

Die Simulation langfristiger Überrenditen

Olaf Ehrhardt, Ralf Koerstein

Postadresse: Humboldt-Universität zu Berlin
Institut für Bank-, Börsen- und Versicherungswesen
Spandauer Straße 1
D-10178 Berlin

Telefon: (030)-2093-5620
Fax: (030)-2093-5666
Email: ehrhardt@wiwi.hu-berlin.de

Die Simulation langfristiger Überrenditen

1. Problemstellung

Untersuchungen der kurz- und langfristigen Aktienkursreaktion auf die Ankündigung von Ereignissen oder auf die Bekanntgabe unternehmerischer Entscheidungen haben erheblich zum Verständnis der Funktionsweise von Kapitalmärkten beigetragen. In informationseffizienten Märkten werden neue Informationen vollständig und ohne zeitliche Verzögerungen von den Aktienkursen widerspiegelt. Die beobachteten Kursreaktionen entsprechen den veränderten Erwartungen der Marktteilnehmer über die zukünftigen Zahlungsströme. Liegen zu einem bestimmten Ereignis, zum Beispiel der Ankündigung einer Fusion, eines Aktienrückkaufes oder einer Kapitalerhöhung, eine Vielzahl von Beobachtungen vor, dann kann aus dem Vorzeichen der durchschnittlichen, um sonstige Einflüsse bereinigten Kursveränderung darauf geschlossen werden, ob die Managemententscheidung im Interesse der Anteilseigner liegt.¹ Wird die angekündigte Maßnahme im Unternehmen umgesetzt, so sollten in einem effizienten Markt im Durchschnitt über alle in die Untersuchung einbezogenen Unternehmen keine weiteren systematischen Kursveränderungen beobachtet werden.²

Seit der vielbeachteten Veröffentlichung von Fama/Fisher/Jensen/Roll (1969) zu Stock splits wurde in den siebziger Jahren eine Reihe von Untersuchungen zu den kurzfristigen Kursanpassungen bei der Durchführung von Finanzierungsmaßnahmen vorgelegt. In den achtziger Jahren standen Untersuchungen zu Ankündigungseffekten im Blickfeld des wissenschaftlichen Interesses, die sich insbesondere auf Veränderungen der Kapitalstruktur, des Ausschüttungsverhaltens oder der Kontrolle von Unternehmen bezogen. Die beobachteten sofortigen Kursreaktionen sind nach Einschätzung von Fama (1991) der klarste empirische Befund für die Existenz informationseffizienter Kapitalmärkte. Der Vorteil von derartigen Ereignisstudien liegt in erster Linie darin, daß aufgrund des kurzen Betrachtungszeitraumes die Schätzprobleme bei der Ermittlung der erwarteten Aktienrendite nur eine geringe Rolle spielen.³ Bei der Untersuchung von Ankündigungseffekten ist jedoch nicht auszuschließen, daß die beobachteten Kursveränderungen durch Über- und Unterreaktionen der Marktteilnehmer beeinflußt sind. Aus diesen Überlegungen heraus entstanden zu Beginn der neunziger Jahre Untersuchungen zur langfristigen Renditeentwicklung in Zusammenhang mit der Durchführung unternehmensspezifischer Maßnahmen.

Empirische Untersuchungen führen zum Ergebnis, daß die Aktienrenditen für Untersuchungszeiträume von drei bis fünf Jahren nach Börseneinführungen (Ritter, 1991), Kapitalerhö-

¹ Vgl. Shleifer/Vishny (1997), S. 746.

² Vgl. Stehle/Ehrhardt/Przyborowsky (1999), S. 2.

³ Vgl. Fama (1991), S. 1607.

hungen (Loughran/Ritter, 1995), Aktienrückkäufen (Ikenberry/Lakonishok/Vermaelen, 1995), Dividendenveränderungen (Michaely/Thaler/Womack, 1995), Fusionen (Mitchell/Stafford, 1997) und anderen unternehmerischen Entscheidungen systematisch von ihren geschätzten Erwartungswerten abweichen. Diese Ergebnisse werden häufig als Über- bzw. Unterreaktionen auf die Bekanntgabe bewertungsrelevanter Informationen interpretiert. Die beobachteten Abweichungen können nach Ansicht von Mitchell/Stafford (1997) zu falschen Schlußfolgerungen führen, wenn allein auf der Grundlage der beobachteten Ankündigungseffekte Implikationen abgeleitet werden.⁴

Die festgestellten positiven oder negativen Überrenditen deuten nach Mitchell/Stafford (1997) darauf hin, daß eine hohe Wahrscheinlichkeit besteht, daß Anleger die Veränderungen der zukünftigen Einzahlungsüberschüsse und Diskontierungsfaktoren falsch einschätzen.⁵ Fama (1998) argumentiert hingegen, daß nach den bisher vorliegenden Untersuchungsergebnissen positive und negative Überrenditen mit gleicher Häufigkeit zu beobachten sind.⁶ Zufällige, betragswertmäßig kleine Abweichungen würden der Hypothese informationseffizienter Kapitalmärkte nicht widersprechen.⁷ Nach Auffassung von Fama (1998) sind die festgestellten langfristigen Überrenditen vor allem auf die verwendeten Berechnungsmethodiken zurückzuführen.⁸ Fama (1998) stützt seine Argumentation einerseits auf die Schwierigkeiten bei der Schätzung der Erwartungswerte von Renditen einzelner Aktien (*Bad model problem*), andererseits auf zusätzliche methodische Probleme bei der Durchführung langfristiger Ereignisstudien.

Die Schwierigkeiten bei der Berechnung langfristiger Überrenditen zeigen sich sehr deutlich in den Ergebnissen bisher vorgelegter Studien. So werden beispielsweise erhebliche Abweichungen in den berechneten Durchschnittswerten gefunden, je nachdem welche Schätzwerte für die langfristige Rendite der Untersuchungswertpapiere genommen werden.⁹ Präzisere Schätzungen langfristiger Überrenditen erfordern Kenntnisse über jene Bestimmungsfaktoren, die die Renditeunterschiede zwischen Aktien erklären. In der gegenwärtigen wissenschaftlichen Diskussion wird intensiv über den Erklärungsbeitrag des aus dem CAPM abgeleiteten Beta-Faktors, der Höhe der Marktkapitalisierung und des Quotienten aus dem Buch- und dem Marktwert einer Aktie diskutiert.¹⁰ Nach Brav/Gompers (1997) und Brav/Geczy/Gompers (1999) spielen die beiden letztgenannten Faktoren im Zusammenhang mit langfristigen Ereignis-

⁴ Vgl. Mitchell/Stafford (1997), S. 1.

⁵ Vgl. Mitchell/Stafford (1997), S. 2.

⁶ Vgl. Fama (1998), S. 284 f.

⁷ Vgl. Fama (1998), S. 287.

⁸ Vgl. Fama (1998), S. 303.

⁹ Auf diese Problematik haben schon Brown/Warner (1980) und Dimson/Marsh (1986) in ihren Simulationsstudien hingewiesen.

¹⁰ Vgl. hierzu den Überblicksaufsatz von Stehle/Sattler/Wulff (1997).

nisstudien eine wichtige Rolle. Ihre Untersuchungen für den US-amerikanischen Markt ergeben, daß die langfristige Underperformance nach Börseneinführungen und Kapitalerhöhungen auf die zu starke Gewichtung von Unternehmen mit einer geringen Marktkapitalisierung und einem niedrigen Buchwert/Marktwert-Verhältnis zurückzuführen ist.

Mit den methodischen Problemen langfristiger Ereignisstudien setzen sich Lyon/Barber/Tsai (1999), Kothari/Warner (1997), Cowan/Sergeant (1997) und Barber/Lyon (1997) auseinander. Mit Monte-Carlo-Methoden simulieren sie langfristige Ereignisstudien, insbesondere die Ablehnungshäufigkeiten der Nullhypothese. Dabei findet man bei den häufig in der Literatur verwendeten Verfahren zur Berechnung langfristiger Überrenditen stark verzerrte Teststatistiken. Diese Verzerrungen werden in erster Linie durch eine schiefe Überrenditenverteilung verursacht, die entsteht, wenn Portefeullerrenditen als Schätzwert für die Rendite einzelner Untersuchungswertpapiere genommen werden (*Skewness bias*). Bei der Verwendung von Portefeullerrenditen ist außerdem zu beachten, daß in Abhängigkeit der Bildungsvorschriften der Portefeulles Umschichtungen innerhalb der Untersuchungsperiode erforderlich werden können. Dies führt in der Regel dazu, daß Unterschiede in der Berechnung der Renditen der Untersuchungswertpapiere und der zur Schätzung des Erwartungswertes verwendeten Portefeullerrenditen bestehen (*Rebalancing bias*). Die Untersuchung der Renditeentwicklung über einen Zeitraum von mehreren Jahren führt in vielen Fällen zu einer zeitlichen Überschneidung der einzelnen Ereignisfenster, die mit weiteren, im Abschnitt 2.1 näher beschriebenen statistischen Problemen verbunden sind (*Overlapping horizons bias*).

Trotz einer intensiven wissenschaftlichen Diskussion der Ergebnisse langfristiger Ereignisstudien konnte keine Einigung über die Ursachen der beobachteten positiven oder negativen Überrenditen erzielt werden. Die offenen Probleme könnten Ausgangspunkt für die Durchführung langfristiger Ereignisstudien für andere nationale Kapitalmärkte sein. Interessant ist in diesem Zusammenhang, ob die beobachteten langfristigen Überrenditen auch in Märkten mit anderen institutionellen Gegebenheiten, bei firmenspezifischen Unterschieden in den untersuchten Stichproben¹¹ und unter Verwendung von Daten festgestellt werden können, die nicht der CRSP-Datenbank entnommen werden.¹²

Bei solchen Untersuchungen dürften sich die angesprochenen methodischen Probleme häufig noch stärker auf die Untersuchungsergebnisse auswirken, da oft noch eine geringere Anzahl von Beobachtungen als in den USA zur Verfügung steht (vgl. Abschnitt 3.3.1). Diese eventuell nur wenigen Beobachtungen können zudem sehr unterschiedlich über den Betrachtungs-

¹¹ So verweist z.B. Ljungqvist (1997), S. 1313, darauf, daß die Unternehmen, die in Deutschland ihre Aktien im Zeitraum 1970 bis 1993 an der Börse einführten, im Durchschnitt größer und älter waren als US-amerikanische Unternehmen.

¹² Bei empirischen Untersuchungen spielt die Datenqualität eine wichtige Rolle. Es ist in einigen Fällen nicht auszuschließen, daß die Berechnungsergebnisse empirischer Untersuchungen durch einen „*Ex post selection bias*“ als Folge unvollständiger Datenbanken (Kothari/Shanken/Sloan, 1995), fehlerhafte Daten (Shumway, 1997) und Verzerrungen der verwendeten Indizes (Canina/Michaely/Thaler/Womack, 1998) auf systematische Weise beeinflusst sein können.

zeitraum verteilt sein.¹³ Die Untersuchung kurzer Zeiträume oder eine Konzentration der untersuchten Ereignisse in bestimmten Zeitperioden, kann zu Verzerrungen der berechneten Durchschnittswerte führen, wenn sich die Bestimmungsfaktoren von erwarteten Aktienrenditen im Zeitablauf verändern. Zu diesen Variablen zählt zum Beispiel der als Size-Effekt bezeichnete Erklärungsbeitrag der Marktkapitalisierung für den Erwartungswert der Renditen.¹⁴ Diese Rahmenbedingungen für die Durchführung langfristiger Ereignisstudien werden zum Anlaß genommen, mit einer Monte-Carlo-Simulation mehrere Methoden zur Berechnung langfristiger Überrenditen für kleine Stichproben zu untersuchen.

Bei den Simulationsuntersuchungen werden zur Einschätzung der langfristigen Renditenentwicklung der Untersuchungswertpapiere sowohl Portefeullerrenditen als auch die Renditen von Vergleichswertpapieren verwendet und der Size-Effekt sowie die Schiefe in der Verteilung der langfristigen Renditen berücksichtigt. Zusätzlich werden im Rahmen der Monte-Carlo-Simulationen die zugrundeliegenden Verteilungen variiert und so ein zeitliches Clustering simuliert.

Im nachfolgenden Abschnitt 2 werden erstens Ergebnisse vorliegender Simulationsstudien zur Berechnung langfristiger Überrenditen skizziert. Insbesondere werden die in der Literatur vorgeschlagenen methodischen Verbesserungen erörtert, die zu weniger verzerrten Mittelwerten und Teststatistiken führen sollten. Zweitens werden die Resultate langfristiger Ereignisstudien vorgestellt, die eine Erklärung der beobachteten positiven oder negativen Überrenditen auf unternehmensspezifische Unterschiede zurückführen. Es wird am Beispiel von Ergebnissen für den deutschen Kapitalmarkt gezeigt, wie sensitiv die berechneten Durchschnittswerte auf Veränderungen der Untersuchungsmethodik bei kleinen Stichprobengrößen reagieren. Abschnitt 3 ist den methodischen Fragen der Berechnung langfristiger Überrenditen im Rahmen von Simulationsstudien gewidmet. Es wird auf die Berechnung der Mittelwerte und Teststatistiken, Details der Stichprobenbildung im Rahmen der Anwendung der Monte-Carlo-Methode sowie kurz auf das verwendete IT-System eingegangen. Abschnitt 4 enthält eigene empirische Untersuchungen für verschiedene Stichprobengrößen unter „idealen“ Testvoraussetzungen und bei Vorliegen eines kalenderzeitlichen Clusterings. Abschnitt 5 faßt die wichtigsten Schlußfolgerungen der Arbeit zusammen.

¹³ Beispielsweise konnten am deutschen Kapitalmarkt zwischen 1960 und 1979 nur 15, zwischen 1980 und 1992 hingegen 172 Börseneinführungen beobachtet werden, vgl. Stehle/Ehrhardt/Przyborowsky (1999), Tabelle 8.

¹⁴ Einige empirische Untersuchungen kommen zum Ergebnis, daß im Vergleich zu den Schätzwerten des CAPM Aktien mit niedrigerer (höherer) Marktkapitalisierung im Durchschnitt höhere (niedrigere) Renditen erzielen. Für den deutschen Markt liegen Berechnungsergebnisse von Oertmann (1994), S. 249 f., und Stehle (1997), S. 245, vor. Die Auswirkungen eines nichtstationären Size-Effektes auf die Untersuchungsergebnisse langfristiger Renditen belegen Stehle/Ehrhardt/Przyborowsky (1999), vgl. die Tabellen 4 und 5.

2. Bisherige Untersuchungen

2.1 Diskussion der Berechnungsmethodik

In den bisher vorliegenden Simulationsstudien wird vornehmlich unter Verwendung von Daten des US-amerikanischen Marktes die Berechnungsmethodik für langfristige Ereignisstudien untersucht. Die Studien kommen zu dem Ergebnis, daß die für die Berechnung langfristiger Überrenditen häufig genutzten Verfahren zu systematischen Fehlern in den berechneten Mittelwerten und Teststatistiken führen können. Die durchgeführten Simulationen zeigen insgesamt zu hohe Ablehnungshäufigkeiten für die Nullhypothese, wobei positive Überrenditen zu selten und negative Überrenditen zu häufig als statistisch signifikant ausgewiesen werden.¹⁵

Die festgestellten Verzerrungen sind vor allem auf methodische Unterschiede in der Berechnung der Endwerte der Untersuchungswertpapiere und den zur Beurteilung der langfristigen Renditeentwicklung benötigten Schätzwerte für die erwartete Rendite zurückzuführen. Der Berechnung der Überrendite liegt in der Regel die Annahme zugrunde, daß die Rendite der untersuchten Aktie gleich der tatsächlichen Rendite eines Vergleichsportefeuilles bzw. der Änderungsrate eines Indexes im Untersuchungszeitraum ist. Bei einer solchen Vorgehensweise werden sich die aus Einzelrenditen berechneten Endwerte der untersuchten Aktien und ihrer, aus Portefeullerrenditen kumulierten Vergleichswerte in den Verteilungseigenschaften im allgemeinen unterscheiden. Aufgrund der daraus resultierenden Schiefe der Überrenditenverteilung führen Mittelwerttests, die eine Normal- oder eine symmetrische Verteilung voraussetzen, zu verzerrten Teststatistiken bei der Überprüfung der Nullhypothese (*Skewness bias*).¹⁶

Ein Vergleich von Einzel- mit Portefeullerrenditen kann ebenfalls zu einer Verzerrung des Mittelwertes führen, wenn systematische Unterschiede in der Höhe der Marktkapitalisierung oder in anderen, für die Erklärung von Aktienrenditen relevanten Bestimmungsfaktoren zwischen den Untersuchungswertpapieren und den in den Vergleichsportefeuilles enthaltenen Aktien bestehen (*Benchmark bias*).

Barber/Lyon (1997) verweisen zudem darauf, daß sich die Zusammensetzung der in den Vergleichsportefeuilles enthaltenen Aktien innerhalb der Untersuchungsperiode verändern kann. Da Barber/Lyon (1997) von einer langfristig geringeren Durchschnittsrendite der neu an einer Börse eingeführten Aktien ausgehen, sehen sie im sogenannten „*New issue bias*“ eine weitere Ursache für systematische Verzerrungen der Berechnungsergebnisse. Fehler in der Einschätzung der langfristigen Renditeentwicklung der Untersuchungswertpapiere können sich ferner

¹⁵ Vgl. Kothari/Warner (1997), S. 314, Barber/Lyon (1997), S. 364, Cowan/Sergeant (1997), S. 33.

¹⁶ Vgl. Barber/Lyon (1997), S. 347.

ergeben, wenn die Renditeschätzungen auf der Grundlage eines gleichgewichteten Portefeuilles erfolgen. Bei einer gleichgewichteten Portefeuillestrategie sind in jeder Periode Umschichtungen erforderlich, um die als Folge der Aktienkursentwicklung entstandenen Veränderungen der Portefeuillegewichte auszugleichen. Hierzu werden Anteile von Aktien mit einer überdurchschnittlichen Renditeentwicklung in der vergangenen Periode verkauft und Anteile von Aktien mit einer unter dem Durchschnitt liegenden Rendite hinzugekauft. Diese, die Durchschnittsrendite beeinflussenden Umschichtungen betreffen die als Schätzer des Erwartungswertes verwendeten Portefeuillerenditen, nicht aber die Renditen der Untersuchungswertpapiere (*Rebalancing bias*).¹⁷

Ein weiteres Problem ergibt sich aus einer zeitlichen Überschneidung der einzelnen Ereignisfenster. Die Folge ist eine positive Korrelation der Aktienrenditen innerhalb einer Stichprobe. Die damit verbundene Beeinflussung der Standardabweichung führt zu einer Verzerrung der Teststatistik (*Overlapping horizons bias*).¹⁸ Dieser Effekt wird noch verstärkt, wenn es in einzelnen Branchen zu einer zeitlichen Häufung der untersuchten Ereignisse kommt. Branchenspezifische Effekte konnten bisher vor allem bei Fusionen und Börseneinführungen nachgewiesen werden.¹⁹

In Auswertung der Simulationsergebnisse von Barber/Lyon (1997) und Kothari/Warner (1997) schlagen Cowan/Sergeant (1997) vor, die Schiefe durch Winsorisation²⁰ der berechneten Überrenditen zu reduzieren und die Auswirkungen einer positiven Korrelation zwischen den Aktienrenditen auf die Varianz durch den t-Test für unabhängige Stichproben zu berücksichtigen.²¹ Häufig wird auch die Verwendung von Vergleichswertpapieren zur Berechnung von Überrenditen empfohlen.²² Der Vorteil dieser Methode ist die Vermeidung des *Rebalancing bias* und des *Skewness bias*, der Nachteil liegt in einer höheren Standardabweichung der beobachteten Stichprobenmittelwerte.²³ Lyon/Barber/Tsai (1999) empfehlen, die Renditen von Buy-and-hold-Vergleichsportefeuilles als Schätzwerte für die Renditen der untersuchten Aktien zu verwenden, die Teststatistiken auf der Grundlage eines schiefeangepaßten t-Tests zu berechnen und die kritischen t-Werte zur Ablehnung der Nullhypothese mit einem Bootstrap-Verfahren zu bestimmen.²⁴ Die Festlegung eines Ablehnungsbereiches bei der Überprüfung der Signifikanz der berechneten langfristigen Überrenditen mit einem Bootstrap-Verfahren

¹⁷ Die Bezeichnungen „*New issue bias*“ und „*Rebalancing bias*“ wurden von Barber/Lyon (1997), S. 347 f., eingeführt.

¹⁸ Vgl. Cowan/Sergeant (1997), S. 20.

¹⁹ Vgl. Andrade/Stafford (1997), Ritter (1984).

²⁰ Unter Winsorisation wird ein Verfahren zur Begrenzung von Extremwerten verstanden. Es werden alle Beobachtungen, die zum Beispiel drei oder mehr Standardabweichungen vom Mittelwert entfernt liegen, auf den niedrigsten bzw. höchsten der verbliebenen Werte begrenzt.

²¹ Vgl. Cowan/Sergeant (1997), S. 12 ff.

²² Vgl. z.B. Barber/Lyon (1997), S. 342 f.

²³ Vgl. Cowan/Sergeant (1997), S. 34.

²⁴ Vgl. Lyon/Barber/Tsai (1999), S. 10 f.

ren wurde bereits von Ikenberry/Lakonishok/Vermaelen (1995) vorgeschlagen.²⁵ Ein solches im Abschnitt 3.7 näher beschriebenes Verfahren führt in den Untersuchungen von Lyon/Barber/Tsai (1999) und Cowan/Sergeant (1997) zu gut spezifizierten Teststatistiken.²⁶

Inwieweit die vorgeschlagenen methodischen Verbesserungen tatsächlich zu unverzerrten Mittelwerten und Teststatistiken führen, ist allerdings nicht eindeutig zu beantworten. Für Stichproben, in denen z.B. einzelne Unternehmen oder Unternehmen einzelner Branchen mit einer höheren Häufigkeit auftreten, beobachten Lyon/Barber/Tsai (1999) systematische Verzerrungen der Teststatistiken für Verfahren, die unter den „idealen“ Testbedingungen simulierter Stichproben unverzerrte Werte liefern.²⁷ Brav (1998) kritisiert andererseits die Anwendung der Bootstrap-Verfahren, da erstens innerhalb einer Stichprobe die Aktienrenditen der untersuchten und zufällig ausgewählten Wertpapiere eine unterschiedliche Querschnittskorrelation und zweitens eine unterschiedliche Varianz aufweisen können.²⁸ Die höhere Varianz der untersuchten Wertpapiere begründen Mitchell/Stafford (1997) damit, daß die betrachteten Unternehmen Entscheidungen zur Investitions-, Finanzierungs- und Ausschüttungspolitik treffen, die Vergleichsunternehmen hingegen nicht.²⁹

In den bisherigen Untersuchungen zum methodischen Vorgehen bei der Berechnung langfristiger Überrenditen wird darauf hingewiesen, daß die Ergebnisse sehr vorsichtig interpretiert werden müssen. Nach Kothari/Warner (1997) ist in den Simulationsstudien zum einen eine Beständigkeit der Existenz von Fehlbewertungen festzustellen, zum anderen ist die Richtung der Verzerrungen sehr sensitiv bezüglich der Eigenschaften der Wertpapiere, z.B. unterschiedlicher Buchwert/Marktwert-Verhältnisse.³⁰ Lyon/Barber/Tsai (1999) empfehlen, die Berechnungsergebnisse für unterschiedliche Zeitperioden und andere Kapitalmärkte (*Out-of-sample*-Tests) auf der Grundlage traditioneller Bewertungsmodelle und empirisch beobachteter Anomalien zu überprüfen.³¹

²⁵ Vgl. Ikenberry/Lakonishok/Vermaelen (1995), S. 189 f.

²⁶ Vgl. Lyon/Barber/Tsai (1999), S. 16, Cowan/Sergeant (1997), S. 35.

²⁷ Vgl. Lyon/Barber/Tsai (1999), S. 20 ff.

²⁸ Vgl. Brav (1998), S. 5.

²⁹ Vgl. Mitchell/Stafford (1997), S. 28.

³⁰ Vgl. Kothari/Warner (1997), S. 335. Ähnlich argumentieren Mitchell/Stafford (1997), S. 17.

³¹ Vgl. Lyon/Barber/Tsai (1999), S. 29.

2.2 Unternehmensspezifische Unterschiede zur Erklärung langfristiger Überrenditen

Einige Studien deuten darauf hin, daß firmenspezifische Aspekte bei der Erklärung der langfristigen Renditen eine zentrale Rolle spielen können. Nach Brav/Gompers (1997) und Brav/Geczy/Gompers (1999) wird die langfristige Underperformance nach Börseneinführungen und Kapitalerhöhungen am US-amerikanischen Kapitalmarkt durch Unternehmen mit einer geringen Marktkapitalisierung und einem geringen Buchwert/Marktwert-Verhältnis verursacht. Für die Erklärung der festgestellten langfristigen Überrenditen sind offensichtlich die gleichen empirisch identifizierten Bestimmungsfaktoren von Bedeutung, die in den Modellen von Fama/French (1993) und Carhart (1997) einen Beitrag zur Erklärung der Preisbildung auf den Aktienmärkten leisten.³²

Teoh/Welch/Wrong (1998, 1999) finden eine Underperformance bei Unternehmen, die vor Kapitalmaßnahmen in verstärktem Umfang Ansatz- und Bewertungswahlrechte wahrnehmen.³³ Für einige in der Stichprobe enthaltenen Unternehmen werden langfristige negative Überrenditen beobachtet, die dadurch entstehen, daß Anleger die Täuschung durch das Management erkennen. In ihrer Studie verweisen sie zudem auf den Aspekt, daß bei langfristigen Ereignisstudien die untersuchten Aktien in der Regel gleichgewichtet werden. Die Berücksichtigung von Unternehmen mit einer geringen Marktkapitalisierung kann folglich zu Berechnungsergebnissen führen, die nicht die Wertentwicklung der von den Anlegern gehaltenen Portefeuilles widerspiegeln.

Ferner enthalten empirische Studien Hinweise, daß zwischen der langfristigen Renditeentwicklung der Aktien nach Börseneinführungen und der Mitwirkung von Finanzintermediären ein Zusammenhang besteht. Positive Auswirkungen auf die langfristige Rendite sind bei einer Aktienplatzierung durch einen Konsortialführer mit hoher Reputation bzw. bei einer Kapitalbeteiligung einer Venture-capital-Gesellschaft zu beobachten.³⁴

2.3 Unterschiedliche Berechnungsergebnisse langfristiger Ereignisstudien am deutschen Markt

Für den deutschen Kapitalmarkt liegen bisher nur wenige langfristige Ereignisstudien vor. Die berechneten Durchschnittswerte unterscheiden sich zum Teil stark voneinander. Als Beispiele sein genannt:

- Ljungqvist (1997) und Ehrhardt (1997) untersuchen die langfristige Rendite nach Börseneinführungen für einen fast identischen Datensatz und vergleichen diese mit der langfristigen Renditeentwicklung eines marktwertgewichteten Portefeuilles, welches aus allen im

³² Vgl. Brav/Gompers (1997), S. 1800, Brav/Geczy/Gompers (1999), S. 21 f., insbesondere S. 29.

³³ Vgl. Teoh/Welch/Wrong (1999), S. 13.

³⁴ Vgl. Carter/Dark/Singh (1998), S. 296; Brav/Gompers (1997), S. 1800.

amtlichen Handel Frankfurt notierten Aktien besteht. Eine unterschiedliche Formel zur Mittelwertbildung über alle einbezogenen Aktien und über die Zeit führt vermutlich zu dem beobachteten Renditeunterschied von 17,56 Prozentpunkten.³⁵

- Bei Verwendung unterschiedlicher Schätzwerte für die erwartete Rendite der Untersuchungswertpapiere variieren die berechneten kumulierten Durchschnittswerte z.B. bei Sapusek (1998) zwischen ca. 5 und -30 Prozent für die ersten fünf Jahre nach der Börseneinführung. Stehle/Ehrhardt/Przyborowsky (1998) untersuchen die langfristige Rendite über einen Zeitraum von drei Jahren. Ihre Durchschnittswerte schwanken zwischen 1,54 und -6,64 Prozent bei der Untersuchung von Börseneinführungen bzw. zwischen -3,17 und -9,01 Prozent bei Kapitalerhöhungen.
- Ljungqvist (1997) findet große Unterschiede bei der langfristigen kumulierten Durchschnittsrendite nach Börseneinführungen für unterschiedliche Betrachtungsperioden, d.h. -1,8 Prozent für IPO-Aktien mit der Erstnotierung im Zeitraum 1970 bis 1987 bzw. -27,2 Prozent zwischen 1988 und 1990.

Die unterschiedlichen Ergebnisse der langfristigen Renditeberechnungen nehmen Stehle/Ehrhardt (1999) zum Anlaß, die bisherigen in der Literatur verwendeten Verfahren zur Mittelwertbildung über alle einbezogenen Wertpapiere und über die betrachtete Zeitperiode intensiv zu diskutieren. Nach den dort dargestellten Überlegungen scheint die als „*Buy-and-hold abnormal return (BHAR)*“ bezeichnete Methode am besten geeignet zu sein. In Stehle/Ehrhardt/Przyborowsky (1998) wird die Frage erörtert, welcher Schätzwert für die erwartete Rendite der Untersuchungswertpapiere genommen werden sollte. Sie empfehlen anstelle der Änderungsrate eines Indexes oder der Rendite des marktwertgewichteten Portefeuilles aller in Frankfurt amtlich gehandelten Aktien die Verwendung von Size-Portefeuillerenditen, da bei einer solchen Vorgehensweise die von Ljungqvist (1997) beobachteten beträchtlichen Unterschiede der durchschnittlichen Überrenditen für einzelne Teilperioden nicht mehr festzustellen sind.

3. Berechnungsmethodik

3.1 Berechnung der durchschnittlichen langfristigen Überrendite

Zur Berechnung der durchschnittlichen Überrendite nach dem BHAR-Verfahren wird in einem ersten Schritt für jedes ausgewählte Untersuchungswertpapier durch Multiplikation der monatlichen Renditen ein Endwert nach 36 Monaten berechnet. Dieser Endwert wird in einem zweiten Schritt mit dem dazugehörigen und auf gleiche Weise berechneten Endwert eines Vergleichswertpapiers oder -portefeuilles verglichen, der als Schätzwert für die Rendite der Untersuchungswertpapiere betrachtet wird. Ein solcher Vergleich wird in der Literatur auch

³⁵ Vgl. Stehle/Ehrhardt (1999), S. 15.

als Konzept marktbereinigter Renditen (*Market-adjusted approach*) bezeichnet.³⁶ In einem dritten Schritt wird die Differenz der berechneten Endwerte gebildet und über alle Werte der Stichprobe das arithmetische Mittel berechnet. Formelmäßig gilt:

$$CAR_{T=36}^{BHAR} = \frac{1}{N} \sum_{i=1}^N \left(\left(\prod_{t=1}^{T=36} R_{i,t} + 1 \right) - \left(\prod_{t=1}^{T=36} R_{P,t} + 1 \right) \right) = \frac{1}{N} \sum_{i=1}^N (W_{i,T=36} - W_{P,T=36}). \quad (1)$$

$W_{i,T=36}$ = Endwert zum Zeitpunkt T einer Kapitalanlage von einer Geldeinheit, die zu Beginn des Untersuchungszeitraumes ($t=0$) in Aktie i investiert wurde;

$W_{P,T=36}$ = Endwert zum Zeitpunkt T einer Kapitalanlage von einer Geldeinheit in ein Vergleichsportefeuille oder –wertpapier für die Aktie i ;

$R_{i(P),t}$ = Rendite des Untersuchungswertpapiers i (bzw. des dazugehörigen Vergleichsportefeuilles/-wertpapier) im Monat t .

Für die Simulationen werden die monatlichen Renditen aller im amtlichen Handel der Frankfurter Wertpapierbörse notierten Aktien für den Zeitraum 1960 bis 1995 verwendet. Für die Schätzung der erwarteten Renditen der Untersuchungswertpapiere dienen erstens die von Stehle/Hartmond (1991) aus der gleichen Datenbasis berechneten und inzwischen zeitlich ergänzten Renditen eines gleich- bzw. marktwertgewichteten Portefeuilles. Zweitens werden aus den Daten Renditen von Size-Portefeuilles berechnet. Dazu werden die Aktien nach ihrer Marktkapitalisierung geordnet, in zehn Portefeuilles mit der gleichen Anzahl von Aktien eingeteilt und dann gleich- bzw. marktwertgewichtete Size-Portefeuillerenditen bestimmt. Bei der Berechnung der Überrendite nach Gleichung 1 wird dem Untersuchungswertpapier jenes Size-Portefeuille zugeordnet, bei dem die Differenz zwischen dem Marktwert des untersuchten Wertpapiers (MW_i) und der durchschnittlichen Marktkapitalisierung aller in einem Size-Portefeuille enthaltenen Aktien \overline{MW}_{SPj} minimal ist, also:

$$\text{Size-Portefeuille-Nr. } j: j = \min_j |MW_i - \overline{MW}_{SPj}|. \quad (2)$$

Der Vergleich der Einzelrenditen der Untersuchungswertpapiere mit Portefeuillerenditen kann zu systematischen Verzerrungen der Untersuchungswertpapiere führen, die ihre Ursachen im *Rebalancing*-, *New-issue*- und *Skewness*-Effekt haben können (vgl. Abschnitt 2.1). Deshalb wird drittens die Rendite eines nach der Marktkapitalisierung vergleichbaren Wertpapiers als Schätzwert für die Rendite der Untersuchungswertpapiere verwendet.

3.2 Berechnung der Teststatistiken bei schiefen Verteilungen

3.2.1 Schiefeangepaßter t-Test

Es wurde schon mehrfach darauf hingewiesen, daß bei der Verwendung von Portefeuillerenditen als Schätzwert für die erwartete Rendite der untersuchten Aktien die Verteilung der

³⁶ Vgl. Brown/Warner (1980), S. 207 f.

langfristigen Überrenditen durch einen positiven Wert für die Schiefe charakterisiert ist. Um die aus einer schiefen Verteilung resultierenden Verzerrungen in den Teststatistiken zu reduzieren, schlagen Lyon/Barber/Tsai (1999) die Verwendung eines schiefeangepaßten t-Testes vor. Die Teststatistik wird berechnet durch:³⁷

$$t^s = \sqrt{N} \cdot \left[\frac{\mu(CAR_{T=36}^{BHAR})}{\sigma(CAR_{T=36}^{BHAR})} + \frac{1}{3} \cdot \gamma(CAR_{T=36}^{BHAR}) \cdot \left[\frac{\mu(CAR_{T=36}^{BHAR})}{\sigma(CAR_{T=36}^{BHAR})} \right]^2 + \frac{1}{6 \cdot N} \cdot \gamma(CAR_{T=36}^{BHAR}) \right]. \quad (3)$$

$\mu(\cdot)$ ist der Mittelwert der Überrenditen aller Untersuchungswertpapiere einer Stichprobe, $\sigma(\cdot)$ die dazugehörige (Querschnitts-)Standardabweichung und $\gamma(\cdot)$ der Koeffizient der Schiefe. $\gamma(\cdot)$ wird berechnet als:

$$\gamma(CAR_{T=36}^{BHAR}) = \frac{\sum_{i=1}^N [CAR_{T=36}^{BHAR} - \mu(CAR_{T=36}^{BHAR})]^3}{N \cdot [\sigma(CAR_{T=36}^{BHAR})]^3}. \quad (4)$$

3.2.2 Bootstrap-Verfahren

Die Simulation empirischer Verteilungen mit einer Bootstrap-Methode und einer darauf basierenden Bestimmung der kritischen Werte zur Ablehnung der Nullhypothese stellt eine weitere Möglichkeit dar, Verzerrungen der Teststatistiken als Folge asymmetrischer Verteilungen zu vermeiden. Das in der vorliegenden Studie verwendete Bootstrap-Verfahren wird nachfolgend kurz skizziert.³⁸

$F(CAR_{T=36}^{BHAR})$ sei die unbekannte theoretische Verteilung der langfristigen kumulierten Überrendite mit der Prüfgröße $\Theta_1(t)$ bzw. $\Theta_2(t^s)$, wobei der Berechnung der t- bzw. schiefeangepaßte t-Wert zugrundegelegt wird. Die Quantile der Bootstrap-Realisierungen erlauben die Bestimmung eines Verwerfungsbereiches für die Nullhypothese.

Die Bootstrap-Methode ersetzt $F(CAR_{T=36}^{BHAR})$ durch eine empirische Verteilungsfunktion $\hat{F}_B(CAR_{T=36}^{BHAR})$ mit der dazugehörigen Statistik $\hat{\Theta}_1(t_1, \dots, t_B)$ bzw. $\hat{\Theta}_2(t_1^s, \dots, t_B^s)$. Die empirische Stichprobenverteilung der Statistik wird durch eine Monte-Carlo-Simulation approximiert. Aus den zur Menge der empirischen Verteilung $\hat{F}_B(CAR_{T=36}^{BHAR})$ zugehörigen Aktien werden in einem ersten Schritt B Bootstrap-Stichproben aufgebaut, indem eine gewünschte Anzahl von Werten mit Zurücklegen gezogen wird. Aus den gezogenen Werten werden langfristige Überrenditen berechnet und zu den Stichprobenmittelwerten $CAR_{t_1}^*, \dots, CAR_{t_B}^*$ (Bootstrap-Verteilung) aggregiert. Für jede Bootstrap-Stichprobe wird dann in einem zweiten Schritt ein Wert

³⁷ Die schiefeangepaßte Teststatistik wurde von Johnson (1978) auf der Grundlage einer Edgeworth-Expansion entwickelt, von Hall (1992) und Sutton (1993) diskutiert. Die Anwendung der schiefeangepaßten Teststatistik bei langfristigen Ereignisstudien geht auf eine Idee von Lyon/Barber/Tsai (1999), S. 9-11, zurück.

³⁸ Die nachfolgenden Ausführungen sind an Efron (1982), S. 1-3, 27-36 angelehnt.

der Statistik $\hat{\Theta}_1(t_1^*, \dots, t_B^*)$ bzw. $\hat{\Theta}_2(t_1^{s*}, \dots, t_B^{s*})$ berechnet. In einem dritten Schritt werden aus den beiden zuletzt genannten Verteilungen die Grenzen der Konfidenzintervalle bestimmt:

$$\Pr[t_{krit.}^* \leq t_b] = \Pr[t_{krit.}^* \geq t_b] = \frac{\alpha}{2} \text{ bzw. } \Pr[t_{krit.}^{s*} \leq t_b^s] = \Pr[t_{krit.}^{s*} \geq t_b^s] = \frac{\alpha}{2}. \quad (5)$$

3.3 Nullhypothese und Stichprobenbildung

3.3.1 Stichprobengrößen bei der Untersuchung bewertungsrelevanter Ereignisse

Die nachfolgende Tabelle 1 zeigt die Unterschiede in der Anzahl der zur Verfügung stehenden Beobachtungen, wenn Ereignisstudien für den US-amerikanischen bzw. für den deutschen Kapitalmarkt durchgeführt werden. Anhand der dargestellten Stichprobengrößen werden für die Simulationsrechnungen in einem ersten Schritt die Anzahl der Beobachtungen pro Stichprobe auf 50 festgesetzt. In einem zweiten Schritt wird die Stichprobengröße auf 150 verdreifacht und die beobachteten Veränderungen analysiert.

Tabelle 1

Anzahl der Beobachtungen bei Renditeuntersuchungen zur Finanzierungs- und Dividendenpolitik für den US-amerikanischen und den deutschen Kapitalmarkt (Beispiele)

| Studie | Untersuchtes Ereignis | Anzahl der Beobachtungen |
|---|---|--------------------------|
| Teil A: US-amerikanischer Kapitalmarkt | | |
| Loughran/Ritter (1995) | Initial public offerings (LR) | 4753 |
| | Seasoned equity offerings | 3702 |
| Michaely/Thaler/Womack | Dividend omissions (A,LR) | 887 |
| | Dividend initiations (A,LR) | 561 |
| Mitchell/Stafford (1997) | Share repurchases, Open market (A,LR) | 2492 |
| | Share repurchases, Tender (A,LR) | 437 |
| | Mergers & Tender offers (A,LR) | 2636 |
| Teil B: Deutscher Kapitalmarkt | | |
| Entrup (1995) | Ausgabe von Optionsanleihen (A) | 69 |
| Gerke/Oerke/Sentner (1997) | Dividendenerhöhungen (A) | 218 |
| | Dividendensenkungen (A) | 46 |
| | Dividendenausfälle (A) | 20 |
| Gebhardt/Entrup/Heiden (1994) | Kapitalerhöhungen aus Gesellschaftsmitteln (A) | 69 |
| Kaserer/Brunner (1997) | Kapitalerhöhungen aus Gesellschaftsmitteln (A,LR) | 181 |
| Stehle/Ehrhardt/Przborowsky (1999) | Kapitalerhöhungen (LR) | 584 |
| | Börseneinführungen (LR) | 187 |

Es bedeutet: A = Untersuchung von Ankündigungseffekten, LR = Untersuchung der langfristigen Renditeentwicklung nach einem bewertungsrelevanten Ereignis.

3.3.2 Stichprobenbildung durch Monte-Carlo-Simulation, Nullhypothese

Wie bereits erwähnt, stehen für die Simulationsstudie monatliche Renditen aus allen im amtlichen Handel der Frankfurter Wertpapierbörse zwischen 1960 und 1995 notierten Aktien zur Verfügung. Um einen „*Selection bias*“ zu vermeiden, werden auch Aktien in die Untersuchung einbezogen, die im genannten Zeitraum ihre Börsennotierung durch Konkurs, Fusion, Going private oder anderen Gründen einstellten bzw. in Zusammenhang mit einem Going public, einem Börsensegmentwechsel oder der Erstemission einer weiteren Aktiengattung aufnahmen.

Die Berechnung langfristiger Überrenditen über einen Zeitraum von 36 Monaten bedingt, daß für die Auswahl des Ereignismonats 0 nur Daten zwischen 1960 und 1992 zur Verfügung stehen. Aus dieser Grundgesamtheit werden 50.000 Beobachtungen mit Zurücklegen gezogen (Monte-Carlo-Simulation). Der Auswahl der Aktien und des Ereigniszeitpunktes (Monat 0) liegt die Gleichverteilung zugrunde. Die 50.000 Beobachtungen werden in 1.000 Stichproben zu jeweils 50 Aktien eingeteilt.

Für jede der 50.000 zufällig ausgewählten Beobachtungen wird eine langfristige kumulierte Überrendite für die dem Ereignismonat 0 folgenden 36 Monate berechnet. Als Schätzwert für die erwartete Rendite der zufällig ausgewählten Wertpapiere werden die Renditen von vier verschiedenen Portefeuilles und eines nach der Marktkapitalisierung vergleichbaren Wertpapiers verwendet (vgl. Abschnitt 3.1). Durch die zufällige Auswahl der Aktien und Ereigniszeitpunkte sollten sich die über einen Zeitraum von 36 Monaten berechneten Endwerte der Untersuchungswertpapiere und die Endwerte der dazugehörigen Erwartungswerte im Durchschnitt nicht voneinander unterscheiden. Die nachfolgend formulierte Nullhypothese bildet folglich die Grundlage zur Überprüfung der Berechnungsmethodik:

$$H_0 : CAR_{T=36}^{BHAR} = 0. \quad (6)$$

Bei einem Simulationslauf werden insgesamt 1,85 Millionen Datensätze verarbeitet (50.000 x 37). Um eine Erhöhung der Stichprobengröße von 50 auf 150 zu erreichen, müßten 5,55 Millionen Datensätze in die Berechnungen einbezogen werden. Aus diesem Grund wird bei der Erhöhung der Stichprobengröße auf eine Bootstrap-Methode zurückgegriffen. Als Ergebnis der ersten Simulationsrechnung liegen 50.000 über einen Zeitraum von 36 Monaten berechnete langfristige Überrenditen vor. Aus diesen Daten wird durch Ziehen mit Zurücklegen eine Stichprobe von 150.000 Beobachtungen aufgebaut, die dann in 1.000 Einzelstichproben eingeteilt wird.

Ein weiteres Problem bei der Berechnung langfristiger Renditen ist die Behandlung von Notizeinstellungen. In der vorliegenden Simulationsstudie wird in einem solchen Fall die kumulierte Überrendite für den Zeitraum bis zur Notizeinstellung berechnet. Dieser Wert wird dann unverändert bis zum 36. Monat fortgeschrieben. Wenn eine Aktie im Monat 0 ausgewählt wird, die im Folgemonat die Börsennotierung einstellt, so stehen keine monatlichen Renditen

für die Berechnung zur Verfügung. Dies führt zu einer Reduzierung der in der Stichprobe enthaltenen Unternehmen um 0,13 Prozent.

3.4 Das EDV-System zur Durchführung der Untersuchungen

Zur Durchführung der Untersuchungen wurden unterschiedliche Standard-Softwareprodukte mit dem Ziel verwendet, die jeweiligen Stärken dieser Systeme in sich gegenseitig ergänzender Weise zu nutzen. Die Verbindungen dieser Komponenten wurden dabei weitestgehend in einer Client-Server-Architektur mit Standardprotokollen und teilweise selbsterstellter Kopplungssoftware der Autoren realisiert.

Die umfangreichen Datenbestände mit zum Teil mehreren Millionen Einträgen werden auf einem zentralen relationalen Datenbankserver (DB2) gehalten und können mit der Standard-Datenbanksprache SQL abgefragt oder modifiziert werden. Komplexe datenbanknahe und geschwindigkeitskritische Prozeduren wurden in ESQL/C-Programmen (in C eingebettetem SQL) implementiert. Als zusätzlicher Datenbank-Client findet eine Standard-PC-Software (MS-Access) Verwendung. Als System für die abschließenden statistischen Auswertungen wurde SPSS (Statistical Package for the Social Sciences) eingesetzt.

Für die Untersuchungen war die Generierung umfangreicher Mengen von (Pseudo-)Zufallszahlen hoher Qualität mit unterschiedlichen, speziell angepaßten Verteilungseigenschaften notwendig. Hierzu wurde das Computer-Algebra-System Mathematica verwendet.

Zur Erzeugung von ganzzahligen Zufallszahlen setzt Mathematica einen zellulären Automaten nach Wolfram ein.³⁹ Rationale Zufallszahlen werden mit einem Subtraktion-mit-Übertrag Verfahren nach Marsaglia-Zaman generiert.⁴⁰ Dadurch können qualitativ gute Pseudozufallszahlen mit (sehr) langen Periodenlängen erreicht werden.

Mathematica ist über eine SQL-Schnittstelle (von den Autoren erstellte Middleware SQLMATH in ESQL/C, basierend auf der Standardschnittstelle Mathlink) direkt mit dem DB2-Datenbankserver verbunden. Datenbankinhalte können so unmittelbar und datentypenverträglich zwischen beiden Softwareprodukten ausgetauscht werden. Relationale Tabellen des Datenbanksystems werden in Mathematica als Liste von Listen abgebildet. Mathematische und statistische Funktionen können auf dieser Standard-Datenstruktur unmittelbar angewendet werden.

³⁹ Vgl. Wolfram (1997), S. 1094, Wolfram (1986).

⁴⁰ Vgl. Wolfram (1997), Marsaglia/Zaman (1991).

4. Empirische Ergebnisse

4.1 Untersuchungen für unterschiedliche Stichprobengrößen

4.1.1 Deskriptive Beschreibung der Verteilung der Stichprobenmittelwerte

In der Tabelle 2 sind die durchschnittlichen Stichprobenmittelwerte, die zugehörigen Standardabweichungen und die Werte für die Schiefe der Verteilung angegeben. Die über einen Zeitraum von 36 Monaten berechneten langfristigen Überrenditen sind hoch signifikant, wenn die Renditen der zufällig ausgewählten Untersuchungswertpapiere mit den dazugehörigen Renditen eines Portefeuilles aus allen amtlich in Frankfurt gehandelten Aktien verglichen werden. Der signifikante Unterschied der berechneten Endwerte deutet auf systematische Verzerrungen der Berechnungsergebnisse hin. Eine unterschiedliche Gewichtung der im Vergleichsportefeuille enthaltenen Aktien führt zu einem, vermutlich auf dem Size-Effekt basierenden Unterschied von über fünf Prozentpunkten bei der berechneten durchschnittlichen Überrendite nach 36 Monaten.

Das negative Vorzeichen bei Verwendung des gleichgewichteten Portefeuilles deutet auf eine zusätzliche Beeinflussung der Ergebnisse durch den Rebalancing-Effekt hin. Zur Aufrechterhaltung der Portefeuillegewichtungen werden bei einer gleichgewichteten Portefeuillestrategie Anteile von Aktien mit überdurchschnittlicher Kursentwicklung in der zurückliegenden Periode verkauft und Aktien mit unterdurchschnittlicher Kursentwicklung hinzugekauft. Diese Strategie läßt (bei Abstraktion von Transaktionskosten) im Vergleich zu einer marktwertgewichteten Portefeuillestrategie auf den ersten Blick eine geringere langfristige Durchschnittsrendite vermuten. Können die Renditeveränderungen im Zeitablauf durch einen Mean-reversion-Prozeß beschrieben werden, so kann eine gleichgewichtete Portefeuillestrategie zu höheren langfristigen Durchschnittsrenditen führen.

Werden die Renditen der zufällig ausgewählten Wertpapiere mit Schätzwerten für die langfristige Rendite verglichen, die den Size-Effekt und somit unternehmensspezifische Unterschiede zwischen den einzelnen untersuchten Aktien berücksichtigen, so können mit Ausnahme des ungewichteten Size-Portefeuilles keine statistisch signifikanten Mittelwertunterschiede festgestellt werden. Die Ergebnisse deuten darauf hin, daß ein häufig bei langfristigen Ereignisstudien verwendeter Vergleich der Renditen der Untersuchungswertpapiere mit Aktienindizes zu ungenauen Schätzungen der langfristigen Renditeentwicklung führt.

Tabelle 2

Mittelwert, Standardabweichung und Schiefe der Überrenditenverteilung einer simulierten langfristigen Ereignisstudie

Die simulierte Studie basiert auf 1.000 Stichproben mit jeweils 50 bzw. 150 Aktien. In einem ersten Schritt werden für die einzelnen Untersuchungswertpapiere und Vergleichsportefeuilles bzw. -wertpapiere über einen Zeitraum von 36 Monaten Buy-and-hold-Renditen berechnet, in einem zweiten Schritt durch Differenzbildung die Überrenditen bestimmt und innerhalb einer Stichprobe zu einem arithmetischen Mittel aggregiert. Mit dem t-Test für abhängige Stichproben wird überprüft, ob sich der Mittelwert der Stichprobenmittelwert-Verteilung signifikant von null unterscheidet.

| Schätzwert für die langfristige Rendite | Stichprobenmittelwert-Verteilung | | | Durchschnittliche Schiefe der Stichprobenverteilung |
|--|----------------------------------|--------|------------------------|---|
| | Mittelwert (%) | t-Wert | Standardabweichung (%) | |
| Teil A: 1.000 Stichproben mit 50 Aktien | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | |
| ungewichtetes Portefeuille | -1,78 | -5,17 | 10,91 | 1,77 |
| marktwertgewichtetes Portefeuille | 3,62 | 10,26 | 11,51 | 1,56 |
| <i>Berücksichtigung des Size-Effektes</i> | | | | |
| ungewichtetes Size-Portefeuille | -0,61 | -1,81 | 10,62 | 1,61 |
| marktwertgewichtetes Size-Portefeuille | -0,21 | -0,62 | 10,62 | 1,62 |
| Vergleichswertpapier | -0,61 | -1,28 | 15,11 | -0,09 |
| Teil B: 1.000 Stichproben mit 150 Aktien | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | |
| ungewichtetes Portefeuille | -1,82 | -9,02 | 6,38 | 3,17 |
| marktwertgewichtetes Portefeuille | 3,61 | 17,19 | 6,34 | 2,92 |
| <i>Berücksichtigung des Size-Effektes</i> | | | | |
| ungewichtetes Size-Portefeuille | -0,64 | -3,26 | 6,23 | 2,96 |
| marktwertgewichtetes Size-Portefeuille | -0,25 | -1,26 | 6,24 | 2,98 |
| Vergleichswertpapier | -0,27 | -0,99 | 8,50 | -0,01 |

Die in der Tabelle 2 enthaltenen Ergebnisse zeigen ferner, daß mit steigendem Stichprobenumfang die Querschnittsstandardabweichung der Stichprobenmittelwerte beträchtlich sinkt, so daß bei einem Stichprobenumfang von 150 Beobachtungen höhere t-Werte ausgewiesen werden. Die Standardabweichung der Stichprobenmittelwerte gibt zudem Aufschluß über die durchschnittliche Abweichung vom „wahren“ Durchschnittswert, mit der im Rahmen *einer* empirischen Untersuchung eines konkreten bewertungsrelevanten Ereignisses gerechnet werden muß. Zudem wird (erwartungsgemäß) eine positive Schiefe der kumulierten langfristigen Überrenditen festgestellt, wenn die Renditen von Einzelwertpapieren mit Portefeuillerenditen verglichen werden. Die Schiefe in den einzelnen Stichproben nimmt mit der Anzahl der einbezogenen Beobachtungen zu, wenn mit Portefeuillerenditen bereinigt wird. Die Schiefe der Verteilung läßt Verzerrungen bei der Überprüfung der Nullhypothese vermuten, wenn parametrische Tests oder nichtparametrische Testverfahren, die die Symmetrieeigen-

schaft voraussetzen, eingesetzt werden. Werden die Renditen von Vergleichswertpapieren als Schätzer für die langfristige Rendite verwendet, so ist die Verteilung der kumulierten Überrenditen symmetrisch. Es besteht folglich auch kein Zusammenhang zwischen Schiefe und Stichprobenumfang.

4.1.2 Ablehnungshäufigkeiten für die Nullhypothese

Die Überprüfung der Nullhypothese für die untersuchten 1.000 Stichprobenmittelwerte erfolgt zunächst unter Verwendung parametrischer Testverfahren (Tabelle 3). Aufgrund der beobachteten schiefen Verteilungseigenschaften werden die Teststatistiken bei Verwendung von Portefeuillerenditen mit einem schiefeangepaßten t-Test berechnet. Werden die Renditen von Vergleichswertpapieren den Berechnungen zugrundegelegt, so kann ein konventioneller t-Test Verwendung finden, da die Überrenditenverteilung nach den in der Tabelle 2 angegebenen Werten für die Schiefe annähernd symmetrisch ist. Die Ablehnungshäufigkeiten für die Nullhypothese sind in der Tabelle 3 dargestellt.

Bei Verwendung von Portefeuillerenditen zur Schätzung der Renditen der analysierten Wertpapiere sind systematische Verzerrungen der Teststatistiken erkennbar. Erstens liegen die Ablehnungshäufigkeiten der Nullhypothese über den theoretisch erwarteten Werten. Zweitens sind die Verzerrungen der Teststatistiken für positive und negative Überrenditen unterschiedlich. Es wird zu häufig eine Underperformance, aber zu selten eine Overperformance ausgewiesen. Drittens sind die Abweichungen der empirisch beobachteten und der theoretisch erwarteten Ablehnungshäufigkeiten in fast allen Fällen bei einem Signifikanzniveau von 1 % statistisch signifikant. Der schiefeangepaßte t-Test kann nach den vorliegenden Ergebnissen die Verletzung der Verteilungsannahmen nicht kompensieren.

Werden nach der Marktkapitalisierung vergleichbare Wertpapiere zur Berechnung der Überrenditen verwendet, so werden bei einer Stichprobengröße von 50 Aktien keine statistisch signifikanten Unterschiede festgestellt. Bei einer Stichprobengröße von 150 Aktien führt ein im Durchschnitt leicht negativer Mittelwert zu einer etwas niedrigeren Ablehnungshäufigkeit für positive Überrenditen, wenn ein zweiseitiger Signifikanztest bei einer vorgegebenen Irrtumswahrscheinlichkeit von einem bzw. fünf Prozent zugrundegelegt wird. Bei einem Signifikanzniveau von zehn Prozent ergeben sich keine signifikanten Abweichungen. Die Asymmetrie bei der Ablehnung der Nullhypothese in beiden Verteilungsenden ist deutlich geringer als bei der Verwendung von Vergleichsportefeuilles.

Tabelle 3

Ablehnungshäufigkeiten für die Nullhypothese unter Verwendung parametrischer Testverfahren

Die simulierte Studie basiert auf 1.000 Stichproben mit jeweils 50 bzw. 150 Aktien. In einem ersten Schritt werden für die einzelnen Untersuchungswertpapiere und Vergleichsportefeuilles bzw. -wertpapiere über einen Zeitraum von 36 Monaten Buy-and-hold-Renditen berechnet, in einem zweiten Schritt durch Differenzbildung die Überrenditen bestimmt und innerhalb einer Stichprobe zu einem arithmetischen Mittel aggregiert.

Zur Überprüfung der Nullhypothese werden die Teststatistiken bei Verwendung der Portefeuillerenditen mit dem schiefeangepaßten t-Test, bei Verwendung von einem nach der Marktkapitalisierung vergleichbaren Wertpapier mit dem t-Test berechnet. Die Abweichungen der empirisch beobachteten von den theoretisch erwarteten Ablehnungshäufigkeiten werden mit dem Binomialtest bei einem vorgegebenen Signifikanzniveau von 1 % überprüft. Signifikante Abweichungen werden mit * gekennzeichnet.

| Signifikanzniveau (%) | Ablehnungshäufigkeiten in % | | | | | |
|--|-----------------------------|------|------|------|-------|-------|
| | 1 | | 5 | | 10 | |
| kumulierte Dichtefunktion (%) | 0,5 | 99,5 | 2,5 | 97,5 | 5 | 95 |
| Teil A: 1.000 Stichproben mit 50 Aktien | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 4,1* | 0,0* | 9,2* | 0,7* | 12,5* | 1,8* |
| marktwertgewichtetes Portefeuille | 1,6* | 0,5 | 2,8 | 4,2* | 4,5 | 8,3* |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 3,0* | 0,1* | 6,6* | 1,3* | 11,0* | 3,0* |
| marktwertgewichtetes Size-Portefeuille | 3,2* | 0,1* | 6,6* | 1,1* | 10,5* | 3,0* |
| Vergleichswertpapier | 0,4 | 0,3 | 2,3 | 1,8 | 5,2 | 4,0 |
| Teil B: 1.000 Stichproben mit 150 Aktien | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 4,0* | 0,1* | 9,6* | 0,7* | 15,5* | 1,4* |
| marktwertgewichtetes Portefeuille | 0,5 | 2,1* | 1,6 | 9,3* | 2,7* | 17,0* |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 3,0* | 0,0* | 7,5* | 1,4* | 11,5* | 3,6 |
| marktwertgewichtetes Size-Portefeuille | 2,7* | 0,0* | 6,9* | 1,8 | 10,4* | 4,2 |
| Vergleichswertpapier | 0,6 | 0,0* | 2,2 | 1,5* | 5,3 | 3,9 |

Nach den vorliegenden Untersuchungsergebnissen entstehen bei der Schätzung langfristiger Renditen unter Verwendung von Vergleichsportefeuilles systematische Verzerrungen bei der Überprüfung der Nullhypothese, wenn parametrische Testverfahren eingesetzt werden. Eine Möglichkeit zur Vermeidung von Fehlern erster bzw. zweiter Art bei den Signifikanztests ist die Simulation einer empirischen Verteilung und einer darauf basierenden Schätzung kritischer p-Werte zur Ablehnung der Nullhypothese.

Dazu wird in einem ersten Schritt jedem Untersuchungswertpapier ein anderes, zufällig ausgewähltes Wertpapier zugeordnet, welches zu dem gleichen Size-Portefeuille gehört. In einem zweiten Schritt werden die langfristigen Überrenditen für einen Zeitraum von 36 Monaten berechnet. Als Schätzwert für die langfristige Rendite werden die vier in der Tabelle 3 ange-

gebenen Portefeuilles verwendet. In einem dritten Schritt wird für jede Stichprobe ein schiefeangepaßter t-Wert berechnet. Auf der Grundlage der Percentile der schiefeangepaßten t-Wert-Verteilung werden in einem vierten Schritt die kritischen Werte zur Ablehnung der Nullhypothese geschätzt. Die Ablehnungshäufigkeiten auf der Grundlage der empirischen Verteilung sind in der nachfolgenden Tabelle angegeben.

Tabelle 4

Ablehnungshäufigkeiten für die Nullhypothese unter Verwendung nichtparametrischer Testverfahren

Die simulierte Studie basiert auf 1.000 Stichproben mit jeweils 50 bzw. 150 Aktien. In einem ersten Schritt werden für die einzelnen Untersuchungswertpapiere und Vergleichsportefeuilles bzw. -wertpapiere über einen Zeitraum von 36 Monaten Buy-and-hold-Renditen berechnet, in einem zweiten Schritt durch Differenzbildung die Überrenditen bestimmt und innerhalb einer Stichprobe zu einem arithmetischen Mittel aggregiert. Die kritischen Werte zur Ablehnung der Nullhypothese werden auf der Grundlage der schiefeangepaßten t-Wert-Verteilung bestimmt. Zur Simulation der Verteilung wird jedem Wertpapier eine andere Aktie aus dem gleichen Size-Portefeuille zugeordnet und die langfristige Überrendite auf identische Weise berechnet. Die Abweichungen der empirisch beobachteten von den theoretisch erwarteten Ablehnungshäufigkeiten werden mit dem Binomialtest bei einem vorgegebenen Signifikanzniveau von 1 % überprüft. Signifikante Abweichungen werden mit * gekennzeichnet.

| Signifikanzniveau (%) kumulierte Dichtefunktion (%) | Ablehnungshäufigkeiten in % | | | | | |
|--|-----------------------------|------|-----|------|-----|------|
| | 1 | | 5 | | 10 | |
| | 0,5 | 99,5 | 2,5 | 97,5 | 5 | 95 |
| Teil A: 1.000 Stichproben mit 50 Aktien | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 0,8 | 0,3 | 3,0 | 1,9 | 5,4 | 5,1 |
| marktwertgewichtetes Portefeuille | 1,0 | 0,5 | 2,7 | 3,0 | 4,7 | 4,2 |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 0,4 | 0,1* | 2,7 | 2,0 | 4,7 | 5,2 |
| marktwertgewichtetes Size-Portefeuille | 0,4 | 0,1* | 2,8 | 1,8 | 4,8 | 5,0 |
| Teil B: 1.000 Stichproben mit 150 Aktien | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 0,9 | 0,9 | 2,3 | 3,5 | 6,0 | 6,2 |
| marktwertgewichtetes Portefeuille | 0,6 | 0,7 | 1,9 | 3,6 | 5,0 | 7,6* |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 0,8 | 0,5 | 3,0 | 4,3* | 6,0 | 7,0* |
| marktwertgewichtetes Size-Portefeuille | 0,8 | 0,4 | 3,1 | 4,2* | 6,1 | 6,7 |

Die Testergebnisse auf der Grundlage der empirischen Verteilung zeigen nur in wenigen Fällen statistisch signifikante Abweichungen zwischen den empirisch beobachteten und den theoretisch erwarteten Ablehnungshäufigkeiten. Zudem sind keine systematisch unterschiedlichen Ablehnungshäufigkeiten für positive und negative langfristige Überrenditen erkennbar.

Für die Durchführung langfristiger Ereignisstudien kann man aus den vorliegenden Untersuchungsergebnissen die Empfehlung abzuleiten, die langfristigen, auf der Grundlage von Portefeuillerenditen berechneten durchschnittlichen Überrenditen mit nichtparametrischen Testverfahren zu analysieren. Die kritischen Werte zur Ablehnung der Nullhypothese bei einer vorgegebenen Irrtumswahrscheinlichkeit können auf Basis von empirischen Verteilungen hinreichend genau geschätzt werden. Die Untersuchungsergebnisse zeigen, daß bei diesen Verfahren die empirisch beobachteten Ablehnungshäufigkeiten für die Nullhypothese sehr nahe bei den theoretisch erwarteten Werten liegen.

4.2 Untersuchungen bei Vorliegen eines kalenderzeitlichen Clusterings

Für die Simulation des Kalenderclustering wird eine Dreiecksverteilung (Simpson-Verteilung) mit einer Spannbreite von 73 Monaten (Monat 0 ± 36 Monate) gewählt, deren Mittelpunkt (Scheitelpunkt) über ein Intervall von Januar 1963 bis Dezember 1989 geht. Es werden 1.000 gleichverteilte Ziehungen der Mittelpunkte mit jeweils 50 Ziehungen unter der Dreiecksverteilung vorgenommen.

Tabelle 5

Mittelwert, Standardabweichung und Schiefe der Überrenditenverteilung einer simulierten langfristigen Ereignisstudie bei Vorliegen eines kalenderzeitlichen Clusterings

Die simulierte Studie basiert auf 1.000 Stichproben mit jeweils 50 Aktien. In einem ersten Schritt werden für die einzelnen Untersuchungswertpapiere und Vergleichsportefeuilles bzw. -wertpapiere über einen Zeitraum von 36 Monaten Buy-and-hold-Renditen berechnet, in einem zweiten Schritt durch Differenzbildung die Überrenditen bestimmt und innerhalb einer Stichprobe zu einem arithmetischen Mittel aggregiert. Mit dem t-Test für abhängige Stichproben wird überprüft, ob sich der Mittelwert der Stichprobenmittelwert-Verteilung signifikant von null unterscheidet.

| Schätzwert für die langfristige Rendite | Stichprobenmittelwert-Verteilung | | | Durchschnittliche Schiefe der Stichprobenverteilung |
|--|----------------------------------|--------|------------------------|---|
| | Mittelwert (%) | t-Wert | Standardabweichung (%) | |
| Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien | | | | |
| ungewichtetes Portefeuille | -1,40 | -3,29 | 13,51 | 1,75 |
| marktwertgewichtetes Portefeuille | 5,20 | 8,69 | 18,90 | 1,65 |
| Berücksichtigung des Size-Effektes | | | | |
| ungewichtetes Size-Portefeuille | -0,56 | -1,34 | 13,11 | 1,54 |
| marktwertgewichtetes Size-Portefeuille | -0,19 | -0,46 | 12,89 | 1,57 |
| Vergleichswertpapier | -0,95 | -1,91 | 15,74 | -0,03 |

Im Unterschied zu den Simulationsergebnissen ohne Vorliegen eines kalenderzeitlichen Clusterings zeigen die Ergebnisse der Tabelle 5 eine etwas höhere Standardabweichung der Stichprobenmittelwerte. Besonders auffällig sind die Abweichungen, wenn die Rendite des marktwertgewichteten Portefeuilles aus allen im amtlichen Handel Frankfurt notierten Aktien als Schätzwert für die langfristige Rendite der Untersuchungswertpapiere verwendet wird.

Tabelle 6

Ablehnungshäufigkeiten für die Nullhypothese bei Vorliegen eines kalenderzeitlichen Clusterings

Die simulierte Studie basiert auf 1.000 Stichproben mit jeweils 50 Aktien. In einem ersten Schritt werden für die einzelnen Untersuchungswertpapiere und Vergleichsportefeuilles bzw. -wertpapiere über einen Zeitraum von 36 Monaten Buy-and-hold-Renditen berechnet, in einem zweiten Schritt durch Differenzbildung die Überrenditen bestimmt und innerhalb einer Stichprobe zu einem arithmetischen Mittel aggregiert. Die kritischen Werte zur Ablehnung der Nullhypothese werden auf der Grundlage der schiefeangepaßten t-Wert-Verteilung bestimmt. Zur Simulation der Verteilung wird jedem Wertpapier eine andere Aktie aus dem gleichen Size-Portefeuille zugeordnet und die langfristige Überrendite auf identische Weise berechnet. Die Abweichungen der empirisch beobachteten von den theoretisch erwarteten Ablehnungshäufigkeiten werden mit dem Binomialtest bei einem vorgegebenen Signifikanzniveau von 1 % überprüft. Signifikante Abweichungen werden mit * gekennzeichnet. Das Kalenderclustering wird über eine Dreiecksverteilung mit einer Spannbreite von 73 Monaten (Monat 0 ± 36 Monate) modelliert. Es erfolgen 1.000 gleichverteilte Ziehungen der Mittelpunkte der Dreiecksverteilung mit jeweils 50 Ziehungen unter der Dreiecksverteilung.

| Signifikanzniveau (%) kumulierte Dichtefunktion (%) | Ablehnungshäufigkeiten in % | | | | | |
|--|-----------------------------|-------|-------|-------|-------|-------|
| | 1 | | 5 | | 10 | |
| | 0,5 | 99,5 | 2,5 | 97,5 | 5 | 95 |
| Teil A: Parametrische Teststatistiken (t-Test bzw. schiefeangepaßter t-Test) | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 7,0* | 0,1* | 13,4* | 2,1 | 17,6* | 4,2 |
| marktwertgewichtetes Portefeuille | 5,3* | 15,7* | 9,5* | 23,3* | 13,5* | 29,0* |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 6,5* | 0,3 | 12,9* | 2,4 | 16,9* | 5,5 |
| marktwertgewichtetes Size-Portefeuille | 5,5* | 0,4 | 11,6* | 2,2 | 15,3* | 5,4 |
| Vergleichswertpapier | 0,4 | 0,3 | 3,0 | 2,2 | 5,7 | 4,4 |
| Teil B: Nichtparametrische Teststatistiken auf Basis empirischer Verteilungen | | | | | | |
| <i>Portefeuilles aus allen in Frankfurt amtlich gehandelten Aktien</i> | | | | | | |
| ungewichtetes Portefeuille | 0,0* | 0,7 | 2,2 | 3,0 | 5,3 | 5,4 |
| marktwertgewichtetes Portefeuille | 0,2 | 1,1 | 1,9 | 2,2 | 4,4 | 4,6 |
| <i>Berücksichtigung des Size-Effektes</i> | | | | | | |
| ungewichtetes Size-Portefeuille | 0,1* | 0,5 | 1,3* | 2,5 | 5,3 | 5,1 |
| marktwertgewichtetes Size-Portefeuille | 0,1* | 0,6 | 1,9 | 2,5 | 5,1 | 5,0 |

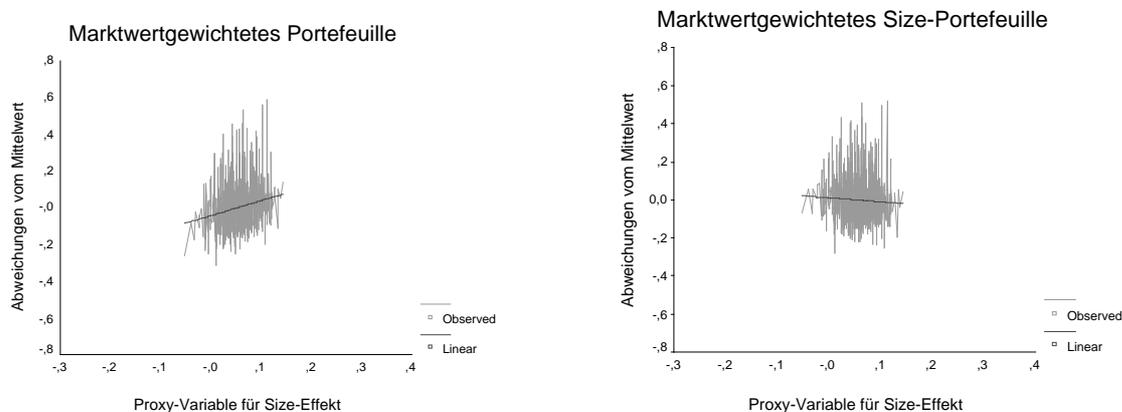
Die Abweichungen der empirisch beobachteten von den theoretisch erwarteten Ablehnungshäufigkeiten nehmen bei einer zeitlichen Konzentration der Ereignisse innerhalb einer Stichprobe zu, wenn Portefeuillerenditen als Schätzwerte der langfristigen Rendite der Untersuchungswertpapiere und der schiefeangepaßte t-Test zur Überprüfung der Nullhypothese Verwendung finden. Die eingesetzten Verfahren auf der Grundlage von Vergleichswertpapieren oder die Bestimmung der Ablehnungshäufigkeiten für die Nullhypothese auf Basis empirischer Verteilungen führen auch hier zu den genaueren Ergebnissen.

Abbildung 1 illustriert, daß die der Schätzung der langfristigen Rendite der Untersuchungswertpapiere zugrundeliegenden Vergleichsportefeuilles in einem unterschiedlichen Umfang den Einfluß des Size-Effektes widerspiegeln. Zur Darstellung des Zusammenhangs wird eine Proxy-Variable für die Auswirkungen des Size-Effektes auf die Höhe der langfristigen Überrenditen benötigt. Im Abschnitt 4.1 wird festgestellt, daß die auf der Grundlage eines gleich- oder marktwertgewichteten Portefeuilles geschätzten Erwartungswerte für die Renditen der Untersuchungswertpapiere zu systematischen Verzerrungen der Ergebnisse führen. Die unterschiedlichen Vorzeichen der im Durchschnitt beobachteten Abweichungen deuten auf einen Erklärungsgehalt des Size-Effektes hin. Als Proxy-Variable für den Size-Effekt wird deshalb auf der Abszisse für jede Stichprobe die durchschnittliche Differenz der auf der Grundlage der genannten Portefeuilles geschätzten Erwartungswerte abgetragen. Als abhängige Variable wird die Differenz des einzelnen Stichprobenmittelwertes vom Durchschnittswert aller 1.000 Stichprobenmittelwerte verwendet. Die langfristige Überrendite der Untersuchungswertpapiere werden auf Basis des marktwertgewichteten Portefeuilles aller in Frankfurt amtlich gehandelten Aktien bzw. auf Basis des entsprechenden marktwertgewichteten Size-Portefeuilles berechnet. Die Regressionsgleichung lautet:

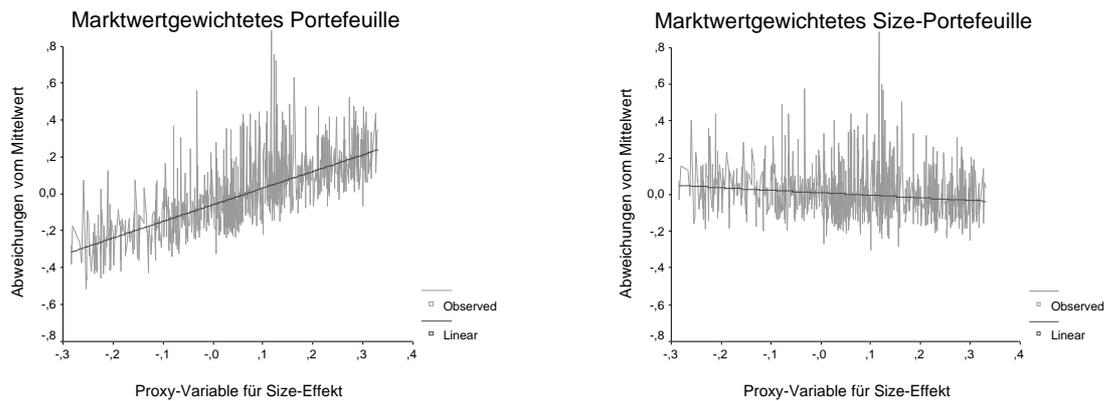
$$CAR_{s,T=36}^{BHAR} - \left(\frac{1}{S} \sum_{s=1}^S CAR_{s,T=36}^{BHAR} \right) = \alpha_s + \beta_s \cdot \left(\frac{1}{N} \sum_{i=1}^N \left[\prod_{t=1}^{36} (R_{s,P(ug),t} + 1) - \prod_{t=1}^{36} (R_{s,P(mg),t} + 1) \right] \right) + \varepsilon_s \quad (7)$$

$R_{s,P(ug/mg),t}$ = erwartete Rendite für das Untersuchungswertpapier i aus der Stichprobe s zum Zeitpunkt t, welche auf der Grundlage eines gleich- (ug) bzw. marktwertgewichteten (mg) Portefeuilles aller im amtlichen Handel Frankfurt notierten Aktien geschätzt wird.

Abbildung 1
Streuung der Stichprobenmittelwerte ohne Kalenderclustering



Streuung der Stichprobenmittelwerte mit Kalenderclustering



Aus der Abbildung 1 ist ein positiver Anstieg der Regressionsgeraden und somit eine systematische Beeinflussung des Vorzeichens des berechneten Durchschnittswertes der langfristigen Überrendite erkennbar, wenn die Renditen der untersuchten Wertpapiere mit der Wertentwicklung aller im amtlichen Handel Frankfurt notierten Aktien verglichen werden. Bei einem Vergleich mit den Renditen von Size-Portefeuilles werden zeitliche Instabilitäten des Size-Effektes auch bei Vorliegen einer Häufung der Ereignisse zu bestimmten Zeitpunkten weitestgehend ausgeglichen. Die Ergebnisse zeigen ferner, daß bei Vorliegen eines kalenderzeitlichen Clusterings der untersuchten Ereignisse die beobachteten Verzerrungen der Ergebnisse verstärkt werden, wenn bei der Schätzung der erwarteten Rendite der Size-Effekt nicht oder nur approximativ berücksichtigt wird.

5. Schlußfolgerungen

Die unter Verwendung von Aktienrenditen deutscher Unternehmen durchgeführte Simulationsstudie bestätigt eine systematische Verzerrung der Durchschnittswerte langfristiger Überrenditen und der dazugehörigen Teststatistiken, wenn die bisher in der Literatur verwendeten „Standardverfahren“ (Aktienindex, t-Test) zur Anwendung kommen. Die deutlichen Verzerrungen der Ergebnisse bei Verwendung eines gleich- oder marktwertgewichteten Portefeuilles aller amtlich in Frankfurt notierten Aktien führen zu der Schlußfolgerung, daß die verfügbaren Aktienindizes (DAX, DAFOX etc.) in den meisten Fällen nicht als Schätzwert für die langfristige Rendite der Untersuchungswertpapiere nach bewertungsrelevanten Ereignissen geeignet sind.

Die Durchführung langfristiger Ereignisstudien erfordert die Bildung von Vergleichsportefeuilles oder die Auswahl von Vergleichswertpapieren nach jenen empirisch identifizierten Bestimmungsfaktoren, die einen Beitrag zur Erklärung von Renditeunterschieden und somit zur Preisbildung auf den Aktienmärkten leisten (Size und/oder Buchwert/Marktwert-Verhältnis).

Werden die Renditen von Vergleichsportefeuilles als Schätzwerte für die Renditen der Untersuchungswertpapiere verwendet, dann ergeben sich aufgrund der Schiefe der Überrenditenverteilung verzerrte Teststatistiken. Eine solche Vorgehensweise erfordert eine Schätzung der kritischen Werte zur Ablehnung der Nullhypothese auf der Grundlage einer empirischen Verteilung. Bei der alternativen Verwendung der Renditen von einzelnen Vergleichswertpapieren entsprechen die empirisch beobachteten den theoretisch erwarteten Ablehnungshäufigkeiten für die Nullhypothese. Empirische Verteilungen der Teststatistiken werden folglich nicht benötigt. Der Nachteil dieses Verfahrens ist in der höheren Standardabweichung der Stichprobenmittelwerte zu sehen. Positive und negative Abweichungen in Höhe mehrerer Prozentpunkte vom „wahren“ Durchschnittswert können insbesondere bei kleineren Stichproben häufiger auftreten und zu fehlerhaften Interpretationen der Untersuchungsergebnisse führen. Ein weiterer Nachteil ist, daß die berechneten Überrenditen für die einzelnen Untersuchungswertpapiere nicht in Querschnittsuntersuchungen verwendet werden können, da ein Einfluß unternehmensspezifischer Ereignisse der Vergleichswertpapiere gegeben ist.

In der vorliegenden Untersuchung führt eine modellierte zeitliche Häufung der Ereignisse zu einer Erhöhung der Standardabweichung der Stichprobenmittelwerte. Zusätzlich ist eine Beeinflussung des Vorzeichens des Stichprobenmittelwertes gegeben, wenn der Size-Effekt bei der Ermittlung eines Schätzwertes für die langfristige Rendite der Untersuchungswertpapiere nicht explizit berücksichtigt wird.

Treten Ereignisse in bestimmten Zeitperioden häufiger auf, werden kurze Betrachtungszeiträume gewählt oder kann nur eine geringe Anzahl von Beobachtungen in die Untersuchung einbezogen werden, so sind aus den vorliegenden Simulationsergebnissen keine Hinweise erkennbar, daß sich Veränderungen für die Empfehlung der zu verwendenden Untersuchungsmethodik ergeben. Vielmehr ist bei der Anwendung der „Standardverfahren“ mit stärker verzerrten Mittelwerten und Teststatistiken zu rechnen. Die Erhöhung der Standardabweichung der Stichprobenmittelwerte sollte zum Anlaß genommen werden, Stabilitätsbetrachtungen für verschiedene Teilperioden, nach Marktkapitalisierung, Buchwert/Marktwert-Verhältnis und anderen Kriterien durchzuführen.

Literatur

Andrade, G. ; Stafford, E. (1997): Investigating the characteristics and determinants of mergers and other forms of investment, Working paper, University of Chicago.

Barber, B.M.; Lyon, J.D. (1997): Detecting long-run abnormal stock returns: The empirical power and specification of tests statistics, *Journal of Financial Economics* 43, 341-372.

Brav, A. (1997): Inference in long-horizon event studies: a re-evaluation of the evidence. Working paper, University of Chicago.

Brav, A.; Geczy, C.; Gompers, P.A. (1999): Is the abnormal return following equity issuances anomalous, Working paper, University of Pennsylvania.

- Brav, A.; Gompers, P.A. (1997): Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies, *Journal of Finance* 52, 1791-1821.
- Brown, S.J.; Warner, J.B. (1980): Measuring security price performance, *Journal of Financial Economics* 8, 205-258.
- Canina, L.; Michaely, R.; Thaler, R.; Womack, K. (1998): Carveat compounder: a warning about using the daily CRSP equal-weighted index to compute long-run excess returns, *Journal of Finance* 53, 403-416.
- Carhart, M. (1997): On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.
- Carter, R.B.; Dark, F.H.; Singh, A.K. (1998): Underwriter reputation, initial returns, and the long-run performance of ipo stocks, *Journal of Finance* 53, 285-311.
- Cowan, A.R.; Sergeant, A.M.A. (1997): Interacting biases, non-normal return distributions and the performance of parametric and bootstrap tests for long-horizon event studies, Working paper, Iowa State University.
- Dimson, E.; Marsh, P. (1986): Event study methodologies and size effect: The case of UK press recommendations, *Journal of Financial Economics* 17, 113-142.
- Efron, B. (1982): *The Jackknife, the bootstrap and other resampling plans*, Pennsylvania.
- Ehrhardt, O. (1997): *Börseneinführungen von Aktien am deutschen Kapitalmarkt*, Wiesbaden.
- Entrup, U. (1995): *Kapitalmarktreaktionen auf Optionsanleihen*, Wiesbaden.
- Fama, E.F. (1998): Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283-306.
- Fama, E.F. (1991): Efficient capital market: II, *Journal of Finance* 46, 1575-1617.
- Fama, E.F.; Fisher, L.; Jensen, M.C.; Roll, R. (1969): The adjustment of stock prices to new information, *International Economic Review* 10, 1-21.
- Fama, E.; French, K. (1993): Common risk factors in the returns of stocks and bonds, *Journal of Financial Economics* 33, 3-55.
- Gebhardt, G.; Entrup, U.; Heiden, S. (1994): Kursreaktionen auf Kapitalerhöhungen aus Gesellschaftsmitteln, *Zeitschrift für Bankrecht und Bankwirtschaft* 6, 308-332.
- Gerke, W.; Oerke, M.; Sentner, A. (1997): Der Informationsgehalt von Dividendenänderungen auf dem deutschen Aktienmarkt, *Die Betriebswirtschaft* 57, 810-822.
- Hall, P. (1992): On the removal of skewness by transformation, *Journal of the Royal Statistical Society* 54, 221-228.
- Ikenberry, D.; Lakonishok, J., Vermaelen, T. (1995): Market underreaction to open market shares repurchases, *Journal of Financial Economics* 39, 181-208.
- Johnson, N.J. (1978): Modified t tests and confidence intervals for asymmetrical populations, *Journal of the American Statistical Association* 73, 536-544.
- Kaserer, C.; Brunner, A. (1997): Die Signaleffekte von Kapitalerhöhungen aus Gesellschaftsmitteln – Eine empirische Untersuchung für den deutschen Markt, *Finanzmarkt und Portfolio Management* 11, 77-99.
- Kothari, S.P.; Shanken, J.; Sloan, R.G. (1995): Another look at the cross-section of expected returns, *Journal of Finance* 50, 185-224.

- Kothari, S.P.; Warner, J.B. (1997): Measuring long-horizon security price performance, *Journal of Financial Economics* 43, 301-339.
- Ljungqvist, A.P. (1997): Pricing initial public offerings: Further evidence from Germany, *European Economic Review* 41, 1309-1320.
- Loughran, T.; Ritter, J.R. (1995): The new issues puzzle, *Journal of Finance* 50, 23-51.
- Lyon, J.D.; Barber, B.M.; Tsai, C.-L. (1999): Improved methods for tests of long-run abnormal stock returns, *Journal of Finance* 54, forthcoming.
- Marsaglia, G.; Zaman, A. (1991): A new class of random number generators, *Annals of Applied Probability* 1, 462-480.
- Michaelis, R.; Thaler, R.H.; Womack, K.L. (1995): Price reactions to dividend initiations and omissions, *Journal of Finance* 50, 573-608.
- Mitchell, M.L.; Stafford, E. (1997): Managerial decisions and long-term stock price performance, Working paper, University of Chicago.
- Oertmann, P. (1994): Firm-Size-Effekt am deutschen Aktienmarkt, *Zeitschrift für betriebswirtschaftliche Forschung* 14, 229-259.
- Ritter, J.R. (1991): The long-term performance of initial public offerings, *Journal of Finance* 46, 3-27.
- Ritter, J.R. (1984): The "hot issue" market of 1980, *Journal of Business* 57, 215-240.
- Sapusek, A. (1998): Empirical evidence on the long-run performance of initial public offerings in Germany, *Banque & Marchés* 34, 38-45.
- Shleifer, A.; Vishny, R.W. (1997): A Survey of Corporate Governance, *Journal of Finance* 52, 737-783.
- Shumway, T. (1997): The delisting bias in CRSP Data, *Journal of Finance* 52, 327-340.
- Stehle, R. (1997): Der Size-Effekt am deutschen Aktienmarkt, *Zeitschrift für Bankrecht und Bankwirtschaft* 9, 237-260.
- Stehle, R.; Ehrhardt, O. (1999): Renditen bei Börseneinführungen am deutschen Kapitalmarkt, *Zeitschrift für Betriebswirtschaft*, forthcoming.
- Stehle, R.; Ehrhardt, O.; Przyborowsky, R. (1999): Long-run stock performance after initial public offerings and seasoned equity issues in the German capital market, Working paper, Humboldt-Universität zu Berlin.
- Stehle, R.; Hartmond, A. (1991): Durchschnittsrenditen deutscher Aktien 1954-1988, *Kredit und Kapital* 24, 371-411.
- Stehle, R.; Sattler, R.; Wulff, C. (1997): Der Size-Effekt am US-amerikanischen Aktienmarkt, Working paper, Humboldt-Universität zu Berlin.
- Sutton, C.D. (1993): Computer-intensive methods for tests about the mean of an asymmetrical distribution, *Journal of the American Statistical Association* 88, 802-808.
- Teoh, S.H.; Welch, I.; Wong, T.J. (1998): Earnings management and the long-run market performance of seasoned public offerings, *Journal of Finance* 50, 63-100.
- Teoh, S.H.; Welch, I.; Wong, T.J. (1999): Earnings management and the long-run market performance of initial public offerings, *Journal of Finance*, forthcoming.
- Wolfram, S. (1997): *Das Mathematica Buch*, 3. Aufl., Bonn.

Wolfram, S. (1986): Random sequence generation by cellular automata, *Advance in Applied Mathematics* 7, 123-169.

Zusammenfassung

Gegenwärtig werden die Ergebnisse langfristiger Ereignisstudien intensiv diskutiert. Dabei stehen die Unterschiede in den berechneten Durchschnittswerten, methodische Fragen der Berechnung langfristiger Überrenditen und die Diskussion der Hypothese informationseffizienter Kapitalmärkte im Mittelpunkt der Betrachtung. Unsere Simulationsstudie überprüft die Eignung der Berechnungsverfahren für kleine Stichproben und bei einer sich im Zeitablauf verändernden Anzahl von Beobachtungen, wenn ein Size-Effekt existiert.

Die wichtigsten Ergebnisse sind: (1) Bei der Schätzung der langfristigen Rendite der Untersuchungswertpapiere sind die von Fama/French (1993) empirisch identifizierten Bestimmungsfaktoren zur Erklärung von Aktienrenditen zu berücksichtigen. Die Verwendung von Aktienindizes dürfte in den meisten Fällen zu deutlichen Verzerrungen der Durchschnittswerte und Teststatistiken führen. (2) Werden bei der Berechnung der Überrenditen Vergleichsportefeuilles verwendet, so sind aufgrund der Schiefe der Überrenditenverteilung die kritischen Werte zur Ablehnung der Nullhypothese auf der Basis einer empirischen Verteilung zu schätzen. (3) Bei der Verwendung von Vergleichswertpapieren sind keine systematischen Verzerrungen der Berechnungsergebnisse zu erkennen. Jedoch ist die Standardabweichung der Stichprobenmittelwerte höher als bei Verwendung von Vergleichsportefeuilles. Dies gilt insbesondere für kleine Stichproben.

Summary

Currently, the results of long-run event studies are being discussed deeply. The focus is on the differences of the long-run performance means, the methods to calculate the abnormal returns and the discussion of the market efficiency hypothesis. In our study, we use simulations to evaluate the validity of the procedures for small sample size and different numbers of the events in several time periods, when there exists a strong size effect.

Our study shows the following results: First, the estimation of the abnormal returns should include the same return pattern as reported by Fama/French (1993). Market-value weighted stock indices are not ideal benchmarks in long-run event studies. Second, using control portfolios as a benchmark results in a skewed distribution of the individual long-run abnormal returns. For that reason, we compute the empirical p-value to get well-specified test statistics. Third, we do not find any misspecification of the mean of abnormal returns and test statistics by using the control firm approach. But the standard deviation of the one-thousand sample means from the overall mean is larger for control stocks than for control portfolios when the sample size is small.

