau_2)}{\sqrt{\widehat{\text{var}}(\overline{CAR}(\tau_1, \tau_2))}} \sim N(0, 1)$$

A.5 まとめ

以上導出した統計量がどういう条件下で標準正規分布 $N(0, 1)$ に従うのか、最後にまとめておこう。まず簡単明瞭なのは十分条件で、それは N と L の両方が大きいことである。このとき、すべ

つの実現値 $x_{k,t}$ を確率変数 x_k のデータと考える。すると、 x_k の標本標準偏差は次のとおり。

$$\hat{v}_k = \sqrt{\frac{1}{M-1} \sum_{t=1}^M (x_{k,t} - \bar{x}_k)^2} \quad \text{ただし、} \bar{x}_k = \frac{1}{M} \sum_{t=1}^M x_{k,t}$$

この式の名分母を $M - 1$ としているが、 $M - 1$ ではなく、代わりに M で除したとしても実用上の差異はあまりない。

^{*10} 自由度 $L - 2$ のステューデント t 分布の分散 $\frac{L-2}{L-4}$ は、 $L \rightarrow \infty$ で 1 に収束する。

ての Z 値は $N(0, 1)$ に従う。なお N はイベント数を、 L はパラメタ推定のサンプルサイズを表している。議論が若干錯綜するのは必要条件の方で、本稿では分散を既知と見なす場合と未知の場合の 2 ケースを取り上げた。

分散既知と考える場合、 $\widehat{SCAR}_i(\tau_1, \tau_2)$ が自由度 $L - 2$ のステューデント t 分布に従うことで、その分散の値を決めるのであるが、ステューデント t 分布に従うには、前で述べたとおり、その元の $y_{i(CAR)}$ が $N(0, 1)$ に従うことを仮定する必要がある。個別イベント i の統計量であれば、必ずしも L は大きな値である必要はない。ただクロスセクション統計量 J の値を計算する際は、 N はやはり大きな値であることが必要であろう。というのは、個々の確率分布の形にかかわらず、中心極限定理により標本平均が正規分布に従うとするしかないからである。従って J という比率に依存する Z 値の場合、 L の値は必要条件にならないが、 N が大きな値であることはやはり必要条件のように思う。

他方、分散は未知でその標本統計量を当てる場合、 L の値が大きいならば、個別イベント $\widehat{SCAR}_i(\tau_1, \tau_2)$ はすべての i で $N(0, 1)$ に従うので、クロスセクション標本平均も自動的に正規分布に従う。従ってこのときの必要条件は L の値が大きいことであり、 N の値は必要条件とはならない。ただし正規分布に近似させるステューデント t 分布の導出に、元々正規分布の仮定が必要なのは注意しておくべきであろう。

参考文献

- [1] Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997. *The Econometrics of Financial Markets*, Princeton University Press. 祝迫・大橋・中村・本多・和田訳『ファイナンスのための計量分析』2003, 共立出版.
- [2] Eckbo, B.Espen, 1983. "Horizontal Mergers, Collusion, and Stockholder Wealth," *Journal of Financial Economics*, Vol.11, No.1-4 (April, 1983), pp.241-273.
- [3] Hayashi, Fumio, 2000. *Econometrics*, Princeton University Press.
- [4] MacKinlay, A. Craig, 1997. "Event Studies in Economics and Finance," *Journal of Economic Literature*, Vol.35, No.1 (March, 1997), pp.13-39.
- [5] Mikkelsen, Wayne H., and M.Megan Partch, 1986. "Valuation Effects of Security Offerings and the Issuance Process," *Journal of Financial Economics*, Vol.15, No.1-2 (January-February, 1986), pp.31-60.
- [6] Schipper, Katherine, and Abbie Smith, 1986. "A Comparison of Equity Carve-outs and Seasoned Equity Offerings: Share Price Effects and Corporate Restructuring," *Journal of Financial Economics*, Vol.15, No.1-2 (January-February, 1986), pp.153-186.